Excerpts from

PROCESSES IN BIOLOGICAL VISION:

including,

ELECTROCHEMISTRY OF THE NEURON

This material is excerpted from the full β-version of the text. The final printed version will be more concise due to further editing and economical constraints. A Table of Contents and an index are located at the end of this paper.

James T. Fulton
Vision Concepts
jtfulton@neuronresearch.net

April 30, 2017

Copyright 2000
James T. Fulton

11 Introduction to modeling of biological phenomena [1]

Don't believe everything you think. *Anonymous bumper sticker*

It is the models that tell us whether we understand a process and where the uncertainties remain.

Bridgeman, 2000

11.1 Background

In presenting the second quotation, Bridgeman characterized the current state of research into the vision process. He said, "The neuroanatomical analyses can tell us **where** to look for a particular mechanism, and neurophysiology can tell us **when** the critical information is being processed, but current techniques leave the '**how**' of the process tantalizingly elusive."[2] The first quotation highlights the difficulties of relying upon intuition, or what is sometimes described as introspection. The goal of this Chapter is to provide a model that can support a wide range of experimental results and lead the way to a broader understanding of the visual system.

There are a number of problems associated with modeling the signal processing functions in vision. One of the biggest is the plethora of simplistic circuit diagrams found in the literature. The second is the extremely high impedance level found in the initial stages of signal processing. These impedance levels are to be expected; they relate to the extremely small size of the elements and they reduce overall power consumption. A third is the reluctance of the biology community (at least until recently) to embrace the fact that there are synaptic junctions that involve only the movement of electrons across the boundary. No movement of chemical ions (much less complex organic molecules with the complexity of a protein) is involved. And finally, the fact that 99.9% of the neurons in the retina operate in the analog (not in the binary pulse) mode puts them in a class separate from much of the vision literature.

It is noteworthy that the field of morphology is involved primarily with static conditions or very slowly changing conditions relative to growth. The field has largely overlooked the subject of neural signaling. One neurobiology text, that is heavily weighted toward morphology lists 14 types of cell-to-cell communications in hierarchal order but does not mention neural signaling based on electrolytics[3]. It is also noteworthy that the literature contains no detailed description of the signals at different positions along the signal path from the input voltage at a neurite to the output voltage at the axon of a neuron.

Some of the material in this Chapter is designed for the actual laboratory researcher or serious analyst. It goes into more depth than required by the typical reader. The chapter is focussed on the modeling of the larger parts of the visual system. The modeling of an individual Activa, an individual synapses and various individual circuits within a single neuron are addressed in **Chapters 8 & 9**. **Section 11.1** can be read selectively as needed by the reader. **Sections 11.2** through **11.8** encompass the various phases of the detailed modeling of the visuals system.

Section 11.1.1 is critically important if one is to avoid the many conceptual traps found in the current literature.
Section 11.1.2 provides many detailed definitions that are required to insure consistency and correlation of the material in adjacent chapters.
Section 11.1.3 concentrates on the high level architecture of the overall neural system, using the non-visual system as a template for understanding the visual system..
Section 11.1.4 provides an internally consistent set of morphological nomenclature for the neural system and
Section 11.1.5 does the same thing for the electrical waveforms normally associated with the neural system.
Section 11.1.6 provides a review of graphical techniques of specific value in modeling the visual system.

---

[1]Released: April 30, 2017
[2]Bridgeman, B. (2000) Neuroanatomy and function in two visual systems. *Behav. Brain Sci*. vol. 23, no. 4, pp 535-536
[3]Shepherd, G. (1994) Neurobiology, 5th ed. NY: Oxford Press. pg. 72 in the 1988 edition

2 Processes in Biological Vision

Section 11.1.7 discusses the subtle differences between data and information rates.
Section 11.1.8 reviews some of the models used in other presentations on the system aspects of the visual process.

This chapter must take note of a number of papers that should be associated with signal processing and signal manipulation but rely exclusively on a mathematical model. Cornsweet & Yellott have provided a paper of this type[4]. While the paper is extensive, it relies upon a concept believed to be totally foreign to the physiology of vision and not found within the realm of conventional optics. They propose an optical spread function whose spatial width is light intensity dependent. Their derivation leads to several comments related to Weber's Law and Mach Bands but no physically realizable physiology. Such exclusively mathematical models will not be addressed in this work.

11.1.1 Clarification of concepts

Before embarking on modeling of biological systems, there are a variety of major concepts that must be discussed before proceeding to the more detailed subject of terminology and conventions. The analysis of biological systems follows the same methodology as for any other "system," whether chemical, mechanical, electrical, hydraulic, etc. In fact, the typical biological system employs elements from each of the above disciplines. The rules of system analysis are well established and precisely documented. However, their development is beyond the scope of this work. For the non-mathematically inclined, Regan has provided a list of principles related to system analysis and a number of useful examples[5]. While tending to bound the field and be very useful to the layman, his seventeen postulates should not be taken as mathematically rigorous. As he states, there are a variety of caveats applying to some of the postulates. For a stronger foundation in the subject of system analysis applied to psychophysics, and a definition of the various modes of psychophysics testing, Green & Swets is an excellent choice[6].

11.1.1.1 Models and Diagrams to aid understanding

To help interpret this work and the literature, it is useful to explicitly define a series of different types of models. The most basic classes of models are the abstract and the functional.

+ **Abstract models** typically present a conceptual block diagram or contain a mathematical equation. The conceptual block diagram is frequently found in pedagogical situations. The mathematical equation is usually intended to bound the performance of some process or mechanism in vision without relating directly to they underlying mechanisms or processes. Both tend to be misleading with respect to the underlying electrophysiology of the system.

+ **Functional models** typically attempt to describe the actual mechanisms and processes of vision and must conform to the electrophysiology of the system to be considered correct.

Computer emulations normally conform to the definition of the abstract model because the algorithms used seldom relate to the underlying electrophysical mechanisms and processes.

Equivalent circuits on the other hand, and especially if sophisticated, attempt to represent the electrophysiological situation using an equivalent topology in a different technology.

Specialists in system synthesis and analysis have long used a series of block, schematic and circuit diagrams to document the configuration and operation of a wide variety of systems. These tools attempt to represent the actual circuit topology of the visual system. The diagrams are used to aid in seeing both the forests and the trees simultaneously. Simultaneously, they provide a medium for documenting systemic relationships. Such systems may be of many types. They include electrical, hydraulic, mechanical, aerodynamic and *biological* types. Although not used widely or consistently in the visual field, an attempt will be made to use them consistently here according to the following hierarchy;

+ Global block diagram

---

[4]Cornsweet, T. Yellott, J. (1985) Intensity-dependent spatial summation *J Opt Soc Am A* vol 2(10), pp 1769-1785
[5]Regan, D. (1991) Spatial Vision. Vol. 10 of Vision and Visual Dysfunction, Cronly-Dillon, J. *general ed.* Boca Raton, FL: CRC Press pp 3-12
[6]Green, D. & Swets, J. (1974) Signal Detection Theory and Psychophysics. Huntington, NY: Krieger Publishing

+ Species specific block diagrams
+ Schematic diagrams
+ Multiple circuit diagrams
+ An individual circuit diagram

A block diagram is usually used to indicate fundamental but general relationships between major elements of the system. They are useful for orienting the reader and defining terms. They usually lack detail and are not indicative of the topology of the system. The schematic diagram is intended to illustrate relationships between the elements of a block diagram. The titles of the elements are frequently chosen to indicate the top level topology of the system and to indicate both feed forward and feedback paths. At the next level of detail are the circuit diagrams. These frequently occur in two classes, higher level circuit diagrams illustrating the interplay of several circuits and the individual circuit diagrams. The higher level circuit diagrams frequently employ high level symbolic representations of individual circuits to simplify the presentation and emphasize the topology of the system at this level. The individual circuits on the other hand use the lowest possible level of symbology to provide a detailed understanding of the actual circuit. The symbols are usually of the most basic type and indicative of the most fundamental circuit elements. Because the symbology is so important to understanding within a profession, the actual symbols are usually standardized by professional organizations within each discipline.

A complete set of diagrams prevents floating models and inappropriate assignment of individual parameters.

A typical mathematical model without a connection to the underlying physiology is exemplified by Purushothaman, et al[7]. The model is largely conceptual. It consists of a series of ordinary differential equations (4) and about two dozen arbitrary coefficients. While they present some useful experimental data, their equations do not account for how the observed phenomena arose.

11.1.1.1.1 Domains of signal manipulation

It is important to differentiate between the spatial, temporal and abstract (or saliency) domains of vision. The bulk of the operation of the visual system prior to perception in the midbrain occurs in the temporal signaling domain (See **Chapters 12 to 14**). Beyond the midbrain, the bulk of the interpretation and cognition occurs in a more abstract domain that will be called saliency space (see **Chapter 15**). Only a very limited amount of signal manipulation occurs in the spatial domain (See **Chapter 2** and xxx [tremor reference])

As a result of defining the functional model of the visual system in considerable detail, it becomes necessary to properly account for the different domains used for signal manipulation as well as the demarcations between them. It is obvious that the optical system of the eyes is able to process information in the spatial position and spatial frequency domain related to object space. Additional processing in this domain is frequently assigned to the retina and/or the rest of the system. However, it is important to be specific. There is no known processing within the retina that does not rely either on time delay or temporal frequency related parameters. The time delay between the arrival of two signals at a given node is a major, and generally, unrecognized parameter in the visual system. When discussing the transformation of spatial information imaged on the retina into temporal domain information, tremor plays a major role, particularly with regard to the perception of complex shapes and relationships. Nearly all apparent spatial signal processing in vision that have been related to the retina can be accounted for by either time delay or the effects of tremor (motion introduced by the visual system to aid in perception and interpretation of primitive shapes in object space).

Chapter 15 of this work develops considerable information about the post retina signal processing employed in vision. This processing involves a variety of relatively independent signaling paths. As signals proceed along these paths, their processing becomes more and more abstract and independent of both the spatial and temporal domains (although the vectors created to represent features may contain a time related tag). By properly interpreting the signal space associated with different locations along these paths, a new perspective becomes available concerning the signal manipulations employed in perception, interpretation and cognition.

Beyond the point of perception, the operation of the brain and the definition of the saliency space and saliency signals associated with memory remains quite enigmatic. This work will not proceed beyond discussing the framework for this processing and the rudimentary methodologies of perception and cognition related to reading and

---

[7]Purushothaman, G. Lacassagne, D. Bedell, H. & Ogmen, H. (2002) Effect of exposure duration, contrast and base blur on coding and discrimination of edges *Spatial Vision* vol 15(3), pp 341-376

primitive shape recognition (See **Section 15.4**).

11.1.1.1.2 Bayesian (or other statistical) mathematics are not required in vision

While a variety of investigators have applied the concepts of statistical mathematics to their analysis of the visual system, this is an inappropriate tool. Except when discussing the operation at the lowest light levels, where quantum noise is a significant factor, the visual system is entirely deterministic. While some of the mechanisms of vision require advanced mathematics to quantify them completely, statistical mathematics does not offer an alternative solution.

11.1.1.2 Proper assignment of signal sources

It is common in the literature to find an author associating a particular characteristic of vision with a functional element other than the element in which the characteristic first occurred. For instance;

+ rods are routinely defined in terms of achromatic vision at the psychophysical level. However, at the electrophysiological level, there are presentations on red rods and green rods but no presentations on the specifics of achromatic rods.

+ it is common to find descriptions of the dark adaptation function which describe the shorter term response function to cones and the longer term function to rods even though the literature shows the same adaptation function to occur in regions of the retina where there are presumably no rods.

+ it is common to find the eccentric cell in Limulus described as if it were a photoreceptor cell and a spectral absorption characteristic associated with it.

+ it is common to find descriptions of the chromatic sensitivity of a ganglion cell because the investigator was limiting his investigations to action potentials.

In order to gain the maximum understanding of the vision process, it is important to describe where in the signal chain a specific characteristic first occurs and properly associate that characteristic with the functional element that creates it.

11.1.1.2.1 Proper assignment of dependent variables

There is a tendency in psychophysics to assign a dependent variable status on a variable when the origin of that variable is not completely understood. This has generally been the case in calculating a dependent variable, velocity, as the ratio of the temporal frequency response of the visual system divided by the spatial frequency response[8]. This assumed dependency has led to considerable confusion with regard to selective frequency tuning circuits within the visual system. The above reference describes the current confusion in this area. It will be shown that temporal frequency is the dependent function, the product of spatial frequency in object space times a velocity. Velocity is an independent variable in this relationship. It may be of exterior origin or due to tremor within the visual system.

11.1.1.3 Adequately quantifying stimuli

Quantifying precisely the photon fluxes applied to the retina and the electronic currents flowing in the retina is relatively difficult at this time because of the limited amount of precise data available in some areas.

As an example of the geometric problems, it is relatively easy to quantify the radiant flux from an external source falling on an area the size of the apparent aperture of the eye. However, since the actual aperture is internal to the optical system, the size measured from the outside frequently is not the actual aperture size. If the radiance originates from an off-axis location, the correct size of the aperture is seldom determined correctly.

Since the terrestrial eye employs an immersed optical system, it is not correct to calculate the photon flux falling on the retina based on simple ray tracing involving similar triangles. The ratio between the apex angles must equal the reciprocal ratio of the index of refractions between the two mediums, air and the vitreous humor.

---

[8]Regan, D. (1991) Op. Cit. pp 53-54

The retina is a mosaic of photoreceptors sensitive to different spectral regions. It is frequently difficult to determine the amount of the radiant flux falling on the retina which is actually intercepted by a photoreceptor of a given spectral characteristics. In the insect and some lower animals, it appears possible to determine the geometrical distribution of the various photoreceptors and then calculate the geometric cross-sectional area for each spectral type of photoreceptor. This has not been done for the human eye. For the human and many other eyes, there is little if any rigorous data as to the percentage of the available area of the retina which is sensitive to a specific spectral band. This percentage may, and probably does vary with location on the retina.

Where really careful measurements have been made of the electronic waveforms generated within the eye due to a given radiance, it is not always possible to determine from the literature explicitly whether the polarity of the signal was positive going or negative going at a specific point upon the application of an increase in flux.

11.1.1.4 The grounding plan

The architecture of the electrical grounding plan of the visual system appears to be closely correlated with the barriers associated with control of chemical diffusion within the tissue. Cunha-Vaz[9] has edited the papers of a symposium on this latter subject. At that time, the importance of the various barriers was appreciated with respect to disease but the symposium was not designed to address the utility of the barriers to the electrical performance of the eye or to the chemical protection of the delicate chromophores.

In most circuitry, it is usually advisable to establish a common electrical reference point for the entire circuit that is known as the circuit ground. Frequently, this circuit ground is connected to a highly conductive metal plate or film that then serves as a common ground for all of the circuitry mounted on, but isolated from, that common ground plane. This approach makes it convenient to connect to the common ground by connecting a wire from a specific circuit terminal to the ground plane rather than using a conductor (wire) to connect the terminal to a distant circuit ground. This design approach frequently encounters problems in high performance circuitry for two reasons. In high performance circuitry, there are frequently electrical signals of vastly different amplitude and impedance that can generate local changes in the voltage of the ground plane due to its finite impedance. Secondly, the finite impedance of the ground plane, particularly at high frequency, can introduce sufficient impedance between two terminals of a circuit or circuits to cause instability. Under these conditions, it is preferred to separate the ground planes associated with the major portions of the circuitry. In many man-made circuits, it is also advisable to provide a separate ground plane for the power conversion circuitry. These individual ground planes are then connected by a single low impedance connection that restricts the flow of undesirable ground currents related to irrelevant signal currents.

The important portions of the visual system of animals are the high sensitivity-high amplification photodetection circuits, the signal processing circuits, the signal projection circuits and the remotely located signal interpretation circuits. The visual system employs a system of dispersed power supplies that are associated with each individual circuit. This approach greatly reduces the instability problems that might be associated with a common power supply. As indicated elsewhere, the region of the retina between the Outer Limiting Membrane and Bruch's membrane constitutes an electrically and metabolically isolated area. The electrolyte of the IPM in this area can be considered the local ground plane for at least the collector lead of the input Activas of the photoreceptor cells. In several papers, the drawings suggest the common emitter leads of the photoreceptor cells are also in electrical contact with the IPM. The remainder of the electrical connections of the photoreceptor cells, along with the neurons of the signal manipulation stage and the ganglion cells, appears to be within the INM.

11.1.1.4.1 The detailed ground plan of the retina

Figure 11.1.1-1 presents the local ground plan of the retina. It shows the retina as divided into a series of electrically isolated regions with currents flowing to and from a large number of neurons within each region. It is the net value of these various currents within an in-vivo region that are usually measured in the laboratory as a voltage across the impedance network shown on the right. The ground connection shown on the right is a global ground associated with the entire system as discussed in the next figure and section. It is difficult to define the physical location of the global ground and the node shown by the black dot. The actual voltage measured is frequently between two "taps" along the length of the right-most impedance. The use of different tap locations

[9]Cunha-Vaz, J. (1979) The blood-retinal barriers NY: Plenum Press

results in measurably different experimental results. This is particularly true with respect to the summation of individual waveforms exhibiting different time delays.

The figure shows one complete "straight through" signal path and a portion of a second in order to illustrate the operation of both the bipolar and the lateral signal manipulation neurons. The photoreceptor cell shown is not defined in terms of its spectral sensitivity. Not shown explicitly are the Activas of the photoreceptor cell found within the Outer Segment and providing the current that forms the Class C waveform at the location shown, circled C. This waveform is usually measured as a voltage due to the summation of the currents from a variety of spectrally selective photoreceptor cells passing through the impedance associated with the IPM. The Activa shown represents the distribution amplifier of the photoreceptor cell. The voltage, circled D, is the voltage of the axoplasm of a single photoreceptor cell. This voltage can be measured individually. Frequently however, a voltage is measured that represents the sum of the currents through the load impedance of the photoreceptor cells multiplied by the impedance associated with the INM. The waveforms of these two recording are necessarily different being that one represents a current across a diode and the other represents the summation of similar currents through a resistive electrolyte. This situation has led to considerable confusion in the discussion of these waveforms in the literature.

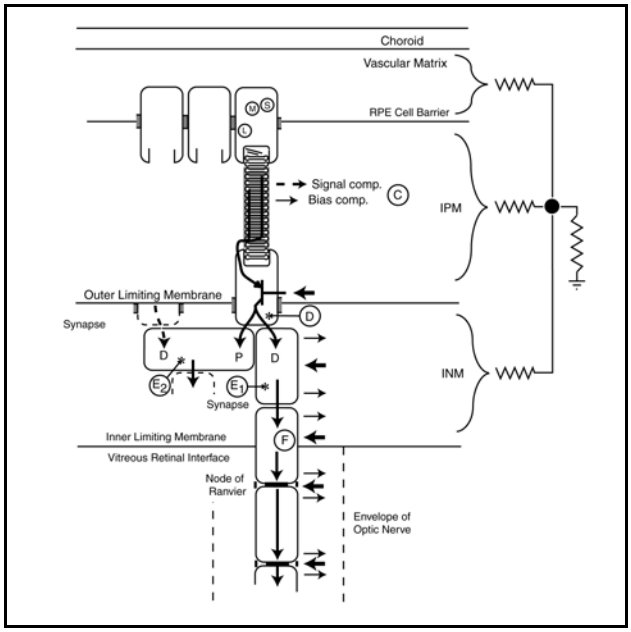

Figure 11.1.1-1 The ground plan within the retina. A series of membranes act as electrical as well as chemical barriers within the eye. The currents associated with the circuits of the eye are confined by these barriers. The dashed arrow represents the signal summed from the currents passing through the collectors of the adaptation amplifiers within the IPM. The circled letters, C through F, describe the <u>class</u> of waveform found at a specific location. D and P represent the dendritic and poditic paths through a neuron respectively. The black dot represents the location of the electrode used to sense the waveforms associated with the retina. The resistances shown are electrolytic and exhibit a time delay, as do the neurons. (See text. Expanded from a concept in H & A fg. 11-3.)

The voltage, circled D, at the pedicel of this cell is shown being passed to the dendrite, D, of the "straight through" bipolar cell and the podite, P, of the lateral (horizontal) cell. This results in the formation of two distinctly different signal waveforms within the S-plane of the retina, circled E_1 and circled E_2. The voltage at circled F is that of the action potential measured within the axoplasm of a ganglion cell. The currents associated with this waveform from a large number of individual cells pass down the optic nerve to the brain. The voltages are individually regenerated at Nodes of Ranvier along the way.

In this figure, the tightly packed RPE cells (not shown individually) form a physical barrier between the choroid artery and the IPM for both chemical and electrical isolation. In this context, the morphologists and clinicians label this barrier the outer barrier, one of the blood-retinal barriers. The ONL and ILM also form electrical barriers within the retina and the ILM may also play a role in chemical isolation between the vitreous humor and the richly perfused tissue of the retina. The purpose of these barriers has not been considered by the morphological community.

If each region of electrolyte between these barriers is considered an electrical ground plane, the electrical circuit elements on the right can be drawn to help explain some of the signals obtained by various LERG procedures. It is also possible to sketch in the electrical circuits of the fundamental signaling path through the retina and into the optic nerve. The individual circuits (neurons) of the fundamental path are shown simplified. The heavy arrows show the <u>non-signal currents</u> entering each neuron or Node of Ranvier. The smaller arrows perpendicular to the signal path suggest the <u>non-signal currents</u> leaving the neurons. Using a notation similar, but not the same as, Neuman's[10], the INM or IPM would be considered a current source for the neurons via the heavy arrows and a current sink via the small arrows. Alternately, the labels can be interchanged if applied to the cells themselves. The heavy arrows

[10]Neuman, E. (1985) XXX Trends in Neurosci. vol. 8, pp. 156-159 Also shown in Heckenlively & Arden (1991), pg. 104

would relate to the cell as a sink.

The labels source and sink are not normally used when there is source of potential in the loop as in the case of the electrostenolytic processes supporting the cells. The convention is that the conventional current flows from a point of higher positive voltage to a point of lower positive voltage. In the case of the neurons and their individual electrostenolytic power sources, the arrows indicate the direction of conventional current flow from and back to an electrolytically homogeneous medium.

> This model reverses the direction of the current at the Outer Segment relative to that shown by Newman. It also eliminates the difference between his current flow to the Outer Segment associated with a vertebrate and that of Laughlin for an invertebrate. The conclusion is that the photoreceptor cells of all animals exhibit the same location of "sinks" and "sources." Finally, it employs electrons and not potassium ions to explain the current flow.

In this simple representation, the signal currents travel along the centerline of the signal path. Note that a probe inserted into the IPM or the INM measures only a net current or a net voltage due to the summation of many individual currents.

> It is one of these net currents associated with an isolated neuron that Hodgkin & Huxley attempted to graphically, and then mathematically, decompose into an inward and an outward current. They then assigned arbitrary names, associated with the concentration of sodium and potassium ions on the two sides of a biological membrane, to these putative currents. It is now recognized that these biological membranes contain a hydrophobic core and are essentially impervious to these ions. **The signal currents through biological membranes are due to the flow of electrons and/or holes, not ions.**

11.1.1.4.2 The large scale ground plan of the visual system

The distance between the brain and the signal processing region of the retina is electrically significant. These two regions effectively use separate electrolytic ground planes. The use of pulse signaling between these two regions reduces the sensitivity of the system to voltage and time delays. However, researchers frequently encounter voltage differences and time delays that are not accounted for adequately.

> Instead of a two dimensional ground plane as found in most man made electronic equipment, the ground planes of vision consist of separate three-dimensional regions of a relatively high conductivity electrolyte.

The overall visual system is best described using four (4) local ground planes, one formed by the IPM and associated with the high sensitivity portion of the photoreceptor cells, one formed by the INM and associated with the signal processing portion of the retina, one formed by the cerebral-spinal fluid (CSF) enclosing the brain, and one representing the remainder of the body fluids as illustrated both schematically and morphologically in **Figure 11.1.1-2**. These regions are interconnected by low, but not zero, impedance inter-ground bridges. These impedances consist of both a resistive and capacitive component. Of even greater importance is the time delay associated with these resistances due to the finite transport velocity of ions within the electrolytes. Precision measurements of visual waveforms must account for these time delays. Since the various regions of individual electrolytes may parallel each other for some distance, it is not always possible to define a specific location for these inter-ground bridges. Many of the bridges to the body fluids can be represented by the conduits formed by the vascular system. However, even the walls of these conduits represent an impedance. Sampling of the voltages associated with the currents in these finite impedance inter-ground bridges have played a significant role in the electronic measurements of vision when they have been located.

The bridge between the IPM and INM is unique in that it is in the form of a distinct lip, or collar, at the extreme edge of the cup formed by the IPM and containing the cup formed by the INM. Although this lip cannot be directly contacted, the current density in the region causes a significant voltage drop. Because of the large area of this cup relative to the thickness of the Outer Limiting Membrane separating the IPM and INM, there is also a significant capacitance between these two regions. There is a similar magnitude capacitance between the IPM and the choroid/sclera due to the thinness of Bruch's membrane.

An electrode placed in contact with the sclera above this bridge with a return electrode placed on some other portion

of the epidermis can detect a voltage representative of this complex electrical network. If the sensed voltage is recorded as a function of time, an ERG, electroretinogram, is obtained. The second electrode is frequently placed on the forehead or near the ear. These locations are frequently labeled in the literature as "indifferent locations." However, these locations are not indifferent with respect to time delay. Each location can be characterized by its distance from the source of the signals it is attempting to record. This subject is also addressed from a clinical perspective but in less detail by Arden in Heckenlively & Arden[11].

Making electrical contact to the neural system of an

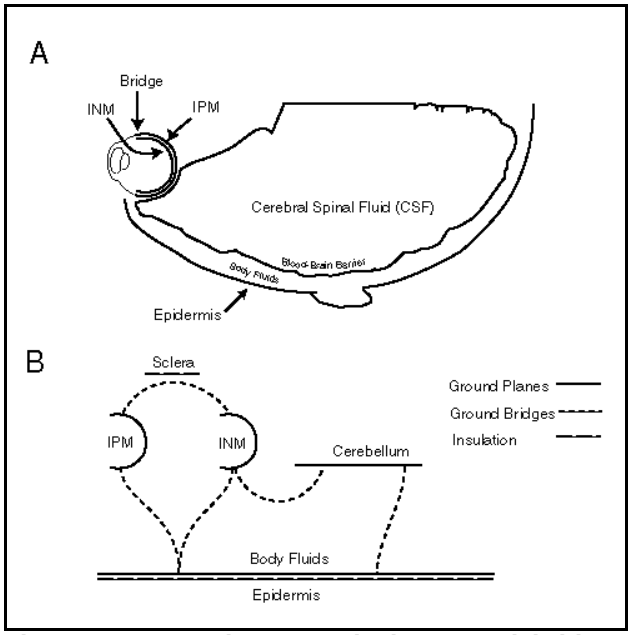

Figure 11.1.1-2 The ground planes and bridges pertinent to the Visual system. The optic nerve is not shown between the eye and the brain. The epidermis is a poor conductor of electricity. This condition is frequently accepted in order to avoid medical invasion of the subject.

[11]Heckenlively, J. & Arden, G. (1991) Principles and practice of clinical electrophysiology of vision. St. Louis, MO: Mosby Year Book, pg. 384-390

animal via mechanical contacts on the epidermis is a poor but convenient method of obtaining technical information about the visual process. Whereas the individual neural signals involve voltages in the 1-100 millivolt range and exhibit very well defined waveforms, most electroencephalograms and electroretinograms involve measured signal voltages in the range of a few microvolts. It is only possible to obtain sufficient current through an impedance to generate adequate signal voltages for recording by illuminating a large region of the retina. This is because a large number of Adaptation amplifiers are excited synchronously. As the collector current from all of these amplifiers are summed and passed through the IPM to INM bridge, a signal of significant amplitude and waveform can be recorded. Otherwise, the recorded waveforms are very noisy.

11.1.1.5 Inappropriateness of Ohm's Law in vision

Another great difficulty in vision research has been the assumption that the eye involves linear resistive elements and Ohm's Law can be used in describing the signal voltages through simple models. Nothing could be further from the truth! Ohm's Law is a simplification of one of Kirchoff's Laws. It is based on the assumption that the resistance (or conductance) of the resistive elements in a circuit is independent of the current passing through it. This condition is not true in vision or the neural system in general! **Ohm's Law does not apply to the vision process in animals** except under the most restrictive conditions. If an experimenter is making measurements over an exceedingly narrow irradiance range, of less than +/-10%, an effective resistance may be determined by measuring the change in voltage related to a change in current. The resulting value only applies for the irradiances specified and an extremely narrow range near it. Kirchoff's Laws on the other hand allow for any kind of impedance, but the mathematics becomes more involved.

The basic fact is that the neurons of a living animal do not employ resistive elements as normally defined and manufactured by man. The basic impedance utilized in living tissue is represented by the diode. The voltage across a diode is not given by the product of a constant times the current passing through it, i.e., the common expression E=I * R; the voltage across an ideal diode is given by the natural logarithm of the current passing through the diode, conversely and more specifically, $I = I_o * (e^{V/V_T} - 1)$ where V_T is related to the temperature and the material comprising the diode.

Because the impedance of neural tissue is not resistive, most of the algebra appearing in the literature is not appropriate to neural or vision processes. Models employing common resistive elements are rarely appropriate.

Linear matrix algebra is not appropriate for the solution of vision problems if the excitation level changes by more than +/-10%. Linear matrix algebra is based on the use of linear homogeneous equations with constant coefficients. Mathematically, the terms are not homogeneous and the coefficients are clearly not constant in neural tissue and vision.

11.1.1.6 Inevitable results of the non-linearity of animal tissue

There are two inevitable results of the fact that animal neural tissue is not characterized by a resistor.

First, as will be shown in the next Chapter, the output of the Photoexcitation/De-excitation process and hence the signal associated with vision is a quantity of electrons (which is normally moving with respect to time and constituting a current). When passed through a neural impedance, the resulting voltage is not linearly related to the charge (current). The voltage is related to the natural logarithm of the current. Thus, in probe experiments and ERG's, the measured quantity is generally not linearly related to the signal; the measured quantity is proportional to the natural logarithm of the signal parameter.

Second, as will be shown in the next following Chapter, the signal channel associated with chrominance information relies on the linear summation of voltages; i.e., the linear summation of logarithmic functions of the underlying signals. In fact the signal channel associated with luminance information also relies upon the linear summation of voltages. *The summing of logarithmic quantities is equivalent to the multiplication of the underlying signal parameters.* Nearly all of the psychophysical and many of the electrophysical experiments in the literature measure parameters resulting from such logarithmic summations, but they do not recognize that fact and proceed to perform linear subtractions in order to separate spectral components of their data. There are two principal results of this failure to recognize the nature of the parameters measured.

> The calculated visual spectrum, based on the arithmatic summation of the absorption coefficients for the individual chromophores, does not match the overall high irradiance (photopic) or low

irradiance (scotopic) spectral performance of the eye. A logarithmic summation of the individual absorption coefficients *does give the measured overall spectra for all irradiance levels!*

The attempts to characterize the absorption coefficients of the individual chromophores by employing the linear subtraction of measured voltages invariably give incorrect spectral characteristics for at least one of the individual chromophores (typically the L-channel). Utilizing logarithmic subtraction, the *correct absorption coefficients can be obtained* for all three (four in appropriate animals) of the individual chromophores. Moreover, all of the absorption coefficients obtained are unipolar. No negative absorption in the blue region is required of the L-channel chromophore.

11.1.1.7 Availability of Complete Photoexcitation/De-excitation Equation

The major focus of Chapter 12 is the development of the Complete Photoexcitation/De-excitation (P/D) Equation of the vision process in animals. It will be shown that such an equation exists, is applicable to all vision systems and is based on conventional physics. The mathematical details of the equation are presented in Appendix A.

The P/D equation provides a closed and complete solution to the visual systems initial response to irradiance. When combined with certain nonlinear processes, it provides the exact solution to what is usually called the generator potential. The solution is presented in a variety of simplified forms to show how it is most easily applied to different experimental problems. In the complete form, it demonstrates both the amplitude profile of the generator potential and the time delay associated with the response in detail. The P/D equation includes a number of primary features. The P/D equation:

+ includes the spectral absorption coefficient of the chromophore explicitly,

+ describes the effect of ambient temperature on both the amplitude and time delay of the photoreceptor of any animal,

+ describes the changes in both the amplitude and time delay of the response as a function of irradiance, and

+ is applicable to each and every spectral response channel of the animal vision system

The P/D Equation also includes a number of secondary features. The P/D Equation:

+ provides the exact equation of both the rising and falling portion of the waveform under any conditions, and

+ reduces, in a special case, to an equation originally suggested by Hodgkin.

11.1.1.8 Obsolescence of the Michaelis Equation and the Logistic Equation in the Vision context

For more than the last half century, the experimentalists have not been able to provide an accurate mathematical equation for the generator potential and have developed a number of “rule of thumb” equations and simple models. Many of the investigators ignored the trailing edge of the generator potential and concentrated on the rising (or leading) edge of the waveform. They frequently applied a simple hyperbolic equation to their data using a variety of names depending on what their own discipline was[12]. Some went further. They adopted Lumped Constant Filter Theory to develop simple electrical models to account for the observed changes in the rising waveforms. Unfortunately, none of these equations and models are able to describe both the rising and trailing edge of the generator potential. The simple models based on Lumped Constant Filter theory required so many degrees of freedom that they could be used by the investigators to match any waveform if only the characteristics of the complete generator function were available. The P/D Equation demonstrates successfully that:

+ it can describe the entire generator potential function under all condition,

---

[12]See Charlton & Naka (‘70), Naka & Rushton (‘66) (‘67), Examples of Forti and of Lamb in Torre, et. al (1990)

+ the actual process does not involve a hyperbolic rising edge,

+ Transport Delay Theory, not Lumped Constant Filter Theory provides the correct solution to the problem, and

+ it can be used to describe the entire a-wave of the typical ERG. In fact, it can be used to subtract out the a-wave, thereby providing the complete form of the b-wave.

11.1.1.9 Filtering of the P/D Equation by proximal circuits

There are rate limiting processes in the signal chain of the retina which become important under high amplitude stimulus situations. These limitations are primarily related to the transport of materials required to keep the various electrolytes and locations of redox processes adequately supplied. Thus, the output versus input characteristic of many neurons exhibits an **S-shaped curve**. However, although the curve appears at first glance to be of arbitrary shape, Kahn & Bockris writing in Gutmann & Keyzer[13] have provided the exact shape of this curve when applied to signaling in neurons. As they point out, the shape is actually invariant and results from the ratio of two exponential expressions.

11.1.1.10 Elucidation of additional ERG waveforms

The conventional ERG represents a composite waveform derived from mixing waveforms from both Stage 1 and Stage 2 of the model. Because of this mixing, it is difficult to find a unique home for the ERG in the context of this model. As will be seen in **Chapter 13**, the P/D Equation actually expands the data obtainable from the ERG. With the P/D Equation of Stage 1 available, it can be shown that there are additional "waves" present in the typical ERG. Specifically, after mathematical removal of the *complete* a-wave from the overall waveform, it is possible to actually see not only the b-wave associated with the in-line processing performed in the Inner Nuclear Layer but also see the additional waves associated with the ancillary processing related to the horizontal cells and the amercine cells. An expanded convention for naming the components of the ERG is provided in **Section 11.1.5.**

11.1.1.11 Exceptions to the Principle of Univariance

There is another principle which needs to be carefully applied in the following material. The Principle of Univariance has been taught widely as a foundation stone of the visual process. At a conceptual level, it can be applied to the vision process with some success. However, it generally fails at the detailed level. To interpret the Principle, it must be rigorously defined.

Naka & Rushton,[14] speak of the Principle of Univariance as "the most powerful postulate in colour physiology."They begin with a description of a photocell and conclude that "the output is of only one kind (univariant) it cannot give simultaneously separate information as to both intensity and wave-length." While the output signal is unidimensional, its value remains a function of two input variables. As Naka and Rushton proceed with their discussion on page 538, they deviate farther and farther from reality.

In the next paragraph they say "The Principle of Univariance may be stated as follows: Each visual pigment can only signal the rate at which it is effectively catching quanta, it cannot also signal the wavelength associated with the quanta caught." This restatement obscures the fact that the variable is a function of two variables. The variable is equally capable of signaling the wavelength it is capturing quanta under conditions of fixed quanta intensity.

They proceed to restate the Principle it in what they call a more definite form: "The signals from each cone depend only upon the rate at which it is effectively catching quanta; it does not depend upon the associated wave length."

This author would subvert the term cone in favor of the more general form. "The signals from each photoreceptor depend only upon the rate at which it is effectively catching quanta; it does not depend upon the associated wave length."

---

[13]Gutmann, F. & Keyzer, H. (1986) Modern Bioelectrochemistry. NY: Plenum Press pp. 1-44

[14]Naka, K. & Rushton, W. (1966) S-potentials from colour units in the retina of fish (*Cyprinidae*) J. Physiol. vol. 185, pp. 536-555

In a second paper in their set (pp 556-586) Naka & Rushton say "But before proceeding to a description of this work, we must mention a fundamental principle that we believe to hold without exception, for our arguments rest upon it." They then open the paragraph titled Principle of Univariance by saying "It is not accurate to say that every quantum absorbed by a visual pigment has the same effect irrespective of its associated wavelength!" [Exclamation point added by this author] Their discussion proceeds into areas they were attempting to explore and understand better.

When discussing color physiology, it is important to note the rate of quanta capture as a function of wavelength is critical to the perception of chrominance in vision. The next two paragraphs will examine the reality of luminance and chrominance detection in vision more carefully.

It is best to consider the Principle of Univariance as an archaic pedagogical tool that obscures understanding of the actual operation of the biological vision system.

11.1.1.11.1 Failure of Univariance with respect to luminousity

This principle seems intuitively obvious. However, **the Univariance Principle is not rigorously correct** when applied to the long wavelength spectral channel and indirectly to the luminous channels of vision. It is based on what can be called classical or Newtonian physics. It assumes the energy of the photons is absorbed "thermally." Quantum physics provides a different postulate for detection processes, such as vision, involving the internal photoelectric effect. Things in the quantum mechanical world are not always as they seem. This is particularly important here where the Photoexcitation/De-excitation process has been divided into three sub-processes. In the first, an electron is raised to a higher quantum energy band due to the absorption of a photon. However, where it is located in that energy band is still a function of the initial energy (wavelength) of the photon. After excitation of a number of electrons into the π^* band of the chromophoric material using photons of different energy, one can actually select which excited electrons will be de-excited by adjusting the energy threshold of the de-excitation process (it is typically a function of temperature in animals). By moving this threshold up or down, one can actually recover a plot of total excitons created as a function of the energy (wavelength) that created them. Therefore, at this stage, the chromophoric material can signal both the number of photons absorbed **and** their energy.

The translation sub-process involved in vision (and described as de-excitation) exhibits a fixed energy threshold (which is a function of temperature). There are two ways that excited electrons can be de-excited from the excited quantum state in the translation sub-process.

1. In the first, a free electron is generated in a "conventional" conductive circuit for each exciton in the π^* band of the chromophore which is de-excited. These electrons will correspond numerically to the number of excited electrons which had an energy exceeding the threshold. Thus, the number of free electrons is univariant with respect to the variables, number and wavelength of the photons in the irradiance. ***The Principle of Univariance is sustained in this case***.

2. In the second case, the excited electrons in the π^* band do not have enough energy individually to generate a free electron in the conductive circuit. However, they are able to sum their individual energies and if the sum exceeds the threshold, one free electron will be released for each group (typically two but sometimes three or four) of electrons. If there are other phenomena occurring wherein the excited electrons can be de-excited thermally or by fluorescence, the output clearly becomes a rate equation which is not linear. Even in the absence of these other modes of de-excitation, the function describing the number of free electrons produced is a convolution of the number of excited electrons present and their energy level. This is not a simple product of the number of photons absorbed. ***The Univariance Principle does not hold for this case***.

The L-channel of human vision is an example of the second case and interesting experiments can be performed to demonstrate this fact. The literature contains many observed examples of this case. <u>The fact that the amplitude response of the L-channel photoreceptors exhibits a square law relationship to the incident signal amplitude in the mesopic region is a direct example of the failure of the Univariance Principle</u>

There is also a third situation of trivial importance: In general, individual photons absorbed by the chromophores of vision which have an energy less than the threshold of the de-excitation process will not generate any free electrons in the output of the overall process. At very low irradiance levels, excitons generated by photons of energy less than the threshold of the translation sub-process can be absorbed by the chromophore but fail to be summed with another exciton before they are lost to the vision process. These photons never result in the generation of a free electron. The Principle of Univariance ***is thwarted completely in this case***.

14 Processes in Biological Vision

In summary, the overall Photoexcitation/De-excitation process in vision does not conform to the Univariance Principle. The Univariance Principle only applies to situations where the incident photons have sufficient energy to generate excitons with an energy greater than the de-excitation threshold energy of the translation process. In general, the L-channel does not meet this criterion. Therefore, taking a que from Naka & Rushton, the principle is restated in a more definite form--

> **Principle of Univariance** --The overall Photoexcitation/De-excitation process in vision is linear and univariant for photon wavelengths corresponding to energies above the threshold energy of the translation sub-process (wavelengths typically shorter than 575 nm). For photons corresponding to energies below the threshold of the translation process (typically at longer wavelengths than 575 nm), the process is inherently nonlinear and not univariant.

This is the same situation found in photography where a change in response with irradiance is commonly found in the long wavelength portion of the spectrum. The term "γ" is used to describe the exponent in the output versus input equation for photography. In both cases, the exponent is not a constant with respect to the photon wavelength. It is approximately one for wavelengths shorter than 0.5 microns and approximately two for wavelengths between 0.5 and 1 micron.

An explicit value for the threshold energy of the translation stage has not been found in the literature. However, an implicit value can be determined from the observed performance of the visual system in humans. Only the L-channel exhibits a change in sensitivity relative to the other channels as a function of irradiance. This would lead one to assume that the threshold energy is in the region of two electron-volts or about 0.6 microns. The actual value of the threshold energy may be determined by observing the long wavelength skirt of the M-channel spectral absorption coefficient as measured electrophysically at the output of the photoreceptor cell. At a wavelength where the energy of the photon becomes less than the threshold energy, the absorption spectrum should show a discontinuity. A similar discontinuity may be found in the short wavelength skirt of the L-channel. In the case of the L-channel, it may be more difficult to discern this change because it will not involve a rolloff in effective absorption. It may only involve a change from a one exciton process to a two-exciton process--i.e., the dynamics of the process may change with respect to changes in irradiance level, similarly to the photographic case discussed in the previous paragraph. Using the data of Sliney in the infrared, it appears the best estimate of the threshold energy of the translation stage is not more than 2.34 eV (0.53 microns).

11.1.1.11.2 Failure of Univariance with respect to chrominance

In the biological retina, an interdigitated array of spectrally specific photoreceptors are used. This differs from the situation in photography where individual layers of color sensitive material are laid over each other. In either case, the sensation of color is determined from the relative number of quanta absorbed per unit of excited area by each spectrally specific class of absorbers (where the excited area is significantly larger than the cross-section area of a single photoreceptor).

Thus the photoreceptors are used in the very mode that Naka & Rushton exempted from their discussion and Principle of Univariance. The output of each spectral class of photoreceptors is proportional to the wavelength of the quanta captured under conditions of constant excitation per unit excited area. It will be shown in Chapter 13 that the signal processing accomplished in the retina subtracts the logarithm of the signal amplitude from the photoreceptors of individual spectral types. This subtraction is accomplished using pairs of spectrally selective photoreceptors. The result is a set of signals that are nearly linearly proportional to wavelength over significant spectral intervals. These signals constitute the chrominance signals of biological color vision. ***The operation of color vision contradicts the Principle of Univariance*** as stated by Naka & Rushton.

11.1.1.12 Emergence of the Biological Transistor

It is amazing to report in the year 2001, there is no mention in the scientific literature of the presence of a device or configuration in the neural system of any animal that has been described as a transistor--**a biological transistor**--when there is so much evidence, *overwhelming evidence*, for such a designation. It can only be explained by many investigators failing to see the forest for the trees. Giving credit where it is due Michael Bennett came very close to

discovering such a device in the 1960's.

Interestingly, in mid 2001, a search of the Internet did not recover a single reference to the term “biological transistor,” except at this authors site.

There is an immense amount of electrophysical evidence for the definition of a biological transistor as an analog to the semiconductor transistor. Some of this evidence will be presented here. More will be presented in Appendix B where the formal properties of the device will be developed and explanations of how the device is used within the neural system will be presented. The evidence is compelling and the acceptance of this reality greatly simplifies the investigator’s task.

During the development of this Theory, the active electrolytic semiconductor device, the Activa, was discovered and awarded a United States Patent, #5,946,185.

Until the 1960’s, it was categorically denied by the leaders of the vision community that the eye employed signaling by way of electrical currents (as opposed to chemical affecters) at synapses. This has changed considerably beginning in the 1980’s. Similarly, there is now a large community of investigators whose work is based on the physical transport of large numbers of complete (and large) ions every second through a membrane surrounding the Outer Segment of each of the photoreceptors of the eye even though there are immense amounts of evidence from other investigations that ***there is no cellular membrane surrounding the Outer Segments***. The Outer Segments are actually extracellular.

11.1.1.12.1 The Activas within the Outer Segment

The Outer Segment in *Chordata* consists of a stack of disks and each of these disks is essentially covered by a liquid crystalline membrane. However, this material is only an organically derived film. The substrate of these disks consists of a structural protein, created in a manner not unlike a fingernail. The substrate is then covered by an organic film by precipitation. The result is a nonliving structure with very unique optical and electronic properties. This structure is in electronic communications with the Inner Segment of the photoreceptor via what are usually labeled the cilia of the photoreceptor cell. However, it is more precise to speak of these as the dendrites of the neuron enclosed within the Inner Segment of the photoreceptor. The interaction between the disks of the OS and the dendrites of the IS is the first instance of a biological transistor in the visual signal path. But, it is not the last.

Each of the dendrites of the IS is laid in one of a set of grooves around the edge of the stack of disks. Each dendrite is in contact with each of the disks while the whole assembly is immersed in what is called the Inter-Photoreceptor-Matrix (IPM). This is basically a conductive media that is also used to transport the chromophoric material before it is precipitated on the disks as indicated above. The result is a very close “junction” at each disk between an organic membrane of the disk and an organic membrane of the dendrite with a thin space between them. Each of these membranes is porous and subject to the laws of diffusion. Interestingly, each is also subject to the laws of electrostatics--and they are in extremely close proximity, on the order of 50 Angstroms. These are the electronic characteristics of a device capable of providing signal amplification, more specifically charge (or current) amplification. By applying the appropriate bias voltages between the various elements of this configuration, an excited electron in the chromophoric material of a disk can generate 1,000 to 10,000 free electrons in the dendritic structure. Note the distinction here; the disk material is of organic origin but it is non vital (living). It is extracellular. The dendritic material is of similar organic origin, but is intracellular and is vital. The presence of the IPM is a necessity as it provides an ohmic connection to the structures immersed in it.

The key to the operation of a transistor is the coupling of the diffusion process between the two regions in the presence of an electrostatic field. However, this introduces another critical requirement. In a semiconductor transistor, it would be self-destructive to employ a diffusion process that actually moved the ions making up the structure during operation. The process would ultimately destroy the underlying physical structure. Instead, there is apparent motion of the ions, by what are called “holes.” While electrons move about in the *conduction band* of this sandwich under the countering forces of electrostatics and diffusion, the “holes” move through the *valence band* of the material subject to the same forces. The velocity of these two different “currents” is in general quite different and the difference constitutes “transistor action.”

The effect to the outside world is that there is an electron current flowing in one direction and a flow of ions in the opposite direction. *In fact, no ions ever move*. They must be present, but *they do not move*! The result of this arrangement of organic membranes to form an organic transistor is truly fantastic. Again, the complete development of this area is found in Appendix B.

The Rhabdom in the eye of *Arthropoda* and *Mollusca* is constructed similarly topologically to that of *Chordata* but differently topographically. It communicates with the rest of the neural signal path in a similar way.

11.1.1.13 Ramifications of the Biological Transistor

As will also be developed in Appendix B, the mysterious purpose of the Nodes of Ranvier is easily and directly explained in the context of a biological transistor. And the critical test of this theory is provided by Tasaki, a leader in the field of nerve excitation. Quoting his 1975 paper[15], "We note that cleaned axons are incapable of developing action potentials when immersed in a medium completely free of divalent cations." Going further, it is proposed here that the actual situation was that axons are incapable of developing action potential when immersed in a non-conducting medium or a medium not capable of providing the proper ohmic contact to the middle layer or (poda) of the biological transistor located in the axon at the Node of Ranvier. The whole purpose of the break in the myelin (insulating) sheath of the axon is to allow such contact with the surrounding media.

> It is a prediction of this work that investigators will find that complete photoreceptor cells will also misfunction if they are first washed and placed in a non-ionic or inappropriate ionic solution.

Going further, the key feature of a biological transistor circuit is the easy creation of a region of negative resistance in the voltage-current characteristic of the device. The simplest possible use of this feature is to recognize that such a characteristic cannot exist over an extended range; it requires a source of energy to create it over even a limited range. Therefore, in a typical situation, this region is bounded on both ends by a region of positive resistance. It is absolutely elementary to create a bistable circuit using a biological transistor that will exhibit two stable states. Hence, the *foundation* for the TWO-STATE THEORY of neural excitation is explained. The only trick is getting the device to switch states in the desired manner. To accomplish this, utilize a biological transistor and;

> adjust the impedance of the electronic element in the poda lead properly and you will achieve an action potential regenerator.
>
> adjust the electronic element in the dendritic and/or axonal leads of the biological transistor and you will produce an action potential with an interpulse time interval which is a direct function of the analog input signal (if it exceeds the threshold level).
>
> adjust the impedances differently and you will obtain extremely high analog current amplification under very low noise conditions (and no spike will be generated).
>
> use two biological transistors in what is commonly called a cascode arrangement and you can obtain extremely high analog current amplification under nearly thermal-noise free circumstances.

11.1.1.14 Rationalizing the confusion over polarity

There has been a long standing, and sometimes acrimonious, debate over whether the polarity of the generator potential observed in *Arthropoda*, *Mollusca* and *Chordata* are the same or different. Unfortunately, until the recent development of residual dye techniques, it was difficult for investigators to know precisely what cells they were actually obtaining electrophysical signals from--and whether these signals represented intracellular conditions or extracellular conditions. Furthermore, capacitive pickup from adjacent cells has frequently been a problem because of the high impedance levels involved.

There has also been the problem of relevance. Ottoson provided a side by side comparison of the output potentials from invertebrate and vertebrate photoreceptors[16]. However, his comparison suffered from several unstated factors. He used the response from an eccentric cell of *Limulus* as a typical invertebrate photoreceptor cell response, whereas most investigators consider this particular cell as a special case of a signal manipulation type cell. He also showed the response to a long pulse of light in the case of the invertebrate and the response to an impulse in the case of the vertebrate.

This problem will be examined carefully in **Chapters 12 & 13** in order to provide a defendable explanation of the

---

[15]Tasaki, I (1975) Evolution of theories of nerve excitation in vol. 1, The Nervous System, Tower, D. editor in chief NY: Raven Press pg. 177-195

[16]Ottoson, D. (1983) Physiology of the nervous system. NY: Oxford Press, pg. 365

actual situation--a necessity if many of the experimental programs under way in neural signaling are to be concluded successfully. In brief, the results of this examination are that all photoreceptor cells in animal vision operate similarly and generate the same polarity signals at the same location. The polarity of the signal sensed near the OS is opposite to that sensed by probing the axoplasm of the photoreceptor cell because of the differential amplifier circuit found in the IS of all photoreceptor cells.

11.1.1.15 Important consequences of the adaptation mechanism

The fact that the adaptation mechanism introduces a zero in the transfer characteristic of the photoreceptor cell means that the intrinsic animal visual system is only sensitive to changes in object space. ***The visual system is inherently blind to fine detail in a stationary scene.*** The higher level molluscs and the higher level chordates have evolved an additional mechanism to overcome this characteristic. They introduce tremor via the ocular muscles to continually move the scene relative to the line of fixation of the eyes. This tremor is discussed in detail in **Section 11.1.2.8.4**.

11.1.2 Terminology and conventions

Some of the electrical terminology that has evolved in the biological field is not easily translated into conventional electronic engineering notation. Before proceeding, the relevant terms and conventions will be examined and some equalities defined.

11.1.2.1 Responses

With vision being a life science, investigatory work in this field necessarily involves biological-based experiments. These biophysical experiments fall into several categories, exhibiting various degrees of precision:

> **Psychophysical**--based on the observed or reported response of an organism to application of a physical stimulus. Psychophysical experiments lead to psychophysical relationships rather than psychophysical measurements. They do not normally produce absolute values. They normally provide null values or suggest trends.
>
> **Electrophysical**--based on the use of electrical probes to intercept signals generated by an organism in response to the application of a physical stimulus. Electrophysical experiments lead to exact measurements but, on the scale of the visual processes, may not involve exact knowledge of what is being measured.

Within the group of electrophysical measurements, there are a number of fundamental types:

> **mass responses**--electrophysical measurements frequently taken at a remote point relative to the point of electrical signal generation, hence representing an average response of many circuits operating simultaneously (and frequently representing a voltage across a resistor in a “ground loop” in electrical engineering terms)
>
> **Electroretinagraphs (ERG)--**mass responses taken by measuring the voltage on the surface of a complete eyeball with reference to the rest of the organism (and hence measuring the voltage on an equivalent impedance consisting of the impedance paths due to the entire optic nerve, the muscle tissue connecting to the eyeball and the liquid medium surrounding the eyeball plus any other impedance leading to the point where the return lead is connected)
>
> **Visual Evoked Potentials (VEP)--**mass responses taken by measuring the voltage between a point on the scalp at the rear of the skull relative to a second point generally at a second point on the skull considered in some sense neutral or *indifferent* for purposes of the experiment
>
> **more specific responses**
>
> **Local electroretinographs (LERG)--**potentials measured with a probe, either extracellular or intracellular, and referenced to a more local reference point than the eye cup. A pair of concentric probes may be used resulting in a return signal lead connected to a point only 5-100 microns from the input signal probe**.**

Extracellular potential--potentials measured with respect to a return signal lead by using probes and frequently involving the capacitative coupling of signals from several adjacent cells. The return signal lead is frequently connected to some nearby structure, the eye cup, the retinal epithelium, etc.

A particularly important extracellular potential obtained with a probe is the potential in the vicinity of the Outer Segments due to the P/D process.

specific responses

Intracellular potential (current)--potentials (currents) measured utilizing microprobes to penetrate a given cell and measure its internal voltage with reference to the return lead. The return signal lead is frequently connected to some nearby structure, the eye cup, the retinal epithelium, etc.

A particularly important intracellular potential obtained with a probe is the axoplasm potential of a photorecepter cell and frequently labeled the generator potential.

The above descriptions surfaces a serious problem in vision research, **the frequently casual treatment of the return electrode.** Quoting the Brussels Symposium ad-hoc committee[17] for VEP research in discussing the Electrode Montage, "The electrode montage and polarity convention (e.g., negative upward) should be clearly stated. The word 'indifferent' should not be used in describing the reference electrode, as it begs the question. The ear electrode, commonly designated a 'reference' is frequently anything but a passive reference in relation to the occipital response." The same can be said about other reference electrode positions. The conventional ERG is a classic case of this problem. The response measured is frequently described in terms of an a-wave, b-wave and c-wave, etc. based primarily on when they occur in time, not where they originate in the visual signal path. Little or no attempt is made to subtract the 'tail' of the a-wave from the start of the b-wave, etc.--although independent signals conforming to the a-wave are now generally available from intracellular recording.

Failure to adequately describe the electrode montage leads to data which is difficult to interpret because of frequent corruption by other unexpected signals. This is true when a probe is inserted into *the retina* a given physical distance and also when a probe is inserted into *a cell* a given physical distance. The signals resulting from such insertions are frequently indecipherable or misleading without a detailed model of the visual system being examined.

When attempting to record electrophysical waveforms, it is again necessary to be very specific with regard to the technique used. Ruben & Walls[18] make an important point in discussing probing the retina to various depths: "The signals within the retina to which the retinal nerves must attend are not necessarily the same as those recorded from the surface of the cornea, a mere 2 cm away. If the retina is penetrated slowly by a tiny electrode, beginning at the vitreous side and progressing toward the pigment epithelium, the polarity of the signal abruptly changes when the inner limiting membrane is crossed. The a-wave here is positive going, followed by a negative-going b-wave, a positive going c-wave, and a negative d-wave off-potential. *Except for the c-wave, all of these components are upside-down* in relationship to those recorded at the cornea." This reversal phenomenon is well illustrated in Figure 2 of Heynen & van Norren[19]. Such reversals have caused long-standing arguments in the literature related to the signal paths in the retina.

Expanding on the above paragraph, Murikani & Pak[20] have provided material comparing the recording of intra-cellular and extra-cellular ERP's. Their actual comment is specific: "As was noted in the preceding section the intracellularly recorded waveform of the early receptor potential is substantially different from the extracelluarly recorded waveform." They go on, with inserts by this author: "The intracellular recording represents a sum of the

---

[17]Halliday, A. Arden, G. Bodis-Wollner, I. Jeffreys, A. Kulikowski, J. Spekreijse, H. & Regan, D. (1977) Methodology of patterned visual stimulation. in "Visual evoked potentials in man: new developments" ed. J. E. Desmedt Oxford: Clarendon Press pg 6

[18]Rubin, M. & Walls, G. (1969) Fundamentals of Visual Science. Springfield, Il: Charles C. Thomas, pg 195

[19]Heynen, H. & Norren, D. van (1985) Origin of the electroretinogram in the intact macaque eye--I. Vision Res. vol 25. pg. 700

[20]Murikami, M. & Pak, W. (1970) Intraculluarly recorded early receptor potential of the vertebrate photoreceptors. Vision Res. vol. 10, pp. 965-975

potential across the cell membrane and the potential drop along the resistance in the extracellular space {leading to the return electrode}, while the extracellular recording represents the latter alone {plus the signals from adjacent sources through capacitive coupling}."

Heynen & van Norren[21] provide an entirely different scenario of utilizing extracellular recording with a concentric microprobe to obtain LERG waveforms. With a difference in tip position of 25-30 microns, they tend to be making differential voltage recordings between layers occupied by the OS, IS, ONL and INL. Furthermore, with an inner tip of 1-2 microns and an outer tip of 10 microns, the outer tip (and probably both since the probes were being retracted during data taking) is clearly integrating signals from a number of nearby cells. Both probes were at relatively high impedance introducing the likelihood of capacitive coupling. The probes penetrated the retina at an unspecified angle relative to the local surface of the retina and were not correlated with the local signal path geometry, although the sites were apparently extrafoveal.

Another area of difficulty is in the concept of linearity. There seems to be a great deal of confusion with regard to the concept of linearity as it is applied to an input stimulus or signal and as it is applied to a transfer function of a physical process, channel or system. Rodieck & Ford[22] have attempted to clarify this situation by separating the two situations. Whether an input stimulus, or any other signal, can be classified as linear depends on whether it can be represented by an equation of the first degree. Whether a transfer function associated with a process can be considered linear depends primarily on whether the process satisfies the Superposition Principle. This Principle states that the response to two stimuli presented simultaneously is equal to the algebraic sum of the responses to the two stimuli presented separately. They list four corollaries to this Principle:

1. The relative time course of the response is independent of the amplitude of the stimulus.
2. The amplitude of the response is proportional to the amplitude of the stimulus.
3. If the stimulus is inverted the resultant response is inverted.
4. Knowledge of the response to a simple stimulus, such as a 'pulse' or a 'step,' permits calculation of the response to any other stimulus.

They then draw a conclusion that is slightly too restrictive: "The response of a linear system to a pulse or step may thus be said to completely characterize the linear system." This statement is also true for the pulse response of a time-invariant nonlinear system.

The VEP technique will not be discussed in this work because of its limited specificity. Ffytche, Guy & Zeki have made some valid comments on this technique[23]. Heckenlively & Arden present the general measurement regime[24]. Recently, Ffytche, et. al. demonstrate that when combined with auxiliary magnetic intervention (mVEP), it is capable of a quite important degree of temporal precision. However, there is still considerable ambiguity in the precise spatial location of the events monitored and the spatial extent of those events. This leads to ambiguity in the specific type of circuit actually involved.

Any experimentalist in vision would be wise to read the introductory remarks of Rubin & Walls[25] before proceeding into the laboratory.

11.1.2.1.1 Mathematical manipulation of the responses

Historically, the mathematics associated with the field of vision research has been less sophisticated than the biological or psychophysical investigations with which it was associated. The equations of vision found in the literature have always tended toward linearity, even in the face of Weber's Law, Fechner's Law, etc. The CIE Chromaticity Diagram of 1931 was based heavily on the concept that color vision in humans involved linear responses to irradiance and that the various visual channels in the eye were linearly summed in order to create the sense of color. There is overwhelming evidence that this is not the case. In 1989, Flitcroft[26] quotes Naka & Rushton

---

[21]Heynen, H. & Norren, D. van (1985) Vision Res. vol. 25, pp 697-707 & 709-715
[22]Rodieck, R. & Ford, R. (1969) The cat local electroretinogram to incremental stimuli. Vision Res. vol. 9, pp. 1-24
[23]Ffytche, D. Guy, C. & Zeki, S. (1996) Motion specific responses from a blind hemifield. Brain, vol. 119, pp. 1971-1982
[24]Heckenlively, J. & Arden, G. (1991) pp. 400-404
[25]Rubin, M. & Walls, G. (1969) Fundamentals of Visual Science. Springfield, Il: Charles C. Thomas, pg 50
[26]Flitcroft, D. (1989) The interactions between chromatic aberration, defocus and stimulus chromaticity:... Vision Res. vol. 29, no. 3, pp 349-360

from 1966: “For each pigment system, every quantum that is effectively absorbed makes an equal contribution to vision.” One can argue with what Naka & Rushton meant by ‘effectively absorbed’, but the fact is that adaptation is a very nonlinear process and photons absorbed under different conditions of irradiance and time have vastly different impacts. Furthermore, the operation of the L-channel at mezotopic light levels does not satisfy Naka & Rushton’s claim. Never-the-less, this Principle of Univariance has been widely pursued and has led to the use of the ‘Silent Substitution’ method in psychophysics. Flitcroft adopted this tenant and the foundation equations for the CIE Chromaticity Diagram to develop a theory based on matrix algebra of linear homogeneous equations. As will be shown in the next few chapters, the basic signal detection function is grossly nonlinear (it is optimized for change detection and not for good color rendition) and the basic signal processing function is based on logarithmic mathematical manipulations (which relates to multiplication in the underlying signals). **The result is that in the vision process, the appropriate equations,** between input and output, **are not linear, are not homogeneous, and are not of the first order**! This is true even for signal amplitude ranges of as little as 2:1.

When the appropriate arithmetic is used, a wide range of measurements and sensations can be explained without making arbitrary assumptions. These include adaptation, the transition from photopic to scotopic vision (in terms of both the change in spectral sensitivity and the loss of color sensation), the change in temporal frequency response with irradiance (and differential irradiance between signal channels) and the time delay as a function of irradiance found in vision.

11.1.2.1 Polarity

The adjective polar is overused when discussing vision. It is used to discuss the incoming optical radiation, the structure of the nerves involved, the voltage at certain nodes and to describe how the “action potential” is generated/transmitted along the nerve structures.

11.1.2.1.1 Polarization of radiation

In this work, the polarization of the incident optical radiation will be discussed in the conventional engineering sense, the polarization of light is described by the geometrical direction associated with the electrical field vector of the radiation.

11.1.2.1.2 Polarization in morphology

When discussing the structure of neurons, virtually every biology text illustrates the types of neurons using a nomenclature developed by the early histologists. An updated version of such an illustration is shown in **Figure 11.1.2-1**. The histologist defined essentially three types of neural structure based on how many structures emanated from the cell body (alias cyton/parikaryon). Thus, they named the neurons monopolar, bipolar and multipolar. Most authors caveat there figures by saying “these forms are not definitive, the geometry of nerves covers a wide continuum in between these three basic types.” In fact, these designations have no functional significance and lead to endless confusion when discussing the related electronic phenomena. There are three primary structures associated with the neuron. In the figure, the circle represents the nucleus or Soma, the black rectangles indicate the active electronic devices, the Activa, and the areas inside the dashed rectangles are the reticulum or electrical conduits of the neuron. Reticulums are usually separated into two classes, those associated with the input structure of the neuron, and enclosed within the neurites, and those associated with the output structure and enclosed within the axon. Note the term neurite is used to describe two types of input structure, the dendrites and the podites. All neurons contain at least one Activa. All Activa have a minimum of three electrical terminals. These terminals are best described in electronic terminology as the emitter or input, the base or common, and the collector or output terminals. More than three terminals can usually be associated with the Activa. Some of these additional terminals are associated with capacitances. The combination of each Activa with the other circuit elements supporting it, besides the neurites and axon, are defined as the **conexus** of the neuron. This conexus is entirely internal to the neuron and is only observable by high magnification electron microscopy (>50,000x) (A) shows a monopolar neuron with its three electronic terminals shown on the surface of the plasma membrane as dark bars. Note the identical functional organization in (B) for the bipolar neuron. Clearly, how close the nucleus is to the functional parts of the neuron is irrelevant. Whereas (A) and (B) show only the dendrite of significant physical size due to arborization, (C) shows the poditic or base terminal also being of significant size. The result is an easily identifiable three-terminal configuration related to the internal Activa. (D) shows the dendritic terminal further arborized.

Neurites frequently arborize significantly, leading to thousands of input connections[27].

Although there is only one nucleus in a given neuron, it is possible for there to be a great many Activa. (E) shows a typical projection neuron, such as a ganglion cell, containing three internal Activas. The sections between Activas are called internodes, or axon segments, or more appropriately conduit segments. The reason is that conduit segments exhibit terminals suggestive of both dendrites and axons.

The biologist also speaks of the neural cell as "polarized," meaning the signal is usually carried in the direction defined as beginning at the dendrite structure and ending at the pedicel (alias bouton/knob/end foot) of the axon structure. This is a relatively trivial use of the word and it should not cause any problem.

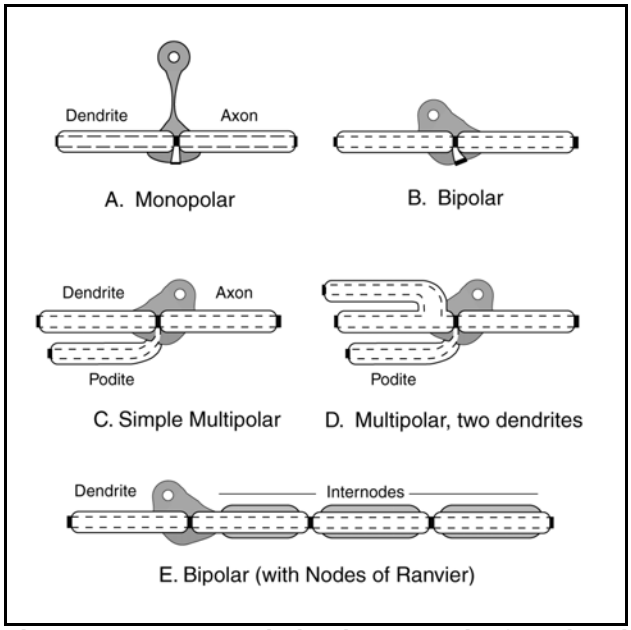

Figure 11.1.2-1 Correlation between the functional and morphological neuron. The nuclei of the cells are indicated by a circle. The Activa within each cell are indicated by black rectangles. The other circuit elements associated with the Activa and completing the conexus or Node of Ranvier are not shown. The reticulum are indicated by areas surrounded by a dashed line. (A), the simple monopolar neuron. (B), the simple bipolar neuron. (C), the simplest multipolar neuron. (D), a multipolar neuron with dual dendritic arborization. (E), A projection neuron containing multiple Activa located at Nodes of Ranvier.

11.1.2.1.3 Potential and Polarization

The term polarization as used to describe an electronic condition in biology is usually used to describe the difference in voltage between the electrolyte on two sides of a membrane making up the wall of a cell. The inside of the cell is normally found to be at a negative potential with respect to an external electrode near the cell wall. The potential is typically 50 to 100 millivolts. This value is normally named the resting potential.

By associating the resting potential with the cell wall instead of the internal electrolyte, many biologists have considered this potential to be similar to the local polarization of a distributed capacitor. Based on the above notion of polarization, a cell wall becomes depolarized when the negative voltage of the internal plasma of the cell is reduced. Conversely, hyperpolarization is taken to be an increase in this negative voltage. This notion has delayed the advance of understanding in the area of neurology. In this work, depolarization and hyperpolarization relate to the relevant plasmas and not to the cell wall.

Figure 11.1.2-2 shows the typical voltage profiles of the three major functional classes of neurons in vision. The profiles focus on the axoplasm of a neuron. There is little data in the literature on the potential of the dendroplasm and podaplasm. However, the voltages of these plasmas are easily deduced from the theory in this work. Some care must be taken, in the case of photoreceptor cells, since some of their plasmas are in close electrical contact with the inter-photoreceptor cell-matrix, IPM, instead of the interneural matrix, INM). Both the photoreceptor and signal processing cells are electrotonic (they process analog voltage waveforms). The projection neurons are phasic and generate action potentials in response to an input signal greater than a specific DC potential.

The reference potential for these figures is the potential of the interneural matrix, INM. In most cases, this potential is significantly different from the circuit ground, V_{BB} of the individual neuron. The circuit ground is usually 10 mV more negative than the reference ground. The intrinsic membrane potential, V_{CC}, is the potential of the axoplasm in the absence of current flow through the membrane. This is not the potential usually reported in the literature. The commonly reported potential is the resting axoplasm potential, V_C. V_C and V_{CC} are approximately equal only in the case of the projection neurons during the interpulse interval. During this interval, the voltage drop related to the

[27] Wassle & Boycott. Physiology Reviews, vol. 71, no. 2 April 1991 pp. 447-480.

current passing through the cell wall is negligible. Under all other neurological conditions, V_C is less negative than V_{CC}. In frame **A** for the case of the photoreceptor cell, V_C is a state variable determined by the average illumination of the cell and the precise state of its adaptation. For this cell, V_C is typically -45 mV relative to the interneural matrix. The vertical bar shows the instantaneous dynamic voltage range of this cell to be approximately 35 mV. This range begins at the dark adapted set point of about -25 mV. In these cells, an increase in light causes the output voltage to approach the intrinsic electrostenolytic potential. This has traditionally been defined as hyperpolarization of the cell. The prefix hyper- is unwarranted in terms of the functioning of the circuit.

The signal processing neurons operate at a different set of nominal voltages than the photoreceptor cells as shown in frame **B**. In this case, the intrinsic electrostenolytic potential is typically -154 mV relative to the INM. However, the Activa is typically drawing current through the axolemma continuously. This results in a continuous voltage drop across the power source. The resulting axoplasm potential is typically near -70 mV and it can vary instantaneously over a span of about 130 mV. Its most negative potential is controlled by V_{CC} and its most positive potential is controlled by V_{CSAT}. V_{CSAT} is the lowest output voltage achievable by the Activa while the collector remains reverse biased. Note the overlap between the dynamic range of the photoreceptor cells and the signal processing cells.

The literature is unclear as to whether bipolar cells accept signals at their poditic terminals and are therefore biphasic in their output. For the following discussion, it will be assumed that only lateral cells can act as inverting amplifiers. This assumption has no impact on the overall theory.

For photoreceptor signals applied to the non-inverting input of a typical lateral cell, the gain of the signal processing cell is ~1.0 and the output at the axon of the signal processing cell is nearly identical in voltage to the input signal from the axon of the photoreceptor cell. However, in this case, the hyperpolarization in the photoreceptor has resulted in a depolarization (a reduction in the instantaneous voltage of the axoplasm relative to the INM) of the second cell relative to its resting potential.

For a signal from a photoreceptor cell that is applied to the inverting terminal of a lateral cell, the situation is different. The synapse between the cells must be of the appropriate size to shift the average DC potential of the signal while maintaining the voltage gain of the combination of the synapse and the inverting input circuit at or near 1.0. After this shifting and amplification, the instantaneous output signal of the signal processing neuron is nearly identical in shape to the input signal but it has been inverted and its DC level has been shifted by about 70 mV. The dark adapted set point of the photoreceptor cell is now represented by a voltage near -140 mV in the signal processing cell and a hyperpolarization in the photoreceptor cell due to an increase in illumination results in a depolarization in the signal processing cell, i.e., the signal is driven toward the resting axoplasm potential.

If two photoreceptor cells are excited equally and their signals are provided to opposite inputs of a lateral cell, the output level of the lateral cell remains essentially unchanged at its resting potential.

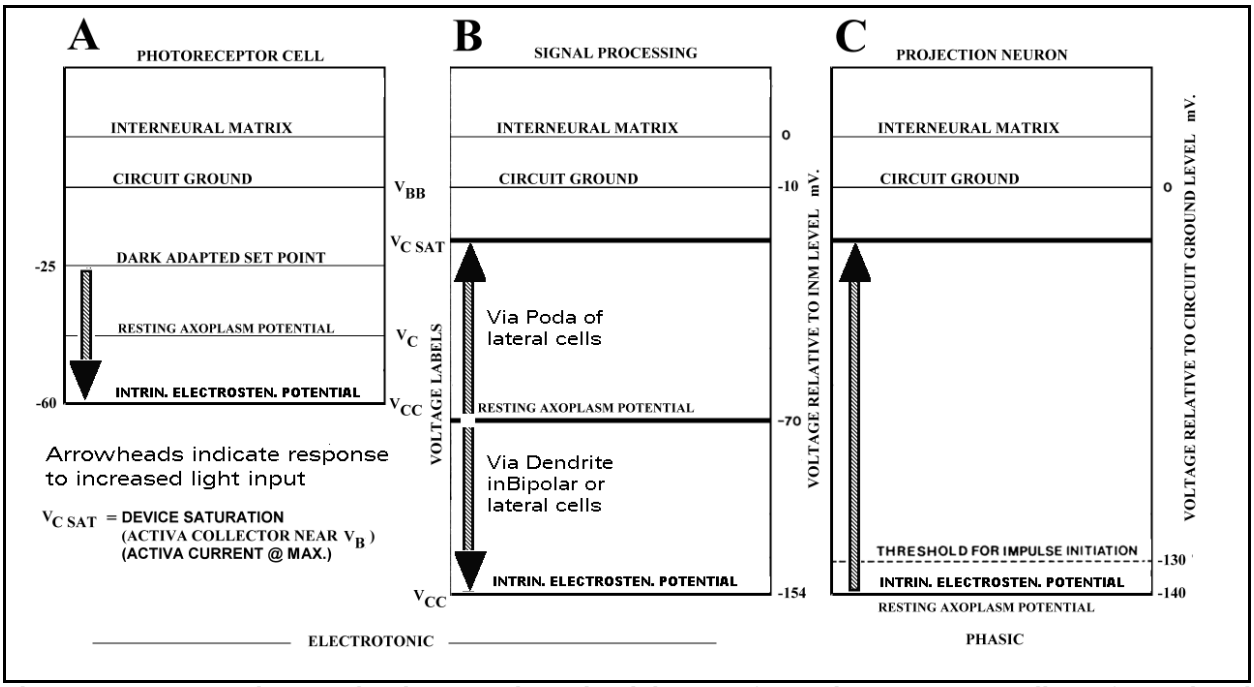

Figure 11.1.2-2 Fundamental voltage regimes in vision. A, for a photoreceptor cell. B, for a signal processing cell. C, for a signal projection cell.

Frame **C** illustrates the electrical profile of the third neuron type, the projection neuron. The potentials applied to the cell are essentially the same as those of a signal processing neuron. However, the cell is used in two diametrically opposite operating modes.

For those projection neurons designed to accept mono phase input signals (typically parasol ganglion cells accepting luminance signals), the internal Activa is normally biased at cutoff. In the absence of an input signal, no current flows through the internal Activa. The resting axoplasm potential is equal to that of the electrostenolytic process supporting that axon. For input signals below the threshold voltage for action potential generation, the projection neuron acts as a linear amplifier with a gain of about 1.0. The output terminal will display an output waveform that is a copy of the instantaneous input waveform. However, it will be at a DC value contained within the voltage range shown as -140 to -130 mV. However, for input signals exceeding a fixed DC potential, the circuit becomes unstable. It becomes a monopulse oscillator if the input signal exceeds the fixed threshold only briefly. Otherwise, it becomes a continuous pulse oscillator. When in the oscillatory mode, all of the generated pulses are of the same amplitude. This amplitude is given nominally by the voltage V_{CC} -V_{CSAT}. It is possible for a measured waveform to exceed this maximum amplitude depending on the size of the capacitances used in the oscillatory circuit. The voltage at -130 mV is labeled as the threshold for impulse initiation to conform to practice in the literature. There is no threshold in the axolemma potential. This level is the equivalent voltage to the actual response threshold voltage in the input circuit of the projection neuron.

For those projection neurons designed to accept biphasic input signals (typically midget ganglion cells accepting chrominance, or other bipolar signals), the internal Activa is normally biased so as to oscillate continuously. The axoplasm potential is typically -140 mV during the interval between pulses and the amplitude of the individual pulses is as described above. The time interval between adjacent pulse peaks is a nearly linear function of the instantaneous biphase voltage at the input terminal of the projection neuron.

As developed elsewhere in this work, all Nodes of Ranvier are also monopulse oscillators employing an Activa circuit identical to those discussed above that respond to monophase input signals. They are normally biased to cutoff and only generate an action potential in response to an input signal exceeding the threshold voltage of its input circuit.

Historically, it has been very difficult to measure the voltage of the dendritic or poditic plasmas. It has been conventional to speak of the potential of the axoplasm as the potential of all of the plasma inside the cell. As seen above, this convention is misleading and untenable.

The intercellular material in any organism is usually a very good insulator compared to metals, xxx 10^{xx} ohms per square (no units are needed here). And cell walls are usually very good insulators, except when disturbed. Although the dielectric constants involved are small, these walls are thin and relatively small numbers of electrons or ions can generate significant voltages.

Most of the biological literature, with its strong leaning toward chemistry, does not consider the transfer of charge along the surface of a membrane. It concentrates instead on the movement of ions perpendicular to the membrane surface. Furthermore, most biological literature does not consider the effect of individual electrons or ions on a membrane, preferring to work in units of moles or gram ions. This will not be adequate here where, in the limiting case, we are examining the signal generated by a relatively small number of photons (i. e. hundreds or less in the limiting cases). In addition, the distinct possibility of individual ions traveling along the surface of a membrane must not be ignored, especially in the case of the retina where distances can be extremely short (10's of microns). Thus a potential difference of one millivolt between the external surfaces at the opposite ends of a one micron long structure generates an electrical field of 10^3 volts per meter, a very significant field strength.

Juusola, et al. has recently provided a limited tabulation of some axoplasm and neurite potentials in general agreement with the above discussion[28].

11.1.2.2 Electrical fields

In the biological field, the vast majority of the older experimental work involved skeletal (alias principal/projection) neurons used to communicate over relatively long distances. More recently, much effort has been applied to describing the actions of signal processing (alias local circuit or /inter-) neurons.

The skeletal neurons exhibit electrical pulses traveling along the axon structures. These pulses are usually described in terms of "action potentials" in contrast to the resting potential. The "action potential" found in the giant Squid axon is found to have a peak amplitude of +95 millivolts. Several investigators have related this voltage to an assumed -60 millivolts resting potential for the axoplasm. Under this assumption, the axoplasm would become positive with respect to the surrounding matrix during the period of the action potential. As seen in **Figure 11.1.2-2(C)**, the axoplasm resting potential of projection neurons is normally -140 millivolts. Based on this value of the resting potential, the axoplasm is depolarized during the action pulse interval. However, there is no reversal of the potential across the axolemma.

Recent neurological literature has been concentrating on the dendritic portion of the skeletal and signal processing neurons. It recognizes that signaling may occur in an analog instead of a pulse mode. In this case, a sustained potential different from the resting potential is described as a "electrotonic potential." Smith indicates these electrotonic potentials (which are usually smaller in amplitude than the potentials associated with action potentials) are usually small depolarizations of a nerve cell membrane and are caused by essentially passive spread of electrical current through the conducting fluids inside and outside nerve cells and their extremities[29] . As indicated above, many of the fluids involved in the signal processing environment may not be highly "conducting" and surface transport must also be considered viable. Frequently, the fluids are liquid crystalline in nature and ionic transport may not be possible. In this case, electron and hole transport may be the method of charge transport through the medium.

Smith also notes that, especially in the retina, "there are very few if any voltage dependent sodium gates present." Furthermore, the electrotonic signals recorded in the signal processing neurons of the retina are both positive going and negative going with respect to the axoplasm resting potential. This work proposes these electrotonic potentials are actually small polarizations and depolarizations of the axoplasm and not of the axolemma.

No report could be found in the literature where an investigator actually measured both the absolute resting potential of the axoplasm and the absolute amplitude of the action potential in the same projection neuron with respect to the surrounding fluid matrix.

11.1.2.3 Signal transfer between neurons

---

[28]Juusola, M. French, A. Uusitalo, R. & Weckstrom, M. (1996) Information processing by graded-potential transmission through tonically active synapses *TINS* vol. 19, pp 292-297, figure 4C

[29] Smith, C. (1989) Elements of molecular neurobiology. NY: John Wiley & Sons. page 257

At the present state of the art in biophysics, the subject of how signals are passed from the axon of one neuron to the dendrite of a second neuron is expanding. Until recently, to imply the transfer of a signal by means other than the physical transport of ions was blasphemous. The existence of so-called gap junctions (alias electrical synapses) between neurons is now becoming accepted. However, their characteristics are not completely defined. Smith gives a brief description of a gap junction. However, the concept will be broadened here to accept the case of current sources as opposed to voltage controlled gates. Smith does define a crucial aspect of a gap junction, a hydrophilic pore created at the junction of two hydrophilic canals, one in each cell. He speaks of this pore being controlled by a voltage. Most morphologists speak of a reticulum inside the axon. It appears the canal is a conceptual version of the reticulum of morphology.

As developed in this work, Smith's concept of two canals connected by a pore can be replaced by two reticulum connected by a three terminal Activa. The Activa consists of the two neurolemma, the axolemma and the neurolemma separated by a liquid crystal of hydronium. The electrical transmission of the Activa is controlled by the potential between the axoplasm within the reticulum and the surrounding fluid. The potential of the orthodromic neuroplasm is inconsequential as long as it is more negative than the axoplasm and the surrounding fluid. Under these conditions, a current is transferred to the neuroplasm in response to the potential of the axon. The result is an all-electronic synapse. No other chemicals are required within the active region of the synapse. The all-electronic synapse is unidirectional because of the required bias potential between the neuroplasm and the surrounding fluid.

11.1.2.3.1 Cascaded Circuits

A problem of paramount importance in the visual system is that the circuits of the signal detection, signal manipulation and signal projection stages (stages 1, 2 & 3) are directly coupled. This method of coupling is not widely used in man-made electrical circuit designs for a variety of reasons. The primary reason is the difficulty of maintaining the signal in the center of the electrical voltage passband of subsequent circuits. For more than three circuit stages, the stability of the circuit usually requires provision of an external feedback loop. Otherwise, the signal is usually truncated at either high or low signal levels as it passes down the signal path.

11.1.2.4 Signal control

Smith describes the "membrane 'gates' or 'channels'" connecting two neurons as dividable "into two categories: those which are controlled by **chemical molecules (ligands)** and those which are controlled by transmembrane **voltage**." He continues: "In fact this division is not absolutely clear cut . . . Nevertheless, for the purposes of exposition, we shall consider each in turn." These words indicate clearly how tentative the science was in this area as of 1989. In fact, the words gates and channels are not synonyms in any context. A voltage may be used to control some other entity which carries a signal but a ligand as used here is clearly a signal carrying entity and not a controller. If the signal were to be carried by electrons instead of ligands (or ions), then it is quite reasonable to think in terms of a current source as generating these electrons as an alternative to a voltage controlling these electrons. The concept of an equivalent current source is always available when one speaks of a voltage controlled current, i. e. Thevinen's Theorem. In this case, with the high impedances involved and the presence of 'channels', a current source may be the more appropriate concept.

11.1.2.5 Signal measurement (Reserved)

This section is provided to collect and summarize the terminology and major concepts discussed in this and previous chapters.

11.1.2.5.1 Voltage polarization (Reserved)

It is important to point out that the topography of synapses associated with the pedicels of photoreceptor cells and the topography of Nodes of Ranvier are essentially convolutes. The signal path is along the axis of a Node of Ranvier and the electrostenolytic reservoirs surround it. In the case of the pedicel, the electrostenolytic reservoir is frequently in the center and the synapses are grouped around it.

11.1.2.5.2 Signal isolation

Padgham & Saunders have illustrated the limited spatial precision used in establishing the circuit used in and ERG

or EEG[30]. They have also illustrated the difficulty in filtering the low level signals normally reported and the resulting poor precision in the elements of the resulting waveforms.

Figure 11.1.2-3 attempts to illustrate the importance of positioning the return electrode carefully while also illustrating the topographical location of the source of various waveforms. In the literature, the placement of this electrode is usually discussed in terms of finding an innocuous, or "indifferent" neutral location, such as the outside of the eye cup, the back surface of the RPE, etc. In fact, these are very poor locations for obtaining meaningful information. The return probe should be located in the INM or IPM as close to the probed location as possible in order to avoid extraneous signals due to "common impedances." When two circuits share a common impedance, the total signal measured by a probe located in one circuit will contain signals related to the other circuit. This is an especially important consideration in high impedance circuits where capacitive coupling as well as resistive coupling can occur. The problem is endemic to experiments in vision research where it is common to present waveforms clearly showing the summation of a low level analog (electrotonic) waveform measured at a high impedance point corrupted by a pulse (action potential) waveform capacitively coupled into the test circuitry from a higher voltage, and lower impedance, signal source. In this case, the extraneous pulse waveform frequently has a voltage higher than the desired electrotonic waveform. This result is easily calculated from the ratio of the impedances involved.

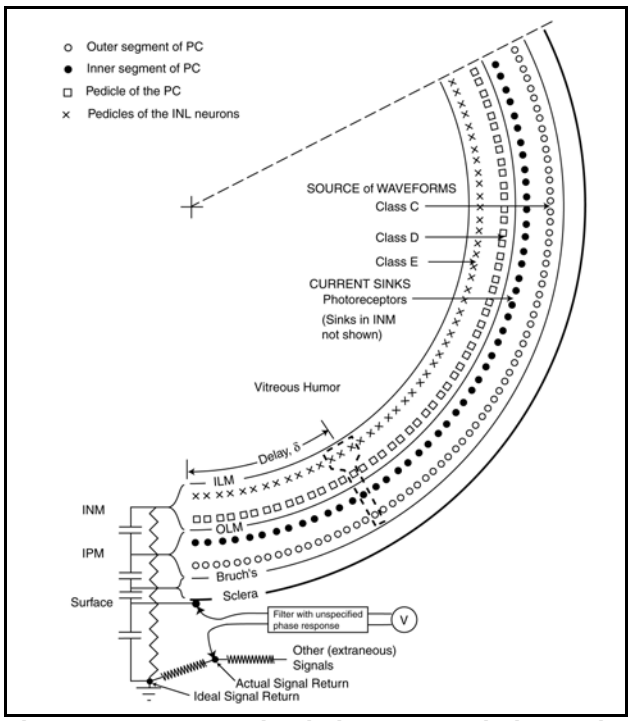

Figure 11.1.2-3 ERG lead placement relative to the topography of the retina. The source of the different classes of waveforms are shown on the right in this cross-sectional view of the retina. The ILM, OLM, Bruch's membrane and the sclera are shown as insulators shunted by conductive impedances on the left. The black dot indicates a signal probe in contact with the outer surface of the sclera. The return probe is shown connected to an "indifferent location." This probe location makes the test set susceptible to extraneous signal pickup.

A related problem endemic to vision research is that of attempting to measure a waveform at a point of extremely high impedance. Frequently, the bandwidth of the original signal is wider than that of the test set used. The waveform obtained is therefore not that of the original signal but that of the test set. This problem is further complicated by the "loading" caused by the presence of the test set probe near or in contact with the signal node of interest. In this case, the resulting waveform may be a complex function of the combination of the test set and the signal under examination. This will be discussed in later sections when it will be pointed out that an open gate field effect transistor cannot normally be used in the testing of very high impedance circuits since the operation of the transistor is not symmetrical with regard to rapid voltage swings about a bias point (a resting potential). In these situations, it is not adequate to purchase and utilize a standard amplifier circuit without demonstrating that the performance of the amplifier alone is well understood and not distorting the signal waveforms.

Baylor et. al.[31] have obtained some very high quality data while operating at the very low signal levels associated with the OS. By making current mode measurements, which are inherently low impedance, their data shows no discernable interference from higher level pulse waveforms generated in the nearby Inner Nuclear Layer.

11.1.2.5.3 Isolation from stimulus generator

[30]Padgham, C. & Saunders, J. (1975) The Perception of Light and Colour London: G. Bell & Sons pp 32-36
[31]Baylor, D. Lamb, T. & Yau, K. (1979) *ibid*

Figure 3 of Baylor, et. al. (1984) shows a particularly insidious type of interference problem[32]. When acquiring data at very low signal levels, it is generally not recognizable. However, when summing multiple data traces to improve the signal-to-noise ratio of the overall data collection, the synchronous noise associated with the stimulus generation circuitry is enhanced rather than suppressed. In this case, a nominally 15 Hz sinewave contaminates the data.

11.1.2.5.4 Decoupling

In electronic circuits, if the individual amplifier circuits are powered by a weak source, the source can be modeled as a voltage source and an associated resistance. The result under large signal conditions (see the following paragraphs for definitions) is that the measured signal will be corrupted by a signal related to the power source. The same situation occurs in the operation of the eye. The power source, the metabolic supply system, is tailored to be "just adequate" under normal conditions. Under large signal conditions, a drift or transient is observed in the measured signal which is directly relatable to the inadequacy of the metabolic system to provide an "absolutely stable" power supply to the photodetection process. In the case of the eye, this large signal problem has three ramifications:

1. The signal performance of an individual signal path will exhibit variable performance under different input illumination levels, including exhibiting a history of previous illumination, due to the marginal performance of the metabolic supply.

2. The signal performance of an individual signal path will exhibit variable performance based on the input illumination appli metabolic supply. This is frequently labeled "cross-coupling" if the signals are of relatively high frequency.

3. The signal performance of an individual signal path will vary with its position in the retina due to the combination of both of the above effects.

Adjacent signal path coupling is a critical factor in the eye. The decoupling is usually inadequate under experimental conditions due to the application of large illumination signals. Because of geometric limitations of the metabolic supply available, the performance of the eye will vary in many ways with respect to geometric position in the retina.

11.1.2.5.5 Adaptation

Because of the relatively poor characteristics of the metabolic supply in the eye, the performance of the eye can exhibit significant temporal changes over short intervals based on the prior illumination history of the eye, i. e. the eye illustrates significant fatigue effects. When these fatigue effects are combined with the decay characteristics related to the photodetection process in the OS, these effects are generally described under the term adaptation. The eye exhibits significant sensitivity variations with time after either an increase or a decrease in average input illumination. These changes vary considerably with position in the retina. The actual circuit used to provide the mechanism of adaptation is unique and based on the unique topography of the photoreceptor cell. This topography introduces a nonlinear mechanism into the collector structure of the first Activa found within the photoreceptor cell. This mechanism causes the circuit to exhibit a very large change in amplification factor as an inverse function of the input stimulus to the circuit. The result is an output that is essentially constant regardless of the input amplitude. The details of this process are presented in **Section 12.5.3**.

11.1.2.6 Signal classification

Because of the dynamic range of the vision system, it is necessary to carefully specify the signal level being applied to the eye, processed by the eye and/or measured in the eye. This is normal practice in the electronics field where the circuits normally perform in a linear mode when small signals are applied to them--and in a nonlinear mode when large signals are applied to them. A convenient criterion in electronic circuits is to classify the signals based on their amplitude as a percentage of the supply voltages. Large signals have amplitudes which are a significant percentage of the supply voltages; whereas, small signals always have an amplitude that is below 10% of the supply voltages.

---

[32]Baylor, D. Nunn, B. & Schnapf, J. (1984) The photocurrent, noise and spectral sensitivity of rods of the monkey Macaca fascicularis. *J. Physiol*. vol. 357, pp 575-607

In the case of vision, the signal circuits will be shown to exhibit an instantaneous maximum dynamic range of about 100:1, two orders of magnitude. The scene contrast applied to the eye is normally less than 6:1 for natural scenes and only approaches 20:1 as a maximum when viewing black ink on white paper. The average natural scene contrast is about 2:1 on an RMS (root mean square) basis.

The eye is normally exposed to low contrast, or small signals, except under abnormal conditions.

Much psychophysical testing is done under abnormal conditions. In fact, much of this testing involves the application of a large signal followed by a threshold test; this involves a combination of a large signal and small signal conditions and any analysis of the resulting data must account for the response of the system to both signals. Because of the transient response associated with all large signals, it is frequently found that the small signal threshold tests are a function of time from the application or cessation of the large signal.

In the case of large signals, the photodetection process in the eye exhibits a logarithmic response. This type of response exhibits interesting properties which will be developed below in later sections of this work.

11.1.2.6.1 “ON” and “OFF” in psychophysics

For those not initiated into the psychophysical community, a brief word is needed about “on” and “off” signals. These terms are not used in their literal sense. The psychophysicist uses these terms to describe the modulation of a baseline or carrier irradiance level. An “on” signal or excursion merely increases incrementally the total irradiance. An “off” signal does not turn off the irradiance. It merely reduces the irradiance level by turning off the incremental value associated with the “on” signal. They then report activity resulting from these changes. Thus, in a large signal situation, they will report activity resulting from an “off” signal which is actually related to the establishment--or reestablishment--of a new, non-zero, irradiance level.

11.1.2.7 Resonance

The term resonance must be carefully defined when discussing the eye. It is used differently in different disciplines and in fact appears in different forms in different functional areas.

The molecular structures found in the eye and utilized for photodetection are found to be quantum-mechanically resonant at wavelengths directly related to the wavelengths of visible light. Mathematically, their operation can be described in terms of second order differential equations.

The overall photodetection/electron excitation process performed by these structures does not involve a resonant process. This quantum statistical process can be described in terms of first order differential equations.

The basic analog signaling structures found in the eye do not exhibit either physical or electronic resonance; they only exhibit resistive and one type of reactive (in all known cases, capacitive) impedances. The responses of these structures are described mathematically by first order differential equations. However, it is possible through feedback of signals within the signaling structures to create electrically resonant circuits. The performance of these circuits must be defined utilizing second order differential equations.

11.1.2.8 Bandwidth and risetime

11.1.2.8.1 Spatial versus temporal frequency response

> Before proceeding, it is important to provide *a cautionary note*. It is common in psychophysical, and some electrophysical, experiments to attempt to measure the spatial frequency response of the eye utilizing temporally varying signals as a substitute. Frequently, bar or checkerboard patterns of variable pitch are caused to reverse in contrast as a function of time. These experiments must be extremely carefully designed and compared to a relevant model if results are to be obtained that can stand close scrutiny. This is because of two things; the complex relationship between spatial features in the object plane and the temporal signals generated in the eye/brain *and* the determination of what is being measured in the experiment--the apparent spatial response measured at the output of a differencing amplifier, such as an amercine cell, can exhibit a different signal depending on the temporal phasing of the chromatic input signals. If the phasing is not carefully specified, confusion will result during the analysis and evaluation of the data.

11.1.2.8.2 Individual circuits

The description of the temporal performance of a signaling circuit can be relatively complex since it may be quite different under small signal and large signal conditions. Under large signal conditions, first order (i.e., non-resonant) circuits exhibit a temporal response consisting of one or more exponential response functions to a step input. The temporal response is usually described by way of the time constant of these exponential waveforms.

The output waveform presented by a first order circuit is quite distorted relative to the input signal. However, this is not of technical significance here. The eye does not make use of the *shape* of the input illumination waveform as a function of time. However, the effect is quite often presented in experimental waveforms and the source of this distortion should be understood.

11.1.2.8.3 Types of temporal frequency filters in vision

The filtering topology and technology used in vision is far more sophisticated than recognized in the literature. The typical animal vison system employs at least four unique types of temporal frequency filters. To fully understand the effect of these filters, it is necessary to have some formal training in filter design. The concepts of poles and zeroes in filters is the key. The properties of the filters found in vision are unusual and only found in sophisticated man-made systems:

1. **A high frequency pole that is a function of illumination** –The quantum-mechanical photodetection/de-excitation process creates a high frequency pole that is sensitive to the illumination level.

2. **A zero at zero frequency** –The adaptation mechanism found in virtually all photoreceptor cells introduces a zero in the frequency response of the visual system at zero frequency.

3. **A conventional high frequency pole**–Many circuits in the visual system exhibit an RC-type high frequency pole due to the effect of circuit capacitance and the effective resistance of the circuit.

4. **A pre-emphasis circuit**–A circuit with a pole in both the numerator and the denominator of the transfer function that are at only slightly different frequencies can cause a narrow peak in the frequency response of the overall circuit.

5. **A sampling induced pole that is a function of the magnitude of a bipolar signal amplitude**–The encoding of bipolar information by the midget ganglion cells introduces a special form of filtering that results in a high frequency pole with a magnitude that is proportional to the signal amplitude.

When a signal is passed through a cascaded series of these different types of filters, the results cannot be accounted for using lumped constant filter theory. It is this fact that makes it so difficult to interpret the papers in the literature. Many authors have attempted to account for the performance of the visual system using multi-stage lumped constant filters. There are, in fact, no multistage lumped constant filters in the visual system and the use of a lumped constant filter as an analog of a dynamic filter is inappropriate.

It has also been common in the literature to consider the basic character of the visual passband a low pass filter. The presence of a zero in the system due to the adaptation mechanism makes this inappropriate. The simplest representation of the actual visual system from the pedicle of the photoreceptor cell on is a bandpass filter with a low pass pole near 0.3 Hz and a high pass pole below 100 Hz (depending on the illumination level).

The performance of these filters will be discussed individually in the following sections and chapters.

The unusual presence of the poles related to the P/D process suggests an unusual methodology when discussing the temporal frequency response of the visual system. Since the rest of the visual system associated with achromatic vision exhibits fixed temporal parameters, it is tempting to treat the output of the P/D mechanism as the stimulus for the electronic portion of the achromatic visual system. This approach allows the electronic system to be described using lumped constant filter theory. In this approach, only the quantum-mechanical stimulus to the electronic system, the A waveform defined below, is a signal with variable temporal parameters that are proportional to the illumination. Although this stimulus cannot be measured directly, the Class B waveform defined below can be measured. Under small and medium signal transient conditions, it is a good representation of the actual Class A waveform.

11.1.2.8.3.1 A bandpass filter combined with pre-emphasis

An area leading to confusion in the visual science involves the combination of a pre-emphasis circuit with an underlying bandpass circuit. This can lead to a circuit that exhibits “peaking” at the high end of the overall response. Some people look at just this peak in the response and speak of it as a bandpass filter while others look at the relatively flat lower frequency portion and describe it as a low pass filter. It should be pointed out that the so-called bandpass portion is merely the peaked portion of the overall filter response. The pole at 0.3 Hz. prevents this overall filter from being described as a low pass filter.

11.1.2.8.3.2 The variable pass band of the P/D Equation

The P/D equation exhibits a low pass filter characteristic where the cutoff frequency is a function of the irradiance. This means that the frequency band of the information passed down the visual signal processing chain varies with the irradiance level. At lower irradiances, the signal bandwidth is reduced. Details of this phenomenon will be found in Appendix A.

11.1.2.8.3.4 Combined variable cutoff low pass and pre-emphasized band pass filters

When the generator function created by the P/D equation is combined with the following circuits that include a pre-emphasis circuit, interesting results occur. At low irradiances, the signal frequency spectrum is limited by the low frequency characteristic of the P/D equation; the overall spectral response of the signal information appears to be that of a low pass filter. If the irradiance level is increased, the limiting frequency response of the P/D equation rises and the signal frequency spectrum becomes wider. The shape of the overall frequency response is now controlled by the stage containing the pre-emphasis circuit. The result is that the high frequency spectrum is boosted relative to the low frequency spectrum. The resulting spectrum is frequently labeled a bandpass characteristic because the investigators are not focused on the low frequency region which is still present. This feature stresses the importance of maintaining close control of the irradiance level on the retina when making spatial frequency measurements.

11.1.2.8.3.5 Combined variable cutoff low pass and simple high pass filters

When the generator function created by the P/D equation is combined with the following circuits that include a high pass circuit, even more interesting results occur. In this case, it is possible for the P/D equation at low irradiance to generate a signal that contains no information in the pass band of the high pass circuit. As the irradiance level rises, the cutoff frequency of the P/D equation rises until it overlaps and exceeds the lower cutoff frequency of the high pass circuit. Signal information now appears in the high pass circuit signal channel. This process appears to be a major factor in the transition from scotopic to photopic vision.

11.1.2.8.4 Scene information modulation by saccadic motion-TREMOR

As discussed above, the intrinsic animal visual system is blind to information in the far field that is not moving relative to the line of fixation. This is a major limitation in the visual system. Without an evolutionary fix for this problem, it is doubtful that the higher animals would have evolved at all. To overcome this limitation, the higher animals have introduced tremor as a mechanism of ocular muscle movement. Tremor, sometime referred to as micro-saccades, consists of a low amplitude signal that is continuously applied to those muscles except during other larger saccades. Tremor has a nominal temporal frequency of 50 Hz and a nominal amplitude of 30 arc seconds (or 1.5 photoreceptor cell diameters). By moving the line of fixation continuously, all of the information in the scene is modulated about this 50 Hz carrier. The effect is to move all of the information related to the scene into the passband of the visual system, 0.3 Hz to a nominal 100 Hz. To recover this information, the visual system must synchronously demodulate the neural signals related to luminance and other signals (such as chrominance) at a later point in the system.

A major advantage of introducing tremor is that the physiological spatial resolution of the visual system is not limited to the pixel size of the detectors. Typically, man exhibits a limiting spatial resolution that is about a six times better than that predicted by the size of the photoreceptors.

An important consequence of introducing tremor to overcome the limitations introduced by adaptation is the fact that large areas of constant irradiance in the scene will still not be modulated by the tremor mechanism. Upon demodulation of the scene information, the brain will still not have information available concerning the luminance or chrominance of these areas. To compensate for this fact, the brain employs a “paint program” to compute the luminance and chrominance as part of the vectorization process. It uses the luminance and chrominance values it receives from points along the perimeter of a constant area, a contrast boundary, to compute these values. This technique is also used to eliminate the blindspot associated with the optic nerve usurping part of the retinal area at its

point of exit.

Failure to recognize the importance of tremor in the signaling system of vision has led to great experimental difficulties in the laboratory.

The fine continuous tremor of the eye has a significant effect on the temporal frequency content of the signals in the neural system. If an edge is projected on the retina and the tremor causes a photoreceptor to continuously move back and forth across this edge, the effect is to modulate the irradiance signal applied to the photoreceptor at the frequency of the tremor. Under this condition, the signal may or may not be passed by the photoreceptor depending on the irradiance level. If it is passed,

1. it may or may not be pre-emphasized by a pre-emphasis circuit.

2. it may or may not be delivered into a high pass circuit channel for further processing.

This process is a major contributor to the color vision capability of the eye under dark adapted conditions. If the intensity is adequate, a point source of colored light will be recognized without difficulty on a black field.

Tremor has recently been introduced into man-made two-dimensional sensor arrays to achieve the increased spatial performance offered by tremor[33,34].

11.1.2.8.5 Cascaded Circuits

Because of the nonlinear performance of first order circuits when used in signal processing, the mathematical description of the bandwidth and rise time of the overall circuit is difficult. The experimentalist usually will just measure the output waveform as a function of the input illumination and describe the overall transfer function as due to an “equivalent” single stage first order circuit. This will frequently leave artifacts in the output waveform unaccounted for.

These circuit characteristics will be developed more completely in the following chapters.

11.1.2.9 Delay

The subject of delay, as used in vision research, is not usually defined clearly. Many authors have tried to interpret the delay times they measure in terms of the delay associated with lumped constant electrical circuits alone. They have largely or completely ignored the idea of transport delay due to the finite and relatively low velocity of transmission of electrical signals over biological pathways. These velocities are frequently in the 10 meter/second range as opposed to the velocity in metallic circuits, which is usually in the 10^8 meter/second range. This difference of 10^7 indicates the effects of low transport velocity must be considered in the calculation of the transport delays found in biological circuits. In most cases, the transport delay will predominate over lumped constant circuit delays.

It is very easy in biological circuits to have a transport delay that is larger than six or seven stages of conventional (copper wired) lumped constant delay. An investigator should always evaluate all contributing factors to the measured delays before beginning an analysis which assumes all delays are due to lumped constant effects.

The definition of delay used in small signal and large signal circuits is usually quite different. In small signal circuits involving sinusoidal signals, it is usually easy to define the delay in terms of the phase shift of a sinusoidal signal measured at one location compared to the same signal measured at another location. This phase shift will usually be a function of the frequency. If the delay times associated with these delays is plotted, the delay may be independent of frequency or it may contain an offset from zero that is independent of frequency. If it does, this constant value is defined as a group delay. A group delay is usually related to a transport phenomenon, not a lumped constant phenomenon. A frequency dependent portion of time delay is usually related to a circuit containing lumped constants. It is not possible to designate the nature of a particular small signal delay without taking measurements at several frequencies.

---

[33]Landolt, O. & Mitros, A. (2001) Visual sensor with resolution enhancement by mechanical vibrations *Autonomous Robots* vol. 11, pp 233-239

[34]Viollet, S. & Franceschini, N. (1999) Visual servo system based on a biologically-inspired scanning sensor *Sensor Fusion and Decentralized Control in Robotic Systems*, vol. II Bellingham, Wa: SPIE pp 144-155

In large signal circuits, such as generally found in the visual system, the delay is usually related to the time between the occurrence of some feature of a waveform at two different places in the overall circuit. If the waveform is sufficiently accurately reproduced as to allow for this measurement, the type of delay involved is invariably of the group delay type, i.e., all frequency components of the waveform are delayed by an equal time interval, and group delay is a transport delay phenomena--not a lumped circuit delay phenomena.

11.1.2.9.1 Transport delay as a tradeoff

It appears and it will be assumed that all of the neural leads related to vision within the retina are unmyelinated. As the axons of the ganglion cells exit the eye, they may become myelinated after reaching the lamina cribrosa[35]. Prince did not further describe this feature. However, it is probable that the myelination extends to the connection of the optic nerves to the midbrain. Myelination reduces the shunt capacitance of the axon and thereby increases the propagation velocity (or transport velocity) of the signal. It is interesting that in the retina where distances are short, lower transport velocity has been accepted in order to achieve increased component packing density. The resultant higher transport delay, that varies with location in the retina, is an integral part of the overall design.

11.1.2.10 Feedback

Feedback involves the reinforcement of an input signal by a portion of the output from any amplifying device. The reinforcement can be positive or negative, or at any phase angle in between. It can also be internal or external (involving a signal path external to the basic circuit). In complex signaling systems, feedback can occur in a variety of scenarios--some not too obvious. In a circuit built from discrete components, it is usually easy to see the external (extrinsic) feedback loops around the major active elements. It is sometimes more difficult to recognize feedback due to common elements shared between two different current paths. This type of intrinsic feedback usually occurs where impedances are placed in the common base lead of man-made and biological transistor circuits. In fact, this type of intrinsic feedback is present in every transistor because of the inherent impedance associated with the base terminal of the device itself. Thus, virtually every active circuit involves internal feedback. The importance of this feedback is only a matter of degree. In most man-made transistor devices, the internal feedback due to the base impedance, r_b, is only important at high radio frequencies and above.

Typical circuits employing internal negative feedback are the lateral cells. Typical circuits employing internal positive feedback are the ganglion cells and the Nodes of Ranvier. There are no clear instances of external positive or external negative feedback in the biological circuits of vision.

11.1.2.11 Charge transport

Electrical charge transport is accomplished quite differently in different forms of matter. **Table 11.1.11** gives a brief summary of the situation:

Table 11.1.11
Charge carriers in matter

		----------Semiconductors*----------		
Particle	Metal	Solid Crystal	Liquid Crystal	Liquid
electron	xx	xx	xx	
negative ion				xx
positive ion				xx
hole		xx	xx	

*The designation semiconductor implies a difference in net conductivity with direction through the material.

Very briefly, charge is transported exclusively by the movement of free electrons in metallic materials.

In the case of liquids, the situation was found to be different, and more complicated, long ago. In dilute solutions,

[35] Prince, J. (1956) Comparative Anatomy of the Eye. Springfield, IL: Charles C. Thomas Pg. 43

not involving a semipermeable membrane or different types of electrodes, the charge is transported by the physical diffusion of both positive and negative ions through the solute under the influence of the externally applied electric field. Depending on the mobility of the various ions present, one ion may be the dominant current carrier. The current can then be described in terms of that ion type alone; i. e., the current can be described as n-type if it is carried primarily by the negative ions and p-type if carried primarily by the positive ions.

If the two electrodes are different, there will be an internally generated electric potential due to the difference in potential of the half-cells associated with each electrode and the electrolyte. The charge will still be carried by both positive and negative ions, and one type may dominate as above.

In discussing a semipermeable membrane, a distinction must be made between a membrane that is semipermeable to different size or chemical class of particles (hydrophilic, hydrophobic, etc.) and one that is semipermeable to electronic charges, i. e., a semiconductor.

If there is a membrane separating the solution into two chambers that is semipermeable based on particle size or chemical class and some of the particles are ions, there will be a partial separation of ions and a resulting non-uniform ion density; this will create an electrical potential generated by the membrane/solution system, which may support or oppose the potential due to the electrode half-cells. The resulting ion current may become considerably more complicated. The mobility of the various ions present can vary between the two bulk electrolytes and the interior of the membrane. It is even possible to describe the dominant charge carrier as a function of the region between the electrodes. The biochemistry of the neurons and vision has traditionally assumed this type of charge transport through biological membranes.

A fundamental biologically based membrane is a bilayer of phospholipids that is not semipermeable to ions or large particles. It contains a hydrophobic core. If the membrane is electrically asymmetrical due to the two layers of the bilayer being different, it is normally semipermeable to electrical charges and is therefore an electrical semiconductor.

For the two semiconductor states associated with the metallic state and the liquid state, the situation is quite different. In these cases, as in the metallic case, the atoms of the material are not free to move under the effect of potential fields, but there are still two types of charge transport. One involves the motion of electrons through the conduction band of the crystalline material. The other involves the apparent motion of a positive charge through the valence band of the crystalline material following the loss of an electron from the valence band at the positive electrode. This motion of an apparent positive charge through a crystal lattice in the absence of any atomic motion is described as a "hole current."

Notice in the above case for a liquid crystal, ionic material must be present but it cannot move since it is locked into a lattice. The current is due to hole movement, not physical ion movement. This effect has been extensively explored and evaluated in the metal semiconductors. It is less well known in the liquid semiconductors. However, it can be easily demonstrated in liquid crystals. The mobility of the electrons and holes can be quite different, resulting in a dominant charge carrier in a given region; also describable as of the n-type due to predominantly electron movement or p-type due to predominantly hole movement. The mobility of these charge carriers is also typically much greater than that of the mobility of ions in equivalent but less concentrated solutions. ***This quantum-mechanical charge transfer mechanism is the mechanism that actually dominates in the biochemistry of neurons and vision.***

11.1.2.12 Reversible Thermodynamics, the non-dissipative resistor

It is not surprising that the reversible nature of the thermodynamic processes used in vision has been overlooked previously. The field of electrolytic chemistry has always been an obscure one, similar to electrical plasmas in physics and electronics. However, it is a critical field that forms the foundation of the neural communication system in the animal. This subject will be explored more fully in the Appendices. The critical fact that must be recognized at the most basic level is that in the thermodynamic equations of an electrolytic system, there is a strong tendency to associate the resistive component associated with the flow of current with a dissipative resistor. ***This is not required!*** The resistive component can be represented by an electrolytic equation relating to the change in state of various atoms in the electrolyte (as when charging a battery). Thus, equation (2-169) of Schwartz in Adelman[36] need not be interpreted as he does:

---

[36] Schwartz, T. The Thermodynamic Foundations of Membrane Physiology. in Adelman, W. jr. (1971) Biophysics and Physiology of Excitable Membranes. NY: Van Nostrand Reinhold

If the reciprocals of the integral conductance, the resistance, to each ion respectively are denoted as

$$1/g_i = r_i = \int dx/F^2u_ic_i;\ i = a, k,$$

Instead, define the unintegrated term for what it is, an exponential function related to the changing of the charge state of atoms in the electrolytic system. An element responsible for this exponential function has a name in electronics. It is a diode. It does not exhibit a constant resistance at any current level. In electrolytics, this diode does not dissipate any power at any current or potential level. It is indicative of an underlying process of electrolysis. It can be, and is in a neurological setting, a thermodynamically reversible process.

The underlying importance of this situation cannot be understated. In biological membrane models throughout history, if each battery and variable resistor in series is replaced by a battery and non-dissipative diode in series, the concept conveyed is much more accurate. **There is no longer a need to find a mechanism that *changes* the variable resistor**.

Because of this unique characteristic of the non-dissipative resistive component in electrolytics, it is important to distinguish this element from its conventional dissipative cousin, *especially* in electro-physics and electro-physiology. The conventional resistor is closely allied to the letter R or r in conventional metallic circuit electronics, where irreversible thermodynamics is involved. Another symbol needs to be found to clearly differentiate the non-dissipative resistor, utilizing reversible thermodynamical principles. The clear choice would be to use the Cyrillic letter for the yau sound, Я. This code is now readily available in desk top publishing within the Word Perfect CyrillicA font by typing small a. It is also available in the Microsoft Cyrillic fong or on ISO Code Page 866.

Since the diode does not exhibit a resistive component that is independent of the current passing through it, it becomes difficult to define a resistive equivalent to it. The normal technique under small signal conditions is to define the differential resistance, $\Delta V/\Delta I$ as the dynamic resistance, R. In the electrolytic (reversible) case, $\Delta V/\Delta I$ =Я. Under large signal conditions, the normal practice is to define $(V-V_o/I-I_o)$ as Я. The location V_o ,I_o is entirely arbitrary but is usually taken as 0,0.

11.1.2.13 Inhibition versus differencing

The term inhibition occurs frequently in the biological literature of vision. The discussions are normally conceptual in nature. At the conceptual level, and particularly in behavioral testing, it is generally used to describe a putative throttling of the signal in one signal channel by a signal in a different channel, frequently in conjunction with a horizontal cell or an amercine cell. It is important to review the mathematical foundations appropriate to the interaction of two signals; in simple situations not involving a feedback loop containing a time delay, the options are a multiplication of two signals or a summing of two signals (which includes the concept of differencing). When electrophysiological tests are performed on *in-vivo* retinas, particularly involving spectrographic tests with more than one chromophoric channel excited, it is found that these so-called inhibitions of one signal by another actually involve the subtraction of one signal from another (a truly antagonistic situation), either to obtain hue information in the chromatic signaling channels or spatial size information in the geometric channels. The result is a situation where either input signal can be described as inhibiting the other (which implies a product relationship); however, the actual situation is that a change in either input signal will affect the output signal linearly (which implies a summing relationship). ***In general, the term inhibition should be viewed with suspicion when found in the visual science literature***. This work has not found a single incidence in the process of animal vision where a signal is actually multiplied by a second signal so as to throttle, or inhibit, the original signal.

11.1.2.14 Use of the term “Neural Networks”

While the use of an expression such as neural network is quite appropriate to the study of both the architecture of the neural system and the circuitry internal to a engine of the central nervous system, it has not been used extensively in these venues. Instead, the title neural network has been usurped by the high performance computational community. It was initially defined conceptually as involving a highly parallel set of nonlinear equations as thought to be found in a complex biological neural network. While this label was originally used to describe a class of networks ostensibly based on real neural networks of the biological sphere; in fact, it merely introducing a different class of nonlinear equations. The nonlinear character of these equations is not found in the biological sphere.

In the context of high performance computing, Jones & Creighton have recently defined a neural network[37]. "A neural network is a set of nonlinear equations that predict output variables (Y's) from input variables (X's) in a flexible way using linear combinations of S-shaped functions." Such networks were contrasted with the more conventional linear set of quadratic polynomials which suffer from two problems. They lack the ability to treat high curvatures verging on discontinuities, they do not treat multiple maximums or minimums well, and they inevitably lead to expressions that go to infinity for large values of the X's.

In the biological sphere, neural networks associated with individual engines of the CNS appear to be highly complex Boolean circuits. A simple biological neural network may accept hundreds of input signals and generate an equally high number of output signals. Some of the input signals may be considered as acting simultaneously as input signals and as control signals within the overall Boolean circuit. The individual operators within such networks appear to consist of simple logical elements. These provide AND, OR, NOT AND (NAND), NOT OR (NOR), limiting and other simple conversion functions. These simple functions that can be assembled in a wide variety of ways to provide both signal processing and short term memory in a fundamentally analog signal environment. The environment of a specific engine operates in a self clocking mode due largely to the time constants of the underlying neurons and the circuit topology adopted to exploit those neurons.

11.1.3 Methods and capabilities of available tools

11.1.3.1 Traffic analysis as a major adjunct to anatomy

When dealing with the peripheral neural system, anatomy can provide a nearly complete mapping of a neural system. However, it does not provide adequate information at the detail level or in areas of very complex circuit paths. Traffic analysis, using the term from military intelligence, is widely used at an elementary level by physiologists to learn more about the detailed circuit diagram of a given neural system. In its basic form, traffic analysis involves observing the volume of signals arriving at a given location by various means as a function of time and attempting to correlate that volume with the potential sources of such signals. Once the coarse relationships between the data arriving at a location and its sources are determined more sophisticated methods can be used to learn more about the traffic. These steps involve a more careful analysis of the origination time and arrival time of the signals to determine if any intermediate relays may have been involved. If a relay is inferred, the analysis is expanded to locate the path(s) involving the relay.

Once the signal traffic has been identified with respect to source, destination and timing, it is possible to attack individual messages in an attempt to determine their structure and information content. Determining their information content generally involves determining the language in which they are encoded. Here, the language may refer to the machine language of the transmission system, the human language of the sender, and any encryption method used to disguise the human language text. As an example there are multiple types of "Morse Code" used to transmit messages by teletype and multiple types of computer code used to transmit computer data. There are also multiple human languages and an infinite variety of encryption schemes.

At the current time, the traffic analyses used in studying the physiology of the central nervous system remains quite simple. Historically, the most common technique has been to insert a probe into neural tissue and then move a source of stimulation within the object field assumed to be associated with that tissue until an electrical response was observed. This technique has produced a large volume of data with respect to the visual systems of many animals. However, this basic technique does not provide information about the path between the stimulus and the observed signal unless multiple probes are used. Multiple probes have only been introduced during the last ten years and the technique remains cumbersome and labor intensive. Most of the data collected has not included the transit time of the signals from the source to the probe location.

Recently, a series of non-intrusive techniques have appeared (MRI, fMRI, PET, magnetic VEP, etc.) that have offered a more comprehensive and less labor intensive traffic analysis capability. Except in the case of the magnetic visually evoked potential (magnetic VEP), these techniques have not provided information concerning the transit time of the signals.

Where the visual signal paths have been determined, attacks on the information content of the signals have been successful where ever the information exhibits a simple spatial relationship to the stimulus in object space. However, as the degree of spatial coherence of the signal with respect to the stimulus decreases, the ability to

---

[37]Jones, B. & Creighton, L. (2003) Using neural networks to model response surface designs Scientific Computing & Instrumentation, April pg 28-31

interpret the signals is reduced. At the current time, man has not determined the nature of the "machine language" used in the neural system to convey abstract (non-spatially coherent) information. Determining the machine language used in the central nervous system is probably the next great frontier in vision (and general neural) research.

11.1.3.2 Spatial and temporal resolution of available tools

Churchland & Sejnowski have provided a useful summary of the spatial and temporal resolution previously achieved using a range of techniques[38]. **Figure 11.1.3-1** extends their figure to differentiate between synapse investigation at the morphological/cytological level and such investigations at the molecular chemistry/crystallography level. The ranges are not meant to indicate theoretical limits. The inclusion of electron microscopy opens a vast area not considered by Churchland & Sejnowski. It allows the definitive description of the synapse and the electrostenolytic mechanism so crucial to understanding the neural system.

[38]Churchland, P. & Sejnowski, T. (1988) Perspectives on cognitive neuroscience. *Science*, vol. 242, pp 741-745

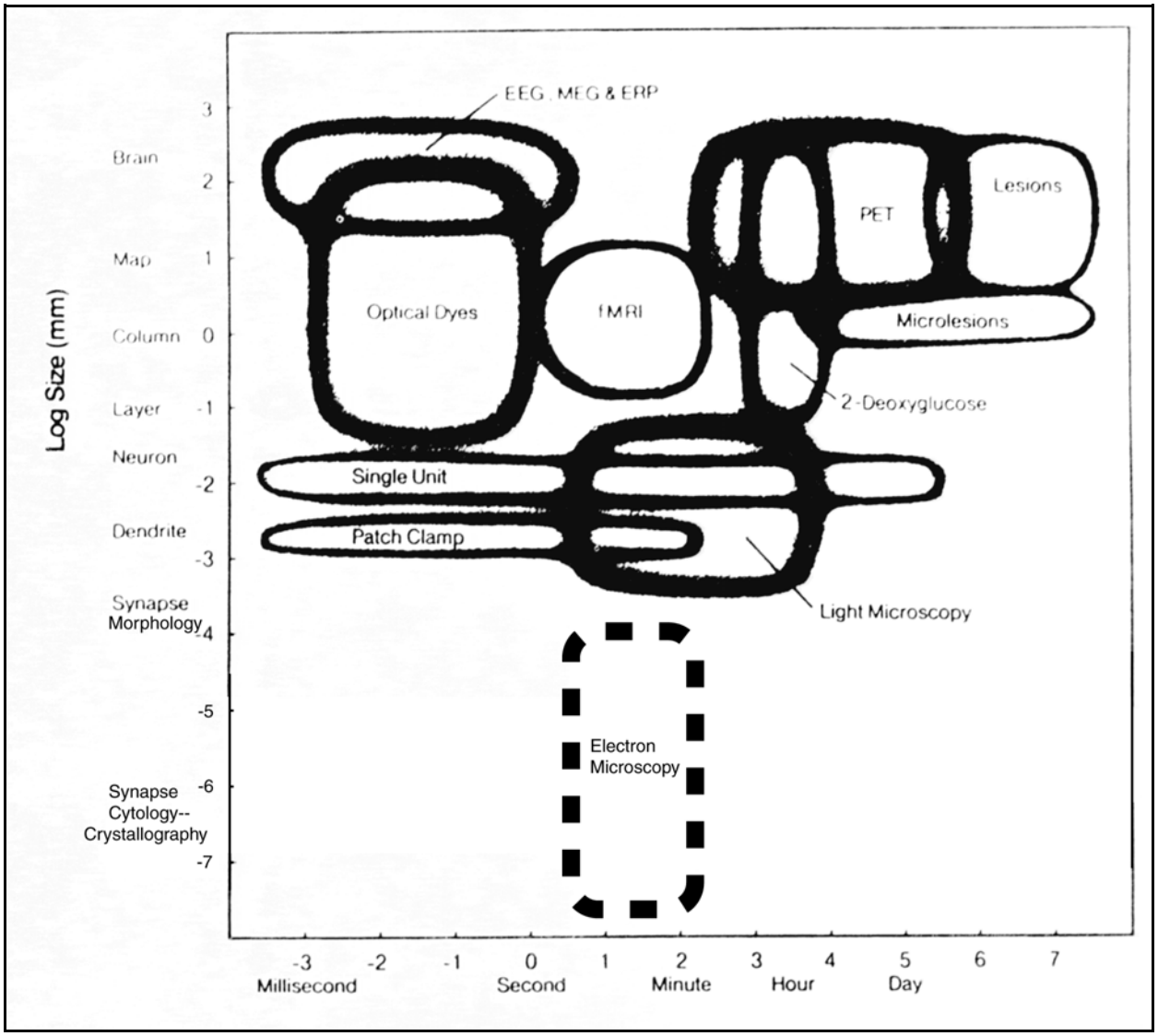

Figure 11.1.3-1 The spatial and temporal resolution range available with current techniques. Modified from Churchland & Sejnowski, 1988.

11.1.4 Architecture of the neural system

The convergence of the work of many investigators has begun to generate confusion in the labels used for many functional structures within the neural system related to vision. It is important to resolve some of these differences to avoid additional confusion.

Recent editions of widely used college level texts have presented introductory level material on the architecture of both the overall neural system and the individual neurons. These works concentrate on the morphology of the neural system and overlook the physiology of that system. While adequate in many ways for the study of simple animals (*Arthropoda* and lower *Mollusca*), these materials do not illustrate the actual architecture of the overall system as found in the higher animals (higher *Mollusca* and *Chordata*).

In the higher animals, the neural system is frequently discussed in terms of a number of operational spheres or

modes. **Figure 11.1.4-1** provides a simple description of the names of these spheres and their primary function.

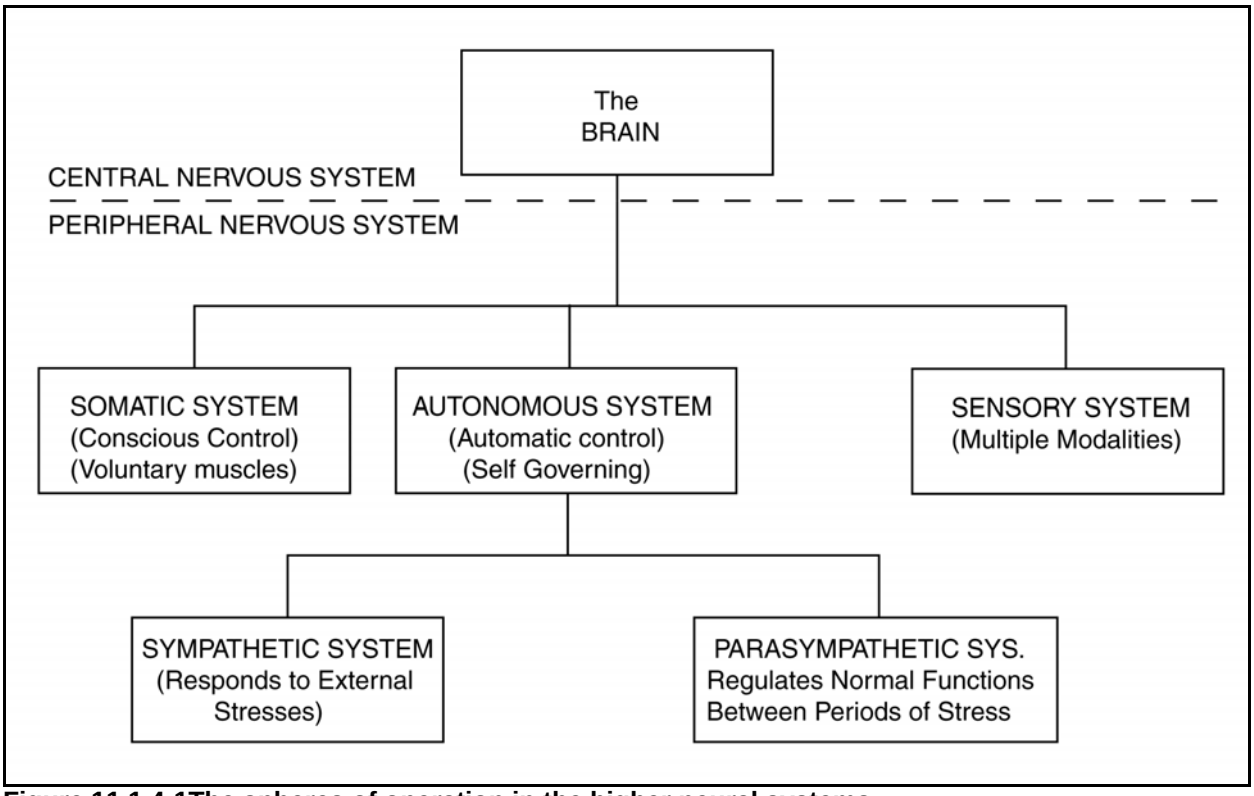

Figure 11.1.4-1The spheres of operation in the higher neural systems.

Many texts overlook the sensory systems in their basic discussion. It is quickly seen that vision does not fit into this figure easily. The oculomotor muscles clearly operate under both the somatic and autonomous categories, virtually simultaneously. The larger saccades can be instigated by volition (conscious control) or due to an alarm (autonomous operation due to an external stress). Microsaccades and tremor are primarily autonomous but do not involve an external stress. Therefore, they would be described as parasympathetic in character.

Torrey has provided a collage of simple neuro/muscular systems that encompasses the simplest neurosensory cell path up through relatively complex reflex arcs[39]. These arcs involve sensory and motor cells interconnected by a signal processing matrix and what is labeled an association neuron. However, there is no discussion of any cognitive element or central nervous system. He also omits any discussion of analog versus pulse signaling over these circuits. Shepherd has presented a similar collage at a single cell level where he speaks of intrinsic and projection neurons[40]. However, the projection neurons are characterized by having a high overall length to nucleus diameter ratio where the length is always less than two millimeters. Similarly, there is no discussion of analog versus pulse signaling over these neurons.

When approached from the signaling aspect of physiology, the architecture of the neural system appears quite different. Circuits (E) and (F) of the Torrey figure are only the low level circuits of a much more complex overall architecture and the so-called projection neurons of Shepherd are only the individual simple analog neurons of the detection and signal processing stages of the neural system. This work defines projection neurons quite differently from Shepherd. This definition is more aligned with the usage by those studying the central nervous system specifically.

---

[39]Feduccia, A. & McCrady, E. (1991) Torrey's Morphogenesis of the Vertebrates, 5th ed. NY: John Wiley & Sons, pg. 406
[40]Shepherd, G. (1994) Neurobiology, 3rd ed. NY: Oxford Press, pg. 64

A grand architecture of the complete neural system can be subdivided into a number of levels based on their functional characteristics. These levels include can be defined by functional names:

0. A primitive neurosensory cell
1. Sensory neural cells
2. Local signal processing neurons
3. Local signal projection neurons (employing analog signaling over distances of less than 2 mm.)
4. Distant signal projection neurons (employing pulse signaling over distances greater than 2 mm.)
5. Motor effector neurons that interface with muscle tissue
6. Glandular effector neurons that interface with secretory tissue

The primitive neurosensory cell was described by Torrey. It is the type of cell that is seen in the most elementary animals where excitation of the sensory cell will cause the animal to contract a nearby muscle and turn away from the stimulus. The detailed signaling physiology of these cells is not well understood.

The above types of cells can be interconnected in a wide variety of ways. These include the simple local reflex arc, the more complex arcs associated with the knee-jerk and other autonomous responses that do not involve the central nervous system. One can build a hierarchy of ever more complex reflex arcs. Each of these arcs represents a more complex expression of a response to one or more stimuli. These responses can all be described in theory using boolean algebra. They are in general reproducible and do not involve time delays that are variable (except with respect to the temperature of the animal). However, at a certain point, the complexity of these arcs begins to involve cognitive processes. At this level, the response of the animal is no longer rigidly reproducible and the time of response may involve significant variation. These are the rudimentary signs of intelligence.

11.1.4.1 The basic neural unit

The morphologist invariably insists that the neuron, as the minimal biological cell must be considered the fundamental neural unit. However, they have not been able to differentiate the minimal biological cell from the minimal biological neuron. On the other hand, Shepherd has held quite recently that the synapse is the fundamental building block of the neural system[41]. These two views are quite far apart. Furthermore, it is quite clear that the synapse is a very complex functional structure.

The problem here is to differentiate the neuron from other cells and to describe the feature that makes the neuron different. There is one fundamental feature that differentiates all neurons from all other cells. There is also a secondary feature that is quite prominent. Both of these features are only discernible using the electron microscope at magnifications of over x100,000. Each neuron contains an active electrolytic semiconductor device that can provide electrical signal amplification. It is called an Activa. The secondary feature is the presence of multiple interior plasma spaces that are isolated from each other by membranes that are in close proximity in at least one area. The Activa is found in this area of very close, but very structured, proximity.

When examining the internal structure of a neural cell, compared to a non-neural cell, it becomes quite clear that the neural structure is an overlay. All of the normal elements of a cell are found in a neuron. However, there are additional elements not found in the non-neural cell. This makes it quite clear that the portion of the neuron in common with non-neural cells provides the housekeeping and growth capabilities of the cell. It is the Activa and the arrangement of the isolated plasmas that constitute the uniquely neural elements of the neuron It will be shown that multiple copies of these features may be found in a single neuron. The photoreceptor cell is a particularly complicated example (**Chapter 12**). On the other hand, the projection neuron connecting a sensor or a muscle to the CNS is a particularly simple example showing considerable replication of the basic neural structure (**Chapter 14**).

Using the projection neuron as an example quickly shows that the synapse is not a basic unit of the neural system. It is only a particular configuration of the basic neural architecture that is shared with all of the Nodes of Ranvier and the internal amplifiers found within the soma. Similarly, looking closely at the projection neuron shows that the neuron itself is not the basic unit. It contains a replicated structure centered on the Activa.

> *The basic unit of a neuron, that separates it from any non-neural cell, is the active electrolytic semiconductor device named an Activa that is formed in conjunction with a series of isolated*

---

[41]Shepherd, G. *Ed*. (1998) The Synaptic Organization of the Brain. NY: Oxford University Press pg 3

plasmas within their individual conduits. The Activa need not occur within a neuron. It is also found connecting the plasma conduits of two adjacent neurons. In this situation, it is the major component of the synapse.

11.1.4.2 A hierarchy of the neural system architecture

As discussed above, the neural system of all animals is designed as a hierarchy. The complexity of this hierarchy is determined by the evolutionary development of the animal and its needs. **Figure 11.1.4-2(A)** illustrates this architecture with respect to ever more complex reflex arcs. Note that it is only at the lowest level that a signal is generated by a sensor and travels in only one direction, and generally to a node before arriving at an actuator mechanism (whether muscular or secretory). All higher level paths are shown as bidirectional. They actually involve two distinct paths, an initial afferent path (upgoing arrows labeled s) and a subsequent efferent path (downgoing arrows labeled m). Note also that all nodes receive signals from and transmit signals to at least four locations. It is the paths leading to the higher nodes that provide the capability of coordinated movement and eventually cognition. In this context, frame (D) of figure 17-1 in Torrey can be considered as illustrating the lowest level reflex arc. There is an afferent neuron and a efferent neuron between the sensory cell and the responding cell. His frame (E) introduce cross-coupling between individual paths and corresponds to the next higher level reflex arc shown in **Figure 11.1.4-2(B)**. His final frame, (F), introduces the concept of an association neuron between two nodes. Although not stated, the assumption can be made that this neuron was processing analog signals. As indicated above, this type of neuron is limited in length to only a fraction of a millimeter.

Note the symbology used in the figure. The nuclei indicated by the black dots within the soma play no functional role in the neural system. They are involved in the support of the signaling function and the health of the cell. The Activa, indicated by the solid arrow within the soma, is the active and primary element within each neuron. ***Without at least one Activa, the cell is not, and cannot function as, a neuron***.

Up to this level of complexity, all signaling employs analog waveforms. This type of signaling becomes energy intensive for lengths greater than a fraction of a millimeters and an alternate, albeit more complex, system is needed. To meet this need, pulse signaling is introduced (recognized by the characteristic action potential). Introduction of this signaling mode requires replacing the single association neuron by a projection neuron *stage* consisting of a signal encoder neuron (typically, the ganglion cell of the nervous system), a signal decoder neuron (typically, the stellate cell of the CNS) and frequently signal regeneration mechanisms between the previous two neuron types (the Nodes of Ranvier). As a result of the introduction of this signal transmission technique and stage, it is useful to extend the concepts in Torrey to include those of **Figure 11.1.4-2(C)**. The shaded area encloses the signal projection stage and involves pulse signaling. Pulses are actually found from the output of the Activas within ganglion cells to the input of the Activas within the stellate cells (the region of darker shading). To the left of the projection stage is the signal processing stage associated with the sensory neurons. To the right is the signal processing stage associated with the effector neurons.

Simple signal projection neurons are also limited in length to about 2mm. Beyond this length, pulse signal regeneration is required to maintain pulse signal fidelity. Nodes of Ranvier are introduced near this spacing all along the axons of signal projection neurons. These Nodes are not shown at the scale of this drawing.

Dowling has presented a conceptual drawing similar to the networks in these figures[42]. However, the details differ from those shown here.

[42]Dowling, J. (1992) Neurons and Networks. Cambridge, MA: Harvard University Press pg 309

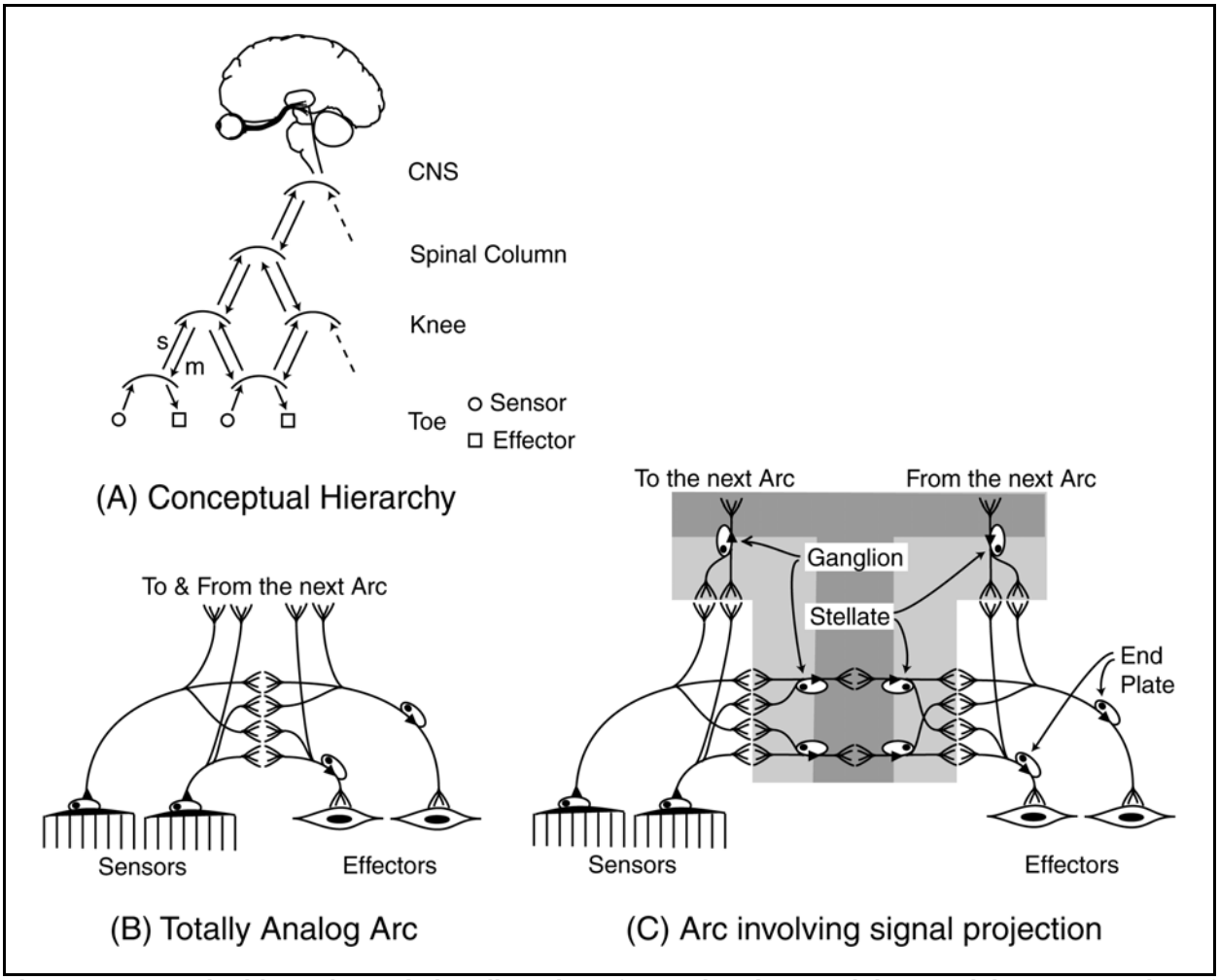

Figure 11.1.4-2 The hierarchy and signaling plan of neural pathways. (A); a partial representation of the typical hierarchy found in the nervous system. Each level can be considered a reflex arc. The lowest level is local (all elements within a fraction of a millimeter diameter) and involves only analog signaling. The higher levels are typically more remote and involve pulse transmission (action potentials) between levels. Within a level, analog signal processing may be employed. All neural signals, except possibly those of taste and smell, are routed through the midbrain initially. The labels on the right are only illustrative. The hierarchy can be applied to any of the numbered nerves of the system. (B); an example of the simplest reflex arc including paths to the next higher reflex arc but involving only analog signal processing. (C); a more extensive reflex arc involving signal projection (pulse signaling using action potential). The encoding ganglion cells and the decoding stellate cells are shown explicitly. This frame can be transformed into a higher level reflex arc by changing the lower labels to "from sensors" and "to effectors" and by removing the label "End Plate." Note the Activas, indicated by the solid arrowheads within each soma. The nuclei shown within each soma are not relevant to the neural signaling process.

11.1.4.2.1 A top level block diagram of the complete neural system

When employed in vision, many of the lower level arcs on the left of the above figure, up to the level of the interface of the optic nerve with the midbrain, are unidirectional. The hierearchy of lower level analog arcs associated with vision are combined and represented by the neural layers of the retina. They represent the signal detection and the initial signal manipulation stages of vision.

Figure 11.1.4-3 presents the visual system expressed in a more general top level block diagram form compatible with the rest of the neural system. The stages of the visual system are shown as defined below. The figure also shows the path of the two servomechanism subsystems associated with the visual system. The first involves a loop that turns at the midbrain and is labeled a reflex arc (RA_M). The second is a loop that turns at the cortex and is also labeled a reflex arc (RA_C). Whereas the first loop is essentially one complex loop serving the muscles associated with the multiple axes of the eye, the second is a broader designation for all high level reflex arcs involving the cortex. These latter high level arcs generally involve the cerebellum. As a result of training, some reflex arcs can be instituted via the cerebrum. However, it is those associated with the cerebellum that are genetically "prewired."

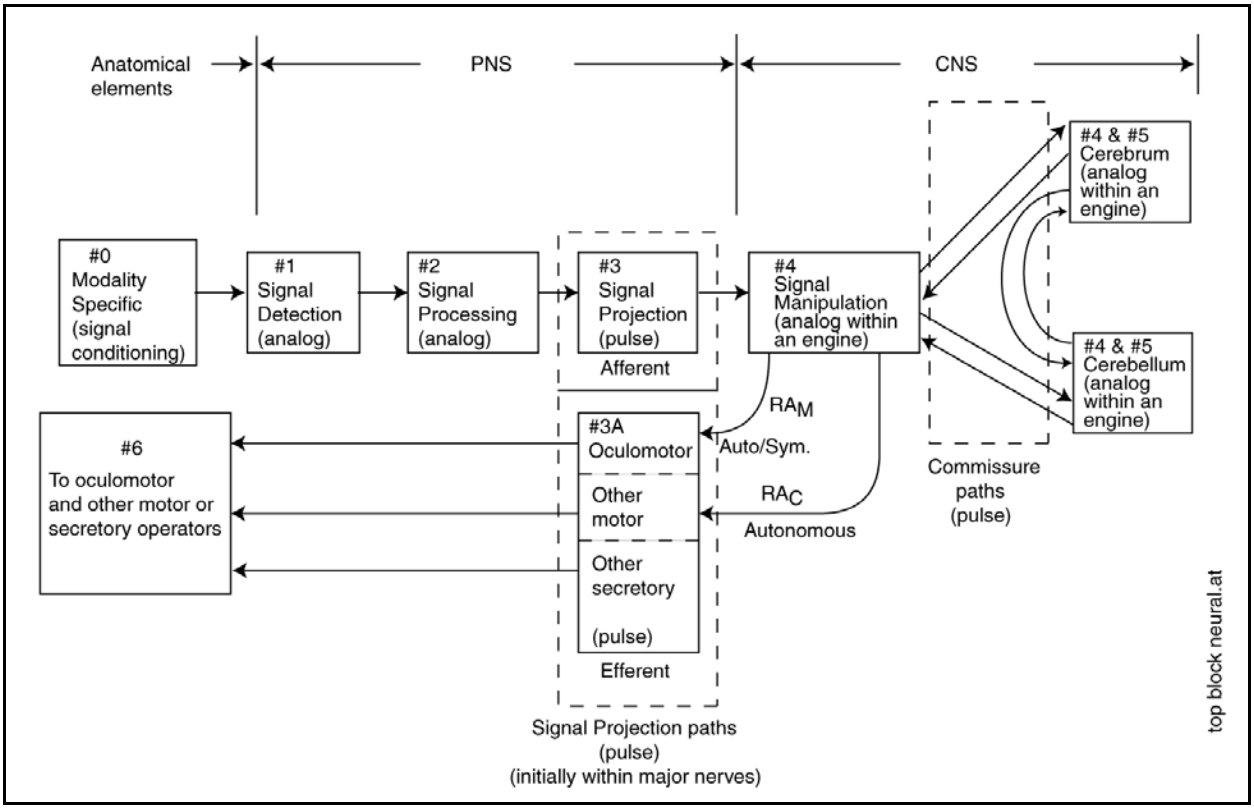

Figure 11.1.4-3 Top level block diagram of the neural system focused on the visual subsystem. The stages defined in the text are shown. The top level reflex arcs are designated RA_M for those turned in the midbrain and RA_C for those turned in the cortex. The role of the midbrain as a "gatekeeper" is emphasized. The parallel roles of the cerebellum and cerebrum are shown. Some reflexes involving the cerebellum are highly autonomous and can not be controlled by the midbrain as suggested by the dotted line. The signals controlling the oculomotor subsystem are a complex mixture of autonomous and sympathetic signals.

Being at the top of the reflex pyramid, all of the high level reflex arcs involving the cortex are able to control large numbers of muscles simultaneously. This allows the animal to close its eyes, turn its head, and crouch simultaneously.

The nature of the signals are described cryptically at the bottom of each stage. In all cases, signal detection and signal manipulation is performed in analog mode. Some short distance signal projection is performed in analog mode but statistically speaking, pulse signaling is the predominant form of signal projection. The arrows at the lower left represent the various nerves proceeding to the muscles and glands of the body.

In this figure, signals can originate in the cortex based on cognition. These signals can not be easily associated with a reflex although they may be reflexive, even after a long period of time. They are basically sympathetic, or at will,

signals.

The efferent oculomotor signals are a complex mixture of autonomous and sympathetic signals where the autonomous signals are also a temporally complex mixture of reflexive and genetically programmed signals.

Stage 1 encompasses the signal detection process subsequent to the imaging of light onto the retina. This stage will end with the presentation of the signals at the pedicels of the photoreceptor cells.

> A **stage 0** can be defined if desired to encompass the physiological optics of vision. This will allow the stages defined here to remain generally compatible with the two and three stage models currently found in the color vision literature.

Stage 2 consists of the signal processing and manipulation carried out between the pedicels and the input terminals of the Activas <u>within the ganglion cells</u> of the retina. Being a functional element in the signal processing stage, the synapses associated with the pedicels will be considered part of Stage 2. Similarly, the neurites of the ganglion cells are also key players in the signal processing of Stage 2.

Stage 3 encompasses the signal projection circuits of the visual system. The signal projection function is highly developed and easily characterized in *Chordata*. It is used wherever it is necessary to transmit neural signals over significant distances (typically a millimeter or more). This includes both the efferent and afferent signal paths of the peripheral nervous system (PNS) and the paths within the brain, or central nervous system, (CNS).

Within the visual system, the projection function is used routinely. It is used to transmit signals over the optic nerve to both the LGN and the pretectum. It also used to transmit signals from the LGN to area 17 of the cortex and to transmit signals from the pretectum to area 5 of the cortex. Within the cortex, projection neurons are used to interconnect the different feature extraction engines and logic units of the brain with the initial motor system command generation centers.

In the nominal case, the projection function extends from the input terminal of the Activa <u>within</u> each ganglion cell to the output pedicel of each stellate cell <u>within</u> the brain that receives enervation from a neuron of the optic nerve. Note: the dendritic structure of the ganglion cell is looked upon as part of the signal manipulation function of **stage 2**, and not as part of the signal projection function of **stage 3**. Such a division highlights the role of ganglion and stellate cells as transition devices between analog and binary (pulse) signaling.

The further subdivision of **stage 3** to accommodate this situation will be addressed in **Chapter 14.**

Stage 4 consists of all of the higher level signal manipulation and processing that occurs subsequent to the arrival of signals at the brain from **Stage 3**. It may be useful to subdivide this stage to clarify the fundamental difference in the signal manipulation occurring in the Pretectum from that occurring in the LGN and that occurring in the cerebral cortex.

Stage 5 consists of command signal generation within the cortex and controlled by the midbrain acting as gatekeeper. This stage comprises the initial efferent signaling generating engines associated with the signals emanating from the cortex and associated with the visual system (such as the oculomotor system and some aspects of the skeletal motor system). As in **stage 4** above, it is difficult to make clear distinctions at this level between the portions of the stage that involve analog signal processing and pulse signal projection within and between areas of the cortex.

Stage 6 consists of the command signal projection paths beyond the cortex.

Except for the very few neurons required to control the muscles within the ocular, the relevant effector signals developed in the brain, and related to vision, pass down through the motor-neuron system. A similar situation is found with respect to the audio neurons as well as those associated with taste and smell.

11.1.4.2.2 The hierarchy within the CNS

Without discussing the details, the brain can be considered responsible for four primary functions:

1. To provide reflex responses of the most complex form at the highest level in the reflex pyramid,
2. To provide the animal with general situational awareness relative to the outside world,
3. To provide highly specific situational awareness relative to specific conditions (such as the color, smell, contour

and texture of possible food items),
4. To provide a degree of memory for certain classes of events and situations,
5. To control communications with other animals based on the mechanisms available to the animal.

The human is noted particularly for its high level of evolution with respect to functions three (contour) and five. This evolution has been highly dependent on the synthesis of a semantic system, language, followed by the development of an external mechanism of memory, writing (and it subsequent ramifications).

The following discussion may conflict with many discussions based only on morphology or only on psychophysics. However, when the work in these fields is melded with the available work in physiology of the neural system, most investigators will agree with the following material.

At the top of the reflex arc pyramid is the brain. In the lower animals, it is typified by what is called the old or mid-brain in humans. It is essentially the top level reflex arc although it may involve considerable cross-coupling between individual signaling arcs. As these cross-couplings have become more complex, the brain has evolved a separate structure, the cerebellum to handle them. This structure continues to be primarily a correlation center for ever more complex reflex actions. It is generally considered the seat of autonomous responses. It does not involve memory or cognition and can be described in terms of, albeit complex, boolean algebra (no variable time delay elements).

Within the chordates and the higher molluscs, a need arose for a higher level of signal processing that included memory as a significant factor. For this purpose, the cerebrum evolved as another extension of the mid-brain. The functional complexity of this structure is so enormous that man has only begun to investigate and understand it. It differs from the cerebellum and the lower structures primarily in its employment of memory. The addition of this capability adds a new dimension to calculations based primarily on boolean algebra.

The addition of memory also adds a new dimension to the cross-coupling between the signals associated with a top level reflex arc. The response need no longer relate only to the present. It can relate to prior responses under similar (or all-together different) circumstances. In the extreme, the response can be the result of cross-correlation only involving prior events. This cross-coupling of signals stored previously in memory can be considered the kernel of cognition.

It is worth noting the recent recognition, primarily through MRI and fMRI techniques, that memory is not focussed in a specific portion of the brain. It is intrinsic to the individual processing centers of the cerebrum and the cerebellum. Because of the overall complexity of these individual centers, they are defined as engines in this work. This name correlates very well with the engines defined in current computer processing. An engine is a major hardware portion of a computer that incorporates memory as well as algebraic manipulation and frequently employs its own specialized software routines.

The cognitive process involves bringing a series of information vectors into the saliency plane of individual engines of the cerebrum and manipulating them.

As a result of this organization of the brain, the mid-brain has evolved to be a gate keeper. It is able to delay or influence a variety of autonomous responses from the cerebellum based on the information received from the cerebrum. It also controls virtually all sympathetic responses of the animal.

As the complexity of the brain has evolved into a mid-brain, a cerebellum and a cerebrum-plus a variety of auxiliary structures- it has become desirable to introduce an additional feature. This feature involves the bifurcation of individual (groups of) afferent signals to allow timely processing of the information by different portions of the brain. This bifurcation is extremely important to the operation of the visual system and apparently the audio system. *This bifurcation has been largely overlooked in morphological studies but it is key to the understanding of these systems*.

In the case of the visual system, the signals from the optic nerve bifurcate upon arrival at the mid-brain. A portion of the signals associated with general situational awareness are initially processed by the lateral geniculate nuclei (LGN), a region of the diencephalon. The most time critical (change related) signals from the retina are stripped off the data stream by the LGN and returned to the Precision Optical System (POS) of the mid-brain immediately. The information required for stereo-optic vision is also stripped off at this point and returned to the POS. The remaining material is then passed to the posterior region (area 17) of the posterior lobe of the cerebrum by the LGN via the geniculocalcerine pathway for further processing. The material presented to the posterior lobe of the cerebrum is further processed by a variety of engines until it is entirely in vector form. This information is then passed to the

cognitive and memory engines in the parietal and temporal lobes of the cerebrum.

The other information from the optic nerve, which is of higher spatial resolution, is more time critical from the perspective of information theory and originates in the foveola, are initially processed by the pretectum or forward portion of the mesencephalon. These signals are combined with other signals, primarily from the vestibular system and completely vectorized prior to being sent directly to one of the cognitive portions (area 7) of the parietal lobe of the cerebrum via the Pulvinar pathway. Although they are visual signals, these high precision, time sensitive signals do not pass through the so-called "primary visual cortex" at the posterior of the posterior lobe of the cerebrum.

It appears that the auditory signals from the ears are also bifurcated and processed simultaneously by two distinct areas of the mid-brain. The auditory system will not be addressed further here.

Recognizing the above role of the midbrain as gatekeeper, the description of the suggested dual signal paths within the brain is slightly different than that in Figure 17-16 of Torrey. In this work, the mid-brain acts as a junction point involved in the receipt and issuance of virtually all signals associated with the CNS and capable of controlling the relay of some of these signals, particularly the efferent command signals to the peripheral system. This activity involves processing two separate classes of signals within two distinct areas of the midbrain. The first class relates to the general situational awareness of the animal. The second involves the specific situational awareness. These signal paths will be developed in more detail in **Chapter 15.** Recognizing this role, it is not clear that any signals from the cerebellum are ever passed through the midbrain to the cerebrum. However, ever and never are hard words to defend in this field.

11.1.4.3 Consequences of introducing memory in the CNS

There are interesting side effects to the introduction of dynamic, as opposed to genetically encoded, memory into the CNS. These can generally be divided into those associated with short term memory and long term memory. It is evident that short term memory is shared with even the bees who are able to remember the path to a food supply long enough to return to the hive and perform their instructional dance for the benefit of other bees and the community as a whole. It is difficult to quantify the presence of long term memory versus genetically programmed procedures in the bees because of their short life span. Are they able to return to the same site some days later based on experience. Nearly all higher animals with a life span of over a few months are known to exhibit long term memory. Recent films have shown the ability of the octopus to return to one of a series of homes over a period of weeks with unexpectedly good precision. It generally does not rely upon intermediate landmarks nor does it take a previously defined path. Long term memory is generally associated with experience whereas genetic memory is provided across evolutionary time intervals.

11.1.4.3.1 Anachronistic consequences of memory in the CNS

With the availability of memory within the signal processing engines of the CNS, these engines are able to call upon experience, as stored in their individual memories, perform computations, and restore the results in other areas of memory and/or generate command signals. This capability can lead to relatively bizarre results. Fortunately, the mid-brain acting as gatekeeper can restrict some of these results. Also fortunately, these results can provide information about the importance of memory in the visual and other processes process (particularly hearing).

A particular example of anachronistic behavior is sleep walking in humans. This phenomenon stresses the capacity of the long term memory within the engines of the CNS associated with sensory system as a whole and the computational scope available to these engines. The engines can generate all of the necessary command instruction to effect the movement of the individual, without benefit of visual ques, over a considerable distance and along intricate paths. Fortunately, the mid-brain in its role as gatekeeper is usually able to prevent this activity from being implemented by the muscles. Sleep walking is frequently interrupted when the person encounters an unexpected obstacle that is not where memory indicated it should be.

A second example is the ability of a human to see ghostly images or hear sounds that clearly are not present. These perceptions are based on the engines of the CNS recalling stored vectors representing these (previously experienced-learned) events into saliency space. The problem arises when the individual believes these computed events (hallucinations) are real.

A more serious situation arises when a human recalls these above computed events from memory, performs computation upon themin saliency space, and then stores the new result in memory. As a result of this process, the individual can progressively lose the ability to separate real events stored in memory from computed but

hallucinatory events. This can lead to a spiral into insanity.

11.1.4.4 Recent fMRI data confirming the distribution of the feature extraction engines of the brain

Within the last dozen years, our understanding of the operation of the visual portion of the brain has increased immensely with the introduction of functional magnetic resonance imaging (fMRI). The name is meant to describe the ability of the imaging device to record activity in the brain in response to images projected onto the retina. A particularly vivid example of this imaging technique was recently published by Logothetis, et al[43]. When combined with the VEP technique, it can be used to determine the time sequence of events relative to the different engines of the brain.

> Such techniques have already demonstrated the arrival of signals in the mid region of the cerebrum **before** the arrival of similar signals in the occipital cortex (the socalled primary visual cortex).

11.1.4.5 Chemical versus electrical foundation of the architecture

As developed in detail in **Chapter 10**, this work does not support a chemical basis for the signaling mechanisms of the neural system. Although the original proposals were for an electrolytic signaling mechanism, the chemical community has pushed a chemical mechanism to the forefront over the last 75 years or more. This competition has not incorporated the newly discovered biologically based electrolytic semiconductor device, the Activa. In the absence of this discovery, the chemical thesis for neural signaling has become ever more fanciful, daunting in the words of one author[44]. However, it has not been able to demonstrate how an electrical signal within a neuron is converted to a chemical signal in the orthodromic interneural space and then back into an electrical signal within the next neuron. Based on the discovery of the Activa and the recognition that the synapse and the Node of Ranvier are simple electrolytic circuit elements based on the Activa, the fundamental electrical mechanism of the neural system becomes evident. In addition, a deterministic description of the neural system, external to the higher levels of the cortex, becomes possible.

Beginning with E. G. Gray in 1959, the electronmicroscopy community began delineating the structure of the synapse at the level of tens of nanometers. The result has been the description of the fundamental synaptic cleft as approximately 30 nm in width and less than 2 microns in diameter. This is the structure defined electrically as an Activa in this work. The discovery of the Activa was memorialized by a United States Patent in 1999. In many synapses, and related structures such as the Node of Ranvier, this fundamental structure is reproduced as an array of fundamental devices in order to achieve higher current carrying capacity. This mechanism provides a lower circuit impedance between orthodromic neural conduits.

The above paragraphs are not meant to denigrate the role of chemistry in a supporting role to signaling. It has a profound role to play in providing the electromotive force energizing the neural system and it may play a significant role in causing the creation of short term and long term memory.

11.1.4.6 Electrotonic versus pulse signaling

The visual signaling system is divided into two classes of circuits. The vast majority of both the CNS and PNS circuits involve electrotonic signaling. Only the neurons of the signal projection stage, stage 3, employ pulse signaling. Although the method of signaling varies, the mechanisms used to generate and pass signals within neurons and between neurons do not. They are electrolytic mechanisms.

It is useful to re-examine the thesis of Bieda & Copenhagen in a recent paper[45]. The paper is based on the conventional wisdom that signaling between neurons involves secretion of chemical neurotransmitters. It notes correctly that transfer of a signal from one neuron to another does not require sodium action potentials. However, it does not demonstrate or reference laboratory tests confirming that either GABA or glycine is required as a signaling

---

[43]Logothetis, N. (2002) Functional MRI probes monkey brain. *Biophotonic International*, April 2002, pp55-56

[44]Shepherd, G. (1988) Neurobiology, 2nd ed. NY: Oxford Press, pp. 75-76 reprinted with different page numbers as the 3rd ed. in 1994

[45]Bieda, M. & Copenhagen, D. (1999) Sodium action potentials are not required for light-evoked release of GABA or glycine from retinal amacrine cells. *J. Neurophysiol*. vol. 83, pp 3092-3095

mechanism between neurons. It merely assumes the conventional wisdom that since these substances are ubiquitous near neurons, the transfer involves these chemical substances. They are correct that these substances impact the transfer of signals because they impact the electrostenolytic power supply to the neurons.

11.1.5 Categorization of neurons by stage and function

In this section, the term function will be used in a much more precise mode than it is usually used in morphology and biology. Where the biologist frequently defines a type of neuron conceptually by the function it "mediates," the function of a neuron will be defined here in one of two more concrete ways. The first is how it modifies the (multiple) signal waveforms presented to it. The second is how it oscillates, either in free-running mode or in response to an instigating signal.

Books traditionally describe an abbreviated set of morphological types of neurons suitable to the purposes of the author. These sets are frequently incompatible. The most fundamental problem with such lists is their focus on the nucleus of the cell as a point of departure. This element has virtually no role in the function of the cell as shown in **Chapters 10** & **11**. As a shining example, the description of a neuron as monopolar, because it has only one appendage on the soma, is completely irrelevant to the function of the neuron. The so-called bipolar neuron actually has three appendages, although one may be difficult to see in the absence of electron microscopy. Similarly, an amercine cell has a fully functional axon, although it may be packaged with one of its neurites for anatomical efficiency. Steriade, et. al.[46], and Pannese[47] provide a variety of morphological images of neurons that can be easily understood based on their electrical functions and electrolytic properties, as developed in **Chapter 8** and this chapter.

A problem arises immediately with regard to the adaptation amplifier within the photoreceptor cell. A decision must be made as to whether a change in amplification associated with the adaptation amplifier should be considered signal processing and be discussed as a Stage 2 function or whether it should be considered an integral part of the Stage 1 function. Considering all of the functions occurring within the photoreceptor neuron as Stage 1 functions leads to a simpler model. With this question resolved, the earlier definitions regarding each stage can be definitized.

Although there are many idiosyncrasies associated with the physical structure of visual neurons, leading to difficulty in determining their proper morphological classification, **Figure 11.1.5-1** defines the more significant functional types to be addressed in the following chapters. A Type Designator is assigned to them for ease of cataloging. The common names are those usually found when discussing the visual system of chordates, probably the most sophisticated of the neural systems. Several names commonly used when discussing non-chordates can be added to the table. The retinular neurons can be listed as receptor neurons and probably contain the same internal circuitry as the photoreceptor cells. In the context of this work, the eccentric neuron of *Limulus* is a stage 3 projection neuron with its input structure performing stage 2 signal processing and its output structure encoding the resulting signals. This suggests it is functionally a ganglion neuron in the above table.

---

[46]Steriade, M. Jones, E. & Llinas, R. (1990) Thalamic oscillations and signaling. NY: John Wiley, pp 132-133
[47]Pannese, E. (1994) Neurocytology. NY: Thieme, pp 5 & 80-116

Class	Type	Common Name	Purpose/Features
In Chordata			
Stage 1 Signal Detection			
Receptor	AT	Photoreceptor (Vibrissal cell)	Adaptation Amplifier; strip line form, exponential internal feedback, collectors in parallel
	AD	(auditory cells) (vestibul. cells)	Distribution amplifier; current to voltage converter
Stage 2 Signal Processing (**within PNS**)			
	AB	In-line/Bipolar	Summing and isolation amplifiers; non-inverting
	AL	Lateral (Horizontal) (Amercine) (Pyramid)	Differencing amplifiers with two inputs; internal feedback
?	AX?	Interplexiform	Putative external feedback amplifier; from amercine to horizontal neurons
Stage 3 Signal Projection			
Signal Encoding	AG	Ganglion	Encoding: analog input to pulse output, internal feedback, oscillatory, action potential generator
Signal Decoding	AR	Stellate	Decoding: pulse input to analog output, capacitor integrator
Sig. Regenerator	AN	Node of Ranvier	Regeneration of action potentials
Stage 4 Signal Manipulation (**within CNS**)			
Signal correlation	AS	Sense amplifier	Sums signals in multi–dimensional correlators, found in LGN, PGN, pulvinar & cerebral cortex
Stage 5 Information Manipulation (**within CNS**)			
Stage 6 Action initiators (**within PNS**)			
Muscle Activation	AM	End plates	Decodes action potentials at high current level
Enzyme secretion	?	?	Secrete enzymes after decoding action potentials

Figure 11.1.5-1 A tabulation of neuron classes and their common names based primarily on the visual system. See text. The existence of an interplexiform neuron supporting external feedback in the retina is questionable. The orderly naming of neurosecretory neurons has not been pursued in this work.

The receptor neurons of Chordata appear to exhibit the full range of neuro-secretory functions available to the neuron. The ability of the photoreceptor neurons, the vibrissal neurons and the hair cells of both the vestibular system and the auditory system to secrete a protein is well established. These proteins form mechanical structures that support the sensory function. It is likely all these neurons exhibit the same internal topology as photoreceptor neurons, as this would explain their ability to vary their sensitivity significantly with stimulus intensity. It is likely that the receptors of taste and smell also secrete proteins.

The horizontal and the large group known as amercine neurons are all lateral cells in that they accept signals from

neurons carrying signals from different spatial locations (which may also exhibit chromatic differences within the retina. They combine this information using a variety of algorithms that vary with their specific role (Chapter xxx). Pyramid neurons are lateral neurons that are found elsewhere in the PNS.

The ganglion neurons of stage 3 exhibit a variety of individual features that will be described in greater detail in **Section xxx**.

The morphological descriptor's monopolar and bipolar do not relate well to the electrical performance of neurons and the morphological term monopolar will not be used in this Chapter. The bipolar name will be used to define a neuron primarily with respect to its in-line location in the retina between the photoreceptor cell and the ganglion cell in a signal path. This neuron is electrically monopolar (its signal waveforms proceed in only one direction from its resting potential when subject to an input stimulus) while it is morphologically bipolar.

Because of the fact that an active device may be formed upon the juxtaposition of any two membranes associated with neural conduits, amplification, in the broad sense, may occur at two different locations in the nervous system. It may occur inside a given neuron and also between two adjacent neurons. The Nodes of Ranvier are examples of multiple amplifiers within a single cell.

The circuit associated with the Activa of Type AN (the N for Node of Ranvier), will be discussed briefly in **Section 3.3 xxx** because of its prototypical role in all signal projection and hybrid neurons within and outside of the eye.

In this section, the term function will be used in a much more precise mode than it is usually used in morphology and biology. Where the biologist frequently defines a type of neuron conceptually by the function it "mediates," the function of a neuron will be defined here in one of two more concrete ways. The first is how it modifies the (multiple) signal waveforms presented to it. The second is how it oscillates, either in free-running mode or in response to an instigating signal.

Books traditionally describe an abbreviated set of morphological types of neurons suitable to the purposes of the author. These sets are frequently incompatible. The most fundamental problem with such lists is their focus on the nucleus of the cell as a point of departure. This element has virtually no role in the function of the cell as shown in **Chapters 10** & **11**. As a shining example, the description of a neuron as monopolar, because it has only one appendage on the soma, is completely irrelevant to the function of the neuron. The so-called bipolar neuron actually has three appendages, although one may be difficult to see in the absence of electron microscopy. Similarly, an amercine cell has a fully functional axon, although it may be packaged with one of its neurites for anatomical efficiency. Steriade, et. al.[48], and Pannese[49] provide a variety of morphological images of neurons that can be easily understood based on their electrical functions and electrolytic properties, as developed in **Chapter 8** and this chapter.

A problem arises immediately with regard to the adaptation amplifier within the photoreceptor cell. A decision must be made as to whether a change in amplification associated with the adaptation amplifier should be considered signal processing and be discussed as a Stage 2 function or whether it should be considered an integral part of the Stage 1 function. Considering all of the functions occurring within the photoreceptor neuron as Stage 1 functions leads to a simpler model. With this question resolved, the earlier definitions regarding each stage can be definitized.

Although there are many idiosyncrasies associated with the physical structure of visual neurons, leading to difficulty in determining their proper morphological classification, **Figure 11.1.5-2** defines the more significant functional types to be addressed in the following chapters. A Type Designator is assigned to them for ease of cataloging. The common names are those usually found when discussing the visual system of chordates, probably the most sophisticated of the neural systems. Several names commonly used when discussing non-chordates can be added to the table. The retinular neurons can be listed as receptor neurons and probably contain the same internal circuitry as the photoreceptor cells. In the context of this work, the eccentric neuron of *Limulus* is a stage 3 projection neuron with its input structure performing stage 2 signal processing and its output structure encoding the resulting signals. This suggests it is functionally a ganglion neuron in the above table.

---

[48]Steriade, M. Jones, E. & Llinas, R. (1990) Thalamic oscillations and signaling. NY: John Wiley, pp 132-133
[49]Pannese, E. (1994) Neurocytology. NY: Thieme, pp 5 & 80-116

Class	Type	Common Name	Purpose/Features
In Chordata			
Stage 1 Signal Detection			
Receptor	AT	Photoreceptor (Vibrissal cell) (auditory cells) (vestibul. cells)	Adaptation Amplifier; strip line form, exponential internal feedback, collectors in parallel
	AD		Distribution amplifier; current to voltage converter
Stage 2 Signal Processing (**within PNS**)			
	AB	In-line/Bipolar	Summing and isolation amplifiers; non-inverting
	AL	Lateral (Horizontal) (Amercine) (Pyramid)	Differencing amplifiers with two inputs; internal feedback
?	AX?	Interplexiform	Putative external feedback amplifier; from amercine to horizontal neurons
Stage 3 Signal Projection			
Signal Encoding	AG	Ganglion	Encoding: analog input to pulse output, internal feedback, oscillatory, action potential generator
Signal Decoding	AR	Stellate	Decoding: pulse input to analog output, capacitor integrator
Sig. Regenerator	AN	Node of Ranvier	Regeneration of action potentials
Stage 4 Signal Manipulation (**within CNS**)			
Signal correlation	AS	Sense amplifier	Sums signals in multi–dimensional correlators, found in LGN, PGN, pulvinar & cerebral cortex
Stage 5 Information Manipulation (**within CNS**)			
Stage 6 Action initiators (**within PNS**)			
Muscle Activation	AM	End plates	Decodes action potentials at high current level
Enzyme secretion	?	?	Secrete enzymes after decoding action potentials

Figure 11.1.5-2 A tabulation of neuron classes and their common names based primarily on the visual system. See text. The existence of an interplexiform neuron supporting external feedback in the retina is questionable. The orderly naming of neurosecretory neurons has not been pursued in this work.

The receptor neurons of Chordata appear to exhibit the full range of neuro-secretory functions available to the neuron. The ability of the photoreceptor neurons, the vibrissal neurons and the hair cells of both the vestibular system and the auditory system to secrete a protein is well established. These proteins form mechanical structures that support the sensory function. It is likely all these neurons exhibit the same internal topology as photoreceptor neurons, as this would explain their ability to vary their sensitivity significantly with stimulus intensity. It is likely that the receptors of taste and smell also secrete proteins.

The horizontal and the large group known as amercine neurons are all lateral cells in that they accept signals from

neurons carrying signals from different spatial locations (which may also exhibit chromatic differences within the retina. They combine this information using a variety of algorithms that vary with their specific role (Chapter xxx). Pyramid neurons are lateral neurons that are found elsewhere in the PNS.

The ganglion neurons of stage 3 exhibit a variety of individual features that will be described in greater detail in **Section xxx**.

The morphological descriptor's monopolar and bipolar do not relate well to the electrical performance of neurons and the morphological term monopolar will not be used in this Chapter. The bipolar name will be used to define a neuron primarily with respect to its in-line location in the retina between the photoreceptor cell and the ganglion cell in a signal path. This neuron is electrically monopolar (its signal waveforms proceed in only one direction from its resting potential when subject to an input stimulus) while it is morphologically bipolar.

Because of the fact that an active device may be formed upon the juxtaposition of any two membranes associated with neural conduits, amplification, in the broad sense, may occur at two different locations in the nervous system. It may occur inside a given neuron and also between two adjacent neurons. The Nodes of Ranvier are examples of multiple amplifiers within a single cell.

The circuit associated with the Activa of Type AN (the N for Node of Ranvier), will be discussed briefly in **Section 3.3 xxx** because of its prototypical role in all signal projection and hybrid neurons within and outside of the eye.

11.1.5.1 Notation for the ganglion neurons

Figure 11.1.5-3 provides a cross-index of the function and names associated with the ganglion cells in different works. The formal references to the various author headings will appear later in this PART. Following common usage in describing the terminus of the P, Q & R channels, the Y-channel signals are shown as projecting to the reticulocellular region of the TRN

Pathway in Optic Nerve	Type of electrical Input to Ganglion	Ganglion name in this work	Ganglion projects to	Names in Selected Papers		
				Properties of Ganglia	Monkey (Leventhal,'81)	Cat (Boycott,'74)
R (brightness)	Monopolar	Parasol	Magnocel.	Y-like	A cells	α cells
Y (direct signals)	Monopolar	Parasol	Reticulocel.		A cells	
P & Q (chromin.)	Bipolar	Midget	Parvocel.	X-like	B cells	β cells

Figure 11.1.5-3 Correlation of ganglion cell names found in the vision literature.

*Enroth-Cugell & Robinson (1966)
+Boycott & Wassle (1974)

After studying the various papers listed above, there is no obvious conflict among these designations. Wassle, et. al. identified the commonality between the X-Y and α−β classifications noted above[50] in 1975. Some of the descriptions of the cells and their output signal specification are only marginally adequate for identification in a

[50]Wassle, H. Levick, W. :& Cleland, B. (1975) The distribution of the alpha-type of ganglion cells in the cat's retina. J. Comp. Neurology, vol. 159, pp. 175-193

system context. There is a problem with the superficial description of these channels given in many works. Van der Heijden gives such a description that suggests the X-cells are only found in the fovea[51]. He also suggests the X channel is associated with sustained signals while the Y-channel is associated with transient signals. These designations appear to reflect the character of the action potentials streams of stage 3 signal projection rather than with the character of the actual information encoded onto these streams.

There is a problem in labeling ganglion cells of the retina as parvocellular or magnocellular because these labels suggest that all parasol cells connect to the magnocellular section of the LGN and omit the parasol type cells connected to the Pretectum.

11.1.6 Notation of principal waveforms of electrophysiology

There are two principal categories of waveforms and features of waveforms used in ERG collection and analysis. There are the fundamental waveforms generated by individual electrolytic processes within the neurons. These are defined in the next Section in conformance with the top level schematic and other models of this work. They are described by an uppercase Class name (Class A waveform, etc.). There are also a variety of labels assigned to various features of the composite waveforms recorded in the ERG by experimenters and clinicians in order to compare and discuss their findings. These labels are the result of empirical investigations and are described by a lower case dashed name (a-wave, etc.). They are also tabulated below.

The clinical community has adopted a standard for clinical electroretinography. However, it is not based on a theoretical foundation and the waveforms are highly stylized. It was originally published in the Arcives of Ophthalmology in 1989, vol. 107:816-819 and also appears in Heckenlively & Arden[52].

11.1.6.1 Component waveforms of the composite ERG and LERG

Based on this work, it is possible to define the specific waveforms making up the observed composite waveforms of the ERG and the LERG found in the literature. This definition is particularly precise because of the development of the suction pipette technique for measuring the currents associated with the Outer Segment of a photoreceptor cell.

The first four waveforms designated by Class, and the Class F waveform are unique. They are annotated in **Figure 11.1.6-1**. The other waveforms involve signal manipulation, including subtraction, scaling and differential amplification, of combinations of the Class D waveforms and are therefore highly variable in detailed structure.

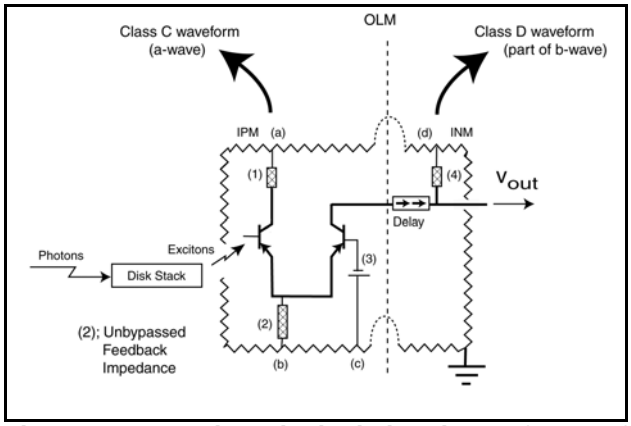

Figure 11.1.6-1 The principal signal waveforms of the photoreceptor cell. The Class A waveform is represented by the jagged arrow and the word "excitons" in the figure. The Class B waveform is measurable only at the open base connection of the left Activa. The Class C and Class D waveforms are in phase opposition and sum to a constant (neglecting the delay). See text.

11.1.6.1.1 The Class A waveform–of the P/D process

The quantum-mechanical exciton flux waveform, generated by the photoexcitation/de-excitation (P/D) process, at the interface between the Outer Segment and the dendrites of the photoreceptor cell (in Stage 1). This waveform is described by the complete P/D equation. The waveform is a direct function of the restricted biological temperature range between the freezing point of water and nominally 50 C.

The photoexcitation/de-excitation waveform (Class A waveform) describes the operation of the chromophoric

[51]Van der Heijden, A. (1991) Selective Attention in Vision. NY: Routledge pp 86-87
[52]Heckenlively, J. & Arden, G. (1991) Op. Cit. pp. 283-288

material of the OS in its in-vivo state. The waveform only exists when the chromophoric material is in functional contact with the dendrites of the IS. Separating the OS from the IS results in a fundamentally different mode of chromophore operation. This waveform can not currently be measured in the laboratory.

11.1.6.1.2 The Class B waveform–at the inner segment, IS

There is a a current waveform, previously un-named, created by the total current at the emitter of the Activa of the adaptation amplifier (in Stage 1). This current creates a negative-going voltage with incident irradiation across the common emitter impedance shared with the Activa acting as a distribution amplifier within the photoreceptor cell. The Class B voltage waveform is measurable as the dendroplasm voltage at or proximal to the colax of the photoreceptor cell. This waveform has not been measured in the laboratory. The charge profile versus time associated with this potential is only present within the emitter-base region of the distributed Activas within the dendrites of the photoreceptor cells.

This "amplified P/D" waveform (Class B') is a description of the electrical signal current through the emitter of the adaptation amplifier. Because of the time constant found in the power source of the collector the amplified P/D waveform, under high contrast (large signal) conditions, is strikingly different from the basic P/D waveform. This is particularly true for long, high intensity illumination pulses. For large long pulses, the adaptation amplifier is driven to saturation. This situation does not occur in the natural world.

11.1.6.1.3 The Class C waveform–at the photoreceptor cell

The Class C current waveform consists of the total current through the surface of the dendroplasm of the photoreceptor cell, distal to the colax, into the IPM (in Stage 1). Probes frequently record a voltage equivalent to this current passing through the impedance of the IPM relative to the local ground plane. This voltage waveform is known as the early potential waveform. The waveform is also negative-going with incident radiation and has led to confusion in the literature since it constitutes a second negative-going signal associated with the photoreceptor cell. The waveform is structurally different from the Class D waveform associated with the pedicel. It occurs significantly earlier and contains components related to the current through the capacitor C_A as well as the recharging current through the dendrolemma related to the electrostenolytic process.

This "amplified P/D with recovery" waveform (Class C* waveform) is a description of the electrical signal current collected within the IPM near the collector of the adaptation amplifier. This waveform is similar to the Class B waveform except it also includes the recharging current that returns capacitor C_A to its original, pre-pulse, value.

The Class C current waveform can be measured by collecting the current emanating from the dendrites associated with the Outer Segment of one or more photoreceptor cells during excitation.

11.1.6.1.4 The Class D waveform–at the photoreceptor pedicle

The voltage waveform at the pedicel of the photoreceptor cell (in Stage 1). Also known as the generator potential. Created by the total collector current of the (second)Activa of the photoreceptor cell acting as a low output impedance distribution amplifier, and the impedance of the axolemma of that cell. The voltage represents the logarithm of the current due to, and is negative-going with respect to, the incident irradiation. This waveform has been widely studied. It is not linearly related to the earlier current waveforms except under small signal conditions.

A summed version of the Class D waveform (the Class D* waveform) is found at the output of the bipolar cells

11.1.6.1.5 The Class E waveform–at the 1st matrix

The Class E waveform is not a unique waveform. It is a family of waveforms created in the processing matrices of Stage 2. It consists of a summation of the waveforms from the lateral cells, primarily from the 1st matrix of the signal manipulation stage in humans. The characteristics of the Class E waveform are controlled by the choice of illumination and the state of the individual adaptation amplifiers in each photosensitive channel. The waveform from each lateral cell consists of the algebraic difference between the voltage waveforms received from two chromatically different types of photoreceptors. Class E waveforms are also created in the 2nd matrix of the signal manipulation stage as part of the creation of the appearance, Z, channels of vision. The encoding of spatial information in the human retina is not a significant function. However, it frequently is in predatory animals. The Class E waveforms as a group are known as S-potentials in the older literature. When recorded from the distal region of the inner nuclear layer (the location of the 1st matrix), they tend to be simple and related directly to the chromatic wavelength of the illumination when recorded using a sweeping monochromatic light source. When

recorded from the proximal region of the inner nuclear layer (the location of the 2nd matrix), they tend to be more complex functions of the total illuminance and the spatial geometry for which they encode.

For purposed of discussion, non-inverted Class E waveforms, relative to the associated Class C waveform, will be defined as Class E_1 waveforms. Inverted waveforms will be labeled Class E_2.

11.1.6.1.6 The Class F waveform–at 2nd matrix

This waveform is actually a variety of waveforms measurable at the output of any of the lateral cells of the 2nd processing matrix. The circuits of this matrix perform a myriad of correlation and integration functions that are similar to the processes carried out in the feature extraction engines of the cortex. Most of the circuits process signals exhibiting time-diversity. This makes their output a strong function of the stimuli applied to the retina. As a result, it is very difficult to interpret probe results from this matrix without a clear understanding of the specific mode of operation of the circuit probed.

Figure 11.1.6-2 provides a caricature of the signals that can be recorded from a single lateral cell of the 2nd matrix. Assume that four photoreceptor cells located at equal spacing (not necessarily adjacent) and in a straight line are excited by an impulse stimulus. Following a specific delay associated with the P/D process, each cell will exhibit a generator potential at its pedicel. These will occur at essentially the same time (except for those cells in the foveola with an exceptionally long axon). Following synapse with other neurons of the INL, and traveling over variable path lengths to reach a neuron of the 2nd matrix, the signals will arrive at that neuron at different times as shown. Assume for the moment that signals 1 & 3 are applied to the non-inverting dendritic input to the neuron and that 2 & 4 are applied to the inverting poditic input to the neuron.

The lower frame shows that three different outputs can be recorded at the axon of the cell, depending on the nature of the stimulus. If all of the selected photoreceptors have the same spectral sensitivity and the stimulus is a field of uniform intensity, The output will resemble the output on the left. The algebraic sum of the signals exhibits ripples that are usually labeled OP's or oscillatory potentials. These will be discussed in a following section under the title OP features. If the stimulus is a semitransparent grating, with an average intensity equal to the uniform field used as a pre-excitation source, and a pitch approximately that of the photoreceptor cell spacing, the response of the neuron will be similar to the center waveform of the frame. This is obviously because the signals from the opaque part of the grating have induced an inverted signal at their pedicels relative to the clear portion. These signals were applied to the inverting input. The right hand waveform is the result of moving the above grating at a specific velocity and in a specific direction under the same illumination conditions as above. All of the signals resulting from the stimulus arrive at the axon of the 2nd lateral matrix at the same time and are summed. The result is a very strong signal specifying the presence of an object moving in object space with very specific characteristics. The direction of the motion is one of these characteristics. ***Reversing the direction of motion will result in a different waveform*** at this neuron. The new waveform will be similar to one of the other two waveforms shown. These are the signal characteristics explored but not explained by Hubel using the cat[53] and employing what is generally

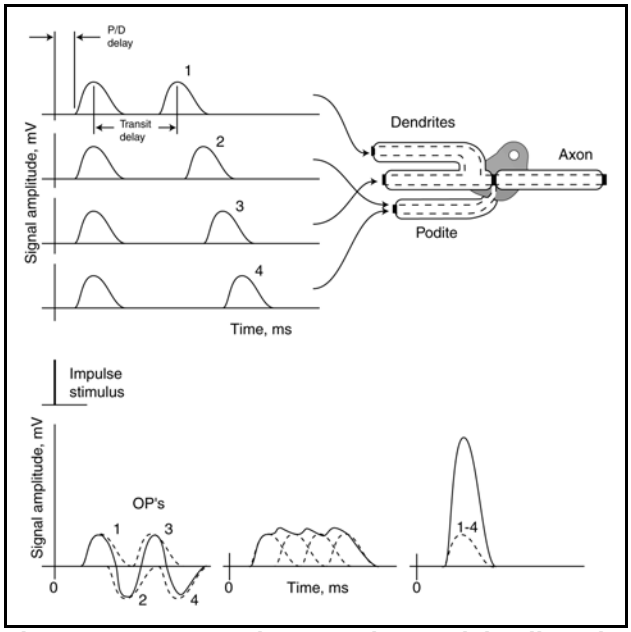

Figure 11.1.6-2 Time and spatial diversity processing in the 2nd lateral matrix. The upper half shows the signals from a line of photoreceptors applied to the dendritic and poditic inputs of a lateral neuron as a function of time. The lower half shows three waveforms that can be recorded at the output of the same neuron depending on the spatial characteristics of the stimulus. See text.

[53]Hubel, D. (1988) Eye, brain, and vision. San Francisco, CA: W. H. Freeman

defined as pattern electroretinography[54]. They exemplify the performance of the 2nd lateral matrix. This matrix is vestigial in humans but is highly developed in hunting animals, including members of the *feline* and the *ave* families. The signal on the right can be passed through to the cerebellum and be used to cause a reflex action extremely valuable in hunting.

The *mechanism* described above, but not the circuit, is the core mechanism underlying all center-surround experiments in the laboratory.

11.1.6.1.7 The Class G waveform–at ganglion output

This waveform is the classic action potential of neurology and is associated with the signal projection stage (Stage 3) of vision. It is a monophasic signal representing the output of either a ganglion cell or a Node of Ranvier. It is created by a monopulse oscillator (in the case of the Node of Ranvier)acting as a signal repeater or as the output of a ganglion cell. The ganglion cell output can take one of two forms, a series of pulses requiring an analog signal above a threshold to initiate the pulse stream or a series of free running pulses with the time delay between the pulses controlled by the net bipolar signal applied to the dendrites of the cell. The free running pulse train is usually associated with a midget ganglion cell. The pulse stream requiring an analog input signal for its generation is usually associated with a parasol ganglion cell.

11.1.6.1.8 The Class H waveform–at the stellate cell

This waveform represents the initial decoding of the Class G waveforms (in Stage 4) to recover the encoded signal from earlier neurons in the signal manipulation stage. It is the result of using a sampled data signal projection system and may contain considerable harmonic distortion relative to the original signal.

The Class G waveform is indicative of the output of a simple time delay pulse signal decoding circuit. This output still exhibits amplitude components that can be correlated to the input pulse train.

11.1.6.1.9 The Class J waveform–following the stellate cell

This waveform is the result of “filtering” the Class H waveform (in Stage 4) to create a signal more closely representing the original signal created in the signal manipulation stage of vision. The recovered signal may be monopolar or bipolar depending on the needs of the cortex.

The Class H waveform is representative of the output of an additional integrator following the initial decoding circuit mentioned above. Although a smooth continuous analog waveform, it may still exhibit inflection points correlated to the input pulse train.

11.1.6.1.10 A complete ERG as an example

Since this work involves developing a model for a single individual signal path and then the aggregating of individual paths to define the ultimate signal path(s) leaving the eye, the “electroretinograph” (ERG) represents a good example of the normal situation found in the laboratory. The ERG is similar to measuring the current drawn from the main battery at a telephone switching office; it is a perfectly valid indication of the health of the overall station but it gives little information about the status of a specific pair of wires connected to a specific subscriber. On the other hand, ERG records provide a surprising amount of information under specific test conditions.

Figure 11.1.6-3. provides a typical ERG recording and the individual waveforms used to create it. The overall waveform is actually a voltage representation of the total signal current waveform as it passes through the impedance due to a ground bridge. This waveform is obtained as indicated above. The first major peak is known as the a-wave and consists primarily of a summation of the Class C waveforms of the photoreceptor cells. The trailing edge of this peak is impacted by the delayed Class D waveform of the photoreceptor cells and the Class E_1 and Class E_2 waveforms of the signal processing neurons. The magnitude and the polarity of the Class E waveforms are a function of the chromaticity of the stimulus. The sum of the Class D, E1 & E2 waveforms can cause the total waveform to go negative in the region of 30 milliseconds delay. If it does, the resultant negative peak is usually described as the b-wave in electroretinography. If it does not, the second positive peak is generally defined as the b-wave. Other -wave designations also become subject to the precise stimulus since they impact the amplitude of the

[54]Heckenlively, J. & Arden, G. (1991) Op. Cit. pp. 291-298

various Class E waveforms that are present.

The individual waveforms can be obtained by measuring the signals produced by a group of similar signaling circuits if proper attention is paid to the ground connection of the test equipment. Conceptually, the photoreceptor waveform can be obtained by measuring the sum of the currents passing between the excited photoreceptor cells as a group and the corresponding group of signal processing neurons **OR** it can be obtained by measuring the return current between the IPM and the INM.

Obtaining a meaningful voltage waveform due to the signal processing stage is more awkward for three reasons. The low frequency currents involved in this stage only circulate within the INM. There are both summing and inverting amplifiers present. Finally, the pulse generators in the ganglion cells share the same ground plane as the signal processing circuits. However, there is a delay within the amplifiers and signal conduits. As a result, the signal waveforms due to this stage can be distinguished in the composite recording. If the voltage of the INM with respect to a truly indifferent location is recorded, a waveform due to the current entering the signal processing circuits will be recorded. It will be followed by a low frequency waveform due to the current leaving the signal processing circuits. The output of either the bipolar or lateral neurons can be minimized in the overall waveform by careful selection of the spectrum of the exciting illumination. In the example, the output of the non-inverting bipolar cells begins at 31 ms. The output of the inverting lateral neurons begins at 37 ms. Both of these signals relate to the initial signal processing matrices. There is also a small signal that begins at 43 ms. It may be related to an additional signal processing path related to the 2nd signal processing matrix or the input stage of the ganglion cells. Summing these signals results in a signal, the solid red line, almost identical to the *complete* experimental waveform labeled 250k in Smith & Lamb[55]. This waveform was unusually well characterized by Smith & Lamb. It was obtained using only blue light, (peak wavelength 450 nm, Lee 195). In contrast to their interpretation of only the leading edge of the waveform, the complete waveform is seen to include the complete a-wave (also known as the generator potential), represented by the dotted blue line, and the later b-wave represented by the polarization sensitive composite of the delayed waveforms.

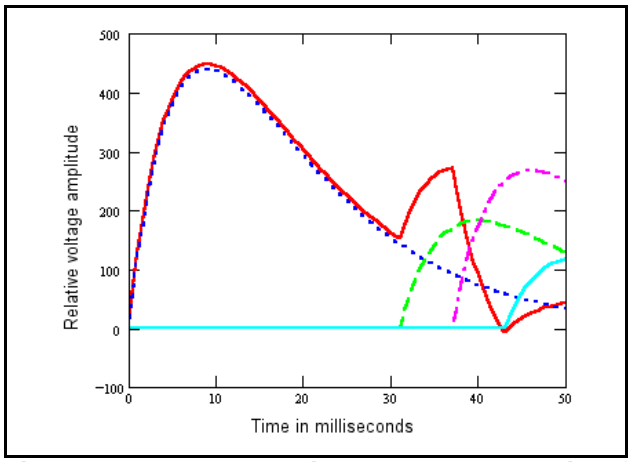

Figure 11.1.6-3 COLOR The ERG as a composite of the photoreceptor and signal processing waveforms. Red solid line, the conventional ERG waveform. Blue dotted line, the isolated waveform of the photoreceptor cells, Class C. Green dashed line, the isolated waveform of the non-inverting signal processing cells, Class E_1. Magenta dash-dot line, the isolated waveform of the inverting signal processing cells, Class E_2. Delayed solid line, an additional signal occurring later in the signal path and related to the non-inverted waveforms. These waveforms do not show the initial 3.36 ms delay due to the transduction process.

Unless signals obtained through electrical contact with the INM are passed through a low pass filter, the recordings are frequently corrupted by the presence of high frequency currents related to the action potential pulses of the ganglion cells.

11.1.6.1.11 Comments applying to the above waveforms

To aid in tracing these waveforms through the circuit diagrams of the following chapters, they will frequently be identified by the above letters enclosed by a circle. When used in the text, the waveforms will be described as a class A or class B, etc. Each of these waveforms exist in the laboratory in two distinctly different forms. One as a result of exciting the system with a long pulse of irradiation and a second resulting from the excitation of the system with an impulse (a pulse of irradiation short compared to any time constant in the system). In the natural world, the visual system never encounters irradiation described by an impulse, except by the observation of a lightening bolt. Recently, man has introduced many sources of very short duration irradiation (such

[55]Smith, N. & Lamb, T. (1997) The a-wave of the human electroretinogram recorded with a minimally invasive technique. Vision Res. Vol. 37, pp. 2943-2952

as strobe lights).

In the laboratory, the use of impulse radiation is key to the interpretation and understanding of the ERG. With this understanding, the ERG in response to longer pulse radiation becomes analyzable also.

In the response of physical, and biological, systems, the amplitude response of a system to an impulse has a unique mathematical relationship to the system. It is defined as the transfer function of the system. When one concatenates a group of individual subsystems, the response of the overall system is easily calculated from the individual transfer functions of the subsystems. The calculation does involve the mathematics of LaPlace or Fourier transformations.

In another situation, exemplified by the ERG, the recorded response is merely the algebraic summation of two or more of the above waveforms. In the summation, each term is modified by a coefficient indicating how much of that term was sensed by the ERG electrode.

The ERG introduces two additional complication. First, to achieve a voltage level above the noise level of the test set, it is normal to illuminate all of the photoreceptors of the retina simultaneously and sum the waveforms due to all of there neurons. Except when examined on a fine time scale, milliseconds or less, the signal paths through the retina (except for some associated with the fovea)exhibit a remarkable uniformity. When examined on a fine scale, differences will be seen reflecting the position of specific neurons in the retina associated with specific photoreceptor cells relative to the position of the test probe. This information can be used to quantify the temporal diversity introduced into the signaling channels of vision as a function of the point of origin of individual signals.

Second, the ERG depends on the capacitive coupling between the electrolytes within the conduits of the neurons and the surrounding electrolytes. The surrounding electrolytes also represent resistive impedances. As a result, the ERG signal is derived from a serial network consisting of at least two sections of RC elements in parallel as shown in [**Figure 11.1.2-3**]. Such a network can pre-emphasize or de-emphasize specific features in the waveforms. This fact must be recognized when analyzing an ERG in detail.

As indicated above, the nature of the group of waveforms labeled Class F waveforms is very complex due to the time diversity technique employed. An interesting LREG experiment could be designed to measure the time dispersal of signals arriving at a specific lateral cell of the second matrix by using a small, narrow spectral band, spot illumination source that could be moved about to illuminate different regions of the retina.

11.1.6.2 Features of the composite waveforms of the ERG

Waveforms of the various class defined above can be combined in a variety of ways as demonstrated in the example of [**Figure 11.1.1-3**] The situation in the literature is more complex because some experimenters have used an impulse (a short flash measured in milliseconds or less) as a stimulus and others have used a long pulse (typically longer than 100 milliseconds). **Figure 11.1.6-4** illustrates the different composite waveforms that may result from these experiments using either type of stimulus. The experimenters have labeled different features of these composite waveforms in an arbitrary way based on their individual experiments without completely recognizing the variability of their waveforms. Because of this practice, correlation between experimenters is awkward. The variability between impulse and pulse stimuli is obvious. The variability as a function of spectral wavelength is more subtle but is shown clearly in the lower frame of the picture where the contribution of the Class E waveform varies significantly. The following sections will define these features of the composite waveform as defined in Heckenlively & Arden[56]. This is the only known text that attempts to define all of these features in a consistent way. Unfortunately, each feature was defined in a separate chapter prepared by a different author. In these sections, the author's name will be given followed by the notation H & A followed by a page number.

---

[56]Heckenlively, J. & Arden, G. (1991) Principles and Practice of Clinical Electrophysiology of Vision. St. Louis, MO: Mosby Year Book

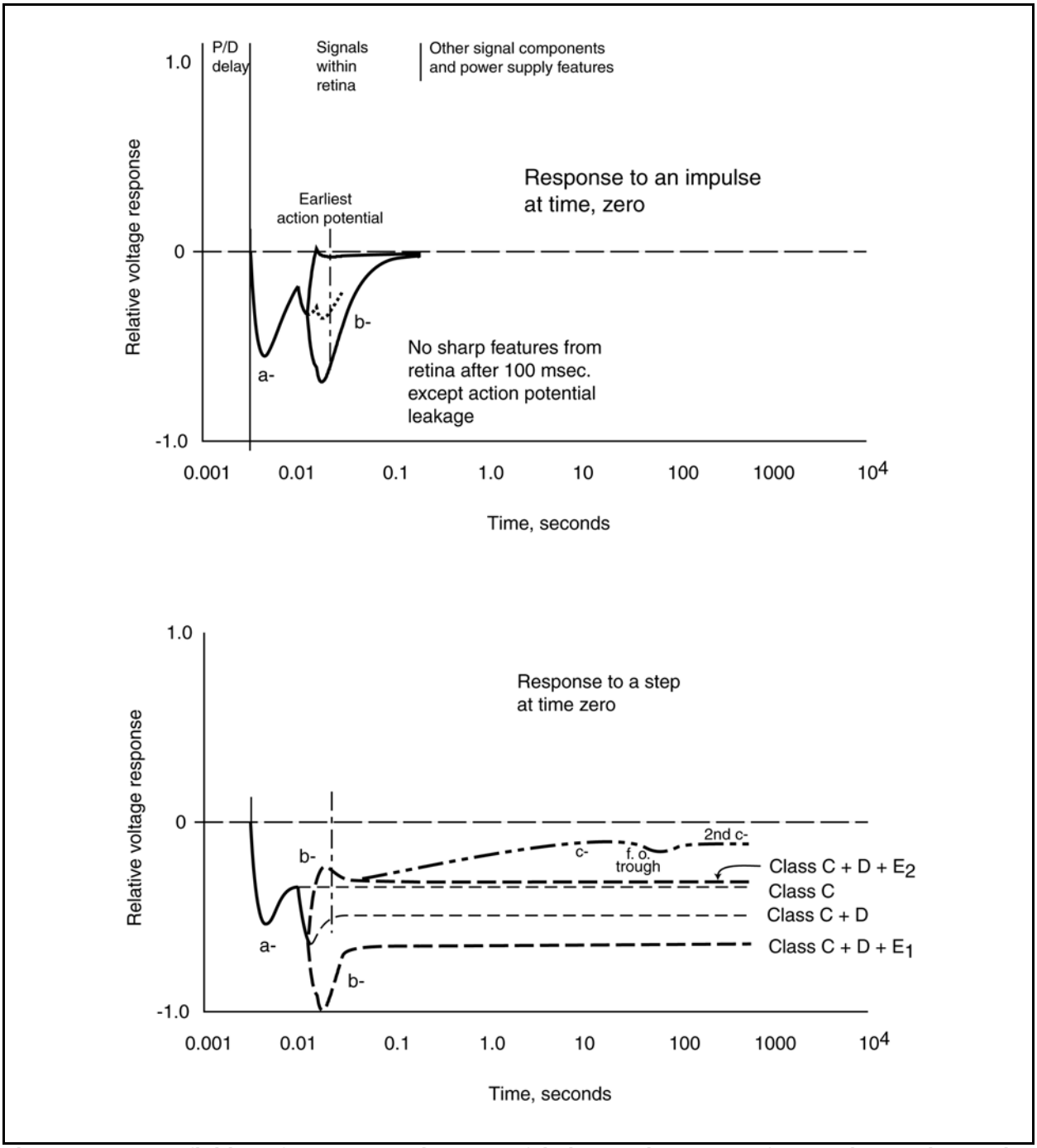

Figure 11.1.6-4 Definition of ERG -waves in terms of their constituent waveforms. The graphs employ a logarithmic abscissa. The upper frame applies to an impulse at time zero. The initial negative going a-wave is formed by the Class C waveform. The following negative going wave is un-named and is formed by the Class D waveform (including the dotted portion). This waveform intersects the Class E waveform. The intersection defines the beginning of the b-wave. The b-wave may be positive or negative going from this point depending on the polarity of the Class E waveform. The ERG in response to an impulse dies out completely within 0.1 seconds. The lower frame shows the same sequence of events for a step change in illumination. Note that each waveform exhibits a steady-state value of about 70% of its peak value (as does the resulting composite). See Text for details.

11.1.6.2.1 The a-wave

In the earliest ERG recordings, it was difficult to tell the initial signal recorded from a transient due to electromagnetic interference. As time went on, this signal was resolved and measured. Using the ERG convention, it was negative going, measured only a few millivolts and its duration was typically less than a few milliseconds. Its amplitude was normally measured from the quiescent baseline to the peak negative deflection. The so-called a-wave came to be defined as this negative going waveform. Some authors have considered the a-wave as only the part of the composite from the baseline to the first peak of the waveform. Others have defined it as extending from the baseline to the trough following the peak. The initial part of the waveform is clearly due to the Class C waveform. Although a small ripple is usually seen just before 0.01 seconds, it is not always resolved. It is due to the Class D waveform.

11.1.6.2.2 The b-wave

The major variation in the composite ERG is usually due to the polarity of the Class E waveform. The polarity can contribute to the formation of a second negative peak or a rapid return of the composite to (or near) the baseline. This situation has led to multiple definitions of the b-wave.

Fulton & Rushton[57] have defined the b-wave, under transient conditions, as the positive going segment extending from the negative peak of the a-wave (the trough of the a-wave) to the peak of the positive excursion. This is the case for an inverting Class E waveform (Class E_2).

Newman & Frishman, in H & A–pg 108 have followed this definition. They discuss the options concerning the source of the b-wave. Their proclivity is to assign it to Mueller cells without strong convictions about the mechanism involved. They also address the alternate view that it might be due to the action of bipolar cells. Their discussion of the origin of the b-wave is based on biochemistry and is not supported here . In this work, the b-wave is dominated by the Class E waveform from a lateral cell in the 1st lateral processing matrix..

Where Newman & Frishman discuss the rod vs cone b-wave, it is only necessary to replace the term rod by M-channel photoreceptor and the term cone by the S- or L-channel photoreceptor to be in agreement with the thesis of this work. The M-channel signal, or b_2-wave, is generally larger and of longer latency. The other channel signals, or b_1-waves are generally smaller and of shorter latency (although latency was not defined in that paper–with a sloping leading edge, larger signals lead to longer latency if one measures to the peak of the waveform). In the first case, they are talking about the long b-wave caused by the Class E2 waveform. In the second case, they are talking about the small positive going peak before the larger negative going peak. This small peak of short duration is due to the Class D waveform associated with a bipolar cell of the luminance processing matrix.

Newman & Frishman, in H & A–pg 108 also discuss a dc-component of the b-wave. They say, “The dc component continues at reduced amplitude in the ERG after the transient response (the b-wave) and then decays at stimulus offset. This component emerges in the dark-adapted ERG at lower intensities than do b- or c-waves.” This component is shown clearly in the lower frame of the figure. Its origin is in the P/D equation, e. g., the Class A waveform. If the stimulus was of sufficient duration, this part of the waveform would be seen to decay back to the baseline due to the adaptation process. For shorter duration stimuli, it appears to be a near horizontal portion of the waveform.

Heckenlively has presented both the scotopic and photopic b-waves based on ERG techniques[58]. There is a significant delay at the lower illuminance levels compared to the higher levels. Bunt-Milam, et. al. have provided some values under pathological conditions[59].

11.1.6.2.3 The c-wave

Investigators beginning with Noell in 1954 have maintained that the c-wave is generated within the RPE because of

---

[57]Fulton, A. & Rushton, W. (1978) Rod ERG of the mudpuppy: effect of dim red backgrounds. Vision Res. vol. 18, pp. 785-792 & 793-800

[58]Heckenlively, J. (1988) Retinitis Pigmentosa. Philadelphia, PA: J. B. Lippincott Co. pg 7

[59]Bunt-Milam, A. Kalina, R. & Pagon, R. (1983) Clinical-ultrastructural study of a retinal dystrophy Invest. Ophthalmol vol 24, pp 458-469

their simple experiment of measuring the voltage between the IPM and the choroid or the choroidal circulation[60]. As discussed in **Section 4.1.2.4.1**, this is not an adequately defined experiment or proposal. It does not demonstrate that the measured potential was related to a "source".

Griff, in H & A--pg 91, used a stimuli with the response measured with a time scale of minutes and describes the c-wave as corneal positive and consisting of two subcomponents, "a corneal negative subcomponent is generated by the neural retina and a corneal positive subcomponent with a similar latency and time-course is generated by the RPE." He points out that if these two subcomponents are equal in amplitude, the c-wave will be absent. He discusses the generally accepted view that the negative going waveform (relative to the cornea) is generated within the neural retina and the positive going waveform is generated within the RPE. These descriptions reflect the limited precision of probing with finite size probes in a very complex environment.

Over periods as long as those used by Griff, the adaptation amplifiers and all subsequent neural stages have essentially returned to the quiescent resting state. The only expected currents would be due to the Class B waveform measured by collecting current from the Outer Segment and the Class B waveform collected at the surface of the Inner Segment. These two waveforms are very closely related and only differ slightly. They are both generated by a single differential pair amplifier. The current measured at the IS exhibits a time delay due to charge transport within the neuron. One current reflects a capacitive component that the other does not. However, this component would not be observed after a period of 100 milliseconds. Griff appears to label the difference due to either the time delay or the capacitance the initial c-wave in his figure 9-2.

This work suggests that the two components enumerated above by Griff and the voltage measured by Noell et. al. are due to a single electrical source physically located in the IPM. This single current source is physically located within the microtubules of the OS and is associated with the emitter terminal of the adaptation amplifier. This current source causes the IPM to become negative with respect to both the cornea and choroid circulation, and with respect to the INM. If the equivalent circuit of this region is drawn, with both the RPE/Bruch's membrane area acting as a high impedance between the IPM and the choroid circulation, and the outer limiting membrane acting as a second high impedance between the IPM and the INM, more relevant electrical properties of this network can be determined. As a result of the change in voltage of the INM, a corneal positive component is measured across the RPE in the terminology of Griff, and a corneal negative potential is measured across the outer limiting membrane. Both of these voltages are of the same polarity relative to the source in the IPM. The fact that the two waveforms are of approximately equal amplitude is in no way coincidental. They originate from the same current source and are both the primary impedance between the IPM and the equipotential (common ground) of the entire body. As a result, the current through each impedance times the value of the impedance is approximately equal to the source voltage due to the photoreceptor dendritic current. Since these two potentials are not employed in the signaling function, they are irrelevant to a detailed description of the visual process. Their amplitude is relevant to quantifying the parameters of the transduction operation.

In this explanation, the trans-epithelial potential (TEP) is not due to any change within the RPE or its individual membranes (apical and basal) but is due to the voltage impressed on the RPE between the source within the IPM and the common ground of the entire body.

11.1.6.2.4 The d-wave

Frishman & Karwoski, in H & A–pg 112, describe the d-wave as an initial deflection in the ERG at light <u>offset</u>!! Their figures are for 0.6-1.0 second pulse illumination. Their graphics show the waveform occurring with varying amplitude and over variable intervals relative to the illumination and to the other elements of the composite waveform. The graphs as a function of penetration retinal depth also show a clear a-wave when near the photoreceptor cells. They describethe d-wave as positive-going in the photopic ERG and negative-going in the scotopic ERG of mammals. In cold-blooded species (frogs), they indicate the d-wave is positive-going when recorded on the cornea or in the vitreous and negative going when recorded in the distal half of the retina. **[their figure 12-1]** They proceed to discuss the controversial aspects of this wave. The duration of their d-wave is similar to their b-wave (that is of the Class E_2 type) under the same conditions.

The fact that the polarity of the d-wave can reverse depending on where it is measured suggests its hereditary quite clearly. It is strongly related to the configuration of the test set. It may be derived from a ground loop current. In

---

[60]Matsuura, T. Miller, W. & Tomita, T. (1977) Cone-specific c-wave in the turtle retina. *Vision Res*. vol. 18, pp 767-775

the context of this work, the d-wave is an artifact of the complex signal processing occurring in the signal manipulation circuits of the INL. or of the re-adaptation process of the photoreceptor cells at the end of the stimulus. Its precise form and timing is dependent on where it originates in space relative to the sensing electrode. Additional data based on flash illumination experiments could clarify the ambiguities related to this feature.

11.1.6.2.5 The e-wave

Karowski, in H & A–pg 115, describes the e-wave as a delayed-off field potential recorded mostly in a very few cold blooded aquatic animals. It has occasionally been recorded in these species trans-retinally. Its primary feature is its occurrence from 2 to over 60 seconds following light offset. It has only been recorded in dark adapted retinas. Karwoski does not provide a graphic representation of the measured e-wave although some of the references may.

With the length of delays indicated, it is not clear how the experimenters relate the putative e-wave to the "off field potential."

11.1.6.2.6 The m-wave

Karnowski & Frishman, in H & A–pg. 118, describe the m-wave as consisting of a slow, negative-going response at both light onset and offset to a small, well-centered spot. They say that "Annular and diffuse illumination elicit complex waveforms, sometimes dominated by the b-wave." It has been described in detail for the cat and the authors suggest the feature appears to originate in the INL near the inner plexiform layer. This work would suggest that these waveforms are formed within the 2nd lateral processing matrix. They are representative of the type of processing associated with center-surround and the spatial orientation of patterns that has been explored most recently by Hubel. This processing involves time dispersal processing of signals on a time scale of milliseconds. The polarity of the signals would be expected to be arbitrary and could easily be as they showed.

11.1.6.2.7 The early and late receptor potentials

The literature contains a variety of references to the early (ERP) and late (LRP) receptor potentials. However, they are seldom defined in detail. Hillman, et. al. have used the term consistently and defined it graphically in a series of papers related to Limulus and the crustacea, *belanus*[61].

The early receptor potential is the feature generated by the Class C current waveform as measured at the OS. It is generally called the a-wave in the ERG. This signal precedes the similar voltage waveform measured (under small signal conditions) at the pedicel of the photoreceptor cell by the transport delay of the signal traveling within the electrolyte of the reticulum of the photoreceptor cell. The distinction between these two waveforms depends strongly on the positions of the test probe.

The late receptor potential corresponds to the b-wave of the ERG. Hochstein, et. al. (in the final paper of the above set by Hillman, et. al.) correctly note that this signal contains two (frequently) antagonistic components, that of the Class D and Class E waveforms. The word frequently has been inserted above since one can obtain b-waves where the only Class E component present is in phase with the Class D waveform.

11.1.6.2.8 The OP or oscillatory potential feature

Karwoski & Kawasaki, in H & A–pg. 125-128, give a brief discussion of this feature. It is described as typically represented by a series of several rapid, low-amplitude potentials superimposed on the b-wave. The number of OP's varying from four to ten (presumably before the next illumination change). The source is not defined precisely but their conclusion is that they originate either in the INL or in the inner plexiform layer (IPL). They support the position that the pulses originate proximally to the photoreceptors based on experiments occluding the central retinal artery nourishing the inner and middle neural layers. They summarize the literature by saying a source within the ganglion cells is unlikely. However, they document a wide variety of OP's found in or near the IPL or ganglion cell layer (GCL). Their final conclusion mentions depolarizing interplexiform cells, presumably within the IPL. These guidelines precisely describe the location of the 2nd lateral matrix defined in this work.

---

[61]Hillman, P. Dodge, F. Hochstein, S. Knight, B. & Minke, B. (1973) Rapid dark recovery of the invertebrate Early Receptor Potential. J. Gen. Physiol. vol. 62, pp 77-86, 87-104 & 105-128

Whereas these features are seldom seen in conventional ERG's they are commonly seen in local ERG's (LERG's). The LERG actually involves a probe penetrating to the layers of the retina. This is understandable when the number of individual signals summed by the ERG is recognized. This degree of summation would average out the signal amplitudes of many of the specialized neurons of the retina. The LERG can probe a much smaller area and isolate individual neurons or only a few neurons.

These OP's of the literature exhibit a variety of pulse intervals and amplitudes, exactly as expected under uncontrolled illumination conditions or no illumination. It appears they can be grouped into two categories, the intrinsic signals at the output of the neurons of the 2nd lateral matrix and stray signals originated by the action potentials of the ganglion cells. The stray signals from the ganglion cells are truly oscillatory in nature but of no technical significance. The designation OP is a poor one for naming the intrinsic signals. Although the signals may exhibit a degree of periodicity, the signals are not oscillatory in character or origin. They exhibit the same periodicity as a string of binary code symbols. A more appropriate name would be periodic potentials, PP's.

In the following paragraph, it is important to differentiate between the frequently observed first ripple in the waveform immediately following the peak in the a-wave and the subsequent ripples after the peak in the b-wave. This first ripple is generally due to the inflection caused by the Class D waveform before the Class E waveform begins to dominate. Subsequent ripples are associated with the set of waveforms defined as Class F waveforms in **Section 11.1.6.1.6**.

Wachtmeister has presented a comprehensive paper on the OP's in the human recorded via the conventional ERG (indifferent electrode attached to the earlobe), primarily at mezotopic light levels, using a 6500 Kelvin pulse source[62]. It bears careful study based on the circuitry proposed in this work. His nomenclature is slightly different and the multiple pulse stimulus protocol complicates the analysis. In his figure 1a, He considers the first major positive peak to be the first OP rather than a sum of both the b-wave and a possible OP. He doesn't consider the OP's to be ripples on the underlying ERG. If his oscillograms are expanded, considerable more structure is apparent than immediately meets the eye and is discussed in his article. The fact that peaks are missing that would be expected to appear if the pulses were due to an oscillation is significant. His peak recorded voltages are in the tenths of a millivolt range, as would be expected using this technique, even though the underlying peak potentials (for a single neuron) are in the tens of millivolts.

Wachtmeister also noted that the "OP's adapt independently of the a- and b-wave to changes in background illumination." This would be expected since the Class C, D & E waveforms exhibit a variable delay relative to the leading edge of the pulse depending on the intensity of the flash stimulus and the prior state of adaptation. However, the Class F waveforms exhibit a nominally fixed delay relative to the Class C waveforms. He noted that the first OP in a series had a frequency considerably lower (about 60 Hz) than that of the dominating frequency of the oscillatory response. Frequency is not an appropriate term here. This statement can be reworded to; the first OP (positive going peak) in a series had a time delay from onset of illumination considerably longer (about 16 ms.) than that of the time interval between subsequent OP features (about 8 ms.). The 8 ms interval is the same as the delay of the peak in the a-wave following the stimulus (This delay includes the initial delay of 3.2 ms due to the P/D process). This situation is consistent with what would be expected under this theory if the OP's were caused by one or more Class F waveform from the 2nd lateral matrix.

Wachtmeister performed a Fourier analysis on the pulses in his oscillograms, usually after discounting the first 18 ms of the waveform by moving the origin to that location. It is difficult to perform a precise Fourier analysis when the pulse train only includes a few pulses. Heckenlively & Arden include a brief discussion pointing out these difficulties[63]. However, a more thorough source should be studied before the indiscriminate use of the readily available software routines.
The Wachtmeister paper and the summary in H & A–pg 322-327, provides a wealth of detailed information concerning the OP's and their relationship to the adaptation state of the human. All of it appears consistent with the position taken in this work. Optimizations for recording the OP's are also listed.

11.1.2.5.9 The PNR–proximal negative response

Karwoski, in H & A–pg 115, describes the PNR as a light evoked field potential than can be recorded in the proximal retina. It consists of a sharp, negative-going transient at both the onset and offset of a small light spot.

---

[62]Wachtmeister, L (1972) *Acta Ophthalm*. vol. 50, pp. 250-269
[63]Heckenlively, J. & Arden, G. (1991) Op. Cit. pp. 237-249

[their fig 13-1]. They stress the spot of light must be centered precisely about the microelectrode tip. The intensity of the spot was not discussed. The use of an annular or diffuse light source elicit complex waveforms that are sometimes dominated by positive-going responses. Dye markings indicate the PNR is maximal in the INL. It has been recorded in virtually all vertebrate retinas examined. The references given suggest that the signal originates in the location of the 2nd lateral processing matrix. It has been observed after sectioning of the optic nerve, under the assumption that this action terminates the functioning of the ganglion cells. It has been observed in many higher mammals, including cat and among primates.

The graphic provided by Karwoski is conceptual. This work would support the conclusion that this is a normal signal resulting from the highly complex signal manipulation within the INL and probably associated with the lateral cells processing spatial information prior to final encoding by the ganglion cells prior to signal projection. In this regard, it would be expected to be more prominent in cats than in man. The effect is directly relatable to center-surround experiments.

11.1.6.3 Features of the multi-focal ERG

The conventional ERG is a global test integrating the performance of a large area of the retina without discrimination. The multi-focal ERG is designed to interrogate and describe the electrical performance of individual sections of the retina. **Figure 11.1.6-5** shows the scheme and presentation from a multi-focal ERG[64].

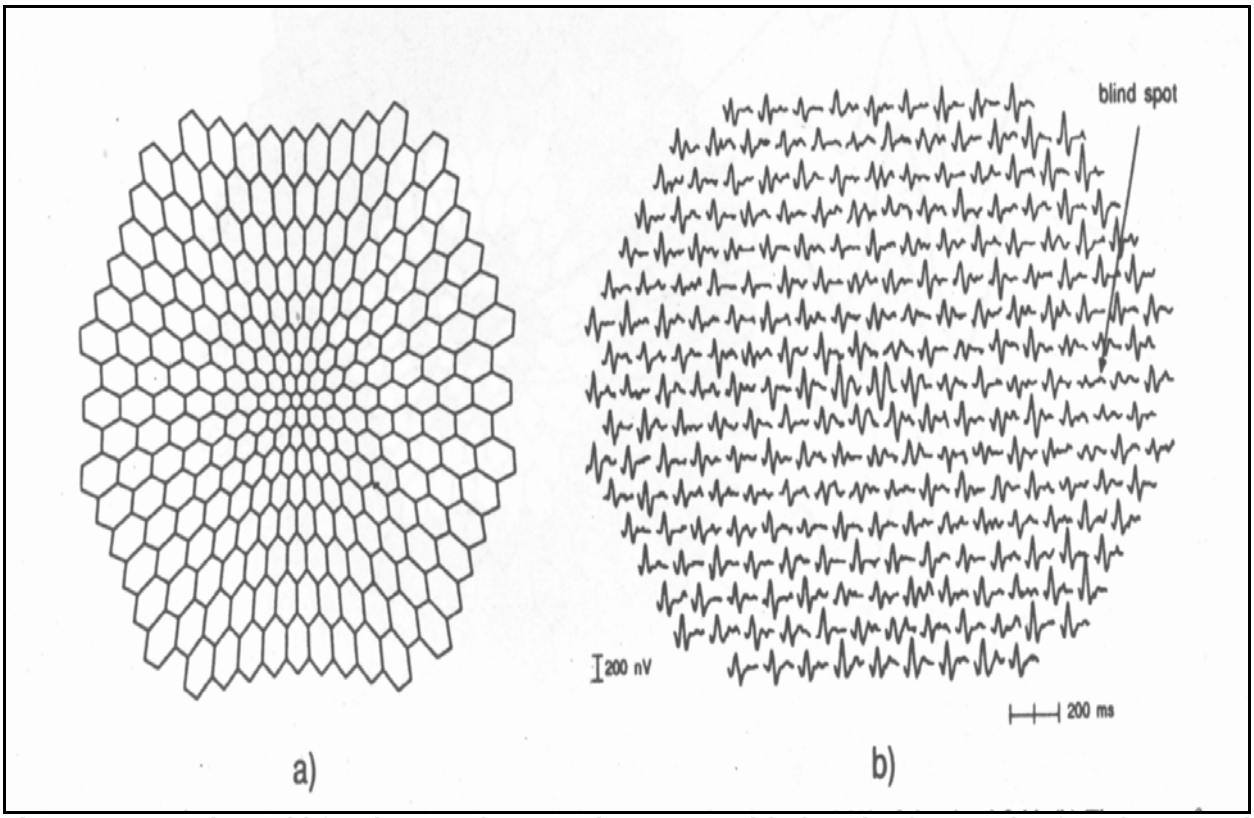

Figure 11.1.6-5 The multi-focal ERG scheme and response. (a) The stimulus matrix of 241 hexagonal elements spanning the central 23° of the visual field. (b) The array of corresponding local lumnance responses derived from a 16 min record. From Sutter & Tran, 1992.

[64]Sutter, E. & Tran, D. (1992) The field topography of ERG components in man: 1. The photopic luminance response *Vision Res* vol 32(3) pp 433-446

11.1.7 Graphical presentations

Although all of the equations required to quantify the visual process are available, there are a large number of variables involved. Many relationships can be understood more easily if graphical presentations are employed. All of the graphical presentations of semiconductor physics can be used profitably. In addition, there are several graphical presentations from photography that are quite useful. One of the most useful graphical presentations employs the multi-quadrant format.

11.1.7.1 Obsolescence of the spectral sensitivity nomograph

Nomographs have traditionally been used in the luminance phase of vision research because the underlying foundation was unknown. The nomograph was a convenient method of bringing some order to the empirically derived data in the absence of a theoretical framework. At the time Dartnall began proposing nomographs to describe the spectral sensitivity of the photoreceptors of the eye, the experimental database suggested a virtually continuous range of different sensitivity profiles among animals. Even in the 1970's, the data base still suggested a continuum of different spectral sensitivities[65]. At that time, there was significant discussion about two points; the considerable scatter of the data relative to the nomographs, and the the systematic deviations of the data from the current nomographs. In hindsight, much of the former was due to poor experimental technique. Much of the latter was due to the inappropriateness of the assumptions upon which the nomograms were based. The empirical data is now showing that there are only a few significantly different spectral responses within the animal community. With the availability of well founded theoretical equations for the chromophores of vision, it is more important to use the actual parameters of the chromophores to calculate their spectral responses. The deviation between the calculated responses and the measured responses are now so small, they actually contribute to a further improvement in both the equations and the test procedures. As will be seen in this PART, the difference between the theoretical and measured values for the absorption characteristics of chromophores of animal vision are now less than +/-2.0 nm. at any wavelength of interest over a irradiance range of 10^7 or greater.

11.1.7.2 Graphs from semiconductor physics

Since most of the operational aspects of the visual process find parallels in conventional solid state semiconductor electronics, it is possible to adapt most of the graphical presentations used in that field. It is only important to realize that most of the circuits of vision involve un-bypassed impedances in the base lead of the Activas. These impedances constitute an internal feedback mechanism. This feedback mechanism frequently perturbs the standard graphical presentations shown in elementary textbooks on solid state electronics. Caution is advised in adapting such elementary presentations.

In most man made solid state electronics, the load impedance found in the collector circuit of an active device is a linear dissipative resistive device. This is generally not the case in the electrolytic semiconductor electronics of vision. The load impedance in vision is generally a nonlinear, non-dissipative resistive device (a diode). The impact of this change is particularly evident in the output circuit of the photoreceptor cell. The diode causes the voltage at the pedicel to be logarithmically related to the current through the distribution Activa of the cell. The graph of **Figure 11.1.7-1** illustrates how this is accomplished. This figure is annotated to show the active, cutoff and saturation regions of the device. For a PNP type device to be active, the collector to base voltage must be negative (the collector must be reverse biased) and the emitter to base voltage must be positive (the emitter must be forward biased). The result is a negative collector current using the standard conventions for these terms.

The basic graph shows the output current at the collector of the Activa as a function of the resting membrane potential and the input current at the emitter. The last parameter is shown parametrically. Note the collector current is essentially identical to the emitter current (typically >99%). This is to be expected in a semiconductor device. To account for the finite load impedance in the collector circuit, a load line is used. The load line reflects the amount of voltage appearing across the load impedance as a function of the current through it. Two load lines are shown in the figure. The straight diagonal line originating at -150 mV represents a linear resistor with an impedance of 3.75×10^9 ohms (3,750 megohms). The curved load line originating at the same point represents a nonlinear diode at 37 degrees Kelvin with a reverse cutoff current of 0.2 pA. The load lines are drawn assuming that the electrostenolytic potential is -150 mV. The quiescent (or resting) axoplasm potential is located at the intersection of the appropriate

[65]Ebrey, T. & Honig, B. (1977) New wavelength dependent visual pigment nomograms. Vision Res. vol. 17, pp. 147-151

load line and the input current at the emitter of the circuit.

These impedance values are extremely high. In the case of a single neuron, it is virtually impossible to measure the voltage in such a circuit with a conventional test set. Only highly specialized test sets such as those employing an open gate field effect transistor (mounted in the probe, not in an equipment rack) can be expected to successfully measure such a voltage without impacting the operation of the circuit[66]. Frequently, even the added shunt capacitance associated with the open gate transistor will cause the circuit to go into oscillation. Fortunately, it is easier to measure currents than voltages in this situation and the relevant currents are easily accessible using an extension of the suction technique developed by Baylor, Lamb & Yau.

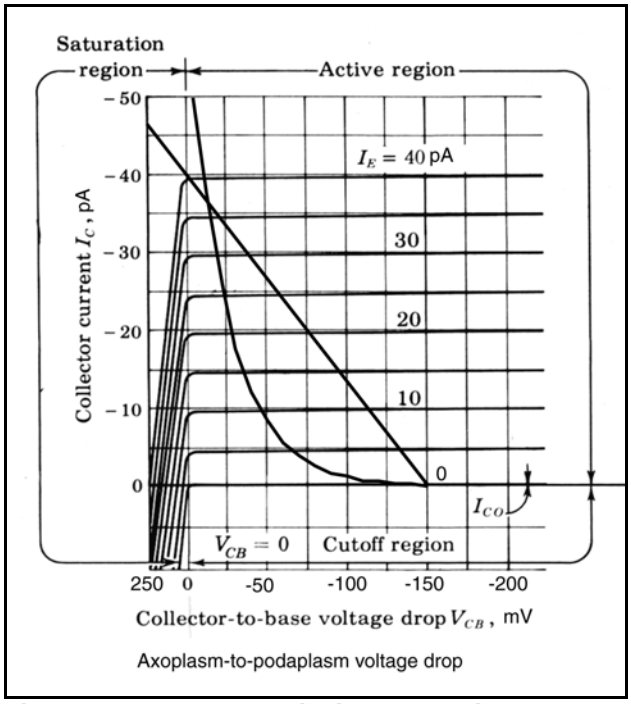

Figure 11.1.7-1 CR Typical common base output characteristic of a P-N-P semiconductor device. The cutoff, active and saturation regions are shown. Two load lines are also shown beginning at zero collector current and -150 mV. Note the change of scale in the saturation region. Compare to Millman & Halkias (1972).

There are several important features to note in the case of the nonlinear load line. First, the quiescent axoplasm potential is a function of the emitter current and the reverse cutoff current of the diode load. The emitter current is determined by the bias potential existing in the emitter to base circuit. This bias is significantly different in the signal detection, signal processing, signal projection, and signal recovery neurons of vision. The nonlinear load line shown is for a diode with a reverse cutoff current of 0.2 pA. For larger reverse cutoff currents, the load line rises more steeply and reaches a given current at a more negative collector to base voltage.

Second, note the *change* in potential between the axoplasm and podaplasm, in the nonlinear case, as the emitter current approaches zero. The potential begins to change more slowly as the current approaches cutoff. It is this characteristic, the axoplasm potential as a function of the dendritic current for a specific load line that defines the intrinsic logarithmic response of the photoreceptor cell.

The maximum current through the various Activa circuits are known to be as high as 40 pA. Torres, et. al[67]. show values between 30 and 40 pA for a newt photoreceptor cell in darkness. They also show a constant current of about half of this maximum value when the same cell was illuminated by a background light.

11.1.7.3 Multiple quadrant graphs

The output characteristic, of the collector (axoplasm) voltage to the emitter current of the distribution amplifier, of the photoreceptor cell can be combined with the input waveform information. When done on a multi-quadrant graph, the output waveform of the circuit is obtained graphically. **Figure 11.1.7-2** illustrates the general graphic form that is widely used to describe multi-variable relationships in many engineering applications, and particularly those involving non-linear relationships. By employing parametric curves in the different quadrants, the effect of as many as 6-8 variables on the ultimate outcome of the situation can be presented at once. This type of graph can be used to illustrate product relationships between variables and amplitude offsets. However, it cannot display summation relationships except in special cases. It is also important to note that a given axis contains only one underlying parameter (or equation). There may be two mathematically related scales, but only one underlying parameter.

---

[66]The Tektronix TekProbe ™ P6243 guarantees an input capacitance of ≤1.0 pF.
[67]Torres, V. Forti, S. Menini, A. & Campani, M. (1990) Model of phototransduction in retinal rods. Cold Spring NY: Cold Spring Harbor Laboratory Press. Pp. 563-573

The highlighted portion of this figure, quadrant #2, is the transfer function defined above. The waveform in quadrant #3 is an illustrative representation of the current waveforms passing through both the emitter and collector circuits of the adaptation amplifier. The bar above the time scale indicates the period of illumination. Because the adaptation and distribution amplifiers of this cell are connected as a differential pair, a rise in the current through the adaptation amplifier results in a decrease in the current through the distribution amplifier. For this example, the distribution amplifier current, I_c, is taken as 20-I_a, the adaption amplifier current, in picoamperes. This is indicated by the dual horizontal scales in quadrant #2.

The amplified P/D waveforms (Class B waveforms) in quadrant #3 are idealized. They emphasize the intrinsic variation in the delay related to the rise of the pulse as a function of illumination level in the P/D function occurring prior to this point. They also stress the fact that all of the pulses in the adaptation amplifier ultimately end at the same point in time but follow different decay characteristics due to the above delay. The amplitude scale in this quadrant is linear for illustrative purposes. It only extends over a limited range relative to the total range of the adaptation amplifier. The time scale is chosen to be long relative to the transport delay in the P/D equation but to be short relative to the time constant of the adaptation amplifier collector circuit. Over a longer time scale, the gain of the adaptation amplifier would change. This would result in a droop in the amplitude of the individual waveforms with time.

Quadrant #1 shows the generator potential waveforms (Class C waveforms) resulting from the input waveforms being applied to the transfer characteristic of the distribution amplifier. There is a finite delay in the signal passing through this circuit. However, it is quite small relative to the delay in the P/D process. It is shown figuratively by the double vertical line near time zero in quadrant #3. The original pulse of illumination is shown beginning slightly to the left of zero to account for this delay. The distribution amplifier does not exhibit a variation in response with time interval as shown. The only change in the output waveforms relative to the input waveforms are a result of the logarithmic compression caused by the transfer function. This situation differs from that represented by Hagins. Hagins[68] assumed that the amplitude versus time function of the leading edge of all the output waveforms was the same and scaled the waveforms in quadrant #1 accordingly. The result was a presentation showing a highly unlikely situation with higher output levels at the photoreceptor pedicel resulting from lower input illumination levels.

A step has been introduced in the illumination generating the highest input waveform for illustrative purposes. The step is completely suppressed in the output waveform due to the cutoff of current through the distribution amplifier. The transfer characteristic also tends to complete the signal normalization process begun by the adaptation amplifier. At the output of the distribution amplifier, the dynamic range is less than 100:1 compared to the scene input dynamic range which may exceed 10^7:1. The highest voltage shown in quadrant #1 is equal to the intrinsic axolemma membrane potential of the photoreceptor cell.

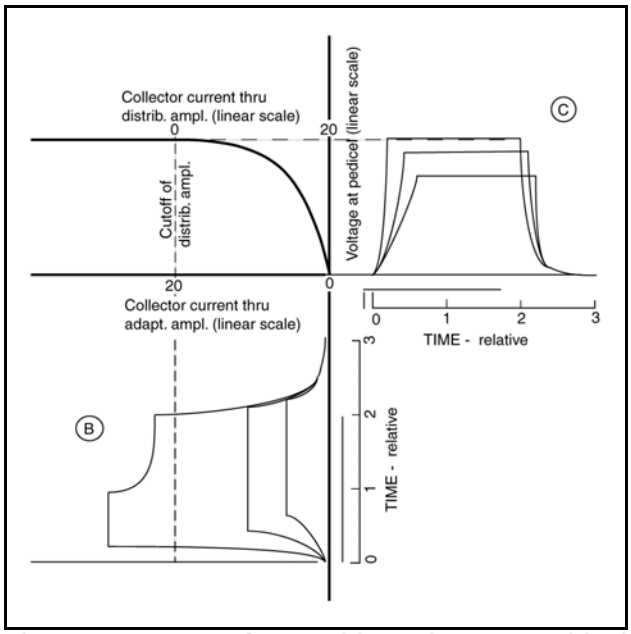

Figure 11.1.7-2 The multi-quadrant graphical presentation.

11.1.7.4 Graphs from photography

Although the visual system operates completely automatically, a presentation known as the reciprocity curve of photography has an important application in vision research. The graph used to illustrate this failure portrays the exposure obtained for various illumination levels and exposure times. In sophisticated presentations, a separate reciprocity curve is shown for each chromophoric layer of the photographic emulsion. The same graphical concept can be applied to vision. It is used in **Section 17.4.1** to describe the overall performance of the human eye as a

[68]Hagins, W. (1965) Electrical signs of information flow in photoreceptors. In Symposia on Quantitative Biology, vol XXX Long Island, NY: Cold Spring Harbor Laboratory. pp. 403-418

function of illumination level.

Figure 11.1.7-3 provides the basic concept behind this graph as used in vision. The horizontal scale extends over the entire illumination range. The vertical scale extends over a much more limited range. This range can be considered the relative voltage range (linear) at the pedicels of the photoreceptor cells or the contrast range (logarithmic) perceived by the animal. The sloping lines represent the transfer function between these two ranges. They represent a 1:1 transfer between the horizontal and vertical logarithmic scales. In photography, they represent the exposure time used to expose the film. Hence, the left scale represents the product of the illumination multiplied by the exposure time. In vision, the sloping lines represent the amplification factor between the input illumination and the output voltage at the pedicel. This amplification factor is adjusted automatically over a very large range primarily by the adaptation process. This range is augmented by the operation of the iris in its secondary role. The object is to keep a selected part, as a minimum, of the contrast range of a scene within the operational range of the visual system. The arrows indicate the change in adaptation amplifier gain required to maintain the contrast range of a scene within the operational range of the visual system. In this case, the object is to keep the solid circle representing the median of the illumination range of the scene in the middle of the perceived brightness range. The amplifier gains shown are illustrative.

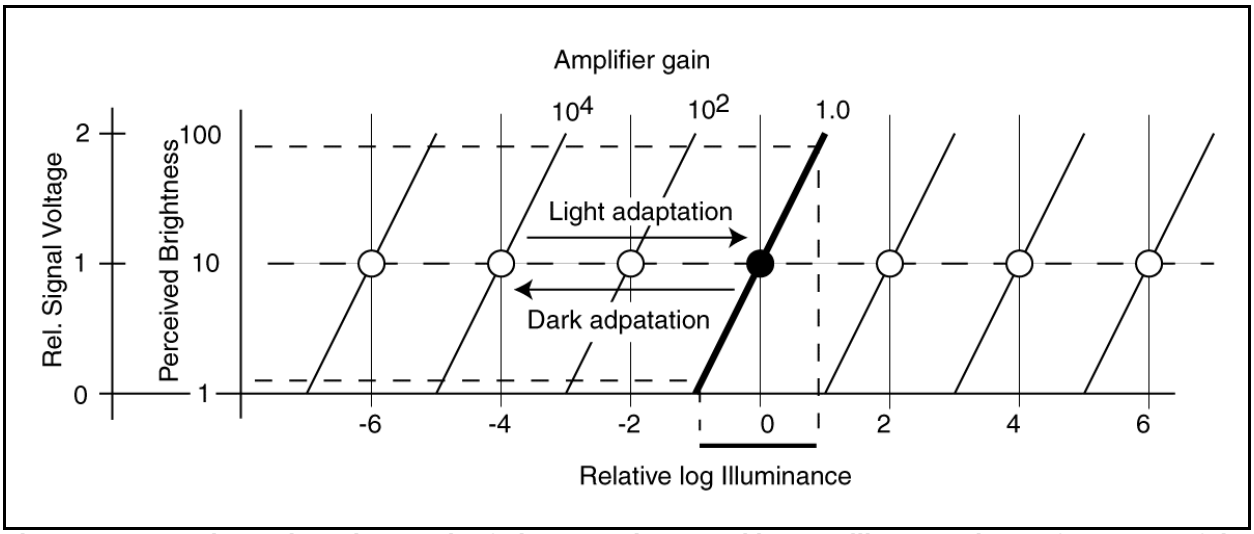

Figure 11.1.7-3 The reciprocity graph of photography. Used here to illustrate the performance of the adaptation amplifier. The diagonal lines are only straight for short and middle wavelength irradiation. By adjusting the adaptation amplifier gain as shown, the signal level impressed on the signal manipulation and subsequent stages of vision is maintained nominally constant regardless of the incident illumination. The circles represent the quiescent signal level in the visual system for a given incident illumination.

11.1.8 Data rates versus information rates

There are a few references in the literature to an authors conceptual understanding of the capacity of the visual system. Estimates of this capacity have generally fallen into the range of 10 to 50 bits per second without any clear definition to what is being discussed[69]. Most of the discussions only relate to the capacity of the foveola related analysis channel based on the type of experiments discussed. Relying only on this channel to define the capacity of the visual system is not completely adequate. Many animals do not have a foveola. Their information capacity is clearly not zero because of this fact.

Most of the discussions in the literature do not recognize, or appreciate, the critical role of the midbrain in the operation of the visual system. Tovee, as an example, speaks of the cerebral cortex by which he means the

[69]Lucky, R. (1989) Silicon Dreams, information, man, and machine. NY: St. Martin's Press. pg 26-31

neocortex to the exclusion of the midbrain[70]. While he presents a variety of interpretations of the structure and operation of the neo-cortex, he does not present any model of the overall visual process to which he can assign individual time intervals that can be accumulated to estimate the actual "Speed of Though." As an example, he does not assign any time delay to the various commissure paths found between the various engines of the neocortex. A more careful assignment would assign time delays to the individual commissure as well as to the various individual engines. In this way, the overall performance could be more realistically discussed. Such assignments require attention to the participation of the midbrain.

The study of information theory has been compared to the "dismal science" of accounting with which it shares a number of features. Although information theory requires considerable precision in the definition of individual concepts and parameters, and requires difficult experiments to obtain precise numerical values for those parameters, the results are critically important to the evaluation of the performance of any system. The text by Lucky was prepared for the mass audience and is not intended to be rigorous in its descriptions of concepts. It merely leads the serious student to other documents.

Lucky does not discriminate carefully between a data rate expressed in bits per second and an information rate, also expressed in bits per second. These are two distinctly different concepts. The data rate is usually used to measure and describe the **capacity of a channel**. The information rate is used to measure and describe the information **content of the signals** transmitted over that channel. This information content may be essentially zero if the channel is not transmitting any meaningful content, i. e., it is not transmitting any signals or is transmitting random noise (truly random noise, not something that is noise like–such as an encrypted signal that is designed to be hidden).

Shannon defined the **capacity of a channel** in great detail and with profound precision during the 1940's[71]. He was working with digital signals over simple channels by the standards of today. However, his results have been extended to include multiple level encoding over both analog and digital channels. Simultaneously, he and his colleagues began the difficult process of determining the **information content** of various signals sent over such channels. This latter work was part of the effort to break the German and Japanese codes of World War II. One of the key aspects of such work is a knowledge of the information content of various character groups generally called words. The information content is not proportional to the number of characters in a word. The information content is a strong factor of the specific characters and their order within a word.

11.1.8.1 The data capacity of the visual system

The data capacity of the visual system varies tremendously within the various signaling paths of the system. As a nominal value, each photoreceptor of the retina can present an analog signal to the following circuitry that can exhibit any one of about 128 discrete levels at intervals of about 0.005 seconds. This corresponds to a channel capacity or data rate of about 128 X 200 or 25,600 bits per second. Various estimates of the number of photoreceptors in a single retina generally fall in the 100 million category. Thus the channel capacity of the photoreceptors of a single human retina, prior to any other processing is on the order of 2.5×10^{12} bits per second or 2.5 million megabits per second. This author would question Lucky's calculation on page 306 for a number of reasons. Photoreceptors do not "fire." They operate entirely in the analog mode. Their maximum passband is closer to 100 Hz than to 1000 Hz. He is correct that there is a considerable amount of manipulation of the raw data signals. In general, the data rate is reduced in order to pass over the optic nerve which consists of about one million individual fibers that can each carry signals at a pulse rate (not an analog passband which is capable of transmitting two data bits per cycle) of up to between 100 and 200 Hz (the higher rate appears to apply to some of the signals from the foveola). The signals transmitted over the optic nerve are in fact monopolar signals carried by action potentials. The encoded signals are actually based on analog signals that may be monopolar or bipolar. The total signaling rate of the optic nerve is then on the order of one million times 100 hundred bits/sec (using a return to zero code) or 100 million bits per second. This is still a substantial rate compared to the above referenced 10 to 50 bits per second. Until just a few years ago, a personal computer could not reach such a high rate.

Once the signals reach the midbrain, they are returned to an analog format for further processing. There are multiple signal paths through the midbrain, some with higher capacity than the optic nerve and some with less. These are described at various levels of detail in Chapter 15. The important point to note that **many** of the individual feature extraction engines of the brain (all operating in analog mode–no action potentials involved) have data processing capacities at least equal to that of the optic nerve. There are a variety of estimates of the number of engines in the

---

[70]Tovee, M. (1998) The Speed of Thought. NY: Springer
[71]Shannon, C. (1948) A mathematical theory of communications *Bell Telephone Journal* xxx

human brain. These numbers usually fall in the 100 to 1000 range but may well be in the 1000 to 10,000 range. Thus the data processing capability of the human brain, calculated by considering all of the individual engines of the brain working in parallel is on the order of one million million bits per second or one million megabits per second. This is a number that the National Security Agency is still trying to use effectively by ganging thousands to millions of individual Pentium class microprocessors.

11.1.8.2 The information capacity of the visual system

Determining a definitive information rate for the visual system is complicated because the importance of profile matching versus raw computational power is not known. It appears patently obvious (but this may be a false assumption) that the visual system operates at a much lower information data rate than the above signaling data rates imply. However, the rates of 10-50 bits per second appears extraordinarily low when considering both the audio and visual (and more limited but useful tactile) capabilities of the human.

11.1.8.2.1 The problem of calculating information content in the analysis channel

This work will not proceed into the calculation of information content. It will only point out that experimentalists have made good progress in computing the information content in individual words but more limited progress in computing the information content of multiple word sentences. The task is made more complex because it involves the human in the evaluation loop. They have still not discovered the rules required to compute the information content of a face, much less the myriad of other objects seen in everyday vision. Experiments have been performed to determine how long it takes to recognize a face or other object. However, the results depend on a great many variables that are difficult to quantify. The information content of audible sounds is still largely uncalculatable. Therefore, the description of the information processing capability of the human brain is still largely conceptual at this time.

11.1.8.2.2 The problem of calculating information content in the alarm channels

A very large role in vision is the detection of changes, particularly sudden changes in the field of view. Such changes frequently relate to a threat to the animal. As will be seen in Chapter 15, a large part of the activity of the midbrain is concerned with recognizing these changes and taking appropriate response measures. These responses can be quite complex and involve much of the skeletal-muscular system. The problem of defining the information content of a specific threat based on trajectory, shape and texture is difficult. Although we allow a pitcher to throw a ball at 90 miles an hour on a trajectory passing within a foot of our chest (and calculate whether it will pass through a box shaped area of about a foot on a side), we take whatever action is possible if he throws it on a collision course with even our foot (which is considerably farther away from the centroid of our body than the target box). In this scenario, the batter has less than one half second to make a very large number of calculations and actually respond. The response time (summing both the neural and muscular response times) is known to take up a significant amount of the total time available. It would appear that information rates considerably higher than hundreds if not thousands of bits per second are involved in this and similar activities.

11.1.8.3 Calculating the data transfer rate of humans

Carpenter & Sutin have illustrated the vastly different signal propagation velocities found in the neural system[72]. **Figure 11.1.8-1** is modified from their figure to show the Activa explicitly within the neuron and to note the difference between the group velocity and phase velocity in action potential transmission. The important point to note is the difference in velocity is related to the mode of signal transmission. The axoplasmic signals propagate by diffusion of charged chemical species within an electrolyte. The action potentials propagate as electromagnetic waves within a coaxial cable type structure. The electrolyte within these structures plays no active role. Carpenter & Sutin suggest the difference in velocity between these two modes is on the order of one million to one. This does not appear to be accurate and may have been due to a typographical error. What they describe as the axoplasmic transport velocity (and is equally applicable to dendroplasmic transport velocity) appears to be closer to 7×10^{-3} meters/sec based on the ERG data of **Section 11.1.6** and the path length data of **Section 17.5.6.1**. This would make electromagnetic transport about 1000 times faster than electrolytic transport. The range in electrolytic transport velocity may be great because it is highly dependent on the diameter compared to the length of the plasma conduit

[72]Carpenter, M. & Sutin, J. (1983) Human Neuroanatomy. Baltimore, MD: Williams & Wilkins, pg 86

and the location of the electrostenolytic process powering the circuit.

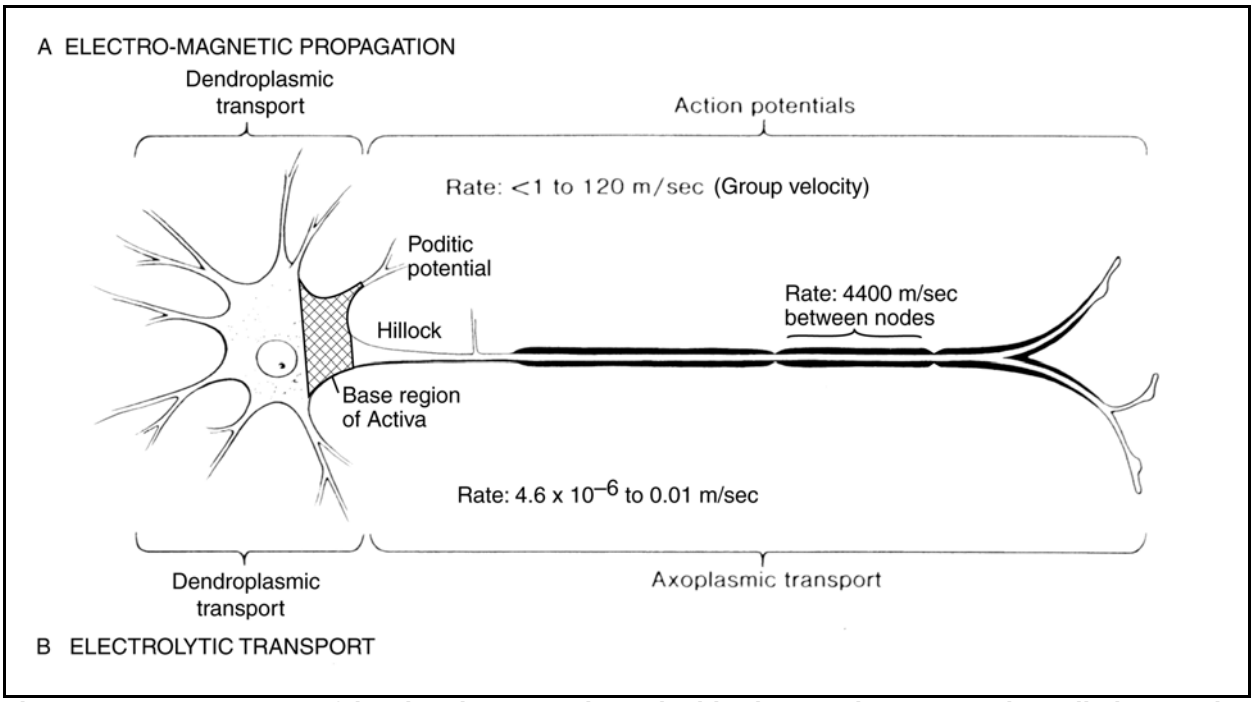

Figure 11.1.8-1 Summary of the signal propagation velocities in neural systems. Three distinct modes of propagation are shown within one neuron. Electromagnetic signal (data) transport is much faster than electrolytic transport. However, the electrolytic transport velocity of Carpenter & Sutin appears excessively slow. See text. The small dimensions of the dendritic tree, compared to the length of the axon, play a negligible role in electromagnetic transport (except in the thalamus). Heavily modified from Carpenter & Sutin, 1983.

The Activa is an analog amplifier found in every soma of every neuron. It is the active mechanism that allows the generation of an output signal of greater power (and potential) than the input signal. The critical factor separating the action potential generator (typically labeled a ganglion cell in stage 3 circuits) from the analog signal amplifier (typically found in stage 2 signal processing circuits), is the relative size and structure of the hillock. The hillock is the source of the large capacitance required to cause monopulse oscillations that are the technical name of the individual action potentials. The hillock also acts as a matching section between the bulk circuit characteristics of the circuit within the soma and the distributed circuit characteristics of the axon. Details of the action potential generating neuron is presented in **Section 9.3**.

A second Node of Ranvier has been shown in the axon on the right. This helps illustrate the difference between the group velocity and the phase velocity of the action potential. Each Node of Ranvier is an active circuit that also contains an Activa.

The ratio between the velocity of action potentials and axoplasmic signals is actually higher than shown if you look at the phase velocity of the action potential signal between Nodes of Ranvier of a single axon. This velocity averages 4400 meters/sec at 37° Kelvin. The average, or group, velocity of the action potentials is reduced considerably because of the time required to regenerate the action potential at each Node. The average group velocity for Nodes spaced at a nominal 2 mm, is 44 meters/sec. This is only one hundredth of the phase velocity. Because of the very small dimensions of the dendritic tree of a cell producing action potentials, there is a negligible delay in the overall signal related to the dendrites.

The slow propagation velocity of electrolytic transport is due primarily to the very low potentials, and therefore low potential gradients found within the neuron between the synapses along the dendrites and the Activa and between the Activa and the pedicle of the axon.

11.1.8.4 Calculating the information transfer rate of humans

There is no adequate model of the information processing system of the human brain or even the visual system at this time. It is conventional to compute the information transfer rate of humans by using simple input rate experiments or simple throughput rate experiments. However, more complex and more revealing calculations might be based on the evaluation of threat conditions in very short time intervals.

The speed records for speed typing (using electronic keyboards) and speed reading are both above 300 words per minute or 1500 characters per minute based on five letter word groups. We won't get into QWERTY versus Dvorak keyboards. While the comprehension rate of a speed reader might be questioned, the throughput rate of a speed typist is more definable. Such a person must read, perceive and transmit signals to the fingers that interpret what he read at the above rate. Using English text, 1500 characters per minute times an entropy of about 3.5 bits per letter[73] corresponds to an information transfer rate of about 87.5 bits per second (achieved while the alarm mode system of the body was still functioning normally). The above is based on the assumption that the typist was retyping random letter groups. If he was retyping recognizable words, the entropy of the letters might be reduced to nearer one bit per letter. This would reduce the above information transfer rate considerably.

Based on the above, and the discussion of the previous sections, the specification of the information *processing rate* of the visual system as 10-50 bits per second appears fanciful at this time. Similarly, specifying such a rate as applying to the information *transfer rate* of humans also seems inappropriate. More work and specific nomenclature are required.

As indicated above, a baseball batter has less than 0.5 seconds to perform a large number of calculations, based on multiple images of the pitcher and the ball, in order to decide whether to swing the bat and also how to hit the ball if the previous decision is positive. However, it is not clear how much of the calculation process relies upon raw computational power and how much relies upon image matching to the memory files (in saliency space) based on experience. Experience (training) clearly plays a large role in the ability of a batter to hit a baseball under professional league conditions.

11.2 Description of the circuits of visual signal processing

It has been common in the past to find very simple electronic models used as analogs for explaining the operation of photoreceptors without significant (i. e. detailed) attention being paid to whether the model fits the actual situation. Usually a very simple voltage based model is assumed without regard to the impedance levels involved. Because of the impedance values found in the eye, it is frequently more appropriate to use a current based model. The current based model is not commonly used in modeling because most people are never exposed to it in early training in electronics. It is frequently taught that Thevinen's Theorem allows the interchange of a voltage source for an equivalent current source; this is true, however, the proper choice is frequently determined by the impedance of the circuit, not just its resistance, i. e. the capacitance as well as the resistance of the circuit must be considered. In addition, the type of circuit involved may make only one type of modeling appropriate.

An additional complication is the quantum nature of light. The eye generates individual electrons in response to individual photons. The quantum nature of these individual electrons is important in the early stages of the operation of the eye. These characteristics are disguised or lost if a voltage based model is used inappropriately in a circuit analog.

In biological material, there are a variety of ways that electronic signals may be communicated from one point to another. These may or may not involve the physical transport of electronic charges between the two points. Most texts discuss skeletal neurons as transmitting their signals through a progressive breakdown in insulating properties resulting in an apparent motion of an electrical "wave" down the length of the axon. However, no breakdown in the dielectric material of an axon has ever been documented under in-vivo conditions. The actual wave is similar to a wave in the ocean. The water actually stands still, only moving up and down vertically as the compression phenomena passes. In other situations, a charge may be present on the surface of a membrane and it may respond to a potential difference and move along the surface of the membrane. This occurs without any significant participation by the metabolic elements associated with the membrane, or in the case of a single cell, without any significant participation by the elements inside the cell.

It is still not widely accepted or taught in neurology that a majority of the neural cells of the eye operate in the

---

[73]Lucky, R. (1989) Op. Cit. pg 137 (or any book on cryptography)

analog mode and communicate with each other through an electronic synapsis (that does not involve the use of a chemical transmitter). However, this appears to be an easily demonstrated fact. In very recent literature, the "gap junction" is described as an electronic synapsis.

The output signal generated by a biological or electronic circuit may be either a voltage (potential) or a current (charge). A critical process in vision is the logarithmic conversion of a current (related to the input photon flux) into a voltage.

It appears that the physical coupling between the elements of the signal path in vision is very tight. The circuits are not easily separated one from the other like they are in most discrete electronic circuits where capacitive or inductive coupling is easily recognized. Because of this fact, it is sometimes difficult to determine what element is being sensed by a probe. Is it the output node of the preceding cell or the input node of the following cell? The question even arises; are these two separate nodes? As will be shown below, these are separate circuit nodes connected via a non-linear active semiconducting device. This device is known as a synapse among the morphological community.

The critical parameter of the active device between an axon and a neurite is its forward current carrying capability, described (ironically) by its maximum reverse current parameter, I_o. This parameter is typically very small. This causes the impedance level at the output of the synapse to be extremely high. As a result, it is extremely difficult to measure either the signal or the impedance level at this node.

Furthermore, the size of any capacitive element in the eye is extremely small. The capacitance may be an order of magnitude lower than that of a man-made probe, i. e. the impedance of the circuit in the eye may be an order of magnitude higher than that of the probe. This makes physical probing of biological tissue a questionable technique, just as it is in the highest density microcircuits.

Many investigators choose to select an equivalent circuit for the input stage of a neuron from among a very limited set; they usually consider a linear summation circuit (with or without a blocking diode associated with it) or a boolean summation device. There are many other possibilities. Considering the input impedance as logarithmic appears critical to the proper understanding of the signal manipulation function in vision.

11.2.1 Nature of the active devices within a neuron

Some investigators have attempted to model the active devices within the visual system using operational amplifiers at the symbolic level. Others have proposed analogs based on field effect transistors at the circuit element level. Still others have proposed junction type transistors, generally without attention to the operating polarity of the devices.

This work has found <u>there is no need to employ an analog to the actual active devices found in vision</u>. All active devices within the visual system have been found to be charge injection type junction devices. They are described as P-N-P type active semiconducting devices. These devices are based on electrolytic rather than solid-state physics. They are formed using liquid-crystalline as opposed to the more widely understood liquid and crystalline states of matter.

The actual devices found within a neuron use the dendroplasm and podaplasm of the cell as signal (current) summing nodes. The axoplasm is used as an output signal voltage node.

There is another interesting aspect of the use of a charge-injection device in a retinal nerve. Such a device has the ability to operate in an analog (electrotonic) mode or a pulse (action potential) mode by merely changing the impedance connecting the circuit to the electrical (metabolic) power supply. This means that there is a single kind of charge-injection device that can be used to form all of the known signal processing neurons in the retina.

In some biological investigations, experiments have been carried out to measure the change in the internal potential of an individual cell, frequently relative to an indeterminate external potential point, without any demonstration that the internal potential is related to the process being studied or that the external potential point is independent of other signals. This can lead to many problems. Further, many times a probe is used that penetrates the cell wall without any demonstration that the act of the probe penetrating the wall of the cell does not affect the potential of the cell or the reaction of the cell. One of the problems of using an inappropriate external potential point as a signal return is the incorporation into the signal displayed on a recorder of waveforms due to an entirely different process. This is frequently the cause of a signal which appears to display both slow analog (tonic) signals and fast binary (phasic) signals combined. This is a real problem with high impedance probes.

A current generator can be replaced by a charge generator, a process involving individual units or quanta. However, a current generator cannot be replaced by a voltage generator in quantum processes. A voltage generator creates a potential field which is not quantized. In the case of modeling a photoreceptor cell, this may be of critical importance. A charge generator can be associated in a meaningful way with the transduction of a photon into a neural signal. Similarly, a noise generator can be incorporated into a model using a charge generator in a meaningful way to indicate the relevant amounts of charge generated due to the input signal and the internal noise of the system.

The selection of an appropriate model for the process being studied is critical; and it is equally important to demonstrate that the test technique is not affecting the results--or that the effect of the test technique is treated as a variable and accounted for.

11.2.2 Neurophysiological-reductionistic models

Uttal has discussed a series of reductionist models of vision popular during the latter half of the 20th Century[74]. Reductionism may be useful in the classroom. However, it is foreign to serious research. He presented four categories of these theories;

+ the single-neuron theories

+ the spatial frequency filter theories

+ the neural network theories and

+ the neuroelectric field theories.

Each of these approaches and the theories included within them is far too simplistic to be of value here. Only one of them will be discussed below. The rest will be discussed as they arise in this PART.

11.2.2.1 The single-neuron theories

The essential premise of this theory, according to Uttal, is that single neurons of the nervous system exhibit highly selective trigger-feature sensitivities to particular aspects of the spatiotemporal pattern of stimuli that limit the range of stimuli to which they will respond. The only comment that can be made here is that this was a generalized theory of an earlier time when the true nature of the neuron was not understood. The vast majority of neurons do not operate in a binary mode, nor can they be considered comparators. They employ a fundamentally analog internal device known as an Activa. When associated with the appropriate additional electrolytic circuit elements, they can be made to perform in many modes, only one of which is associated with the capability defined in the single-neuron theory.

As Uttal notes, and this author agrees, “there is not the slightest hint of any rationalistic overtone in any single-neuron theory of perception; these theories are all strongly empiricistic.”

11.3 Baseline electrical analogs for vision

A very simple transistor circuit of the grounded base type appears to provide an attractive choice for the basic building block to be used to describe the majority of neurons. It:

+ is charge based (not field based) and preserves the quantum nature of the signal.

+ can also be made to operate in a number of different modes without involving external feedback or complex external circuits.

---

[74]Uttal, W. (1981) A taxonomy of visual processes. Hillsdale, NJ: Lawrence Erlbaum Associates. pp. 125-127

+ can provide “voltage gain”

+ can consume no power in the quiescent state.

+ can be built with multiple independent input circuits

The basic transistor device is shown in **Figure 11.3.1-1** along with several simple circuit variations. The basic device is shown in **Figure 11.3.1-1(A)**. It is known as a three-terminal device and it can be configured to accept an input signal at any terminal. It consists of three distinct regions of material. The central region is of the n-type in semiconductor physics. It exhibits an excess of electrons at the lattice level. The two adjoining materials are of the p-type. They exhibit a shortage of electrons at the lattice level. By establishing a suitable junction between these pairs of materials and minimizing the thickness of the n-type material, an active electronic device is formed. In the solid state, it is known as a transistor. In the electrolytic state, it has been named the Activa. An electrical connection to the n-type material is known as the base lead. Since the device is symmetrical, a connection to either piece of p-type material can be considered the input connection. It is conventional to show the input connection on the left, label it the emitter, and attach an arrowhead to the input lead at the point of contact with the bulk material. The emitter, emits carriers that pass into the base material. The remaining terminal is considered the output terminal. It is called the collector. It collects the carriers emitted by the emitter.

In the absence of biasing voltages, this device is inert and its electrical properties are indeterminate. When the proper voltages are applied to the device, it exhibits valuable electrical properties.

For this application;

+ the base is shown connected to ground and the input will be applied to the Emitter lead.

+ in **Figure 11.3.1-1(A-C),** the device will be considered a charge transfer device, not a current carrying device, although the derivative of charge with respect to time is a current.

+ in **Figure 11.3.1-1(D),** the device will be considered a current carrying device.

Figure 11.3.1-1(B) shows the device connected in a minimal configuration. The capacitance shown is the total capacitance at this point; it includes the capacitance of the output circuit of the transistor shown and the input capacitance of any following circuit. With a reverse bias (negative voltage as shown) applied to the collector through an impedance and no voltage applied to the emitter, no power is consumed by the device. However, it is operational. If a negative charge is applied to the emitter, the charge will be transported to the Collector and appear as a charge on the capacitance at the output. The second emitter shown dotted is to indicate that this device can be made with more than one emitter. The emitters will act independently. However, the total charge at the output will be the sum of all charge entering the device.

Figure 11.3.1-1(C) adds a reverse bias to the input circuit. The total circuit still consumes no power. Now, if a charge is applied to the emitter, nothing will happen. If the bias voltage on one of the inputs is reversed and made positive, the device will now transport a continuous current from that emitter to the collector. This current can be increased or lowered by varying the charge on the capacitor as before. The circuit now consumes power because an average current is continuously flowing through the resistor located in the output lead.

Figure 11.3.1-1(D) illustrates several important additions. Depending on the value of the different electrical elements, the circuit can perform in a variety of modes. If an impedance is placed in the base lead and the input capacitance is made large, the transfer characteristic of the device can be changed dramatically. By adjusting these values and the input bias, the circuit can be made to amplify or oscillate. It can generate a single pulse or generate a string of pulses. The width and time between these pulses can also be controlled by adjusting three parameters, the impedance, Z_e, the total capacitance, C_e plus C_o, and the applied field (the total voltage equal to the algebraic sum of the bias voltage and the voltage across the impedance, Z_e).

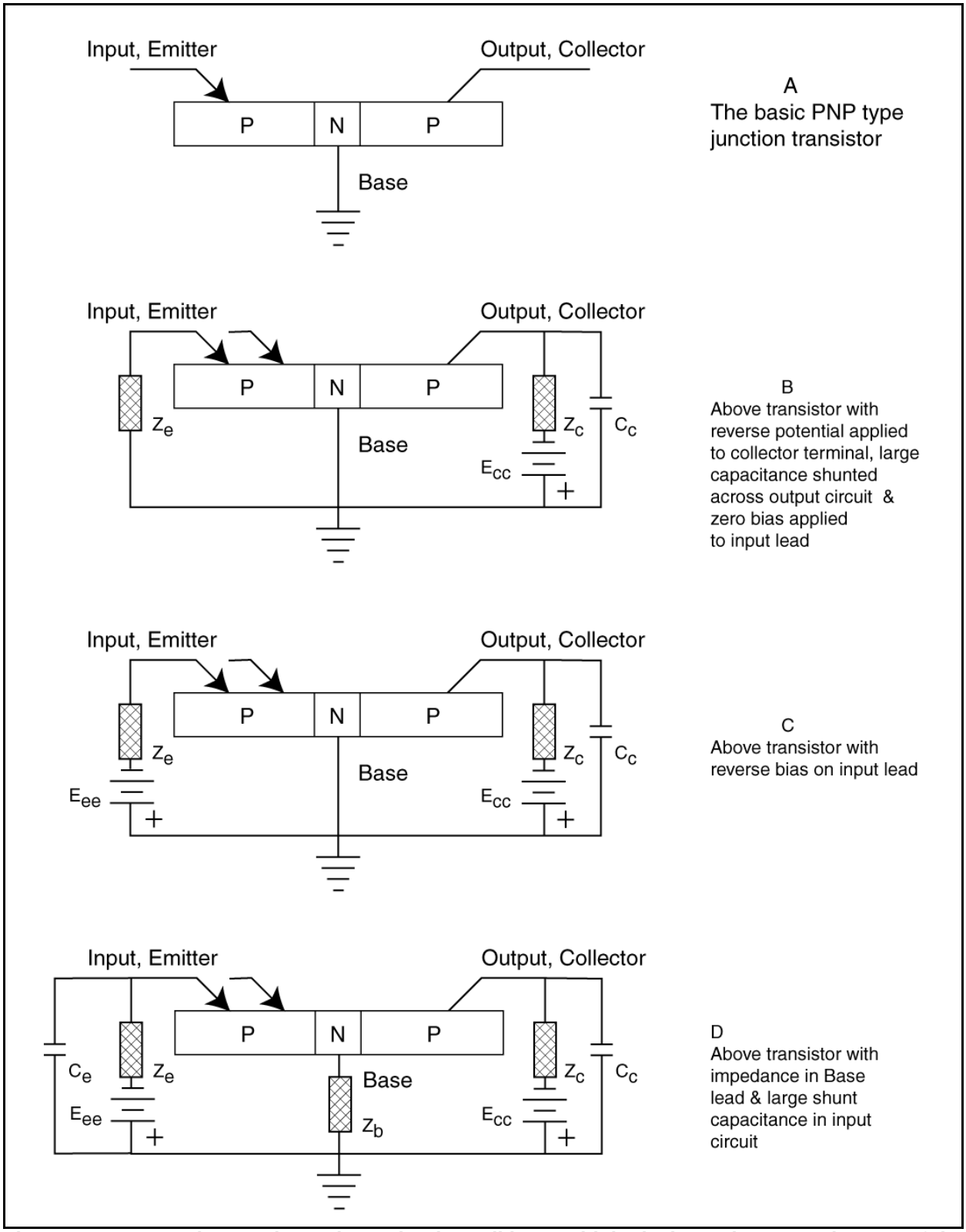

Figure 11.3.1-1 Fundamental transistor circuits. All known biological transistors, Activa, are of the PNP type. See text for discussion of symbols

The nature of the impedances, Z, will be discussed later. Be cautioned however. They are not resistors.

All of the data unearthed to date supports the description of the circuits in the retina in terms of a single biological activa of the P-N-P type. This activa has three functional ports, although the dendritic port may be replicated, as shown in (**D**). The associated input and output circuit elements may make the overall circuit associated with a given biological activa exhibit multiple inputs and outputs. Since the majority of the circuits are directly coupled, it becomes increasingly hard to define boundaries of a given circuit group as the complexity of the retina increases.

11.4 Top level power distribution/consumption

The source of power employed in operating the neural system is seldom discussed in the literature. It is usually ignored or assumed to be a normal metabolic responsibility of the Soma of the cell. The neurological system is actually powered by an electrostenolytic process associated with several individual specialized regions of the plasma membrane of each neuron. Components within the Soma, which play a major role in chemical synthesis, play no role in electrostenolytic process associated with neural signaling.

In the absence of a proper model, it has been common in the literature to encounter analyses of specific processes based on thermodynamic principles. Unfortunately, these have not been the thermodynamic principles used in the neuron and vision. Man normally builds machines based on irreversible, or dissipative thermodynamics involving a Carnot cycle. The neural system of animals has evolved into a system that does not employ significant dissipative elements and does employ completely reversible electrochemical reactions. Any analyses found in the biological literature should be examined very carefully if it is using irreversible thermodynamic principles. The neurological system in animals employs *reversible* thermodynamic principles exclusively.

With the signal paths in the retina fairly well correlated to the histological layers, there is another aspect associated with the signal circuitry of interest. How is the system energized? How is the power necessary to generate and process these signals supplied? One of the technical characteristics of a direct coupled biological transistor chain is that it can be provided power at a number of different points in the chain. For a number of reasons related to total power consumption, crosstalk and individual amplifier stability, it is frequently best to adopt a specific power distribution plan in such circuits. There are at least two approaches to determining the power distribution scheme used in the retina of an animal. Wong-Riley and associates have looked at a number of retinas to determine where metabolic byproducts of power consumption (conversion) are concentrated. Heynan et. al. have looked at the location of electrical current sources and sinks.

Wong-Riley et. al.[75] have demonstrated through sectioning that there are areas of high metabolic activity identifiable through cytochrome oxidase (CO) staining. They relate dark staining to high metabolic activity. Their results indicate: "In all species examined, the most intense oxidative enzyme staining was observed within the inner segments of rods and cones. In this layer, cones were more reactive than rods . . . The photoreceptor outer segment layer and outer nuclear layer were only lightly reactive." They proceeded to investigate and stated: "Within the inner third of OPL and the outer half of the IN, however, darkly reactive cells and processes were observed. Based on their location, these cells were identified as horizontal and bipolar cells." They proceeded: "In each species, retinal ganglion cells were usually moderately to darkly reactive for CO . . . Each animal exhibited slightly different patterns of RGC staining."

In their discussion, the above authors develop the following points:

> "The intense staining of inner segments . . . suggests that a large portion of the energy derived from oxidative metabolism may be utilized for the maintenance of ionic gradients.
>
> The selective dark staining of the large mitochondria in cone pedicels, in contrast to the light staining of the smaller and less frequently observed mitochondria within rod spherules, implies that there is a fundamental difference between rods and cones in the localization of oxidative energy metabolism.

---

[75]Kageyama, G. & Wong-Riley, M. (1984) The histochemical localization of cytochrome oxidase in the retina and LGN of ferret, cat & monkey------ Journ. Neurosci. vol. 4, no. 10. pp. 2445-2459

Horizontal cells. The dark staining of the primary horizontal cell processes and somata located within the inner OPL and outer IN is evidence that these cells are metabolically quite active.

Bipolar cells. The moderate to dark staining of smaller, more vertically oriented cells in the outer IN suggests that many bipolar cells are also metabolically very active.

Amercine cells. Although amacrine somata are characterized by a sparse distribution of lightly reactive mitochondria, the distal amacrine cell processes (in IPL) may exhibit a level of oxidative enzyme staining that is moderate to dark . . . This pattern of staining is consistent with the hypothesis that peripheral processes of neurons may serve as functionally segregated active components of a neural circuit or simply as distal sites of intense synaptic interaction.

RGCs. The presence of moderate to dark staining within RGC dendrites indicates that dendrites contribute to the overall intensity of staining within the IPL. The moderate to dark staining within most RGC somata may be accounted for by their moderate to high tonic maintained discharge rates. . . ."

These authors did not demonstrate that the process they are tracking is directly related to the Soma or other internal elements of a cell related to metabolism. By differentiating between the metabolic and electrostenolytic processes of a cell, their statements appear to support the analyses of this work. As discussed elsewhere, they did not show a causal relationship between the level of illumination and the generation of stainable compounds at any point in the visual process. It is also interesting that the density of staining of small areas of the OS was much lower than that of equivalent areas of the IS. However, the OS is a much more distributed structure than the IS. If the cumulative staining associated with all of the OS were compared to the cumulative staining of all of the IS, a different result might have been obtained.

Heynen & van Norren[76] probed the in vivo retina of macaque monkeys. Their technique was essentially extracellular and sampled a highly heterogeneous structure. However, their results provide a useful map of the resistance level at different levels in the retina and the location of a number of current sources and sinks. Their work indicates a major current source in the area of the outer segments of the photoreceptor cells, and a major current sink in the area of the outer nuclear layer--and that these two areas have a high mutual correlation coefficient. Regions located closer to the vitreous fluid did not show a high correlation to either of these areas. They associate these two areas with the generation of the a-wave (or receptor component). Their data also showed a source-sink profile which they associated with the b-wave. In this case, the current sink was centered in the region of the ONL and the current source was centered in the region of the IN. Their data was not stable enough to indicate anything about the ganglion layer.

The above work provides a fairly clear and detailed description of the power distribution system in the retina. Recognizing the reversible techniques used in vision, the term power consumption will be replaced here by power conversion. This is more appropriate since very little of the power converted at a given location would be consumed there in any event. In many cases, the chemical energy converted into electricity at one point will be converted back to chemical energy at another location.

Photoreceptor cells. It is quite clear that the photoreceptor cells are the sites of significant power conversion in the retina as might be surmised. It is quite clear that the IS exhibits the most concentrated site for the conversion of that power (and that the power is used to provide a circulating current between the IS and the OS represented by the source-sink relationship found by probing). The high power conversion in (or near) the mitochondria associated with the pedicels probably relates to the supplying of power to the axoplasm for use by the Activa within the neuron.

Horizontal cells. There is evidence of power conversion in the structural "processes" of these cells, i. e. the soma is not the only location of power consumption.

Amercine cells. There is significant evidence that there is considerable power conversion associated with the non-soma elements of these cells, i. e. an Activa not physically or topologically associated with the soma.

Retinal ganglion cells (RGC's)--There is significant power conversion related to the dendrites of the

---

[76]Heynen, H. & Norren, D. van (1985) Origin of the electroretinogram in the intact macaque eye--II Vision Res. vol 25, no 5, pp. 709-715

RGC's. Overall, the RGC's are not major power consumers, as expected for analog to pulse encoders.

Overall, the above characteristics continue to suggest that:

the photoreceptor cells are power amplifiers, i.e., the output impedance of the activa is lower than the input impedance. All of the output current is applied to an impedance that creates a voltage node at the pedicel.

the horizontal and amercine cells are high impedance signal amplifiers consuming relatively little power per cell.

the bipolar cells are primarily high impedance buffer amplifiers consuming relatively little power per cell.

the RGC's are efficient analog to binary encoders which consume relatively little power.

The above conclusions are in agreement with the conclusions drawn elsewhere in this work**.**

It should be noted that the description of a major current source in the area of the outer segments of the photoreceptor cells, and a major current sink in the area of the outer nuclear layer relates to the external characteristics of the photoreceptor cells. Internally, the current would flow from the region of the outer nuclear layer (or IS) to the region of the OS. The above description applies to a conventional current, and not the actual electron flow which is in the opposite direction.

11.5 Principle Signal Sources

In probing the retina, it has been possible to localize the origination point of a number of waveforms. The descriptions, and background given by Rubin & Walls[77], although dated, make a good starting point. The a-wave originates in the region of the photoreceptor inner segments (IS), at least 100 microns from the vitreous side of the (primate) retina. The b-wave is found to originate in or near the IN, typically at a depth of 75-100 microns. It is frequently described as a transient that is associated with a sustained response which is frequently called the DC component. Both of these components are associated with the IN. The a-wave and b-wave are of opposite polarity. The descriptions of these waves lacked clear definition even in the literature of the late 1960's. The c-wave and d-wave defy concise verbal description and will be left to section 13.8.

The important point here is that the initial signal generation occurs in the photoreceptor cells and the principle signal processing functions occur in the IN. This provides a starting point for, and a logical break between Chapters 12 and 13.

11.6 Functional and tabular models of vision

The current literature lacks a framework adequate to accommodate all of the data available concerning vision. Providing such a framework is the goal of this Section. Because of the complexity of the overall visual system between various species and phyla, it is necessary to define several different components of the overall framework. The following paragraphs will address this task.

Engineering has developed a framework for describing electro-optical scanning and imaging systems[78]. This framework includes many aspects applicable to the visual system. Based on this framework, the visual system can be classified clearly by hierarchy.

- The system is of two stages in the higher animals, with both a wide field and a narrow field capability.
- The system is a reticle type scanner, with the reticle equal to the aperture of each photoreceptor.
- The system employs a multi-dimensional (time, altitude and azimuth) correlation tracker.
- In the higher animals, the combination scanner/correlator employs a nutating translation mechanism (tremor) in

[77]Rubin, M. & Walls,G (1969) Fundamentals of visual science. Springfield, Il: Charles C. Thomas pg. 190-196

[78]Wolfe, W. & Zissis, G. (1978) The Infrared Handbook. Wash. DC: US Gov't Printing Office, Chap. 17

two (nominally orthogonal) dimensions.

• In the higher animals, the tracker employs a two-speed closed loop servomechanism for pointing.

• The system supports a general purpose, distributed, and re-programmable computational facility (the brain).

• The computational facility can be divided into two major portions,
a largely repetitive but trainable (housekeeping) portion and
a more adaptive cognitive portion.

• The computational facility is supported by a variety of short and long term memory facilities.

This work will address all of these classifications. However, they will frequently be addressed using the terminology more commonly found in the biological and psychophysical literature.

11.6.1 Background

Many authors have presented many models of the visual process and defined 1st, 2nd . . . nth stages as it fit their purposes. It appears advisable to repeat that procedure here so the commonalties and differences between these models can be tabulated. Without exception, the models in the current literature are partial, if not floating, models. That fact highlights the purpose of this work. There is a great need for a global, internally consistent model of the visual system in animals. Only with a global model available is it possible to discuss new laboratory results in the proper context. Subsequent discussion can then be bilateral based on an adequate common foundation. To achieve a consistent set of models in the general sense, it is useful to define an internally consistent hierarchy of levels that can be used to discuss individual portions of the overall process. In this work, the following hierarchy will be used:

Global Model
Functional Block Diagrams, one or more
Top Level Schematic
- Motor related Schematic (involving control circuit functions)
- Vascular Schematic (involving metabolic and electrical power functions)
- Electronic Schematic (involving signaling for both alarm and analysis functions)

Electronic Circuit Diagrams (containing more than one complete circuit)
- Large signal circuit diagrams
- Medium signal circuit diagrams
- Small signal circuit diagrams

Individual Circuit Diagrams (showing components and/or morphology)

Many of these different levels can be represented by both diagrammatic and tabular renditions. To aid in these presentations, terminology drawn from the Standards promulgated by a variety of scientific and engineering organizations will be used. This will be done to minimize the needless growth in new terminology related to a specific discipline.

The interpretation of the visual process is greatly complicated by the degree of divergence and convergence found within even a primitive visual system. To aid in interpretation, this work explored very simple paths essentially free of these two dimensions. With a basic path defined, it was then possible to understand the more complex paths. It was generally found that the divergence of signal paths within the retina was for the purpose of parallel signal processing. This signal processing generally resulted in a significant convergence in the signal paths in terms of the ultimate number of neurons appearing in the optic nerve. However, the information carried by most of these signal paths was more sophisticated in structure than that found in the original or parallel signal paths. This efficient, but complex, coding is interpreted in the brain through a similar degree of signal path divergence. The initial divergence in the brain is merely to support signal decoding. The divergence involved in subsequent interpretation occupies a major portion of the brain that is closely coupled to the decision making portion. Following Interpretation and decision making, the multitude of signal paths converge, leading to the control portion of the motor system of the animal
.
This model takes a multiple parallel path approach and attributes virtually all signal processing, except temporal processing, to the eye. The brain is primarily involved in interpretation of the spatial and temporal features of the luminance, chrominance and appearance vectors provided to it. The spatial processing in the brain does include stereo-graphic manipulation. However, as indicated elsewhere herein, this stereo-graphic manipulation is in vector

space on a virtual map. It does not appear to involve the overlay of separate maps in image space. Although the visual pathways do involve short term temporal characteristics, passbands and delays, longer term temporal characteristics appear to involve changes that involve the short term memory of the brain.

To aid in interpretation, this work will first present the minimal paths found within the visual processes of all animals. Additional paths within the eye, generally related to the concept of divergence, will then be introduced to explain the operation of more advanced visual systems. The natural convergence resulting from signal processing will become obvious as the presentation continues. The complexity of the signaling transmitted via the optic nerve will also become clearer.

11.6.2 The Global Model

To aid in the correlation of this work with earlier authors, a diagrammatic global model is presented in **Figure 11.6.2-1** . This model is annotated to aid in the correlation of this work with that of other authors. It also includes undetailed support functions indicated by the labels metabolic support and electrical support that will be discussed later.

The series of vertical arrows on the left aid in interpreting the visual system from the most global perspective. This perspective has been arbitrarily labeled the Optical Environment and includes the optical as well as the cognitive elements of the system. The stages are given alphabetical designations and numeric sub-designations. These stages represent gross designations related to the visual system. A functional block diagram will be presented below to aid in understanding this figure.

The series of vertical arrows on the right, arbitrarily labeled the Signaling environment, omit the mechanical and vascular aspects of the visual system. They provide more detailed designations of the visual system. In this context, the stages are numbered numerically with alphabetical sub-designations.

Many of the left and right designations share a common element. As an example, Stage C_1 and C_2 on the left divide the Stage C circuits into those based on analog and pulse signaling. The equivalent designations on the right share these same characteristics. All Stage 1 and Stage 2 circuits involve analog signals. Stage 3 on the other hand, involves pulse (action potential) signals. Stages 4 and higher are also believed to be dominated by analog signaling paths.

In this model, Stages A and B includes three groups of elements; the elements associated with forming the optical image, the elements associated with the photodetection process and the signal path up to the output of the photoreceptor cell. Stage C includes all of the signal processing, signal projection and signal evaluation aspects of the visual process. Stage D is labeled administrative optics. This stage relates to a framework created by man, completely external to the visual process, but very important to the understanding the visual process.

A series of numbers is displayed on the figure correlating designations assigned by Polyak to some of the regions described by the chain of boxes in the center of the figure.

The right-hand designations are related to the signaling functions of the visual system exclusively. In that context, the motor and vascular aspects of the visual system are considered secondary.

11.6.2.1 The Signaling environment of the Global Model

Stage 1 of the global model involves two primary functions associated with the photoreceptor cells, photodetection and signal amplification. The second stage involves all signal processing between the output of the photoreceptors and the input to the encoding circuits (the ganglion cells). The third stage involves the encoding/decoding process which is shared by structures at each end of the optic nerve.

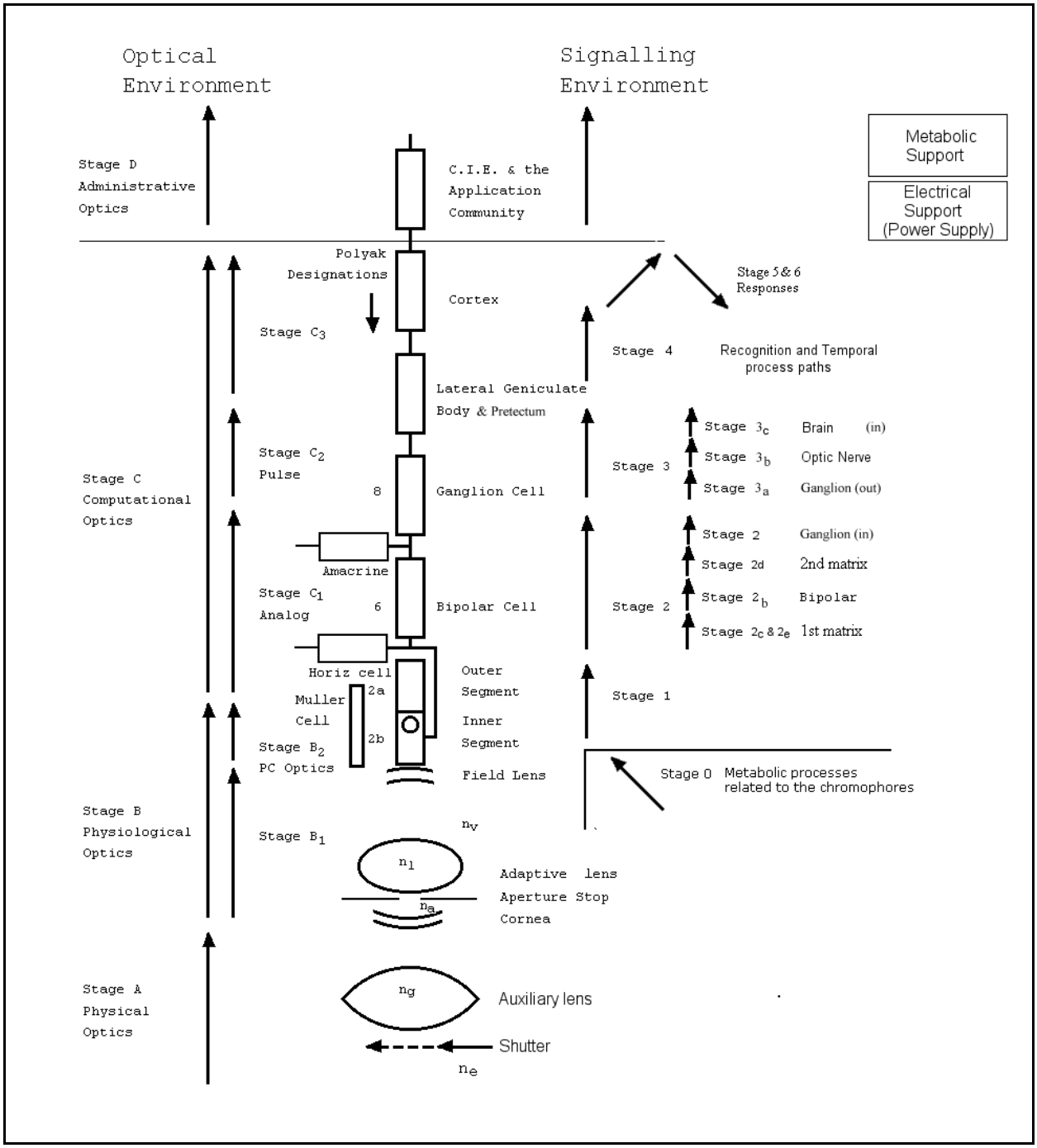

Figure 11.6.2-1 Global model of the vision process in animals. The labels on the left may be correlated with many documents in the literature. The labels on the right correlate with this work. The two boxes in the upper right indicate important support functions. They support all of the signal related neurons. See text for further discussion.

The fourth and higher numbered stages involve all signal manipulation within the brain. There are a few visual signal paths not included in this description. Some are treated separately, i.e., blink reflex signals, etc. Others, such as the diversity encoding channel, labeled stage 2 but without a subscript, are more difficult to illustrate. Some of the major elements of the visual system are shown deviating from the dominant vertical signal path of the figure.

This is done to suggest lateral signal processing between parallel signal paths within the retina.

The ganglion cells play a role in vision of similar complexity to the photoreceptor cells. Their neuritic structures play a key role in two distinct functions; the diversity (analog)encoding of the signals received from different locations in the retina and the analog to binary signal encoding of all of the signals to be transmitted to the brain. Their initial axonal structures also play a critical role in the analog to binary encoding process. Their extended axonal structures, including the Nodes of Ranvier, play a major role in the routing and transmission of the signals to the brain. Because of these multiple roles, it is difficult to assign a role to the ganglion cells based strictly on morphology. In this work, the high frequency signal paths through the neuritic parts of the ganglion cells will be considered part of the signal manipulation functions of stage 2. The lower frequency signal paths through the neuritic parts of the ganglion cells will be considered part of the analog to binary signal encoding process. All parts of the ganglion cells orthodromic of the base region of the Activa within the cell will be considered part of the signal projection process.

Within the 1st stage, photodetection is best subdivided further as follows;

+ stage 1A--the transduction process in the OS
+ stage 1B--the translation process between the OS and the IS and the amplification process in the IS.
+ stage 1C–the adaptation (variable gain amplification) process

Within the 2nd stage, initial signal manipulation, this model includes;
+ stage 2A--the straight through path (foveola to the Pretectum)
+ stage 2B--the luminance channel (other than the straight through path)
+ stage 2C--the chromatic channel
+ stage 2D--the geometric channel (appearance related path)
+ stage 2E--the polarization channel (present in selected species)
+ etc.

Within the third stage, signal projection, this model includes;
+ stage 3A--the encoding of the visual information (Ganglion cells)
+ stage 3B--the routing and transmission of the encoded information (Optic nerve)
+ stage 3C--the decoding of the visual information (some of it synchronously in conjunction with the commanded eye motions).

Within the fourth stage, main signal manipulation, this model includes;

++ servomechanism signal extraction
+ stage 4A–initial signal extraction in the Pretectum, extraction of precision signals related to the spatial content imaged on the foveola. Signals are separated into servo-signals and detailed information signals. The servo information is passed to the POS. The detailed information is passed to area 7 of the cerebrum.
+ stage 4B–initial signal extraction in the LGN–extraction of guidance signals required to overlay the images from the two eyes, stereoopsis.

++ image information merging & vectorization
+ stage 4C–foveola information extraction
The resulting information is passed to area 7 of the cerebrum.

+ stage 4D–merging of ocular images (luminance in magnocellular region)
+ stage 4E–merging of ocular images (chrominance in parvocellular region)
+ stage 4F–merging of ocular images (polarization in pseudo-parvocell.)
+ stage 4G–merging of ocular images (appearance)
The resulting information is passed to area 17 of the cerebrum.
+ stage 4J–activities in area 17, 18, 19 etc of cortex.

++ high level feature extraction and cognition
+ stage 4M–cognition of the *specific* situational awareness information ---area 7 (parietal lobe)
+ stage 4N–cognition of the *general* situational awareness information-----area 22 (parietal lobe)
+ stage 4P–highest level cognition (Frontal lobe)
+ stage 4R–autononous cognition (Cerebellum)

Within the fifth stage, initial command generation
+ stage 5A–sympathetic in parietal lobe
+ stage 5B–autonomous in POS
+ stage 5C autonomous in cerebellum

Within the sixth stage, command implementation
+ (reserved)

Tabulation of the individual functions becomes increasingly difficult after stage 3. Stage 4, 5 & 6 are listed primarily as a coarse framework to be expanded and revised later. Stage 5 will be explored more fully in **Chapter 15.** All of the designations are compatible with the more detailed figures of this work.

11.6.3 The Functional Block Diagram

The Global Model is not able to illustrate the parallel signal paths or divergent/convergent signal paths associated with vision. The divergent paths are only conceptually indicated by the lateral, i. e., horizontal and amercine, cells associated with Stage 2. **Figure 11.6.3-1** presents a Functional Block Diagram better able to illustrate the parallel and divergent paths as well as the control oriented feedback paths associated with vision. It is still unable to present the signal paths associated with the spatial matrixing associated with the appearance channels and the diversity encoding matrix. These signal paths are less well understood but will be explored in following paragraphs of this Section.

The upper vertical column (rows 1 thru 4) presents the functional blocks common to all aspects of vision and can be considered stage 0 in the nomenclature of the previous section. These include establishing a line of sight, forming an image, micro-scanning the image and using the eyelid to suppress imaging during gross changes in the line of sight. These functions are all controlled by servo signals received from the oculomotor nerves and generated in the Pretectum.

Rows 5 thru 7 represent stage 1. They represent the spectrally selective photodetection channels available and widely used in animal vision, with one channel further subdivided into a pair of polarization sensitive channels. The photodetection channels are labeled; UV_2-, UV_1-, S-, M- and L-. These channels support sensing in the ultraviolet, short, medium and long wavelength regions of the electromagnetic spectrum between about 300 and 750 nm. Large terrestrial chordates cannot utilize the ultraviolet portion of this spectrum because of the opaqueness of the cornea/lens combination in this region. Aquatic chordates do not suffer from this limitation. Most arthropods do sense ultraviolet radiation but do not use the long wavelength spectral band. The molluscs use different combinations of spectral bands depending on their familial adaptations.

There is data in the literature that would suggest that other channels besides the UV channel may be subdivided in order to provide additional polarization data. Menzel & Backhaus[79] and others have provided cross-sections of the ommatidium of a honey bee that suggest polarization sensitivity in the S- & M - channels as well.

Each of the above channels contains a photon to exciton transducer, an exciton to electron translator, and an adaptation amplifier. Each photoreceptor cell contains a complete set of these functional elements. The only functional variation between photoreceptor cells is the type of chromophore employed in the transducers (and in some species, the orientation of the UV channel transduction apparatus). As a result of the adaptation amplifiers in the photoreceptor cells, the visual signaling system following the photoreceptor cells operates at an essentially constant mean signal amplitude level.

Rows 8 and 9 represent stage 2, the initial signal manipulation in vision. The primary interconnections between the photodetection channels and the channels of the signal processing circuits are found in row 8. The signal leads in this area constitute the outer plexiform layer (OPL) of the chordate retina. The heavier lines denote the paths leading to the signal summation circuits and the resultant luminance channels of vision. The lighter lines denote paths leading to the signal differencing circuits associated with the polarization and chrominance signal channels.

Row 9 includes the signal processing circuits of animal vision, other than the spatial and diversity encoding processes. Functions in this row generate all of the possible luminance, chrominance, and polarization signal paths found in the visual system of an animal. The polarization channel is labeled N. The chrominance channels are

---

[79]Menzel, R. & Backhaus, W. (1989) Color vision honey bees: phenomena and physiological mechanisms. Chapter 14 in Facets of Vision, Stavenga & Hardie *ed.*. NY: Springer-Verlag

labeled O, P, & Q. The luminance channel is labeled R and employs signal summation in logarithmic space. All of the other signal processing channels, N thru Q, employ signal differencing in logarithmic signal space. The negation process indicated in the chrominance channels is only shown symbolically.

Rows 10 through 15 represent stage 3 of the visual system. The spatial and diversity encoding matrix of row 10, and the decoding in row 14, performs an important role. It decreases the number of neural paths required in the optic nerve while insuring that a single point failure along the optic nerve will not cause a major visual system failure. Except for the foveal signal paths that go around the spatial and diversity encoding mechanism, as indicated by the arcs between row 9 & 11 and between row 13 & 15, a failure of an individual or group of adjacent neural paths in the optic nerve causes negligible impact on the visual system.

Rows 10 thru 13 and the decoding functions in row 14 are believed to be chordate specific. They provide functions required to accommodate the large angular rotations of the ocular globes in the more advanced members of this phylum. The primary task of these functions is to achieve the convergence of the large number of signal paths associated with the photoreceptors of the retina into the much smaller number present in the optic nerve, apparently using techniques that can be decoded in the brain without serious information loss.

Following signal formation, each of the biphasic polarization and chrominance signal channels are encoded into an action potential pulse stream. The monophasic luminance signal is also encoded into an action potential pulse stream. However, there is a thresholding function preceding the encoding process in the luminance channel. The above encoding is accomplished by the ganglion cells. Each ganglion cell incorporates the functions of rows 10 thru 13.

A question remains as to whether the signals received at row 14 are decoded to analog form in the LGN or whether the information from the two eyes is merged while still encoded in the binary pulse domain. It appears most likely that it is decoded, manipulated in the analog domain, and then re-encoded for transmission to the visual cortex. If this is the case, it is reasonable to consider the first function in the LGN, decoding to be considered part of stage 3 and the signal manipulation part of stage 4. The re-encoding, retransmission via an association fiber and subsequent decoding at a different engine within the cortex is a common feature of the system.

The broken line from the LGN in row 14 to the cortex in row 16 represents the magnocellular and parvocellular pathways to area 17 of the cortex.

Row 15 includes the decoding and manipulation of the information from the foveola in order to create the servo signals required to operate the mechanical mechanisms of each eye. These servo signals are also pulse encoded prior to transmission back to the eye and its surrounding tissue. The broken line to the cortex in row 16 represents the pulvinar pathway to area 7 of the cortex.

There appears to be a correlation between the functional types of neurons in the visual system and the morphological types found in the literature. Where this condition appears to exist, labels have been placed under the functional boxes. Thus, the midget ganglion cells appear to be specialized to meet the requirements of the signal differencing channels. The parasol ganglion cells appear to be specialized to meet the requirements of the luminance channel. The bipolar cells appear to be optimized for signal summation. However, they may also be used as simple impedance changing and signal isolation devices. Further correlations may be demonstrated by future workers.

Many animal species only use a selected set of the channels in the above figure:

+ Many animals of *Arthropoda* exhibit sensitivity to polarized ultraviolet light as well as the S- and M - regions of the visual spectrum. These animals use polarization channel, N, and the two chrominance channels, O & P, as well as the luminance channel, R.

+ Many animals of *Mollusca* exhibit sensitivity to ultraviolet light as well as the S- and M- regions of the visual spectrum. Only a few animals of this Phylum have been studied. Whether they exhibit polarization sensitivity or sensitivity in the L-region of the spectrum is uncertain.

+ A very large number of animals in *Chordata* exhibit tetrachromatic vision although reports of sensitivity to polarized light were rare until recently (**Section 11.6.3.1.2 xxx**). These animals, sensitive to UV, S, M, & L channel light, are primarily aquatic and reports suggest they frequently lose their UV sensitivity upon maturation. Terrestrial tetrachromats appear to be limited to the smaller animals in the phylum. The majority of terrestrial chordates are trichromatic due to the spectral limitations of the optical system.

The cross connections shown in the center of the diagram, in row 8, are illustrative of the first level of signal interconnection found in the visual system. These connections serve two purposes, to achieve maximum sensitivity to a broad spectral range of illumination and to encode the spectral information included in the received photon stream. Maximum sensitivity to photons is achieved by summing the signals received from the individual narrow spectral band photoreceptors. This process is shown mathematically as occurring in the luminance channel labeled R. Simple cells of the bipolar type are capable of performing the summation process associated with the R channel. Encoding of the spectral information available is achieved using a multichannel approach. The difference between each pair of signals from the photodetection channels is processed in a separate signaling channel. These channels are labeled O, P, & Q in the figure. The lateral cells associated with this processing are generally labeled horizontal cells. These cells are capable of performing signal subtraction due to their specialized input structure. There are two separate input structures associated with each lateral cell. One of these structures in each cell is capable of reversing the electrical polarity of the signal applied to it.

The second level of signal interconnection is related to the spatial and diversity encoding subsystem. This subsystem is believed to primarily involve signals from the circuits in row 9 as they proceed toward the circuits of row 11 & 12. This encoding technique is the primary method of reducing the very large number of discrete signal channels before the signals pass through the optic nerve.

After diversity encoding, the signals are converted from analog to pulse form in the ganglion cells of the retina. In the case of at least some of the R channel signals in some animals, the signal is first passed through an additional level of signal processing before analog to pulse conversion. This additional level can include signal pre-emphasis in the frequency domain and/or signal thresholding in the amplitude domain. Following conversion, the pulse stream from the ganglion cells of the R channel is ready for transmission to the brain. This pulse stream uses pulse time delay encoding. In the absence of illumination, no pulses are transmitted by the R channel. Pulses with ever decreasing time intervals between them are transmitted as the illumination level increases above the threshold. The signals in the R channel related to the non-foveal area of the retina are transmitted to the LGN portion of the brain. The signals in the channel related to the fovea are transmitted to the Pretectum (and possibly the LGN as well).

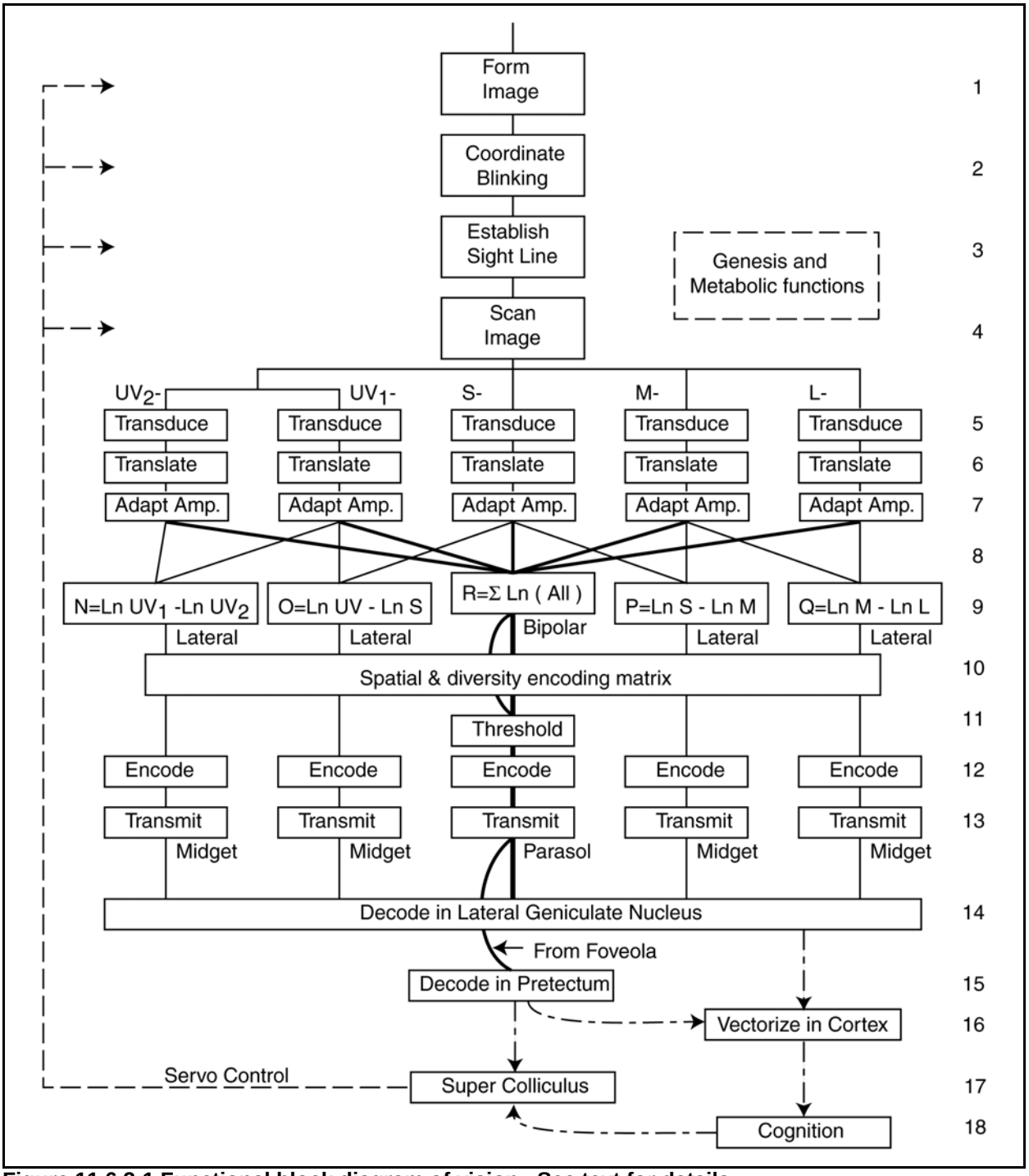

Figure 11.6.3-1 Functional block diagram of vision. See text for details.

The multiple difference signals associated with the O, P & Q channels are also converted from analog to pulse form in the ganglion cells. However, the subsequent processing is stylistically different. In the absence of any difference signal, the ganglion cells associated with each of these channels are biased to generate a pulse stream of nominal temporal spacing. The biphase signals resulting from the differencing process are used to vary the pulse to pulse

spacing in these channels. The spacing can be increased or decreased relative to the nominal spacing in order to transmit the biphase signal information. It is clear from the literature that the chromatic signals associated with the non-foveal O, P, & Q channels are passed to the LGN. It is not totally clear whether such chromatic information associated with the fovea is passed to the Pretectum. It may only be passed to the LGN.

11.6.3.1 Other signaling channels

11.6.3.1.1 The appearance channels

There is another level of signal processing found within the eye but not well illustrated in the above figure. This is the retinal signal processing associated with the spatial content of the image sensed by the eye. It is generally associated with the appearance (or geometric) processing channels, labeled Z_1, Z_2, etc. for purposes of this work. Details relating to the appearance channels are beyond the scope of this work. Only a few remarks applicable to this arena will be provided.

It is important to differentiate between any signal processing related to appearances in the spatial image performed in the retina from that performed at the higher cognitive centers. The magnitude of appearance signal processing within a retina is highly species and family dependent. It appears to be minimal in the human species relative to other animals. However, it is believed to be highly developed in other animals. It appears most developed in the hunters of the animal world. These animals appear to be sensitive to a variety of special geometric features beyond those associated with reproduction of their own species. The resulting signal processing performed in the retina can be considerably more complex and involve more space than that involved in the chrominance channels.

The signal processing associated with the appearance channels involves an additional level of signal matrixing. There is an unresolved question as to whether the signals applied to this processing are derived from the raw photodetection channels, UV, S, M, & L or whether they are derived only from the luminance channels, R. Logic applied to the known cross-sectional structure of the retina would suggest that the Z channels are fed from the R channels. The logic is based on the desire to have the appearance signals derive from achromatic signals and to use the most sensitive channels in the retina. Both of these goals are achieved by using the R channels. Furthermore, there is a group of lateral cells found posterior to the Inner nuclear layer. These are the amercine cells. These cells are known to employ a wide variety of highly arborized input structures. Therefore, the signals directed to the appearance matrixing circuits are drawn in the above diagram at the output of the individual lateral cells of the R channels associated with the entire retina.

It is particularly difficult to determine the exact signal paths leading to the appearance matrix. This is because the R signals contain elements of the UV, S, M, & L signals. Only very careful experiment design can lead to explicit information about this ambiguity. Historically, most experiments designed to evaluate the spatial performance of the human retina have employed light sources that were not carefully matched to the absorption spectrums of the individual chromophores. The result was data that failed to successfully separate the spatial information from the chromatic information. In addition, many of these experiments were subject to temporal delays in the signal processing channels that resulted in chromatic distortion of the results depending on the time involved in experiment execution.

There is an additional facet to the appearance processing. If some of the matrixing involves inputs from the lateral cells related to luminance matrixing and some involve inputs from bipolar cells between the lateral cells and the ganglion cells, a time delay is introduced into the equation. The result is an output that is sensitive to the time of excitation of certain photoreceptors relative to nearby photoreceptors. This feature can generate a velocity sensitive output within the appearance matrix. Such an output signal can be very useful in creating an alarm signal within the retina.

All of the above signals are processed further in the brain in order to *perceive* the environment of the animal and to control its responses to that environment. The perception process involves both the generation of responses to alarm signals and the completion of the interpretation/recognition process related to complex geometric shapes.

11.6.3.1.2 The polarization channel(s) in *Arthropoda* & *Mollusca*

Individual clusters of photoreceptor cells found in both *Arthropoda* and *Mollusca* have frequently been shown to contain pairs of photoreceptors sensitive to the same spectral wavelength. The orthogonal orientation of the photochromic material in the Outer Segments of each pair, with respect to each other and with respect to the Poynting vector of the incident radiation, results in a unique property. The output signal from each photoreceptor is a function of the polarization of the incident radiation. The output signals from each of a pair of photoreceptors are orthogonal

to each other relative to polarization. By subtracting these two signals in a signal processing channel essentially identical to a chromatic signal processing channel, a signal is generated that is proportional to the angle of polarization received by that detector cluster. Since the sky is known to exhibit a complex polarization map when viewed from a given point, an animal is able to triangulate his orientation by using the polarization signals from a group of polarization signal channels strategically placed within its total field of view.

11.6.3.1.3 The polarization channel(s) in *Chordata*

Reports appear in the literature concerning polarization sensitivity in *Chordata*. However, most of the reports are anecdotal. Little data accompanies the reports regarding the cytological and chromophoric structure of the retinas involved. Frequently the reports relate to the periphery of vision which suggests the conventional photoreceptors of *Chordata* are being used in a non-optimal mode (not pointing directly at the pupil of the eye).

Parkyn & Hawryshyn have provided good measurements of the sensitivity to polarization in the UV for many salmonid fish[80]. However, the experiments were more exploratory than applied research. Their electrophysiological data is described as from multiple units within the optic nerve of these fish. The size of their electrical probe, and the voltages recorded, suggests the data may represent capacitively coupled data with the probe located in the INM rather than penetrating a specific neuron. The relatively small changes in observed signal amplitude with polarization angle are not compatible with the much larger signal amplitudes associated with tests involving *Arthropoda* and *Mollusca*. They reported an unusually large second harmonic content in their measured responses as a function of polarization angle. This harmonic content may also suggest the polarization sensitivity related more to non-optimal operation than to a difference in the structural arrangement of the chromophores of photoreceptors reflecting a polarization sensitivity.

No photoreceptors optimized for polarization detection have been reported among members of *Chordata*.
Two polarization phenomena, Boehm's brushes and Haidinger's brushes are reported in the human vision literature. However, they are only encountered intermittently and are believed to be largely artifacts of internal reflections resulting in low level polarization sensitivity. Waterman[81] describes these potential situations in detail.

11.6.3.2 The relative complexity of the functional block diagram

The above Functional Block Diagram is illustrative of the visual functions in any animal. However, it is not able to illustrate the magnitude of the complexity introduced into the chordate visual system in order to provide the mechanical flexibility and dynamics found in the chordate eye. **Figure 11.6.3-2** shows the relative numbers of signal paths found within various regions of the visual system. The large degree of convergence and divergence is primarily to accommodate the mechanical flexibility required between the eye and the cranium. This figure has been expanded from that of Dowling[82] to indicate the special handling of the optical signaling channels associated with the fovea. All of the numbers are for the human eye. Dowling has said that approximately 40,000 neurons of the optic nerve are associated with the 20,000 photoreceptors of the central fovea (apparently the foveola following Hogan's nomenclature). Other authors have noted that some of the neurons of the optic nerve bifurcate and group the resultant neural branches into a separate bundle that proceeds to the Pretectum.

Weiskrantz has published a simple chart showing the neural paths from the retina to ten different terminal bodies[83]. He did not distinguish between sensory and motor paths or between direct and indirect paths. Some of his labels may be redundant. He notes that the 10-15% of neurons not going directly to the LGN is larger than the total number of neurons in the auditory nerve.

It appears that these branch neurons deliver signals related to the foveola and/or fovea directly to the Pretectum. No information has been found concerning how many neurons may be involved in this branch. The Pretectum derives information from these channels and uses it to instruct the Superior Colliculus on where to reposition the optical line of sight. The Superior Colliculus converts these instructions into neuro-motor signals and sends them to the muscles

---

[80]Parkyn, D. & Hawryshyn, C. (2000) Spectral and ultraviolet-polarization sensitivity in juvenile salmonids: a comparative analysis using electrophysiology *J Exp Biol* vol. 203- pp 1173-1191
[81]Waterman, T. (1975) Natural polarized light and *e*-vector discrimination by vertebrates *In* Evans, G. Bainbridge, R & Rackham, O. *eds*. Light as an Ecological Factor: II Oxford: Blackwell Scientific Publications.
[82]Dowling, J. (199X) pg. 350
[83]Weiskrantz, L. (1997) Consciousness Lost and Found. NY: Oxford Press. pg 128

controlling the eye. It is not known whether the chromatic or appearance channel signals related to the foveola are delivered to the Pretectum. It is not clear whether the signals from the fovea also proceed directly to the LGN or whether the pertinent information from the Pretectum is relayed to the LGN via the Pretectum. Evolution and logic would suggest that only the luminance channels related to the foveola would branch from the optic nerve and proceed to the Pretectum. The luminance channels contain a majority of the information related to the high contrast edges in an image. This logic also provides a possible explanation of why it is so difficult to read red letters on a blue background. The L- and S-channel components are of low amplitude relative to the M-channel in the luminance channel. [xxx rewrite to incorporate the pulvinar into the PGN/pulvinar couple]

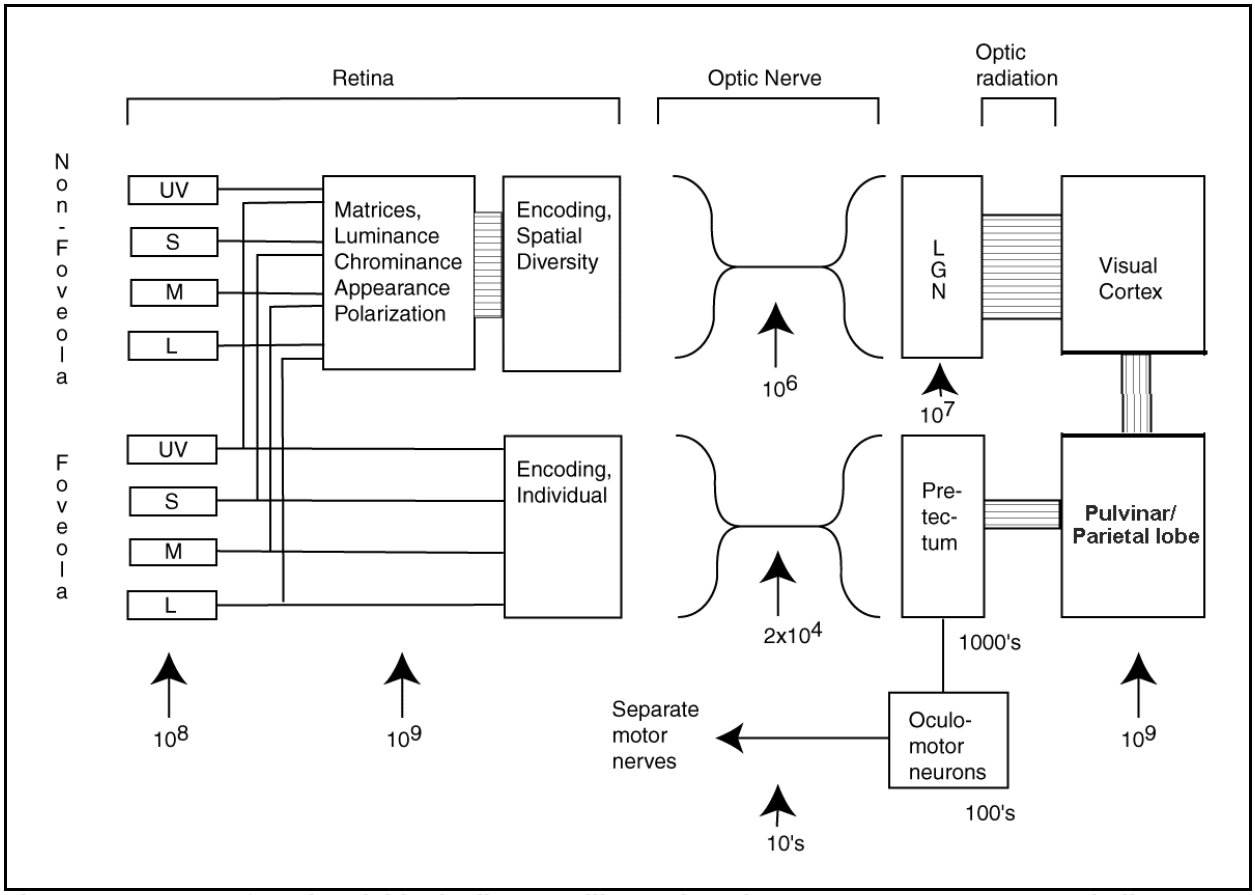

Figure 11.6.3-2 A functional block diagram illustrating the gross convergence and divergence introduced into the chordate visual system to accommodate the pointing capability of the eye. Most of the convergence in the retina results from the spatial and diversity encoding. The signal path from the pretectum, specifically the pre-geniculate neuron (PGN) engine going to the oculomotore neurons constitutes the Precision Optical System (POS). See text. Compare to Dowling, (19XX).

Many chordates do not exhibit a fovea. In these animals, the channels related to the fovea would be expected to proceed directly to the LGN. When appropriate, a fovea evolves in a species along with an appropriate branching of the signals related to the fovea or foveola in order to connect directly with the Pretectum.

11.6.3.2.1 Possible diversity encoding schemes

Since the visual system is fundamentally a change detector, it is possible to employ a simple binary encoding scheme to indicate the location of any individual change in signal intensity imaged onto the retina. Such a simple scheme would only require 30 signal paths. However, the need for a more robust design would raise this number by as much as a factor to ten.

The high degree of convergence should not be equated to the use of a very intricate encoding scheme to compress the information or to a loss of information in proceeding from the retina to the brain. For the purpose of signaling an alarm to the brain, simple binary encoding of the spatial location of a photoreceptor is adequate. Such binary encoding of 10^9 photoreceptors (every photoreceptor in a nominal retina) only requires 30 individual signaling channels in the optic nerve. To provide a redundant signaling system with some error correction capability, the number of individual channels might be quadrupled. There are more than enough neurons in the optic nerve to transmit all of the relevant information following the signal processing discussed in this Chapter.

There are frequent references in the literature to the direct connection between individual photoreceptors in the fovea, or at least the foveola, and individual neurons in the optic nerve. It is not clear that these are exclusive connections. Such a correlation would be consistent with neurons related to the fovea diverting to the Pretectum. It would also be consistent with these particular neurons not using binary encoding with respect to the spatial location of the individual photoreceptors in the fovea. The question remains open as to whether the foveal photoreceptors are included in the binary encoded spatial map of the retina for alarm purposes while simultaneously sending neural signals in non-spatially encoded form directly to the Pretectum.

11.6.4 The Top Level Schematic

The top-level schematic diagram of vision is the most complicated illustration of the overall process available at this time. It is an expansion of the above Functional Block Diagram. This diagram is able to explicitly illustrate the separate matrixing functions found in animal vision, those associated with the luminance information, the chromatic information (the 1st lateral processing matrix) and the appearance information (the 2nd lateral processing matrix). **Figure 11.6.4-1** presents such a figure tailored for the higher chordates. This figure omits the polarization processing matrix found in some of the short wave trichromats and tetrachromats. It summarizes all of the signal channels, the major vascular channels, and the most important motor channels found in the visual process. The figure would not be significantly different for other highly evolved families of animals. By showing a series of reference platforms, the figure is able to support a variety of discussions concerning how the overall visual system performs in conjunction with the motor system of the animal. These features also aid in the explanation of the visual system, motor system, and brain interfaces.

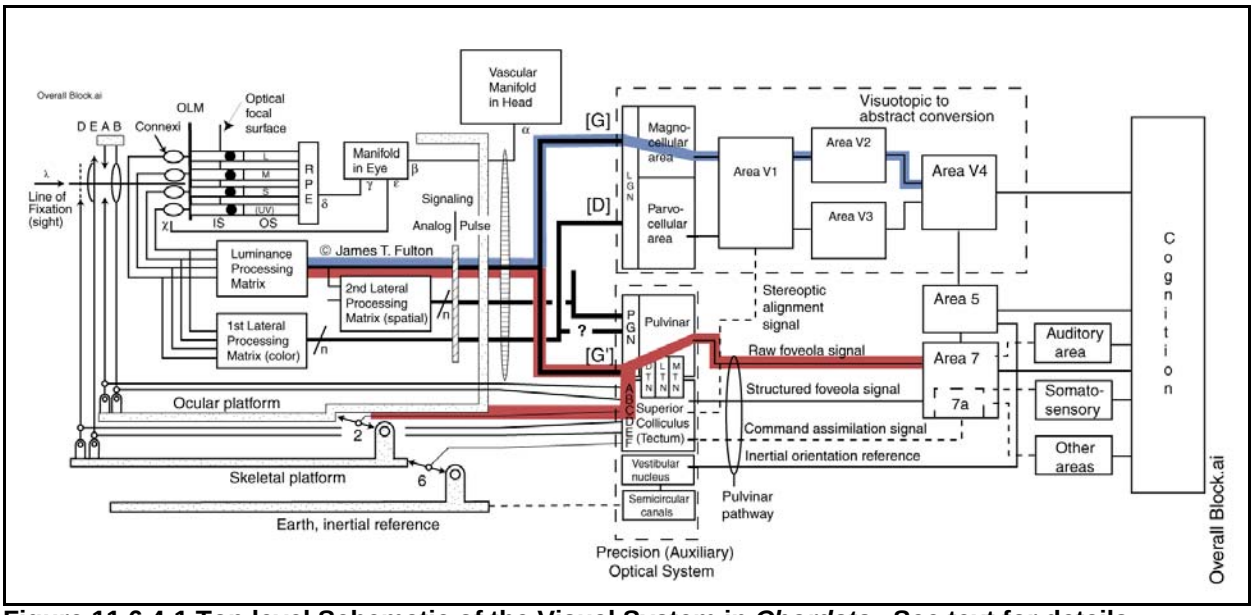

Figure 11.6.4-1 Top level Schematic of the Visual System in *Chordata*. See text for details

It is preferable to present summary descriptions of the platforms and systems peripheral to the Signaling System before proceeding.

11.6.4.1 The Motor System and Platform

The motor system and platform are both shown in generic form. The platform system begins with the earth as the primary inertial reference. Nearly all animals rely upon the force of gravity to provide a primary reference. This reference is often correlated with a visual image of the local horizon for added precision within the overall signal processing and perception chain. The skeletal platform is shown in its simplest form. This form is adequate for discussing animals with body mounted eyes. For more complicated animals, additional motor paths must be provided in conjunction with the structural freedoms associated with the rotation of the head relative to the thorax and the eyes (Ocular platforms) relative to the head. In the more advanced chordates, there are two separate head mounted optical shutters, labeled D & E in this figure. There are also two separate motor circuits within the ocular platforms labeled A & B in this figure. They are the iris motor circuit and the lens motor circuit. Not shown in this figure are the three pairs of muscles and their associated motor circuits positioning the line of sight of the ocular platform relative to the head of the animal. This ocular positioning system is more complex than usually presented. It serves two distinct functions. It must establish the line of sight in all animals having any structural freedom between the eyes and the body. It must also provide the fine scanning motion required in any animal in order to "image" its surroundings.

In most chordates, the second optical shutter, D, is a thin transparent layer and is used as a shield against physical damage to the cornea as the animal moves through its environment. In many chordate animals populating the air-water interface, the second optical shutter has evolved into a lens that can be introduced into the optical path under autonomous control by the animal as it changes its optical environment. The second shutter is considered vestigial in humans.

11.6.4.2 The Cortical System

The cerebral cortex of a higher animal is obviously extremely complex. It will be reviewed in some detail in **Chapter 15.** Since the visual system, along with the motor system and the cognitive capabilities of the animal play such a crucial role in the life of most animals, the interface of the cerebral system with the visual system is also quite complex. The figure shows the brain divided into three primary functional blocks. These blocks are related to memory, motor functions, and higher cognition. Further discussion of the cognitive functions of the brain are beyond the scope of this work. The motor system has historically been shown as including the Pretectum although its function has not been completely understood.

The memory subsystem related to vision has traditionally been divided into two major portions, the Magnocellular region and the Parvocellular region. It appears that the luminance information from the retina is transmitted primarily to the Magnocellular region and the chrominance, polarization, and other information are transmitted primarily to the Parvocellular region. This is particularly true of the non-foveal information.

It appears that the luminance information from the foveola of the retina is passed directly to the Pretectum portion of the Motor region of the brain. The Pretectum is closely connected to the Superior Colliculus that controls motor signal generation in the brain. This is logical since a considerably shorter time delay is obtained in the signal loop between the sensory portion of the fovea and the muscles controlling the image projected onto that fovea. Such a signal loop is also considerably "tighter" from a servomechanism perspective. It is not clear whether the chrominance information related to the fovea is passed to the Pretectum. There is a matrix of possibilities regarding the foveal information. The most likely appears to be that the luminance information is passed to both the Magnocellular region and the Pretectum while the chrominance information, from the 1st lateral processing matrix goes only to the Parvocellular region. It is possible that the foveal luminance information goes only to the Pretectum and is then passed to the Magnocellular region. However, this option is in conflict with a smooth evolutionary path that does not incorporate a fovea in lower animals. The signal paths shown within the brain are highly conceptual.

11.6.4.3 The Vascular System

The importance of the vascular system to both the static and dynamic performance of the visual system is not usually appreciated. To focus on its importance, the portion of the subsystem serving the photodetection portion of the eye is shown by two manifolds and three major flow paths. The portion serving the rest of the retina and the brain are conventional and not of interest here.

Even a partial failure associated with any of the elements supporting the photodetection portion can seriously impact the overall performance of the visual system. Besides the widely taught role of the vascular system in supplying the visual system metabolically, it plays a crucial role in supplying the electrical power used by the visual system. This role involves an extension of the metabolic role. Whereas, the metabolic role involves providing a variety of bioenergetic materials that are consumed in the normal metabolic processes and result in the creation of various glutamate materials as waste products, there is an additional role involving these same materials. In the normal

metabolic role, enzymatic chemical reactions are employed to provide the necessary energy in chemical form. In the additional visual role, the same materials are employed in an electrostenolytic reaction that releases free electrons for use in the electrical signaling activities of the visual system. The requirement to support the necessary electrostenolytic reactions places an additional responsibility on the vascular beds within the eye. This manipulation of free electrons introduces additional constraints on the vascular system with regard to resistivity. One of these constraints plays an important role in the generation of electroretinograph images.

There are a number of time constants associated with the ability of the vascular system to provide adequate bioenergetic material to the visual system. Whereas these time constants, measured in seconds or less, are of little importance to the metabolic operation of the visual system, they are critical to the electrical operation of the visual system. They play a primary role in the adaptation process of the eye and in the signal matrixing functions. Under conditions of vascular fatigue, a wide range of visual impairments can be expected. Some may be quite serious, such as progressive tunnel blindness (occurring for a few seconds prior to complete blindness).

11.6.4.3.1 The chromophore regeneration subsystem

The chromophoric regeneration subsystem is much more complex than previously discussed in the literature. Once generated and deposited on the surface of a disk, the chromophore has a life expectancy in humans of 12 weeks. During this period, the chromophore may be quantum-mechanically excited and electronically de-excited repeatedly. This cycle can be completed and repeated within milliseconds. It does not require any chemical or stereo-chemical change in the chromophore or its auxochromes to accomplish this signal generating cycle. No physical transport of the chromophores is involved in the sensing of the de-excitation process. The regeneration process is developed in detail in **Section 4.6** and **Chapter 7.**

11.6.4.3.2 The electrical power supply subsystem

The electrical power supply subsystem is a unique system due to the great dispersion of the individual electrostenolytic processes throughout the visual system, usually found in at least three different and distinct locations on the plasma membrane of every neuron. The properties of this subsystem are discussed in detail in **Section 4.7.4**.

11.6.4.4 The Signaling System

The conceptual debate in the literature between the Hering and Young-Maxwell schools of color are indicative of the separate signaling paths used in biological vision for the luminance and chrominance information. Neumeyer, et. al. has also suggested a similar separation in goldfish[84]. In fact, this separation (which consists of more than just two paths) is a characteristic of all biological vision.

Both Tilton and Boynton explored alternate concepts of signaling during the 1970's. These conceptual studies involved a number of difficulties. They both relied upon the assumption that color signal processing occurred subsequent to the creation of the luminance channel signal and that this signal was based on linear summation of signals from the individual spectral channels. They also relied upon the assumption that the CIE photopic and scotopic luminosity functions were accurate representations of the visual response of the human visual system.

Boynton attempted to derive a top level signaling plan for the retina that supported a Hering approach to color vision[85]. While he reported the spectral response of the individual chromophores of vision correctly in one part of his book (L- channel peak near 625 nm, page 222) , he proceeded to derive a different set of spectral responses at the conceptual level based on the attempt to satisfy a top level concept of Hering color space (page 212). The result was not defendable at the detail level even though it is universally taught to psychology students and psychophysics experimenters to this day. The approach employs linear mathematics and did not include any change in relative gain between the photosensing channels. It basically assumed that the long wavelength spectral response could be calculated by subtracting the scotopic luminosity function from the photopic luminosity function in linear space. The result was a long wavelength spectral channel with a peak response near 580 nm (page 213).

---

[84]Neumeyer, C. Wietsma, J. & Spekreijse, H. Separate processing of color and brightness in goldfish. *Vision Res.* vol. 31, pp 537-549

[85]Boynton, R. (1979) Human Color Vision. Holt, Rinehart & Winston

Tilton took a different approach based on logarithmic mathematics and did introduce variable gain between the channels prior to the summation of the spectral signals, at least at the introductory level of his discussion[86]. However, he also chose to treat the two CIE luminous functions as precise and accurate test data rather than as highly smoothed data assembled from different laboratories under primitive conditions by the standards of even the late 1970's.

Figure 11.6.4-2 provides the top level signaling architecture determined by this work. It was formulated to accept, with only minor modification, and is based on a compilation of, a large part of the creditable work in the literature. The architecture is tetrachromatic and can be applied to any visual system by adjusting for the presence or absence of specific chromophoric channels. The dashed vertical lines subdivide the system into the distinct signaling stages of vision defined above.

It is important to note that all of the summing and differencing occurs after logarithmic conversion and before the formation of the luminance channel signal. The L-channel spectral response cannot be determined by subtracting something related to the chrominance channels from the luminance channel signal, R-. And as discussed elsewhere, neither can the L-channel response be determined by a linear subtraction of the scotopic luminosity function from the photopic luminosity function. This is especially true of the smoothed CIE standards for these functions. The scotopic luminosity function is identical to the photopic luminosity function with the L-channel gain reduced to zero. However, one must use the actual descriptions of the functions as found in the R-channel and recognize their logarithmic relationship.

---

[86]Tilton, H. (1977) Scotopic luminosity function and color-mixture data. J. Opt. Soc. Am. vol. 67, no. 11, pp 1494-1501

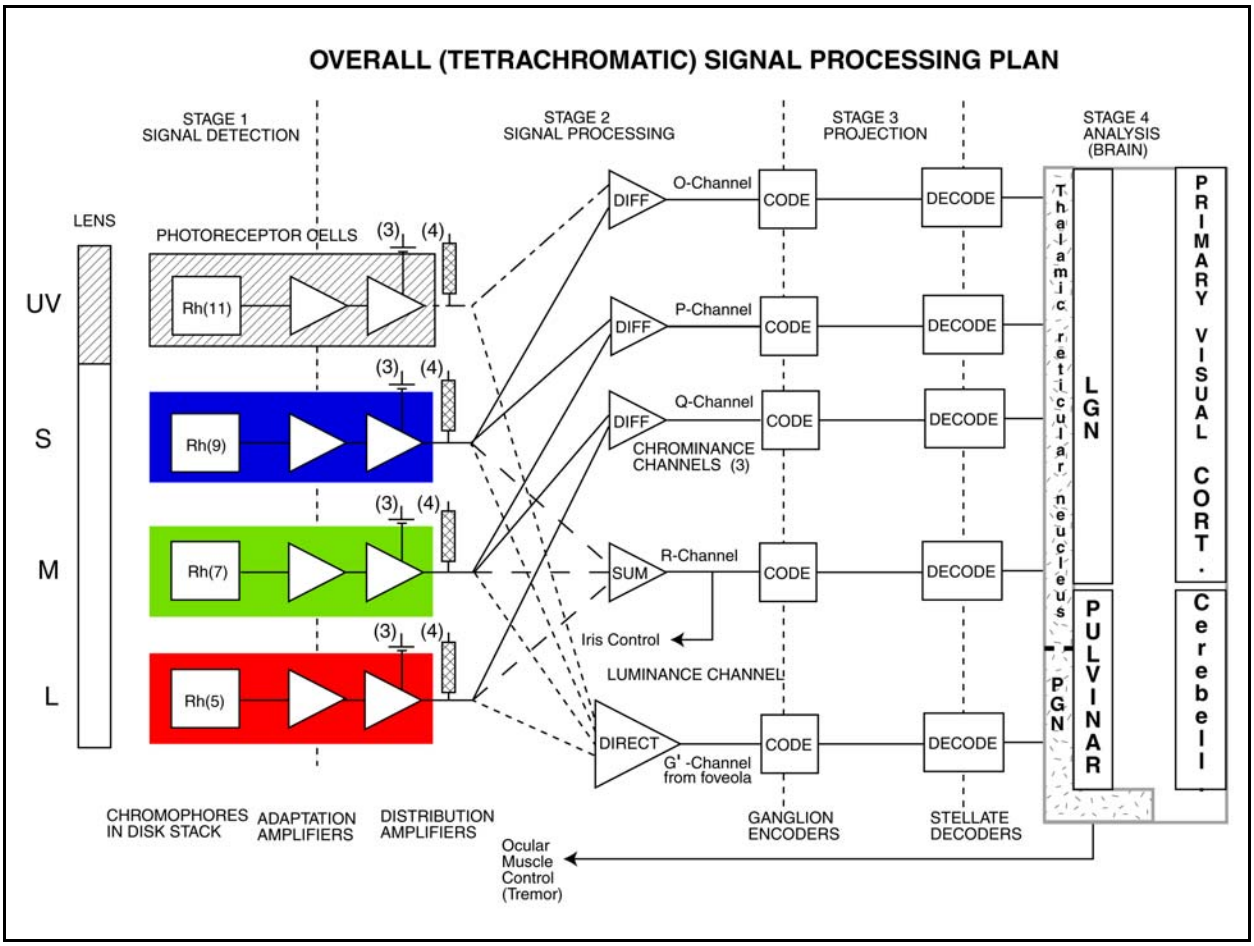

Figure 11.6.4-2 Overall signal processing architecture of the tetrachromatic visual system. A portion of the lens is shown hatched because of the different absorption mechanisms involved in the ultraviolet spectral region. The impedances (4) are diodes that perform the logarithmic conversion of the signal currents of the photoreceptor channels into the signal voltages of the signal processing channels.

The three most important relationships in the visual system are shown in the following equations. These equations apply within the adaptation amplifier ranges of all of the photosensing channels of the visual system (a definition of the photopic range of vision). Outside of this range, the appropriate gain constants must be introduced into these equations.

$\mathbf{R(\lambda) \approx \log \int \#_S \, d\lambda + \log \int \#_M \, d\lambda + \log \int \#_L \, d\lambda}$ (a sum of logarithms representing each spectral channel)

$\mathbf{P(\lambda) \approx \log \int \#_S \, d\lambda - \log \int \#_M \, d\lambda = \log [\int \#_S \, d\lambda \,/\, \log \int \#_M \, d\lambda]}$ – A log of a RATIO

$\mathbf{Q(\lambda) \approx \log \int \#_M \, d\lambda - \log \int \#_L \, d\lambda = \log [\int \#_M \, d\lambda \,/\, \log \int \#_L \, d\lambda]}$ – A log of a RATIO

$\#_X$ = excitation(λ) x absorption(λ) <u>in photons per spectral band</u>

Gain constants omitted

R(λ), P(λ) & Q(λ) are scalar numerical values

This architecture and these equations are fundamentally different from those of both Boynton and Tilton. However, these equations do precisely describe the performance of the visual system at both the gross and detailed levels. Specifically, there is no S - (L + M) signal and the luminance signal, R-, does include an S-channel component. The L- and M- channels are not linearly summed anywhere in the visual system to form a yellow or *y* signal. The chrominance channel consists of two separate and distinct signaling channels, the P = log S - log M channel and the Q = log M - Log L channel. The P, Q and R channels are all treated as if they are orthogonal to each other within the visual system. This is the origin of the complete three dimensional color space. By plotting only the P and Q channels orthogonally, the New Chromaticity Diagram for Research is obtained as will be illustrated in **Chapters 16 & 17**.

11.6.4.4.1 Luminance signaling

The primary luminance signaling channel consists of the various spectrally sensitive photoreceptor channels combining (by summation in the logarithmic domain) in the luminance processing matrix to form the luminance signal. Numerically speaking, essentially all of the photoreceptors of the retina participate in forming this signal. However, the photoreceptors of the foveola appear to skirt this process, at least partially, as will be addressed below. The luminance signals are passed along to the magnocellular portion of the LGN, to the primary visual cortex and then to the feature extraction engines of the cortex prior to cognition. This signal pathway is labeled the sum or **S**-pathway in the figure but is better known in the literature as the M-pathway. The reason for this change will be explained in **Section 15.2**.

It appears that the signals from each of the photoreceptors of the foveola are passed directly to the pretectum, associated with the thalmus but independent of the LGN. These signals travel over the path marked **S'** in the figure, for alternate sum pathway. These signals are used for precision analysis of the fine detail presented to the foveola. This function is performed by a servo loop as shown that incorporates the photoreceptors of the foveola, the individual elements of the auxiliary optical system shown, as well as the oculomotor system. The auxiliary optical system also receives reference information from the vestibular nucleus and the semicircular canals of the inner ear. The vestibular nucleus and the semicircular canals form a servo loop in their own right. This subsystem operates as a first order servo loop designed to provide an inertial reference. It provides precise information concerning the rate of change of the position of the head in earth-oriented inertial space. Such a first order system is not able to provide absolute position angle. To provide this absolute reference relative to the earth (or the local spacecraft environment), it is proposed that the pretectum or super colliculus receives orientation information from the primary visual cortex that has been processed via the S-pathway. See **Section 15.2.3.3.4.**

11.6.4.4.2 Chromatic signal separation

The retina consists of an array of photoreceptors having different absorption spectra. Sub-arrays of the retina can be defined for each spectral group of photoreceptors. However, the statistical properties of these individual arrays and their relation to each other is not known in detail at this time for any animal. The detailed structure and completeness of each of these spatial chromatic filters has not been determined in human. There is some data in lower chordates[87].

Each array senses the chromatic component of the scene imaged on the retina at a resolution determined by the parameters of the array. Therefore, the incident light is separated into a series of individual relatively coarse spectral images by spatial filtering. This information could form the basis of the signals to be transmitted to the brain. However, it does not. The visual system is not an imaging system. It is a change detection system. The chromatic information associated with each small area of the retina is transmitted to the brain on a pixel by pixel basis over channels that are not directly related to the luminance information associated with the same pixel. These channels employ different encoding techniques that sometimes introduce unusual second order effects.

As in the case of luminance signaling, the primary chrominance signaling channel(s) consists of the various spectrally sensitive photoreceptor channels combining (in the logarithmic domain) in the 1st lateral processing matrix to form two or three chrominance signals. Two chrominance channels are functional in humans and other large chordates capable of trichromatic vision. In many other animals, there are three chrominance signals serving

---

[87]Douglas, R. Bowmaker, J. & Kunz, Y. Ultraviolet vision in fish. *in* Kulikowski, J. Dickinson, C. & Murray, I. (1987) Seeing contour and colour. NY: Pergamon Press, pp. 601-616

tetrachromatic vision. Numerically speaking, essentially all of the photoreceptors of the retina participate in forming these signals. As in the case of the luminance channel, it is not clear whether the spectral information of the photoreceptors of the foveola participate in this process. These signals may pass directly to the pretectum without combining within the retina. Each chrominance signal is formed by taking the difference between the signals from two different spectral signals provided by the photoreceptor cells. Being formed in logarithmic space, each difference signal represents the ratio between the intensity of the two spectral signals in object space. The chrominance signals are passed along to the parvocellular portion of the LGN, to the primary visual cortex and then to the feature extraction engines of the cortex prior to cognition.

The individual chrominance signals are independent of each other and are treated as if they were orthogonal within the brain. However, the signals contain a common mathematical term and are not truly orthogonal. This fact becomes important when discussing the chromaticity diagram describing the performance of the visual system.

11.6.4.4.3 Polarization signaling

In those animals capable of sensing the polarization of the light projected onto their retina, the signal information is encoded as it is in the chromatic channels. The data is forwarded to the brain as a biphase signal representing the ratio between the intensity of two orthogonal polarization components of the incident light. Little is known about the processing of the signals obtained from the photoreceptors. The value of the information associated with each pixel is available for transmission. However, the volume of data may be reduced in the signal manipulation stage of vision.

11.6.4.4.4 Appearance signaling

The visual system supports two different modes of processing appearance information. The information from the foveola is passed directly to the Pretectum without participating in any signal manipulation within the retina. This information is processed immediately <u>and extracted in vector form</u> for further transmission to area 7 of the parietal cortex.

The non-foveola information appears to be processed within the retina in a manner similar to chrominance and polarization information. However, the arborization of neurons associated with appearance information appears to be quite extensive. The size of this arborization can introduce considerable time diversity into the signal processing stream. As a result, the Class F waveforms produced by this signal processing are extremely complex. Little specific information is available concerning the signaling transmitted to the brain from the retina concerning the appearance of individual features within the scene imaged on the retina. It is known that the 2nd lateral processing matrix is rudimentary or vestigal in humans. What is known about appearance signals has been obtained primarily from experiments with the cat. The appearance signals are believed to be formed in the 2nd processing matrix, the most posterior portion of the inner nuclear layer of the retina. What is not known specifically is whether this processing matrix, (frequently associated with amercine type cells but containing many other types) receives all of its input signals from bipolar cells representing luminance information or whether it also receives information directly from the photoreceptor cells or from the output of the 1st processing matrix.

Most caricatures in the literature suggest that the 2nd processing matrix only receives signals from bipolar cells via interconnections in the outer interconnect layer of the retina. This situation would suggest that the appearance signals are based primarily on spectral data that has already been processed into luminance signals. Whereas this scenario introduces no problems in the animal visual system, it does introduce considerable difficulties into laboratory experiments designed to interpret these signals. Unless considerable care is taken with the spectral characteristics of the test stimuli, the results will represent a very complex calculation based on both the spectral and spatial characteristics of the stimulus. This situation is the fundamental mechanism causing the inconsistencies in the literature concerning many center-surround experiments. Appearance signal generation occurs late in the signal manipulation chain and the resulting signals can appear very complicated. This is particularly true if the time-diversity encoding feature, which indirectly introduces a spatial-diversity feature) of the visual system is not recognized.

11.7 Circuit diagrams of vision

It is common to find simple circuit diagrams in the vision literature. It is uncommon to find precise statements about the applicability of these diagrams. They are invariably “floating models” based on the assumption of circuit linearity. When electronic test circuitry is attached to the actual visual circuits, it is invariably done in a less than precise manner because of a lack of knowledge concerning the design of the biological circuitry under study. This

Section will present some of the pertinent design information and the applicable circuit diagrams. The diagrams presented in this section begin with the detailed, large signal, diagrams and progress to a selection of simpler diagrams of less general applicability.

It is important to note that the large signal diagram is not an "equivalent circuit." It is the actual circuit employed in animal vision presented in human linguistic and graphical notation. This circuit can be used to accurately predict the performance of the visual system of a generic animal. It can be tailored to more accurately represent the visual system of a specific animal by making minor circuit modifications to account for specific features resulting from evolutionary adaptation. As discussed elsewhere in this work, an "equivalent circuit" can be constructed using man-made components that will faithfully reproduce the performance of the biological eye. However, the equivalent circuit must be constructed of carefully selected components and operated at the appropriate electrical voltages if it is to emulate the biological eye over even a fraction of its dynamic range. Each component must exhibit (frequently nonlinear) characteristics that are the same as those of its biological counterpart.

11.7.1 Background

To describe the actual circuit topology of the visual process, many of the procedures used in electronics are available. The first step in developing a circuit diagram is usually to determine what signaling environment is involved. What is the dynamic range of the amplitude of the signal and what is the frequency of the signal spectrum? In high performance circuits, it is also important to determine the "grounding plan."

11.7.1.1 The grounding plan

See **Section 11.1.1.4** for background information in this area.

11.7.1.2 Effects of temperature

Temperature plays a much larger role in biological electronic circuits than in typical man-made circuits for two primary reasons. First, man-made circuits tend to use circuit elements (other than semiconductor devices) that are not significantly affected by temperature variations between zero and 100 degrees Celsius. The conductors (wires) used for interconnection are virtually immune to temperature change. In the case of semiconductor devices, special "balanced" circuitry is frequently used to reduce the effect of temperature on the performance of the overall circuit.

In biological circuits, the situation is more difficult because essentially all of the circuit elements and interconnections are either electrolytic or semiconductive in nature. The wires used in biological circuits are invariably insulating conduits filled with an electrolyte. Both the conduit and the electrolyte are of organic origin. They frequently operate over a restricted biological range and exhibit a change in electrical parameters with temperature that is much greater than found in inorganic circuitry. The predominant circuit element in the biological circuits of vision is the semiconductor diode and its progeny, the active electrolytic semiconductor device or Activa. Both of these devices are sensitive to temperature according to the basic diode equation. The result of these factors is that biological circuits will not operate as designed beyond the biological temperature range and they exhibit significant changes in performance over the biological temperature range. The temperature must be accurately recorded and specified when making measurements related to vision in animals. In warm blooded animals, this problem is somewhat simplified. However, it is important to specify the temperature of the specimen (not just a nearby substrate) within 0.1 Celsius in accurate work.

11.7.1.3 Large versus small signal conditions

To describe the operation of any visual system completely requires a model capable of processing the very large dynamic ranges associated with vision. Such a model is known as a large signal model. For certain experiments involving only a limited dynamic range, a considerably simpler "small signal model" can be used. In conventional electronics, a small signal is one of less than 10%, and preferably 1%, of the amplitude of the nominal signal. These are the levels usually used to quantify the inter-modulation distortion in a common high fidelity audio amplifier. These values have been selected based on the relative linearity of most man-made electronics. In the case of vision, conditions are somewhat different. The visual system is inherently nonlinear based on the variety of diodes employed at voltages near their effective threshold voltage. It is useful to consider three categories of circuit diagrams;

+ the large signal diagram capable of representing any input signal faithfully.

+ a medium signal amplitude equivalent circuit diagram capable of representing any input of less than 3:1 peak-

to-peak amplitude variation.

+ a small signal amplitude equivalent circuit diagram capable of representing only a low amplitude, less than 10% peak-to-peak signal amplitude, input.

Visual experiments frequently vary from common electrical experiments. Instead of impressing a biphase signal on a fixed carrier, they frequently employ a fixed background and only add a mono-phase signal to obtain the total illumination. In both these experimental procedures, the peak-to-peak signal determines the applicability of the above equivalent diagrams.

The reason for defining three categories is that only the small signal equivalent circuit can be assembled using only linear man-made parts. Any experiments involving peak-to-peak signal level changes of 3:1 or greater will require that the mathematical equivalent circuit include nonlinear circuit elements. Any man-made circuit constructed to emulate the visual system over a peak-to-peak signal range of 3:1 or greater must be assembled from nonlinear parts or circuits that emulate nonlinear parts.

11.7.2 Large signal circuit diagrams

The topology of the fundamental large signal circuit diagram of the visual process is remarkably simple. This is due primarily to the level of sophistication employed in the design of individual circuits and individual circuit elements. The complexity is associated mainly with the level of parallelism involved and the signal processing related to spatial correlation of signals in the retina. The spatial correlation is performed primarily to reduce the number of individual signaling channels required within the optic nerve. **Figure 11.7.2-1** shows a complete circuit diagram for the fundamental circuits of a highly evolved animal eye such as found in small chordates. This circuit includes all of the circuit elements required to describe and emulate the performance of the animal eye. Terrestrial chordates normally are unable to sense ultraviolet light. In these animals, the UV and O signal channels are absent or vestigial. The parameters associated with each Activa can be summarized by a type designation as presented next to each Activa. These specific Activa types are tabulated elsewhere in this work.

It is important to note that the illumination used in an experiment must be specified in terms of its impact on each individual spectrally selective photodetection channel. Otherwise, one cannot expect to understand precisely the signal conditions at subsequent portions of the overall circuit.

It is also important to point out that only the complete fundamental circuit diagram can properly represent the operation of the animal eye under scotopic, mezotopic, photopic AND hypertopic conditions. This is because of the various nonlinearities in the system that come into play at different stimulus levels in different channels. This variation in the nonlinearities of the system requires careful documentation. When employing medium and small signal equivalent circuit diagrams, it is mandatory that the illumination level, to which they apply, be stated.

When discussing the dynamic performance of the eye, the operation and impact of the aperture stop known as the pupil must be considered. The role of the pupil is not that usually portrayed in the literature. The change in the area of the pupil, about 16:1, has a negligible effect on the operation of the eye compared to the 3500:1 dynamic range of the adaptation amplifier in the photoreceptor cells of chordate vision. The pupil remains at its normally minimum diameter throughout the hypertopic and most of the photopic regimes. By constricting the diameter of the optical bundle passing through the highly aspheric lens-cornea system, it provides the highest possible resolution (and contrast) image to the retina throughout these regimes. Over this wide illumination range, it is the adaptation amplifier that limits the signal amplitude within the signal processing functions of the eye. The pupil appears to begin opening at the lower illumination ranges of the photopic range in order to extend the illumination range of the eye. At light levels below the lower range of the photopic regime, resolution is sacrificed in order to maintain sensitivity. The diameter of the pupil is the mechanism that controls this tradeoff.

It is the large dynamic range provided by the combination of the adaptation amplifier and the pupil that maintain a virtually constant average signal amplitude in the circuitry of the visual system subsequent to the adaptation amplifiers. The system is able to maintain full sensitivity from the top of the photopic illumination regime to an illumination level approximately 2×10^{-5} lower. Below this level, the performance of the visual system begins to degrade significantly because the signal level in the signal processing circuits is reduced. The effect on the chrominance and appearance channels is most obvious as discussed below.

The following paragraphs will first discuss the operation of the eye based on the individual circuits represented in the overall large signal circuit diagram and based on their operation as a group. Following this discussion, simpler

equivalent circuits will be discussed that can be used under properly controlled laboratory conditions. The mathematical performance of the eye will be presented in Chapter 17.

11.7.2.1 Discussion of the large signal circuit diagram

As indicated earlier, the overall diagram can be divided into three major regions based on the grounding plan. These are:

+ the circuits within the IPM, basically the photodetection circuits that do not employ free electrons and the Adaptation amplifier circuits of photoreceptor cells.

+ the circuits found within the INM, including the distribution amplifiers of the photoreceptor cells found within the IS, all signal processing circuitry of the retina, and the signal encoding circuits (ganglion cells) of the signal projection subsystem located prior to the optic nerve.

+ The signal de-encoding circuits of the signal projection subsystem and the remainder of the higher level circuits of the brain found within the CSF.

Separate ground symbols have been used in the diagram to emphasize these divisions. The interconnecting bridge impedances between these ground planes are not shown.

The overall diagram can also be divided into three similar regions based on the treatment of the various signals. In this case, the complete photodetection/photoreceptor cell/first synapses complex can be considered the detection, amplification and primary signal distribution portion of the eye. The remainder of the signal processing and the signal encoding within the INM can be considered the second region of the eye. The third region remains the signal decoding and interpretation functions found within the CSF.

Under this second interpretation, **Figure 11.7.2-1(A)** describes the first region of the eye.

> Each of the rectangular symbols in the overall diagram represents a composite element consisting of a diode and an electrostenolytic electrical source that can be considered either a voltage or current source (in either case, a battery). This element also includes a capacitor in parallel with the diode and battery. However, it can be considered negligible in most cases. When this capacitance is significant, it is shown explicitly in the diagrams to follow. The capacitor, C_a, shown bridging the diode-battery of the adaptation amplifier, is a simplified representation of the capacity of the vascular system to support the electrical needs of the adaptation amplifier.

The circuit applies to each of the spectrally selective photodetection circuits of a given eye. There may be four of these present in a single eye. In each case, the photodetection process is shown as a loop containing two diodes for purposes of discussion. At the quantum physics level, the two diodes merely describe the excitation and de-excitation process in a single excitable molecular structure. De-excitation of the chromophores of vision does not occur spontaneously at biological temperature. There must be a de-excitation mechanism provided. This mechanism is provided by the open base connected Activa shown within the dotted box. This Activa can participate in an energy transfer between the chromophore and the hydronium crystal forming the base of the Activa. This hydronium crystal has a minimum activation energy of approximately two electron-volts. This minimum activation energy has a profound impact on the performance of the L-channel.

The diode symbol next to this Activa normally represents a light-emitting diode. In this case, it can be considered a *phonon*-emitting diode. It is analogous to a light-emitting diode with a threshold equal to 2.0 electron-volts. However, the energy is transferred to the base region of the Activa as a *phonon*.

The Activa contained within the dotted box is a unique type. It is described as a type AT Activa. It represents the sum of the individual Activas associated with the individual microtubules found in the crevices of the OS disk stack. For purposes of this figure, a single Activa is shown representing this multitude of individual lower capability Activas. The performance of this amplifier is critically dependent on the impedance in the collector lead of this amplifier. The dark and light adaptation characteristics of the eye are primarily a result of the interplay of the illumination level applied to the channel and the impedance of the power source supporting this adaptation amplifier.

The adaptation amplifier and the distribution amplifier, both contained within the photoreceptor cell, are shown connected as a differential pair with a common impedance between the two emitters and the local ground. The total current through this common impedance remains essentially constant in such a circuit. As a result, the signal current

passing through the distribution amplifier is essentially a mirror of the signal current passing through the adaptation amplifier. However, the output impedance of the distribution amplifier is considerably lower than that of the adaptation amplifier, allowing it to support a large number of 1st stage generic synapses. Each synapse, four are shown, can be considered as a low impedance connection between the collector voltage node of the distribution amplifier and the input circuits of the following signal processing stages. The impedance is low only in the orthodromic direction. In the opposite direction, it is extremely high. From this perspective each synapse acts as an isolation amplifier. Although all of the 1st stage generic synapses are shown as type AS, they exhibit a variety of reverse saturation current (RSC) values. These different values control the actual current level passed to the subsequent signal amplifiers for a given voltage potential at the output of the distribution amplifier.

Figure 11.7.2-1(B) describes the "R" channel, the *luminance* signal processing and encoding region common to all eyes.

> In the case of the simplest eyes, such as those of *Limulus*, several of the functions shown here are still performed by a single cell. This condition is probably indicative of an early evolutionary stage. By electrically separating the functions in later stages, it is probable that a higher degree of performance can be achieved in each.

The first bipolar cell accepts signals from all of the available photodetection channels described in **(A)**. The level of current received from each channel is controlled by the reverse saturation current (RSC) of the 1st synapse associated with that channel. These signals are summed in current space. The signal is then passed to the parasol ganglion cell by the 2nd generic synapse. The ganglion cell is prefixed by the expression parasol because it appears that there is a correlation between the morphological form of the ganglion cell and its actual function. The parasol cells appear to be electrically biased so that they exhibit a threshold level relative to any input signal. Below this level, no action potentials are generated by this *voltage controlled* mono-pulse oscillator.

A possible capacitor is shown bridging the 2nd generic synapse. There are a number of graphs in the literature suggesting that there is a pre-emphasis circuit in the luminance channel of many animals. This circuit provides a boost to the high frequency components in the luminance channel. The actual effect is to cause a slightly higher chance of the signal exceeding the parasol ganglion cell threshold level if it is caused by motion within the incident scene. The circuit appears to provide an enhancement to the alarm function to be discussed below. Such an enhancement was not found in the experimental literature of human vision.

The ratio between the capacitance of the two capacitors shown associated with the type AG Activa of the parasol ganglion cell may play two roles. It may play a role in the threshold performance of the circuit while also establishing the interpulse time interval found in the action potentials of the luminance channel as a function of illumination.

Figure 11.7.2-1(C) describes the *chrominance* signal processing and encoding region found in all but the simplest eyes. In most eyes, this region is replicated to provide either two channels (three-color-vision) or three channels (four-color-vision) signal recognition and interpretation. To simplify the overall diagram, only one chrominance channel is shown that can be used to describe the O, P & Q channels. Each channel receives input from two photodetection channels. The O channel receives input from the UV and S channels. The P channel receives input from the S and M channels and the Q channel receives input from the M and L channels. In each case, one of the inputs is applied to a non-inverting terminal of the type AB Activa. This is the dendritic input. The other input is applied to the inverting terminal of the same device. This is the poditic input. This differential amplifier provides not only signal inversion but also signal amplification to the poditic signal. Here again the reverse saturation current of the 1st stage generic synapse plays a role in determining the relative amplification of this signal. Later analyses will show that this value is selected so that the amplification factor of both the dendritic and poditic signals is the same. The resulting output signal is a biphasic analog signal. This signal is passed through the 2nd generic synapse to the midget ganglion cell. The prefix midget appears to correlate the morphology of this cell to its unique electrical performance. This circuit does not exhibit any threshold to input signals. In fact it is a free running oscillator generating action potentials with a nominal time interval in the absence of any input signal. The input signal, being biphasic, is able to cause this time interval to increase or decrease. The signal information is encoded by this variation in pulse interval.

> The presence of a continuous stream of action potentials from these cells in the absence of illumination is a frequent source of confusion in the literature. The frequency of these pulse streams will decrease in the presence of predominantly S- or L-channel illumination. This characteristic causes these chromatic signals to be mistakenly labeled OFF-Luminance signals.

A fourth type of signal processing occurs in many animals. It involves the processing of two separate UV signals taken from photoreceptor cells with their rhabdomere arranged orthogonally. By employing the same circuitry as shown in **Figure 11.7.2-1(C)**, a fourth *polarization* signal processing channel can provide this information to the CSF.

Figure 11.7.2-1(D) describes the *appearance* signal processing and encoding region also found in most eyes. The scope and significance of this region do not necessarily correlate with our general view of the evolutionary status of the animal. In the human, this area is not highly developed.

The fact that some of the neurons in the optic nerve are found to display a constant train of action potentials while others do not in the absence of any illumination of the retina is simply explained by their role relative to the R channel versus other channels in the eye.

It appears that the timing between the various signaling channels discussed above is maintained to a large degree by the constant length of the neurons between the LGN and the Optic Lobes of the cerebellum. This goal of constant signal path length in spite of variable distance to be covered results in the fan shaped optical radiation found between the LGN and Optic Lobes.

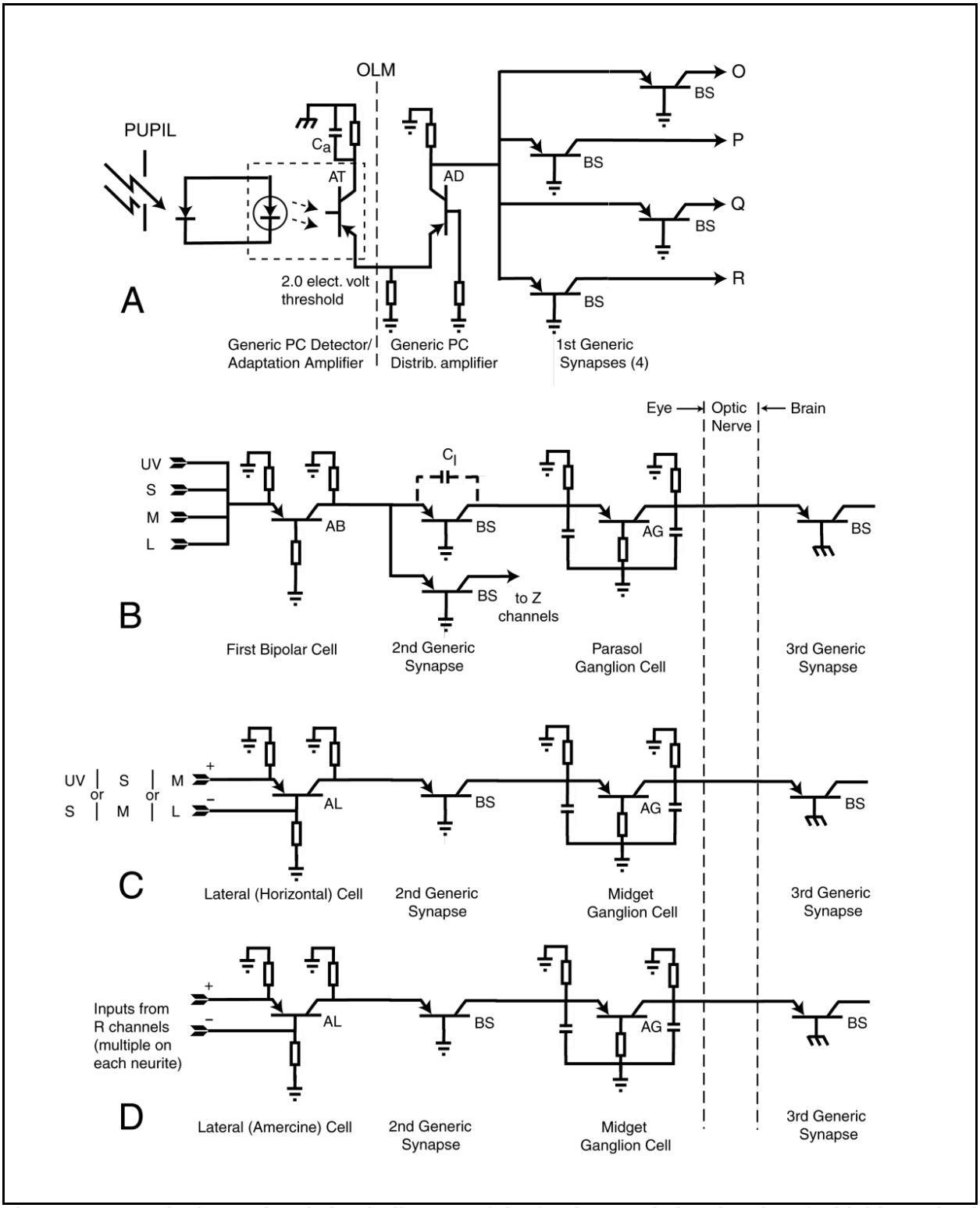

Figure 11.7.2-1 The large signal circuit diagram of the fundamental signal paths of a highly evolved animal eye. (A) Generic photodetection module, (B) Generic luminance channel-- R, (C) One of three generic chrominance channels--O, P & Q. (D) Generic appearance channel--Z_n.

Figure 11.7.2-2 describes, only in conceptual terms, the signal decoding circuits prior to any distribution amplifiers, found within the brain. These circuits can provide two fundamentally different outputs. One output is an analog output that can be interpreted in two different ways by subsequent circuits. For luminance information, the analog signal is interpreted as a monophasic signal indicative of scene illuminance. For chromatic information, the signal is interpreted as a biphasic one centered about a quiescent value occurring in the absence of signal information in that channel. Little is known about the cognition related to the analog signal generated by the appearance channels. This appears to be due to the minor role played by the appearance channels in humans. Little information is available from psychophysical experiments involving the human although some information may be available from work with cats.

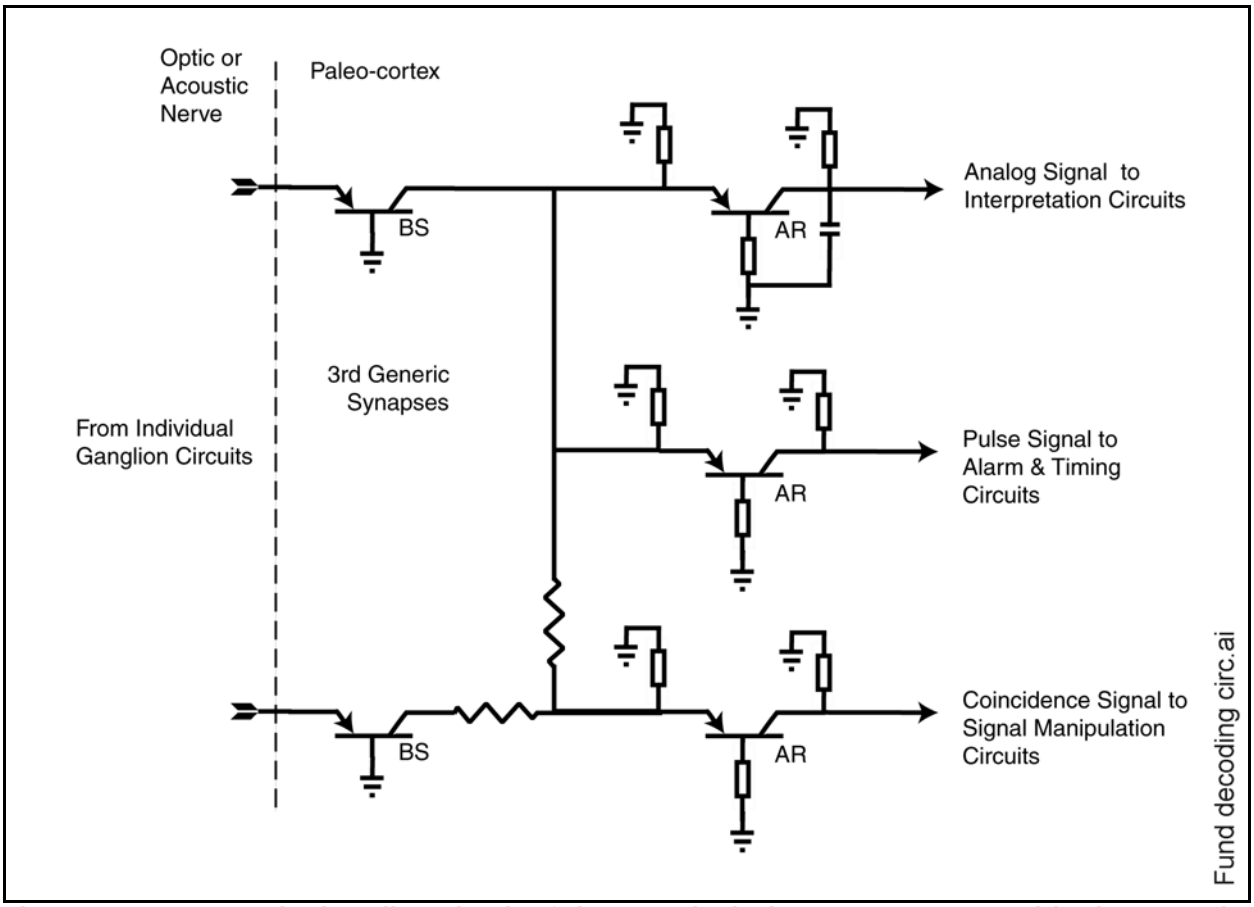

Figure 11.7.2-2 Generic decoding circuit of the neurological system EDIT. Used in the LGN, the Pretectum and the cortex.

The second output is essentially an initial alarm signal and is found primarily if not exclusively in the luminance channels, particularly the (non-foveal) luminance channels terminating in the Magnocellular region of the brain. Such an alarm signal plays a critical role in causing the Superior Colliculus to redirect the optical line of sight of the eye. This redirection brings the image of any potential threat to the animal into the foveal area of the retina.

In analogy with other areas of the brain, particularly the striated cortex, it appears these decoding circuits are represented morphologically by stellate cells. The stellate (multi-cornered or star shaped) cells are optimally suited for the decoding process. They appear in profusion in both layer 4C of the striated cortex[88] and in the layers of the LGN. They generally exhibit both a dendritic arborization(s) extending from the cell opposite to the short axon and

[88] Noback, C. (1967) The human nervous system. NY: McGraw-Hill pg. 233

a poditic arborization extending from an area between the dendrite(s) and the axon. This poditic structure is known as the basilar dendrite in the literature of the striated cortex. In the context of the striated cortex presented by Noback, the nearby pyramid cells are projection neurons which may have axons as long as three feet. These are not likely to be decoding neurons. They are more likely re-encoding projection neurons with axons leading to other areas of the brain or the more remote areas of the neural system. By comparing the work of Noback and that of Hubel[89], considerable information about the initial organization of the brain can be gleaned. Note however that their presentations are not to scale. The vertical dimension has been expanded for clarity. In general, the cortex has a surface area of about 2.5 square feet but is only 1.5 mm. thick in the area of the visual cortex to 4 mm. thick in the motor cortex. Note also the duplication of terminology found in the morphology literature (pg. 232 of Noback). The pyramidal cells with axons labeled projection efferent fibers or association efferent fibers are functionally equivalent to the ganglion cells of the retina. These are all projection neurons that generate action potentials. They typically have axons longer than 1 mm. Those labeled association efferent fibers only travel distances shorter than the length of the lobe of the cortex in which they originate. When they re-enter the same lobe, they are labeled association afferent fibers. Those labeled projection efferent fibers are longer and travel between distinct lobes of the brain. They change names to projection afferent fibers upon approaching their target locations.

Note the similarity of all of the neural circuits of the visual system. There are only three basic circuit types; the unique circuit of the sensory (typically photoreceptor) cell, the simple synapse circuit, and a third, more elaborate, circuit. This latter circuit can be tailored in three ways to satisfy a variety of requirements. By using both the dendritic and poditic neurites as signal inputs, the circuit can be made to take the difference between two input signals. By providing multiple synapses at any neurite, input signals can be summed at this neurite. By adding capacitors at critical points, the circuit can also perform signal processing, signal encoding or signal decoding. By adjusting the potential difference between the emitter and base of the Activa, the circuit can be made to function as a threshold detector, an analog integrator, a free running oscillator or as a mono-pulse oscillator. The necessary capacitance is easily implemented by varying the exposed (unmyelinated or otherwise protected) surface area of the neurite or axon. The ratio of Activa terminal voltages can be affected by a change in the amino acids employed in the two relevant electrostenolytic processes (various acidic amino acids, including GABA) and/or by the relative RSC's of the diodes involved.

11.7.2.2 Gross dynamics of visual performance

It is possible to establish the quiescent voltages and currents applicable to the overall circuit diagram, found in the eye of a specific species. It is then possible to discuss the dynamic operational capability of that eye in some detail. These parameters do not appear to vary greatly among species within a phylum. Probably the greatest variation in these parameters occurs due to the temperature of the specimen. This Section will discuss the parameters of the human eye as an example.

The above parameters can be tabulated for the "no illumination" condition. Under this condition, no circuit is under stress and the gain of the adaptation amplifiers is at maximum. No action potentials appear in the optic nerves related to the R channels. Trains of action potentials appear in the optic nerves associated with the chrominance and appearance channels. Each of these pulse trains exhibits a quiescent pulse-to-pulse interval. These intervals need not be identical. The decoding circuits of the brain are able to accept and adjust to virtually any quiescent pulse interval.

The human eye exhibits at least four operating regimes that can be related to changes occurring in the operating regimes of the various circuits shown in the overall circuit diagram. These regimes have been expanded to five in this work. They are labeled (in order of illumination level) the scotopic, mesopic, photopic, hypertopic and hybrid illumination regimes. All but the hybrid regime assume a constant photon flux per unit bandwidth radiation input to the eye. As discussed in Chapter 2, an illuminant (F), defined by a 7053 K black body, is ideal for physical experiments involving the human eye. It provides an equal photon flux per unit wavelength over the spectral range of interest in experimenting with long wavelength trichromats such as the human. It is possible to use different illuminants, corresponding to different color temperature sources. However, the non-uniform excitation of the various spectral channels of the eye must be accounted for.

> Since the chromophores each exhibit a similar absorption characteristic with wavelength, all of the chromophoric channels are excited nearly equally if an equal photon flux per unit wavelength source is used. Such a source stimulates all of the chromophores nearly equally and leaves each of

[89]Hubel, D. (1988) Eye, brain, and vision NY: W. H. Freeman

the chrominance channels in a quiescent state. Such a source will appear as a "white" source at any intensity level within the photopic range. Sources at other color temperatures vary significantly in perceived color as the intensity is adjusted.

Since these illumination regimes describe a performance envelope that may be controlled by multiple underlying processes, their correlation to the performance of individual circuits in the overall circuit diagram is not unique. There is overlap. However, the correlation between a specific circuit, or circuits, and a specific regime is good.

Very little material has appeared in the literature describing the overall sensitivity of the visual system to light. **Figure 11.7.2-3** reproduces a figure from Barlow with additional annotation[90]. A similar figure was reproduced by Frumkes in 1990[91]. While the Frumkes paper is insightful, he introduced deletions and changes to Barlow's figure that were not well supported. This work does not support the delineation of rods and cones in the original figure.

Barlow described the figure as "a curve of the human increment threshold taken under conditions *carefully selected* to show up to 5 types of orderly behavior:" [emphasis added] The data applies to a point ten degrees from the fixation point. His paper did not describe the direction from the point of fixation where the stimulus was applied. He provided no data points or error bars and the graph appears to be drawn freehand. There are several problems between the original paper and the original figure. The label on the lower axis also appears to need interpretation since the background light was specified as orange. Frumkes added a degree of precision 25 years later by specifying the wavelength of the background light as 600 nm. The background may have been specified in Barlow's mind as equivalent to the photon flux at 507 nm. The measured values apply to the intensity of the source and not to the flux received by the eye. The condition of the pupil was not discussed. It appears that some type of fixed pupil was used. By incorporating the action of the pupil, the overall dynamic range of the system would be addressed by an additional factor of about sixteen, more than a factor of ten.

Section 17.2 and supporting material in **Chapter 16** will develop additional definitions with regard to human vision that appear to apply to this figure. This notation brings the historical names for the phenomenological regions of perceived illumination intensity into alignment with the underlying mechanisms that cause them. The resulting notation has been added along the top of the figure. This material will introduce a number of new interpretations applicable to this graph. It will show that the sensation of color is lost at a level within the mesopic intensity range. It will also show that color constancy is only maintained within the photopic range, the only area where Weber's Law applies.

The material will also show that the phenomenon described as dark light is in fact the limit due to the noise performance of the stellate cells in the CNS (or the establishment of a hard threshold within the CNS to de-emphasize this noise). This threshold is independent of the stimulus intensity. The performance in the mesopic range is the only performance limited by photon noise in the stimulus. This range shows the slope of one-half associated with this condition. Within the photopic range, the performance of the system is limited by the dynamic range of the signal processing circuits rather than the incident photon noise. In the hyperopic range, performance is further limited by saturation effects associated primarily with the output amplifier in the photoreceptor cells. There are additional details regarding these ranges that will be discussed below and in later chapters.

Section 17.2 and **17.3** will also demonstrate that the test stimuli used by Barlow introduced signals into both the luminance and chrominance channels of vision. Thus, the figure does not relate only to the luminance threshold, particularly at high illuminance levels.

---

[90]Barlow, H. (1965) Optic nerve impulses and Weber's Law Symp Quant Biology, Vol. 30, NY: Cold Spring Harbor Laboratory pp 539-546

[91]Frumkes, T. (1990) Classical and modern psychophysical studies of dark and light adaptation *Chapter 7 in* Leibovic, K. *ed.* Science of Vision, NY: Springer-Verlag

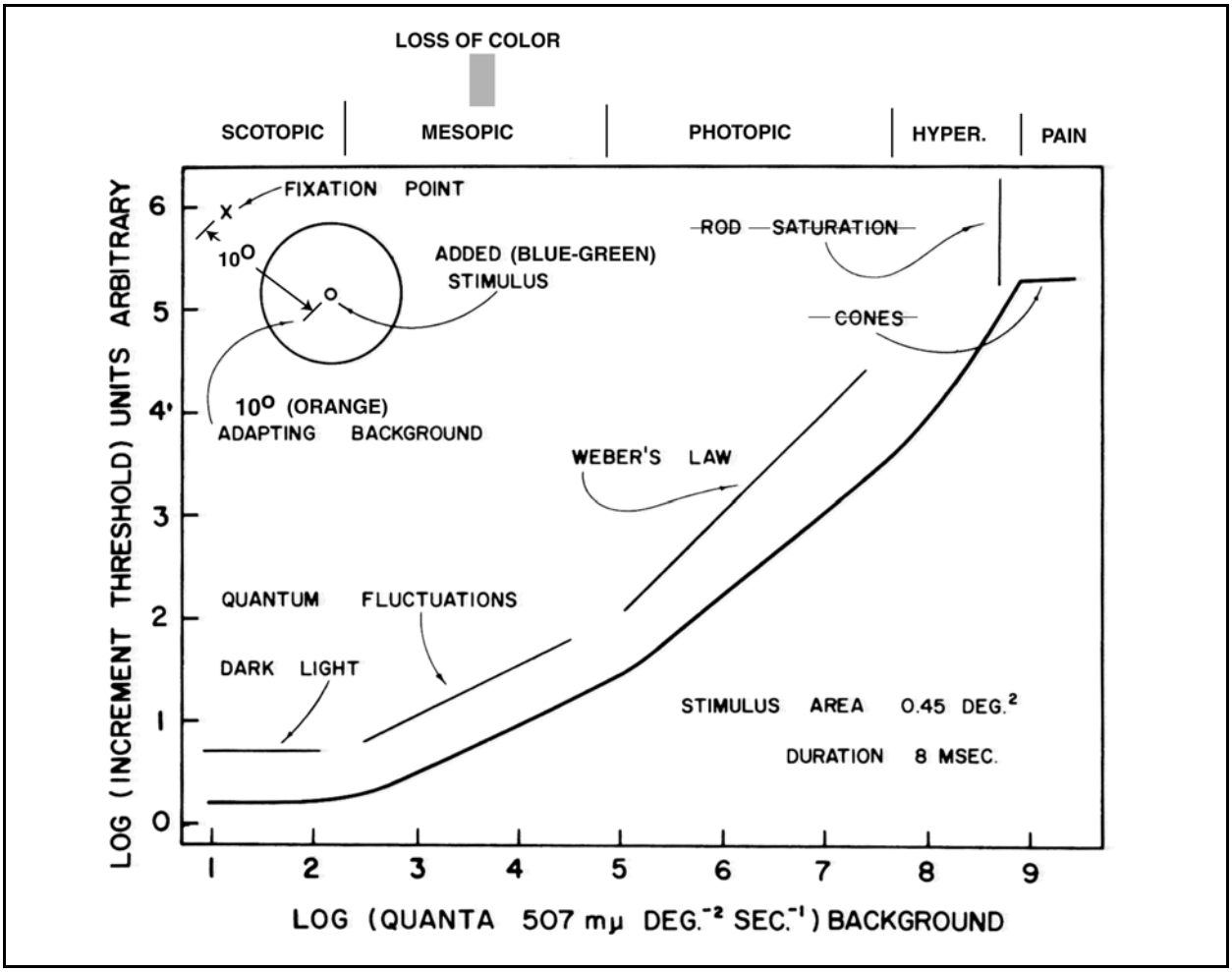

Figure 11.7.2-3 Incremental threshold of a human subject versus a chromatic background. See text for discussion. Modified from Barlow, 1965.

11.7.2.2.1 The photopic regime

The photopic illumination range is the most widely studied. It is characterized by stable color rendition over a large intensity range. This is caused by the stabilizing action of the adaptation amplifiers. These amplifiers exhibit a high degree of negative internal feedback that increases with input signal intensity. They operate near maximum gain at the low intensity end of this range and near unity gain at the high intensity end. For 7053 K illumination, each photodetection channel is equally excited and the output signal levels of the adaptation amplifiers track each other and remain essentially constant. The chrominance channels remain at quiescent levels. If an auxiliary light source of narrow spectral width is introduced, the effect of this additional radiation will result in different signal levels at the output of the photodetector channels and disturbance of the quiescent levels in the chrominance channels. A "color" corresponding to the auxiliary source will be reported by the subject.

The large dynamic range associated with the adaptation amplifier circuit s is obtained primarily from the quantum mechanical operation of the liquid crystalline chromophoric material coating the disks of the outer segment (**Section 7.4.2**) augmented by a poorly regulated collector power supply. There is no external feedback loop and consequently no amplifier in that feedback loop. The time constant associated with the capacitor, C_a, creates a low pass filter in the circuit determining the voltage of the collector terminal. Therefore, the average adaptation amplifier output signal remains near its quiescent value regardless of slow changes in the incident illumination level. Any rapid changes in the incident radiation applied to a photoreceptor cell will be amplified by the adaptation amplifier

since the average voltage on the collector cannot change rapidly because of the above time constant.

If the *average* incident illumination level changes rapidly, a more complex situation is involved. This situation will be examined in the absence of any rapid changes in the incident radiation associated with fine detail in the scene. There are two situations, associated with the general labels dark adaptation and (the less frequently used) light adaptation.

If the light level applied to the retina is raised rapidly from a dark level, the base (poda) terminal of the adaptation amplifier is unable to provide sufficient current to the de-excitation of the chromophoric material. As a result, the chromophoric material becomes significantly more transparent (appears bleached) and exhibits a significantly lower quantum efficiency. This change in quantum efficiency occurs virtually instantaneously and is observable by recording the light adaptation characteristic of the retina.

If the light level applied to the retina is lowered rapidly from a high level of illumination, the chromophoric material will continue to be de-excited by the deactivation mechanism within the current capability of the base (poda) terminal. However, this current capacity is subject to the quality of the collector (axon) power supply of the adaptation amplifier.

For a light level applied to the retina that is lowered rapidly from an intermediate level of illumination, the impact of the collector (axon) power supply will be negligible.

The complete dark adaptation characteristic of the retina incorporates both the medium level and high level cases. The medium level case is represented by a simple exponential curve. The high level case is represented by the more complex “exposine” function (**Section xxx**).

The quality of the collector (axon) power supply parameters vary with location over the surface of the retina.

To obtain the complete adaptation characteristic of the retina, the light and dark adaptation characteristics can be combined (**Section xxx**). To obtain the complete adaptation characteristic of the entire eye, the effect of the pupil of the eye must also be incorporated into the calculation.

As the illumination level decreases to the point where the adaptation amplifiers are operating at or near maximum gain, the aperture stop of the eye (the pupil) is commanded to begin opening. The increase available in the area of the pupil, nominally 16:1, allows the visual system to continue operating in the constant signal amplitude mode as far as the signal processing system is concerned. However, the resolution performance of the eye is degraded in this region.

11.7.2.2.2 The mesotopic regime

The lower extreme of the photopic regime merges with the mesotopic regime. At this juncture, the adaptation amplifiers are operating at maximum amplification (negligible negative internal feedback) and cannot compensate for any difference between the currents generated in the base of the adaptation amplifier. The signal currents generated in the base of the S- and M-channels continue to track each other and decrease linearly with illumination level. However, the current in the L-channel drops more precipitously due to the square law nature of the de-excitation process associated with the L-channel chromophore. The result is a loss in long wavelength sensitivity in the visibility function and a loss in long wavelength chromatic saturation. These effects are presented in detail, both graphically and mathematically, in Chapter 17. The failure of the chromophoric detection channels to track each other at the transition between the mesopic and photopic regimes is characterized by the Purkinje Effect.

After the adaptation amplifiers reach maximum gain, the signals applied to the luminance channels begin to decrease in amplitude. This results in the signals in the chrominance channels also decreasing in amplitude. They begin to approach the quiescent signal level. The result is a continuous decrease in chromatic saturation in this regime. At some point, the chromatic interpretation circuits in the brain are unable to discern an accurate chromatic state because of the low deviation from the quiescent state of the action potentials arriving at the brain. The interpretation circuits report a quiescent condition. They continue to report this condition at lower illumination levels. When the chromatic interpretation circuits report a quiescent condition, the subject reports an achromatic scene. The achromatic condition is the hallmark of the scotopic regime.

11.7.2.2.3 The hypertopic regime

If the illumination level is increased above the normal maximum of the photopic regime, two separate conditions

arise. First, The adaptation amplifiers as a group are unable to reduce their gain below unity. The total current through the adaptation amplifiers attempts to exceed the constant current through the common emitter impedance. This causes the distribution amplifier to be cutoff. Under this condition, the luminance channel transmits a very high signal to the brain and the signals in the chrominance channels are also very large. It appears the cutoff of the distribution amplifiers is reported to the brain via the pain reporting neural system. In this regime, it appears the chromophoric material associated with the M-channel detector (the liquid crystalline material of the disks) also begins to saturate. The result is a loss in signal amplitude in the M-channel relative to the other two channels. Due to the logarithmic summation in the luminance channel, This loss is reported by the subject as an enhancement of chromatic sensitivity at either the 494 nm or 600 nm, or if the instrumentation is sufficiently precise, both wavelengths. This situation is characterized by the Brezold-Brucke Effect. The actual effect is totally related to the luminance channel. It results in an *apparent* increase in chromatic saturation of the scene.

11.7.2.2.4 The scotopic regime

As discussed in the paragraph on the mesopic regime, beginning at the lower edge of the mesopic regime, the interpretation circuits report a quiescent condition. They continue to report this condition at lower illumination levels. When the chromatic interpretation circuits report a quiescent condition, the subject reports an achromatic scene. The achromatic condition is the hallmark of the scotopic regime. Under this condition, **all** of the photodetection circuits are still active and the adaptation amplifiers are operating at maximum gain. However, the output signal levels of the photoreceptor cells are too low to affect the chrominance circuits.

11.7.2.2.5 A hybrid regime

If a scene exhibiting a color temperature different from illuminant (F) at 7053 K is presented to the subject, the experimental results may appear confusing. However, by properly evaluating the performance of the various circuits in the overall circuit diagram, the experimenter can interpret the subject's observations without difficulty or serious error. The primary difficulty is in the requirement to integrate the total photon flux detected by each chromophoric photodetection channel separately. If the scene color temperature is still that of a black body but significantly lower than 7053 K, the excitation of the S-channel will be significantly reduced compared to the other channels. The data obtained will show a significant loss in short wavelength response in all of the visibility functions, corresponding to the above regimes. The new Chromaticity Diagram for Research will also show reduced performance in the short wavelength region relative to what would be expected for an eye irradiated by illuminant (F). Some of the older data in the literature was taken using illuminants with a color temperature as low as 2400 K. This data is so deficient in the short wavelength spectrum as to be useless for scientific purposes.

Although discussed elsewhere, the relatively slow recovery of the gain of the adaptation amplifiers following intense scene illumination provides an excellent method of determining the precise spectral characteristics of the chromophores of vision. By selectively bleaching the retina significantly and then immediately making spectrographic scans, the precise spectral response of the individual chromophoric channels can be obtained psychophysically. This work was done with considerable precision by Wald in 1964[92].

11.7.3 Medium signal circuit diagrams

As discussed above, the medium signal circuit diagram can be used for peak-to-peak signal excursions of up to 3:1. It is not appropriate for larger excursions. **Figure 11.7.3-1** provides a medium signal circuit diagram for the visual system. This circuit is derived from the overall circuit diagram by treating all amplifiers as exhibiting a fixed current gain. This gain is unity for all synapses as indicated by the paired arrows between a diode and a current source at several places in the diagram. The gain is a fixed value for the adaptation amplifiers within the photoreceptor cells, but it need not be the same for each amplifier. Because of the variation in gain among the various photoreceptor cells as a function of luminous intensity, the gains must be specified for each luminous regime of interest. In developing the medium signal equivalent circuit diagram, it is important to recognize that the transfer function between the photodetection channels and the signal processing channels involves a current to voltage conversion. Care must be taken to interpret this conversion properly. The diode associated with the collector impedance of the distribution amplifier in (A) is important in this conversion.

---

[92]Wald, G. (1964) The receptors of human color vision. Science, vol. 145, Sept. 4, pg. 1009

The medium signal level circuit diagrams are luminance-regime specific and must be properly labeled. In all but the hypertopic regime, the signal transfer function associated with the chromophoric material will be a constant for all chromophoric channels. In the hypertopic regime, the individual levels associated with the equivalent current generators must be specified for each spectral band. Similarly, in all but the mesopic regime, the signal transfer function will be a constant for all chromophoric channels. In the mesopic regime, the level associated with the L-channel must be specified relative to that of the other channels. In the scotopic regime, the L-channel signal level is defined as zero.

The signal level in the signal processing channels (B) & (C) is always large relative to the characteristic voltage nV_t associated with the individual diodes. Therefore, the impedance of the individual diodes varies significantly within the dynamic range of a medium amplitude signal input. Because of their individual reverse saturation currents, I_0, two of these diodes in series may or may not be represented by an equivalent pair of resistors in series.

The photodetection portion of the medium signal equivalent circuit is shown in (A). The same circuit is used for each of the chromophoric channels, UV-, S-, M-, & L- for the hypertopic, photopic and mesopic illumination ranges. This portion is modeled as an alternating current source proportional to the variation in the photon flux incident on the appropriate chromophoric photodetector and a diode. The current from the source is applied to a low impedance diode. The voltage across this diode (between the node and ground) is the parameter normally measured in the laboratory. The voltage at this node is used to provide the same signal to each of the signal processing channels, O, P, Q, R, & (if appropriate) Z. The first synapse used to deliver this signal to the signal processing channels is shown in (B) and (C) to simplify the presentation.

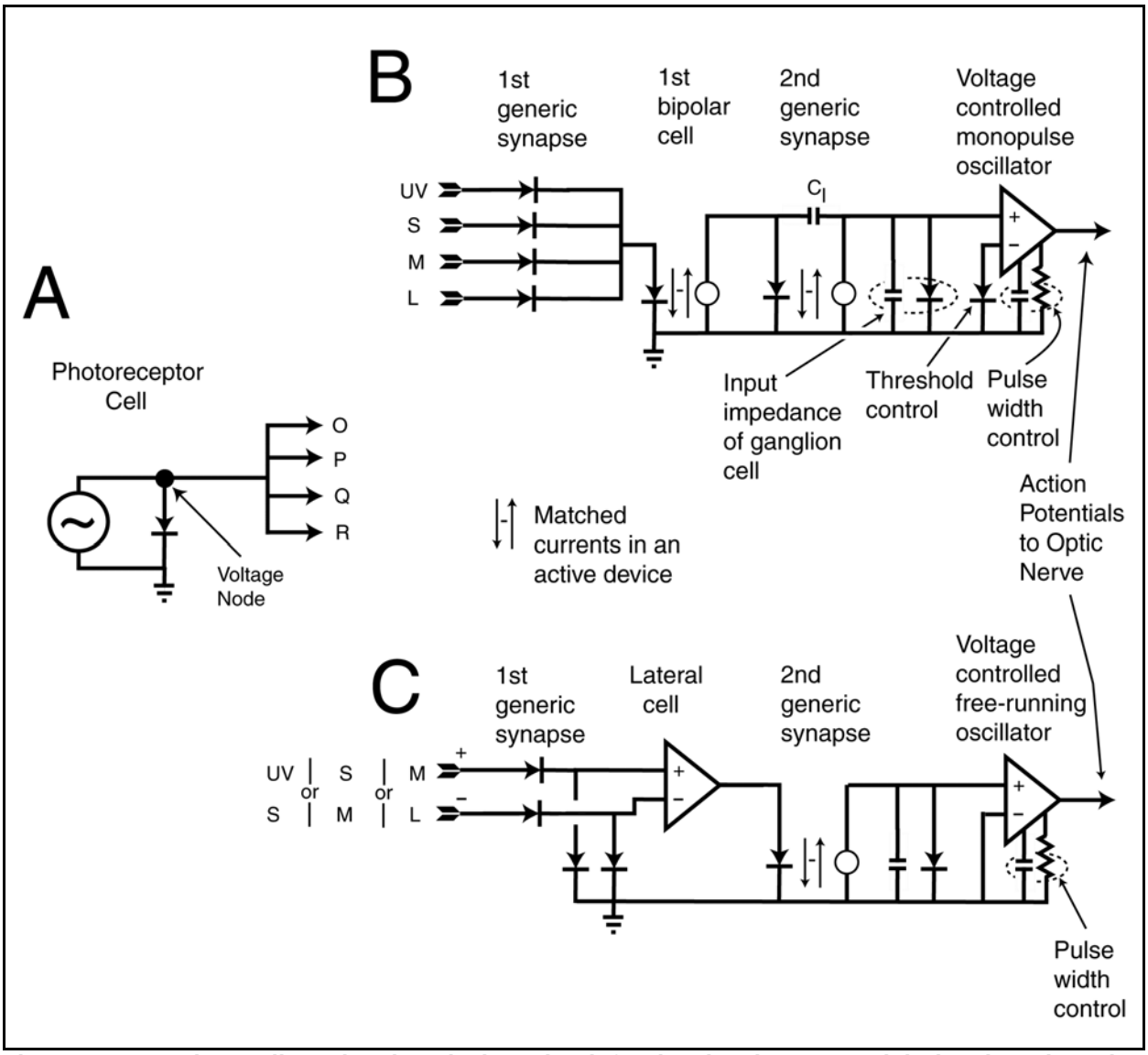

Figure 11.7.3-1 The medium signal equivalent circuit for the visual process. (A) The photodetection portion. (B) The luminance channel, R. (C) A channel representing either the O, P, Q, or (if appropriate) Z channels.

(B) portrays the medium signal equivalent circuit for the luminance channel, R. The first synapses connecting to the luminance channel, R, are shown as simple diodes. However, using two diodes in series as a voltage divider is awkward. The current through each series pair of diodes varies non-linearly with the node voltage. As a result, the current waveform passed through the 1st bipolar cell is a distorted version of the waveform at the voltage node of (A). The amount of distortion depends on the amplitude of the signal at the voltage node. This distortion is easily recognized in the signal waveforms recorded by **XXX**[93].

The second generic synapse is shown as a diode and a current source with equal currents in each to stress the

[93]XXX

isolation between the input and output impedance of the device. In the absence of the capacitor, C_l , the input and output impedances of this device are completely isolated. The impedance of the second generic synapse and the input impedance of the parasol ganglion cell form a lead-lag network if there is a significant capacitance, C_l, in parallel with the synapse. Otherwise, the circuit is a simple lag network. When a lag network, it typically has a high frequency roll-off that is higher than required by the visual system and it does not affect the performance of the visual system. When the circuit is characterized as a lead-lag network, the high frequency components of the signal may be amplified more than the low frequency components. Here again, the two diodes in series, both shunted by capacitors, results in a complex transfer function that is a function of the average signal level in the luminance channel. The average signal level in this channel does not change greatly while the eye is operating in the photopic range. However, the instantaneous signal level may vary considerably, affecting the lead-lag network calculations, and must be recognized. The parasol ganglion cell operates as a voltage controlled mono-pulse oscillator that exhibits a threshold controlled as indicated conceptually. Once the input signal exceeds this threshold, the oscillator will generate a series of action potentials with a pulse-to-pulse spacing determined by the height of the signal above the threshold. The width of the individual action potentials is determined by the pulse width control.

(C) provides a medium signal equivalent circuit applicable to each of the chrominance or appearance channels of an eye. The UV- and S- signals apply to the O channel, the S- and M- signals apply to the P channel and the M- and L- signals apply to the Q channel. The inputs to a Z channel are drawn from groups of R channels as discussed for the overall circuit diagram. As in (B), the 1st generic synapses are shown as simple diodes for convenience. The synapses associated with both the non-inverting and inverting signals are in series with the diode representing the input impedance of the lateral cell. As in the R channel, this series combination can introduce waveform distortion into these channels. However, this distortion is minimized in the output of the lateral cell due to the differencing operation. The 2nd generic synapse acts as a simple current repeater type impedance isolation device. In the chrominance and appearance channels, the midget ganglion cell operates as a free-running (continuous) oscillator with the width of each pulse determined by the pulse width control. The pulse to pulse spacing can be increased or decreased depending on the amplitude of the signal voltage created by the current from the 2nd generic synapse integrated on the input impedance of the cell.

11.7.4 Small signal circuit diagrams

As discussed above, the small signal circuit diagram can be used for peak-to-peak signal excursions of about 1.1:1. It is not appropriate for larger excursions unless a sacrifice in precision is acceptable. A variety of small signal equivalent circuits can be derived from the overall circuit diagram depending on the goal of the investigator. Lacking values for the various circuit elements, a conservative approach is taken in **Figure 11.7.3-2** that preserves certain function relationships between the impedances present. The approach is to replace each real diode in the medium signal equivalent circuit with a real resistor of equivalent impedance at the signal level present. However, replacing the real diodes with real resistors could lead to erroneous results in two principal cases. Making such a replacement for the real diodes in the summing network of the luminance channel could lead to current flow between the voltage nodes of the various photodetection circuits. A similar situation is present in the chrominance and appearance circuits when they are expanded to reflect the arborization present in the real case. Therefore, each real diode has been replaced by a resistor in series with a perfect diode with $V_t = 0$ & $I_0 = 0$. Such a diode acts as a perfect rectifier. It prevents unrealistic current flow within the summing network of the small signal circuits.

To avoid similar misinterpretations, the 1st bipolar cell and the 2nd generic synapse have been described by an unreal circuit element. As shown, this is an active device consisting of a resistor and a current generator operating as a pair. The current generator produces a current equal to the current through the resistor. This is done to preserve the independence of the preceding and following circuits from an impedance perspective. If the actual impedance values were known, it is possible that some of the circuit elements shown could be combined. An even simpler small signal equivalent circuit would then be obtained.

In both the medium and small signal equivalent circuits, the lateral cells are represented by transimpedance amplifiers to preserve the proper impedance relationships. The pulse width control circuits are also shown conceptually. The exact shape of the action potentials cannot be obtained using an RC time constant.

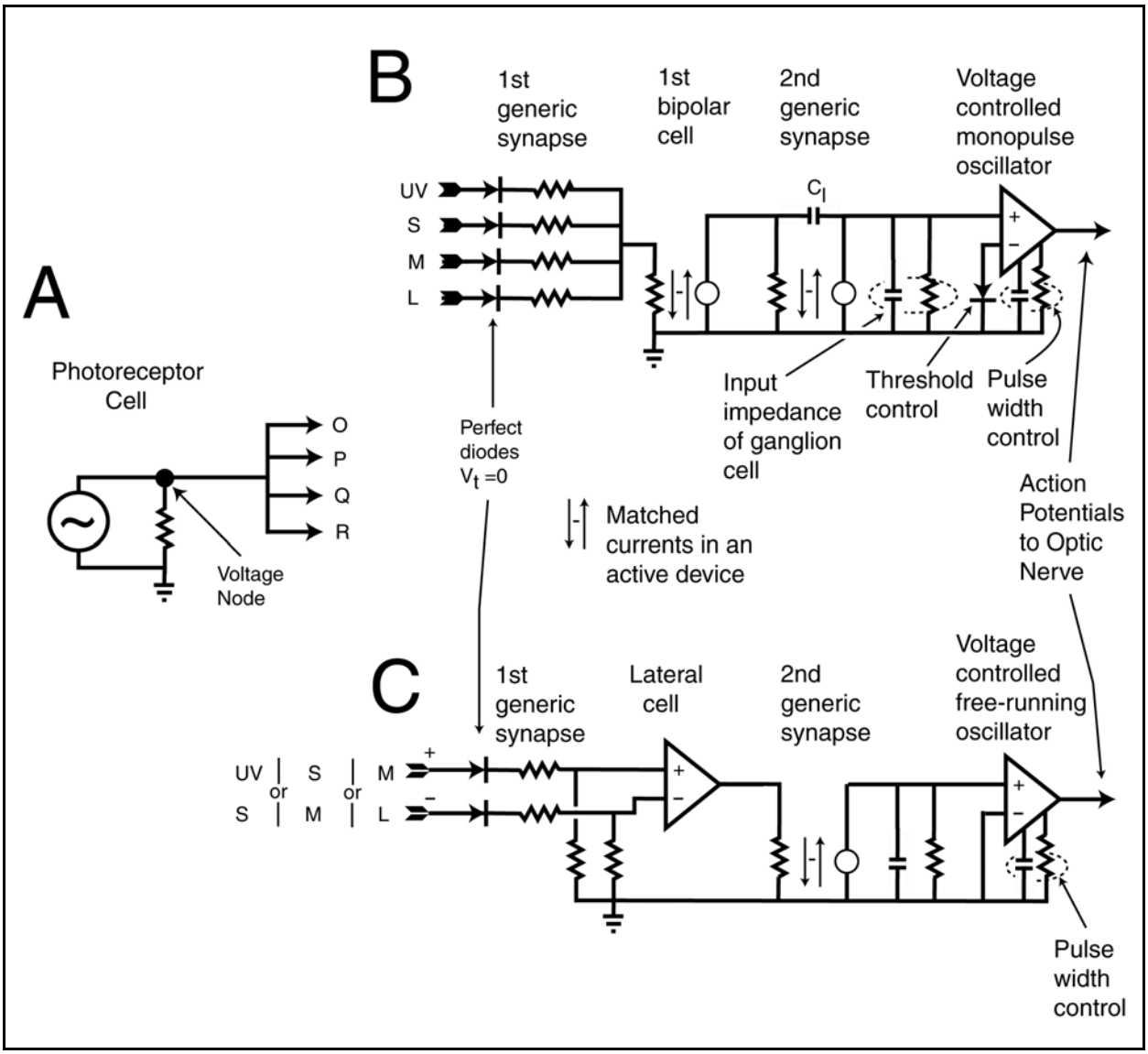

Figure 11.7.3-2 The small signal equivalent circuit for the visual process. (A) One of four photodetection circuits. (B) The nominal luminance channel circuit. (C) The nominal appearance and (one of three) chrominance circuits. The lateral cell equivalent amplifier is of the transimpedance type.

11.7.5 Noise model circuit diagrams

The development of noise models is a specialty within electrical engineering. To develop an appropriate model requires intimate knowledge of the system being modeled. Although appropriate to develop this topic in Chapters 16 or 17, it will be discussed here because of its character as a modeling tool. The discussion will rely upon a number of parameters found within the Appendix on the Standard Eye. Some of these parameters are developed in later chapters of this work.

Few models describing the noise in the visual system have appeared. There is more data than there are models to

relate to that data. Rose attempted to migrate a model, based on his television background, to vision in 1977[94]. It lacked a detailed model and was not well received. Laughlin presented a broad review of the potential noise performance of the visual system, again lacking a detailed model of the system[95]. Data showing the square root relationship between input intensity and measured noise levels is presented for both the ultraviolet and mid wavelength channels.. He introduces the term "shot noise" but does not define it. Shot noise is traditionally associated with the random emission of thermally generated electrons from a hot cathode. The term is frequently used colloquially but will not be used here. The analysis frequently relies upon analogies with chemical kinetics and concludes that the transduction function is inherently noisy. He also concludes that the work of understanding the noise performance of vision is yet to be completed.

An appropriate noise model can be adapted from the above small and large signal models. Aspects of both models are needed because of the quantum noise associated with the input excitation. Because of the unique situation where the long-wavelength photosensing channel employs an additional mechanism not found in the other channels, it is important that this channel be analyzed separately.

11.7.5.1 Defining the model

Defining the noise model requires considerable familiarity with the operating mechanisms of the visual system. Care must be taken to separate the noise measured electrophysiologically from that observed psychophysiologically. Noise at the electrophysiological level within the visual system involves a series of dynamically changing physical circuit parameters. Noise at the psychophysical level introduces additional computational mechanisms that are not rigorously definable at this time.

The simplest situation to address involves illumination at an intensity level near the bottom of the photopic range. In this region, all of the adaptation amplifiers are operating at a nominally fixed (maximum) gain. In addition, the load on the metabolic/hydraulic system is minimal. Therefore the overall gain of the photoreceptor cells can be considered fixed. Under these conditions, the overall signal processing architecture shown in [**Figure 11.6.4.2**] can be used as a point of departure. However, it is important to bound the level of spatial signal integration within the retina. Therefore, only signals from one quadrad (or triad in the case of a trichromat) of adjacent photoreceptors will be considered.

To arrive at the correct noise model, it is necessary to consider all potential noise sources. There are three categories of these sources.

+ thermal noise sources that occur between the terminals of a resistor.

+ current noise sources related to steady currents through an impedance.

+ quantum statistical noise normally associated with the photoelectric effect.

There are others, but these are the most important ones.

It is also necessary to consider all methods of photon detection. There are three categories of these processes.

+ thermal detectors where the heating of the material by light causes a signal

+ photoconductive detectors where the incidence of light causes a change in conductivity (without heatind).

+ photoelectric detectors where a photon causes the emission of an electron from the ground state of an atom.

The operation of the visual system is dependent on the latter mechanism. The chromophores act as liquid crystalline photoelectron emitters. The electrons are released internally instead of into free space as in a photomultiplier tube. The equivalent solid state device is frequently labeled a solid state photomultiplier because of its high sensitivity to light.

The chromophores of vision act as photoelectric detectors, similar to a photocathode. A quantum of light must have

---

[94]Rose, A. (1977) Vision: Human vs Electronic, *in* Barlow, H. & Fatt, P. Vertebrate Photoreception, NY Academic Press, pp 1-13

[95]Laughlin, S. (1981) in Autrum, et. al. ed. Handbook of Sensory Physiology NY: Springer-Verlag, pp 215-221

at least a specific energy to cause excitation of a single electron. The energy bandgap of the chromophores is relatively large, typically 2.0 electron-volts. There is no resistive element in parallel with the chromophore nor is their any current flow through the chromophore in the absence of excitation. Like photomultipliers, the chromophores act as nearly noise free detectors. Like photomultipliers, it is the first amplifier stage following the detector that usually determines the noise performance of the system. In the case of vision, the first amplifier operates without any shunt impedance between the input terminal and ground. The amplifier also exhibits a energy bandgap of nearly 2.0 volts. Hence, this stage also operates as a nearly noise free circuit[96]. There is no thermal or photoconductive noise associated with the retinal portion of the visual system.

It is also important to consider the impact of amplification. Usually, amplification reduces the effect of noise generated in circuits subsequent to the amplification, i. e. the primary sources of noise are usually found prior to the first significant stage of amplification. This is the case in vision, particularly under scotopic, mesotopic and low photopic conditions.

Because of the above conditions, the noise performance of the visual system is dominated by the noise associated with the photons in the excitation, except under the lowest light level conditions where it will be shown to be dominated by the noise associated with the stellate (decoding) cells of the brain. These cells employ a resistive element in combination with a capacitance to decode the pulse signals (action potentials) transmitted from the retina.

The high degree of negative internal feedback associated with the adaptation amplifiers introduces another condition into the visual system. This feedback has the effect of blocking the DC component of the signal but passing the AC component due to illumination. As a result, the noise (AC) amplitude in the circuits following the adaptation amplifiers is directly related to the illumination (DC) level before these amplifiers. The AC noise level at the pedicle is a function of the square root of the illumination (except in the case of the long wavelength channel as discussed later).

Finally, the simplest situation occurs if only the root-mean-square (RMS) amplitude of the noise at a given location is considered. Later, both the temporal and spectral frequency aspects of the noise can be considered.

11.7.5.1.1 The “RMS noise” model

Figure 11.7.3-3 shows a model based on that figure and the above conditions. The circuitry added on the left consists of a signal source and a noise source for each spectral channel. In each case, the noise source is a random source. The intensity of the noise source is proportional to the square root of the intensity of the signal source. In all cases, except for the long wavelength channel, the transduction process is linear and the signal and noise levels at the pedicles of the photoreceptor cells are linear representations of the input conditions. The square law nature of the long wavelength transduction process leads to a different relationship in this channel. The noise source added at the output of each decoder circuit of stage 3 is related to the integrating sub-circuit associated with the stellate cell of that circuit. The noise performance of the feedback circuit at the bottom of the figure is not addressed in this figure. See **Section 7.3** for an analysis of this circuit.

For all but the Long wavelength channel, the signal source i[xxx show signal noise signal level as a function of intensity in each channel.]

This model properly describes the noise component of any signal as it appears at the pedicles of the photoreceptor cells under photopic conditions. However, there are two other limiting elements in the visual system that become important at scotopic illumination levels. In the case of the luminance channel, the ganglion cells of Stage 3 may exhibit a threshold that prevents transmission of very small signal voltages. Second, recall that at the lowest excitation level the noise performance of the stellate cells of the brain ultimately becomes the limiting element in the visual system.

[96]The first amplifier of the photoreceptor cell can be considered to operate as a Class C amplifier instead of the Class A amplifier found in most man-made first amplifiers.

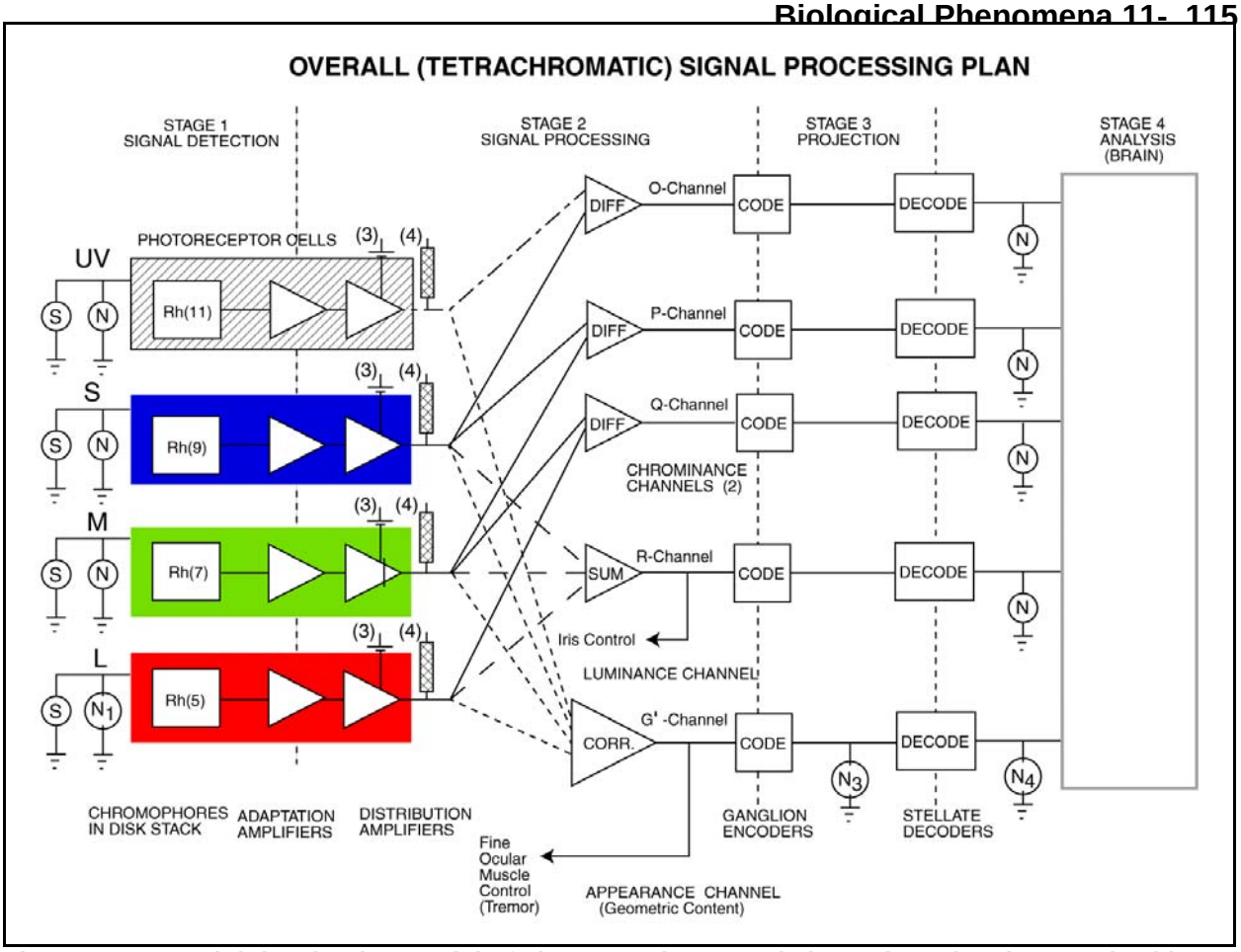

Figure 11.7.3-3 High level noise model of vison. Each spectral channel can be characterized by an external source of excitation and an external source of noise. The noise within the individual signaling channels is a function of the signal processing algorithms. Xxx The noise performance of the feedback loop involving fine ocular muscle control is not developed in this figure.

Copenhagen, Ashmore & Schnapf have provided measured noise data on the *in-vitro* turtle eye, *Chelydra serpentina*[97]. They made measurements in the dark, using 650 nm light and using "bright light." They did not provide a comprehensive model on which their work was based. Although their analyses are not compatible with the model of this work, their data can be interpreted using this model. Their figures 2 & 3 are subject to a totally different interpretation if the variable gain of the adaptation amplifier is accounted for.

11.7.5.1.2 The band-limited noise model

Additional information can be obtained if the above noise model is expanded to include the temporal frequency limiting mechanisms within the visual system. There are three primary frequency limiting factors in the visual system:

+ The intensity controlled second time constant of the photoexcitation/de-excitation process.

+ The time constant of the output circuit of the adaptation amplifier.

+ The poorly defined time constant limiting the pass band of the visual channels orthodromic to the adaptation amplifiers.

---

[97]Copenhagen, D. Ashmore, J. & Schnapf, J. (1983) Kinetics of synaptic transmission from photoreceptors to horizontal and bipolar cells in turtle retina. Vision Res. vol. 23, pp 363-369

The fact that the photoexcitation/de-excitation process involves an intensity dependent time constant is the reason that many authors have suggested that the "rods" react more slowly than the "cones." The proper interpretation is that all photoreceptor cells respond more slowly to low intensity than to high intensity light. This time constant is typically 0.525 seconds and changes inversely with the light intensity. As it is reduced, the apparent bandwidth of the process is increased.

The time constant of the adaptation amplifier output circuit is typically 3 seconds.

The bandwidth of all of the subsequent circuits of vision combined is typically given as between 30 and 100 Hz. The bandwidth of individual circuits may vary significantly from these values. This is particularly true for the circuits of the foveola that may individually reach 200 Hz at their half-amplitude point.

Copenhagen, Ashmore & Schnapf have shown the results of this cascade of filter characteristics in their Fig 4 & 5 for the turtle. Note how they used different normalization values in the two figures and introduced considerable smoothing. These distort the interpretation of the data. The use of difference spectra without acknowledging the change in adaptation amplifier gain is also awkward. However, the graphs of measured data do show three distinct regions. The first is falling with frequency as a function of $1/S^2$, one is falling as a function of $1/S^4$ and one falling at a higher but unquantifiable rate. This is the expected situation in any cascade of electronic filters. Because of the normalization used, little more can be said without redrawing the figures.

11.7.5.1.3 Noise and the signal processing stage

In the above discussion, the noise level was defined at the pedicles of the photoreceptors. Two points are important when this noise is transferred to the neurons of the signal processing stage. First, the amplitude of the signal and noise transferred to the individual signal processing neuron depends on the impedance of the synapse relative to the input impedance of the neuron. There is no amplification by the synapse. The second point concerns whether the signals are summed or differenced. While the signals may be subtracted from each other in horizontal cells, it is important to note that the noise is stochastic and that from different photoreceptors is uncorrelated. Therefore, it is always summed in quadrature, regardless of whether the signals are subtracted. This situation places a penalty on the differencing paths. They suffer from a higher relative noise level than does the luminance channel. On the other hand, the correlation process performed in the geometric channel of the POS is able to enhance the desired signal relative to the noise level. However, correlation is a much more complex mathematical operation than addition or subtraction.

11.7.5.1.4 Noise at the dendritic input to the PC's of the outer segment EMPTY

Based on the model of adaptation developed in this work, **Section 7.4.2**, the noise spectrum is not directly related to the stimulating light. A very careful mathematical model of the noise at the dendrites of the inner segment of the PC is required.

11.7.5.2 The spectral and noise experiments of Fyhrquist et al.

Fyhrquist, et. al. have recently provided data on the noise performance of the photoreceptors of the European toad, *Bufo bufo*[98]. They didn't rely upon any stated model but did employ the technique of Baylor, Lamb, & Yau (1979a) as modified by Fain, Lamb, Matthews, & Murphy (1989). IN this technique, the stimulating light was applied perpendicular to the axis of the outer segment. The resulting absorption spectrum is that of the omnidirectional portion of the retinines, rather than the more important operational spectrum associated with axial stimulation. Their recordings show a clearly narrow spectral band (110 nm wide) spectrum with the peak at 502 nm which is not suggestive of the broad spectrum expected of a "rod."

Their temporal recordings of individual photoreceptors in total darkness are clearly not due to white noise. They appear to exhibit considerable popcorn type (1/f type) noise with amplitudes of 1 to a few pA. Their uncalibrated dim flashes generated single pulses of peak amplitudes of about 18 pA. They attribute their noise pulses to "dark" isomerizations. "we recorded a value of about 0.020 sec -1 for *B. marinus* by visually counting presumed 'dark' events." Their data is not useful in the absence of a schematic model.

---

[98]Fyhrquist, N. Govardovskii, V. Leibrock, C. & Reuter, T. (1998) Rod pigment and rod noise in the European toad, *Bufo bufo*. *Vision Res*. vol. 38, pp 483-486

11.8 Comparison with other models

The literature was searched for other comprehensive models of the visual system (of any animal) and the results were disappointing. Only models of limited scope were found. Less comprehensive models necessarily fall into the category of "floating models" unless they carefully indicate ***all*** of the parameters affecting the input to that model and ***all*** of the output parameters provided by that model *and required to interface with or complete an overall model*. It is virtually impossible for an author working with a model of limited scope to meet these conditions.

The literature search was focused on signaling models that conform to a morphological model of the visual process. It is only of limited value to define a mathematical, or a digital computer, model that provides answers that are equivalent to measured values within a given statistical tolerance. By using enough degrees of freedom in an equation, any desired waveform can be produced; however, the resultant equation will have no relation to the real world problem.

There is a large collection of floating models in the psychophysical literature that have resulted from individual exploratory investigations. These are annotated and discussed in **Section 11.8.5.4**.

Figure 11.8.1-1 presents a summary block diagram from this work at several levels of detail along with a model by DeValois & DeValois discussed below. The top of this chart provides a series of labels found generally in the literature describing the various morphological regions of the visual system.

Note should be taken of the (s) added to some of the common names for the structural elements of the retina. In many species, these layers consist of a number of sub-layers. The number of sublayers in a given layer, particularly in Stage 2, may give some information about the functional capability of the visual system in a particular species. In human, the inner nuclear layer consists of at least three distinct sub layers. The most distal consists of the horizontal cells of the 1st lateral matrix. The most proximal layer consists of the amercine cells of the 2nd lateral matrix. In between these two lies the bipolar cell layer associated primarily with the luminance matrix.

Sterling reviews the morphological situation in some detail using nomenclature that is considered archaic in this work[99]. In Polyak's retinal cross-section of the human retina, the outer plexiform layer is clearly subdivided in two parts. Between them is a row of synapses leading to the cells of the 1st lateral matrix.

Below the top labels in the figure is a series of descriptions of the visual system that can be correlated to the above labels. These labels are segregated according to the descriptions shown on the right. The process descriptors are in general agreement with the top labels. Similarly, the histological descriptors are in general agreement with the top labels. This work chose to subdivide the visual process into four distinct first level stages and then subdivided these stages into distinct sub-stages. The evolution into sub-stages has proceeded further than the work of DeValois & DeValois. The use of four primary stages in this figure instead of three is to provide emphasis to the functionally distinct encoding process of the signal projection process. In this context, the fourth stage could also be divided into two stages to emphasize the decoding process prior to the additional signal manipulation processes found in the cortex. Whether the projection signals are decoded back to analog signals in the LGN before re-encoding for further transmission, will not be addressed in this work. It is likely that the projection signals received at the Pretectum are decoded.

11.8.1 Morphological descriptions

The most highly annotated morphological description of the visual process probably consists of the cross section of the human retina prepared by Boycott and Dowling[100] and the mapping of those signal paths through to the cerebrum by Hubel (except for his omission of the Pretectal region and its relationship to both the foveola and the Superior Colliculus)[101]. Both of these presentations concentrate on the human eye. This work is in agreement with both of these works but goes farther and would suggest expansions in notation in some areas that would support these extensions of the visual system. To this end, a more highly annotated version of the cross-section of the retina has been prepared and the maps of Hubel and others have been extended.

---

[99]Sterling, P. (1983) Microcircuitry of the cat retina. Ann. Rev. Neurosci. vol. 6, pp. 149-185
[100]Boycott, B. & Dowling, J. (1969) Organization of the primate retina: light microscopy. Phil. Trans. (Biol.) Vol. 255, pp. 109-184
[101]Hubel, D. (1988) Eye, brain and vision. NY: W. H. Freeman & Co. Pg. 38, 60 & Chap. 5

It is suggested that there is an asymmetry in the Boycott and Dowling figure that can be corrected and will lead to better understanding. There is a unique process occurring in the middle of their inner plexiform layer that deserves more emphasis. By annotating this zone as the inner interconnect layer and renaming the portion of the inner plexiform layer nearest the photoreceptor cells, a situation parallel to that found in the region between the inner nuclear layer and the outer nuclear layer would result. Based on the ganglion cell layer, the designations would

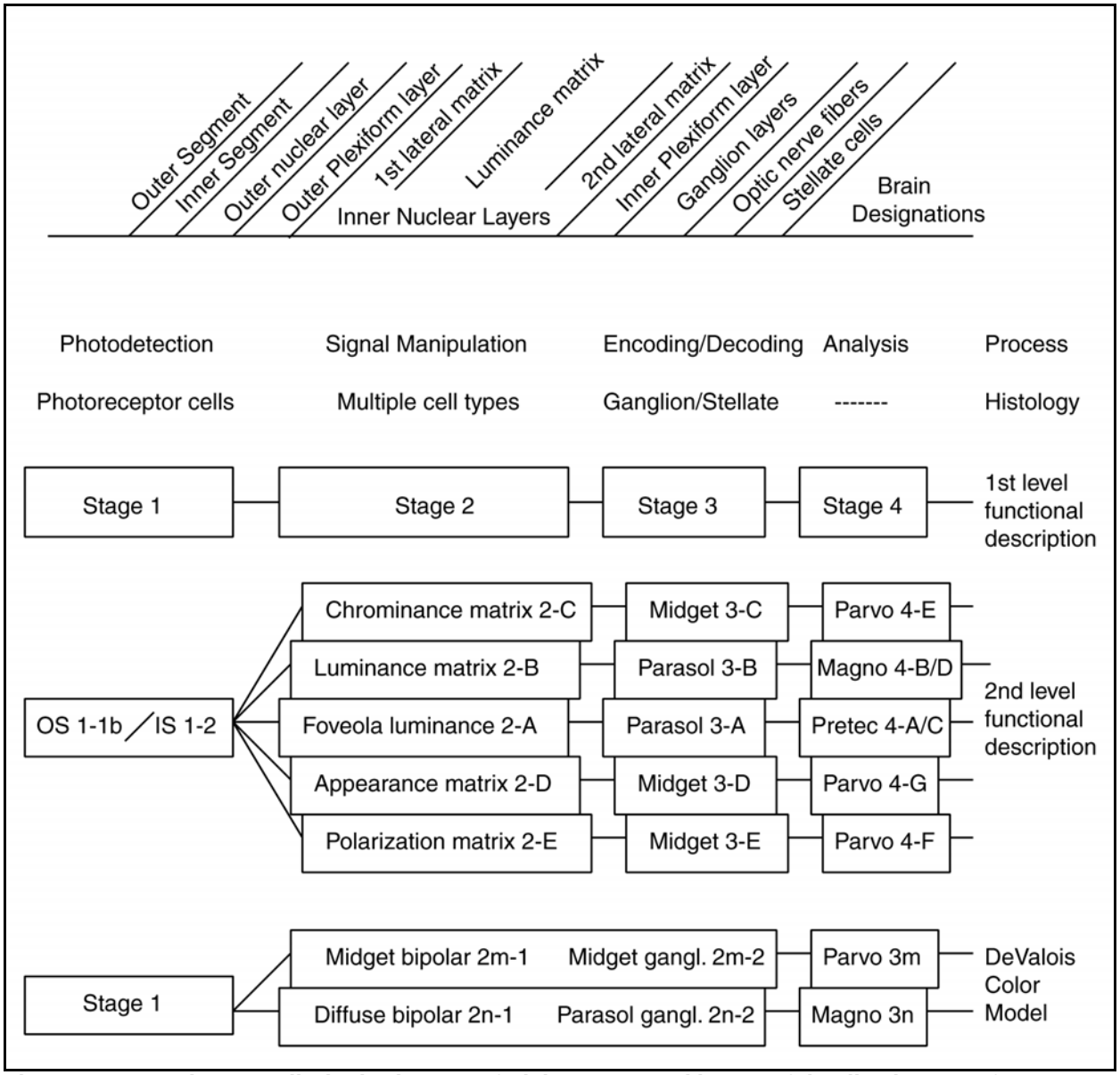

Figure 11.8.1-1 The Overall Block Diagram of Vision at several layers of detail. There are four stages in the afferent portion of the complete visual system. Following stage 1 (signal sensing), a large number of potential parallel pathways are introduced. The human visual system does not employ all of these pathways. Stage 3, (signal projection) also introduces several sophisticated encoding techniques to optimize the signal projection task. Stage 4 (signal manipulation in the brain) is only shown conceptually. The 1996 model of DeValois & DeValois is shown at the bottom of the figure. Their model did not address the signal projection stage of vision and treats the signal manipulation stage in the brina as their stage 3. See text.

become the inner fiber layer, the inner interconnect layer, the inner plexiform layer and the inner nuclear layer. The corresponding layers, beginning with the inner nuclear layer, would be designated the outer plexiform layer, the outer interconnect layer, the outer fiber layer and the outer nuclear layer. Whereas the pedicels and spherules are physically prominent in the newly named outer interconnect layer, the neuritic terminals are far more numerous and deserve equal billing from a functional perspective. Functionally, this layer is actually dominated by the first layer of synapses. The second layer of synapses dominates the inner interconnect layer.

The inner nuclear layer can be subdivided into three sublayers that each has a distinct functional significance, albeit a less than distinct boundary because of topographical considerations. The first sublayer is associated with lateral (horizontal) cells involved in the formation of the chromatic signal channels. The second sublayer is involved with the bipolar cells creating the luminance signal channels. The third sublayer is dedicated to the lateral (amercine) cells associated with the appearance channels. It is likely that this sublayer is also involved in forming the polarization signal channels in animals capable of processing this information. This last sublayer, and the inner interconnect layer, is less well developed in humans than in many other animals.

The direct connection between the photoreceptors of the foveola and the Pretectum and Super Colliculus play a major role in the human visual process. They may play an even more important role in many of the great hunters, the hawks, great cats and possibly sharks. These areas deserve more attention in the research literature. To recognize the functional significance of these areas, it is suggested they be added to Hubel's figure on page 60. The resulting set of visual pathways would be compatible with the results of Yarbus (shown on page 80 of Hubel) and others who have studied the rapid movement of the eyes in the process of interpreting the finer elements of the scene. They would also give more meaning to the physiological tremor and minor saccades of the visual process.

11.8.2 Sub-stages of the Overall Block Diagram

The Overall Block Diagram has been subdivided to clearly define the most obvious, but not necessarily all of the signaling channels of human vision. These channels can be grouped according to the type of signals they transmit.

11.8.2.1 Monophasic signaling channels

Channel A represents probably the most important channel within Stages 2, 3 & 4. of this diagram. It is the luminance channel associated with only the foveola and characterized by individual neural paths from either individual photoreceptor cells or individual outputs of the luminance matrix to the Pretectum. This channel is incorporated into the feedback circuit used to achieve perception of fine detail within the image of object space. These subchannels transmit monophasic signals related to the instantaneous signal amplitude at the voltage nodes of the photoreceptor cells. They employ parasol type ganglion cells.

Channel B is also a luminance-oriented subchannel. It conveys luminance signals to the LGN from all areas of the retina. Also transmitting monophasic information from the luminance matrix, it employs parasol type ganglion cells. These signals go to the magnocellular section of the LGN. They provide a "vector map" of the luminance associated with the instantaneous field of the visual system. This map is merged with the similar map from the other eye and then passed to the visual cortex via the optic radiation.

Only stages 2, 3 & 4 are shown in this figure. There are additional definable stages within the cortex. A failure in subchannel A does not lead to "blindness" although it does lead to an inability to perceive the details of an image required for recognition of a face, etc. Conversely, a failure in subchannel B may lead to "blindness" accompanied by a curious ability to avoid running into objects in the visual field.

11.8.2.2 Biphasic signaling channels

Channels C, D & E of the figure can be considered overlays on the fundamental visual process. They provide enhancements to the visual capability of the animal. These channels transmit biphase signals produced by differencing circuits in the retina. These differencing circuits may involve lateral or bipolar (in line) neurons. The channels employ midget type ganglion cells and deliver their signals initially to the LGN.

Channel E delivers a signal that represents the polarization of light received by specific areas of the retina, frequently defined by individual retinula.

Channel C is itself divided into sub-sub-channels in order to deliver chrominance information to the LGN. Although there are only two sub-sub-channels in humans and other primates, there are in general three sub-sub-channels in

tetrachromats.

Channel D provides appearance information extracted at the retinal level. This channel is not well defined in humans or other animals. It may consist of a series of sub-sub-channels. Studies have indicated significant development of these channels in some aggressive hunters such as the feline family.

11.8.2.3 Special foveola channels

Within the very core of the fovea, the foveola, there appear to be photoreceptors with direct connection to the parasol cells that transmit their signals to the Pretectum and then on to Area 7of the cortex (bypassing Area 17 completely). These signals play a critical role in the servomechanism controlling the eye muscles as well as providing very fine detail about the area of the object plane at the point of fixation. This information is transmitted via the Pulvinar pathway directly to the higher centers of the visual cortex concerned with detailed recognition, such as the recognition of writing symbols.

11.8.3 Other overall block diagrams

The development of other Overall Block Diagrams of the visual process has been hampered by less than adequate understanding of the overall processes involved in the total visual process. Most investigators have only addressed a portion of the overall process, usually overlooking the physiological optics and the motor portion of the system.

To explore all of the previous models of the visual process in the literature is beyond the scope of this work. Most of the widely discussed models have been prepared by psychophysicists. To avoid excessive controversy, many of these models are discussed in detail in the appendices. The intent below is to merely highlight some of their principle assumptions or proposals. All of the following models were less comprehensive than required but are of interest:

1. A Standard Model[102] of the visual process has permeated the Vision Community for an extended period, re-emerging periodically. It remains highly conceptual.

2. Guth[103] published a paper entitled “Model for color vision and light adaptation” which he defined as a multi-zone color model. The model is largely intuitive and does not present or relate to any physiological model of vision. He assumed three, apparently linear photoreceptors and assigned a set of coefficients to a matrix that manipulated the photoreceptor signals into six signal paths. The model introduces noise as a limiting factor prior to the gain control mechanism that precedes any signal matrixing. The model had nonlinear gain control, two opponent-color processing stages and neural compression late in the visual pathway. He then proceeded to elaborate his computer program until it gave a reasonable match to a selected set of measured data. It provided little insight into the operation of the visual process; however, it does provide a good tabulation of earlier abstract models.

3. Graham & Hood[104] provided a model based on their psychophysical background which attempts to properly predict properties related to light adaptation. They limit their presentation to the region of Photopic foveal vision. However, they limit their model to a single spectral channel, avoid the spatial frequency domain and only provide for a single temporal frequency channel. Based on these ground rules, they attempt to provide a comprehensive merger of similar models they describe as based on traditional aperiodic stimuli and on traditional periodic stimuli. At the end, they point out their merged model is deterministic and contains no probabilistic processes. They review the many open questions before their work could be correlated with actual physiological conditions and morphological substrates.

4. Van Hateren[105] provided a theory in 1993 involving a single channel deterministic model that claimed to predict the results of many spatiotemporal contrast phenomena but he did not indicate how it related to the physical processes of vision in animals

---

[102]See Introduction in DeValois & DeValois below
[103]Guth, S. (1991) Model for color vision and light adaptation. J. Opt. Soc. Am. A, vol. 8, no. 6, pp. 976-993
[104]Graham, N. & Hood, D. (1992) Modeling the dynamics of light adaptation: the merging of two traditions. Vision Res. vol. 32, no. 7, pp. 1373-1393
[105]Van Hateren, J. (1993) Spatiotemporal contrast sensitivity of early vision. Vision Res. vol. 33, no. 2, pp. 257-267

5. DeValois & DeValois[106] in 1993 reiterated their 3 stage approach to color vision (only) and gave some background about a wide range of earlier one and two stage models. All of these models have been labeled the Standard Model by their respective authors in their respective times. It appears DeValois & DeValois associated their 1st stage with the overall process from photodetection to pick-off of signals from the "straight through" path. The second stage consisted of the chromatic signal path only and only the first stage of differencing (i.e., "three pairs of cone opponent and 1 pair of cone non-opponent signals"). Their third stage occurred in the cortex and was essentially a transformation of axes in color space to "create a pure RG & YB color axes and separate the luminance information from the color information." Their model appears at the bottom of [**Figure 11.8.1-1**] where the subdivisions of m and n were added by this author for convenience based on the text provided by the above authors. At the end of their paper, they did mention a 4th stage related to color-selective complex cells in the brain.

They make an interesting gratuitous comment to justify their approach claiming limited genetic information forces the visual process to follow their description. Since the equation relating the genetic code to the visual structure is unknown, there is no way to know if the genetic information is in any way limited. The whole structure may be defined by a simple equation containing a two-stage fractal code of only a few terms.

Their model remains highly conceptual, although it refers to its foundation in anatomy, and appears to be a linear one. It assigns an arbitrary ratio between the number of L, M, and S photoreceptors and an arbitrary spectrum to each photoreceptor type. It goes on to specify an arbitrary spatial arrangement of photoreceptors and that each photoreceptor is contacted by precisely four bipolars.

They recognize the difficulties of treating the second stage, signal processing in the retina, due to the presence of simultaneous and combined chromatic and spatial encoding which they describe as confounding the signals. They also discuss the related problem of "unconfounding" these combined signals after they reach the brain. Although they do not differentiate their second stage into substages, they do make some remarks leading in that direction. They indicate that midget bipolars contact midget ganglion cells which in turn feed into the parvocellular geniculate layers, while diffuse bipolars contact parasol ganglion cells which in turn feed into the magnocellular LGN layers. Wilson also notes the confusion over using the abbreviations M & P in this context[107].

6. Livingstone & Hubel[108] have presented a number of papers suggesting models of the signal processing in primate vision based on their work with the brain itself (my stage 4). Interestingly, their discussion implies that a majority of the chromatic signal differencing activity occurs in the parvocellular layers of the geniculate region of the brain. This is in opposition to most of the literature which assigns this function to the Inner Nuclear Layers of the retina (my stage 2). They also make this implication without discussing the encoding of the data from the retina (in my stage 3). This encoding into a relatively small number of neurons in the Optic Nerve seems to assure that most of the color differencing is done before the signal is transmitted over the Optic Nerve. It would appear a more explicit paragraph in their paper would indicate that much of the chromatic signal differencing occurs in the INL, including the convergence of many photoreceptor cell outputs before this differencing, and the resulting signals are processed in the parvocellular region at a more perceptual level. They do follow the same physical description as DeValois & DeValois with regard to two types of ganglion cells connecting to separate regions of the brain; their "A" cell is comparable with DeValois's parasol ganglion cell and their "B" cell is comparable with DeValois's midget ganglion cell. They relate the parvocellular region with color information processing and the magnocellular region with luminance information processing. However, they are quick to point out that there is conflict in the literature about the "X" and "Y" labels when applied to different species by different investigators.

It is unfortunate that neither the work of DeValois & DeValois nor of Livingstone & Hubel explicitly address the process of Signal Encoding/Decoding in Vision. This area is critical to the understanding of the vision process and is denoted by Stage 3 in this work. None of the above models addresses the subject of the pathway between the foveola, Pretectum, and area 7 either. Hopefully, the more elaborate complete set of signaling paths defined in this work will help overcome these shortcomings.

7. Dacey has provided a schematic of the visual pathways with apparent disregard for the cerebral cortex. The

---

[106]DeValois, R. & De Valois, K. (1993) A multistage color model Vision Res. vol. 33, no. 8, pp. 1053-1065
[107]Wilson, H. (1997) A neural model of foveal light adaptation and afterimage formation. *Visual Neurosci*. vol. 14, pp 403-423 (pg 407)
[108]Livingstone, M. & Hubel, D. (1988) Segregation of form, color, movement, and depth: anatomy, physiology, and perception. Science vol. 240, pp. 740-749

schematic has several unique features. First, within its scope, it is in very good agreement with this work[109]. Second, it is very recent. Third, it omits the cerebral cortex from the signaling pathways related to perception. **Figure 11.8.1-2** shows Dacey's figure without its extensive caption. The caption uses a long "run-on" sentence to relate many structures in a very brief and incomplete manner. There are several major point related to this figure. His pregeniculate nucleus may be equivalent to the perigeniculate nucleus (PGN) of this work. Alternately, his pretectal area (PT) may be equivalent to the PGN of this work. He does not provide sufficient detail to resolve this point. The figure needs to be expanded to incorporate a variety of major ganglion (stage 3) neural paths identified in this work. It does not show the extensive pathways (Meyer's loops) extending from the LGN to the visual cortex (areas 17) and back to the thalamic reticular nucleus (TRN).. It also omits the signal paths from the nuclei of the accessory (precision) optic system to the muscles controlling rotation of the ocular globes, i.e., the oculo-motor system does not end at the MTN. Neither does he identify the thalamus and the thalamic reticular nucleus enclosing much of the thalamus, the LGN/PGN and the inferior pulvinar (IPm). The description of the lateral geniculate nucleus (LGN) in the inset is the same as shown in **Section xxx**. However, it does not show the partial gray matter layers that can be defined as P7 and P8. The layers labeled koniocellular are white matter forming the ganglion neural pathways to area 17 via Meyer's loops.

By illustrating the neural pathways back to the ocular muscles, the complete closed loop servomechanism of the precision optical system (POS) would be defined. By illustrating the ganglion neuron pathways to area 17 and back, one of the major information extraction functions of the visual system would be highlighted. By illustrating the ganglion neural pathways from the thalamic reticular nucleus to the parietal lobe, the place of the saliency map within the visual system would be highlighted.

Dacey discusses the midget and parasol ganglion cell pathways and their connection to the parvocellular and magnocellular pathways of the optic nerve in considerable detail. He identifies the unique "private line" circuit consisting of one photoreceptor, one bipolar neuron, one midget ganglion neuron and a single axon extending into the optic nerve. His subsequent discussion is less specific and does not include any diagrams of the connections of the optic nerve to the thalamic and pulvinaric elements of the diencephalon. He does provide a list of seventeen histologically identifiable ganglion neuron types in the macaque retina. The role of many of these are listed as unknown. He does identify a variety of bistratified ganglion neurons and a variety of monostratified ganglion neurons. If accurately identified, the first are capable of forming analog voltage or current differences while the second are incapable of performing differencing. While the degree of arborization of each strata suggests the complexity of the summation function associated with each strata, the shapes of the arborization are not functionally relevant. His page 282 provides some statistics. He notes that "approximately 90% of the ganglion cells project to the LGN." He also notes that"most if not all ganglion cell types in primate project to the LGN and/or the SC." In the context of this work, the ganglion neurons pointing toward the SC are actually terminating in the PGN (his pretectal area) along the brachia of the superior colliculus. Thus, from a functional perspective, all of the 17 histologically identified ganglion neurons proceed directly to the LGN or PGN along the optic nerve. Over 90% of these (closer to 99%) of these proceed to the LGN with the remainder proceeding to the PGN. He finally notes that a majority of his histologically defined ganglion cell types are present at very low densities (1% to 5%) relative to the midget and parasol families.

In the context of this work, the midget ganglion neurons are focused on the chrominance (parvocellular) channels of the optic nerve and the parasol ganglion neurons are focused on the luminance (magnocellular) channels of the optic nerve.

The text of the Dacey paper is worth additional study in relation to other sections of this work.

---

[109]Dacey, D. (2004) Origins of perception: retinal ganglion cell diversity and the creation of parallel visual pathways *In* Gazzaniga, M. *ed*. The New Cognitive Neurosciences, **2nd** Ed. Cambridge, MA: MIT Press Chap 20

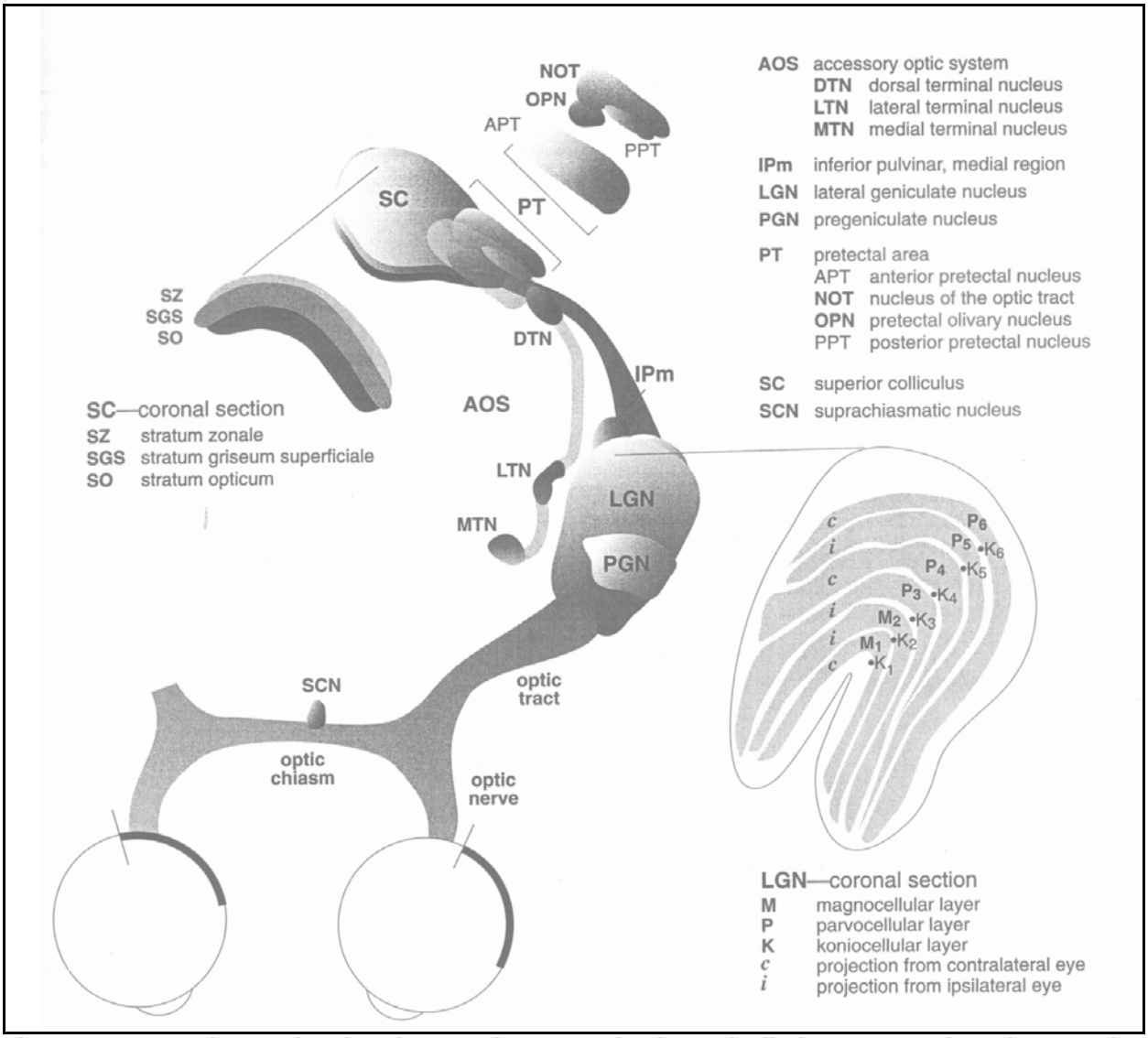

Figure 11.8.1-2 Primate visual pathways shown semi-schematically by Dacey. The primate retina projects to at least 15 histologically distinct brainstem sites. Either the pregeniculate nucleus, or the pretectal area, appears to be the perigeniculate nucleus (PGN) of this work. See text. See Dacey for full discussion of the figure. From Dacey, 2004.

11.8.4 Other circuit diagrams

Although there is a growing awareness of both the importance of the liquid crystalline state of matter and the importance of the semiconducting properties of cells in the biological literature, this awareness has not been reflected in the appropriate circuit diagrams yet. The circuit diagrams in the literature typically do not adequately circumscribe the illumination conditions to which they apply. They also fail to recognize the diode as a critical element in the operation of the visual process. As a result, every circuit diagram in the literature employs a resistor-based architecture. Frequently the resistor is shown as a variable without adequate explanation concerning the mechanism controlling the variable. Hence, all of these circuit diagrams must be described as analogs of some portion of the visual process. They are not, and few authors claim they are, actual circuit diagrams describing the in-vivo visual process. A similar situation is found in the area of computer simulations based on inadequate

understanding of the actual circuits involved. These computer simulations usually employ a sufficient number of independent variable that any desired response can be obtained.

Recently, Usui and associates[110,111] have presented two ionic current models of neural cells based on a fusion of the chemical kinetic analysis of transduction by Torre, et. al. and the excitable membrane model of the neuron from Hodgkin, Huxley & Katz. The results are computer simulation models of simple neurons that are not related in any way to the topology or electrolytic conditions within the neuron.

It is not possible to properly describe the transmission line characteristics of an axon using a network made up only of resistors and capacitors. To transmit signals with a ratio of a pulse-to-pulse interval divided by a pulse rise time on the order of 1000:1, a broadband circuit is required for signal transmission. As a minimum, the serial element in the T or Π network used to represent the axon must contain an inductance. Any coaxial transmission element exhibits both an inductance and capacitance.

It is claimed that the Overall Circuit Diagram presented in this work is indicative of the actual fundamental processes occurring in the visual process of animals. Although there is always a need for more definition, particularly at the component level, this overall diagram is adequate to construct a very sophisticated working model of the visual process in virtually any animal. It is also adequate to trace the signal paths in a living animal and to design more sophisticated test equipment with which to trace these signal paths.

11.8.5 Other conceptual descriptions

11.8.5.1 Detection theories

Little substantive advancement has been made in the theory of the photodetection (both transduction and translation prior to amplification) has occurred since the work of Wald and his associates in the 1930's through 1950's. Their work culminated in the isomerization theory of photodetection. It was developed prior to the development of liquid crystalline theory, solid state physics, excitation of molecules (resulting in the development of the laser) and many other relevant technologies. It is discussed extensively in **Chapter 5** of this work. The isomerization theory is not supported by this work.

11.8.5.2 Chromatic theories

11.8.5.2.1 Trilateral color theories

Up until now, the most widely accepted color theory, based on Young and Helmholtz, has relied on a tristimulus foundation. As shown in **Chapter 17.3**, the New Chromaticity Diagram for Research provides a theoretical foundation for the tristimulus theory which is not pretty. The visual color space in humans is not trilateral. It is based on two color difference channels that are orthogonal in quadrants 1, 2, & 4 relative to the achromatic point. In quadrant 3, the axes of these channels are superimposed. Although it may be possible to calculate loci on the new diagram corresponding to the old x-y space or the newer u,v space, they are irrelevant. The R, G, & B spectra used to compute the old chromaticity diagrams do not adequately characterize the peak absorption coefficients of the actual chromophores.

It also points out the limitations inherent in the opponent theory. While the theory defines the Hering axes as orthogonal, the underlying processes are not. The fundamental processes related to color vision are only orthogonal in quadrants 1, 2, & 4 relative to the achromatic point. The resulting fundamental axes exhibit a distinct bend in this chromatic space.

Tilton attempted to provide an alternate foundation that led to different conclusions than those of the generally

---

[110]Kamiyama, Y. Ogura, T. & Usui, S. (1996) Ionic current model of the vertebrate rod photoreceptor. Vision Res. vol. 36, pp. 4059-4068

[111]Usui, S Ishihara, A. Kamiyama, Y. & Ishii, H. (1996) Ionic current model of bipolar cells in the lower vertebrate retina. Vision Res. vol. 36, pp. 4069-4076

accepted concepts of the trichromatic theory[112]. His analysis was conceptually similar, and can be considered a precursor, to that presented here. However, it was prepared based on the idea that the scotopic and photopic luminosity functions were unique functions related to either the rods or the cones. However, he did not define the characteristics of an individual rod or cone. In fact, he proceeded to redefine the scotopic luminosity function in terms of the response of a unspecified group of rods (a rod domain) and the photopic luminosity function in terms of the response of an unspecified group of cones (a cone domain).

Tilton's analysis also involved the difference in the logarithms of the standard photopic and scotopic luminosity function over the entire spectrum. This analysis shows that these differences should be computed separately for the short and long wavelength portion of the spectrum.

11.8.5.2.2 Opponent color theories

Up until now, the opponent color theory of Hering has relied on a psychophysical foundation. As shown in **Chapter 17.3**, the New Chromaticity Diagram for Research provides a theoretical foundation for the Hering axes. It also points out the limitations inherent in the opponent theory. While the Hering axes are defined as orthogonal, the underlying processes are not. The fundamental processes related to color vision are only orthogonal in quadrants 1, 2, & 4 relative to the achromatic point. The resulting fundamental axes exhibit a distinct bend in this chromatic space.

11.8.5.3 The glutamate and Calcium cascades

Two entirely chemical approaches to explaining the generation of a potential (or a current) by the interaction of light with the photoreceptor cells are common in the literature of the last thirty years. These theories rely primarily upon the abundance of these materials in the vicinity of the observed activity. They are both based on the assumption that there is a plasma membrane surrounding the disks of the outer segment of the photoreceptor cell. Such a membrane has never been recorded at magnifications exceeding 50,000x. Conversely, there is excellent imagery showing that **such a plasma membrane is not present in a variety of chordates**. See **Section 4.3.5.** The presence of such a membrane would also cause significant problems in explaining the process of phagocytosis of the disks without any disturbance to the putative parent cell membrane, and the parent cell, enclosing them.

In the absence of any documentation of such a membrane and the availability of the much more integrated electrolytic approach presented in this work, these approaches are not supported here but are discussed briefly below. As noted at www.4colorvision.com., the author offers a nominal $1000 reward for any reproducible imagery of the presence of such a plasma membrane surrounding the disks of the outer segment.

11.8.5.3.1 The putative glutamate cascade

In recent years, there has been a flurry of activity attempting to describe the photodetection process from a strictly chemical perspective. This has been based on the ubiquitous nature of products of the glutamic acid family in the vicinity of neurons. The hypothesis has been presented that the glutamates participate in a cascade (?, avalanche) effect in order to obtain signal amplification in the area of the Outer Segments of the photoreceptor cells. No mathematical foundation or recognized chemical reaction has been proposed to show how a glutamate cascade actually creates an electrical signal related to the photon flux incident on the chromophores of vision. This work recognizes the ubiquity of the glutamate family[113] but assigns a different role to these chemicals.

In the English vernacular, the word glutamate describes a member of the glutamic acid family. In more precise chemical English, the word describes a salt of glutamic acid. This salt may be inorganic or organic. There is no specific chemical listed in the Merck Index (Vol. 12.1 1996) under the designation glutamate. In advanced texts in biochemistry, members of the glutamic acid family are found participating in a variety of metabolic cycles designed to provide energy (in chemical form) to the organism. The expression glutamate is frequently used to describe various intermediaries in these cycles that are related to glutamic acid.

It is not easy to follow the family relationships among organic materials of five or more carbons because of the great diversity of form possible. Glutamic acid can be described as α-amino glutaric acid. Glutaric acid is the double

---

[112]Tilton, H. (1977) Scotopic luminosity function and color-mixture data J.Opt. Soc. Am. vol. 67, no. 11 pp 1494-1501

[113]Sterling, P. (1983) Microcircuitry of the cat retina. Ann. Rev. Neurosci. vol. 6, pp 149-185

acid of valeric acid. Since glutamic acid is symmetrical, with a carboxyl group on each end, it can just as easily be described as γ-amino glutaric acid. Glutamic acid is one of the simplest amino acids showing a family relationship to the dicarboxylic acid family. Aspartic acid has one less methylene group and can be described as α-amino succinic acid.

This work assigns the glutamates to a different role, as intermediaries. They participate in an electrostenolytic process similar to energy pathways not confined to a (known) structural substrate. In this application, the electrostenolytic process is uniquely associated with a cell wall as a substrate. These electrostenolytic processes provide free electrons that participate in conventional electrical circuit operations. The consumption of material in the electrostenolytic process, ending in the production of (and/or possibly the consumption of) various glutamates, is directly related to the passage of electrons through the collector of the adaptation amplifiers. The axolemmas associated with the collector of these Activas are found in intimate contact with the IPM and in very close proximity to the chromophoric material of the Outer Segments. The glutamates are only indirectly associated with photodetection.

11.8.5.3.2 The Calcium cascade EMPTY

[xxx see pg 8 by Danielli in Weissmann & Claiborne as a more realistic model than that of Bezanilla and others on pag 40 of Yudilevich, et. al. 91.]

11.8.5.4 Performance based psychophysical models

There are a great variety of fragementary models in the literature proposed as a result of individual exploratory investigations, or a group of such investigations. Wilson has recently annotated a variety of these and provided an extensive bibliography[114]. His discussion (page 81) with regard to spatial filtering is particularly interesting when he defines the conflicts between the proposed psychophysical models and many proposed electrophysiological models. By recognizing there are no discrete spatial frequency **or** temporal frequency bandpass filters in the visual system, additional progress can be made in these fields. This is also true of the so-called line element models of psychophysics. Most of the experimental data can be interpreted more readily as a result of the scanning of the two-dimensional foveola due to tremor and the subsequent signal processing in a more abstract saliency space (see **Chapter 15**). The two-dimensional foveola, of some 23,000 photoreceptors, is able to perform integration of individual photoreceptor responses over a significant area that the line element models cannot address. Wilson addresses extensions of the line element approach into more complex abstract models (page 82). On page 83, Wilson discusses application of the Nyquist Limit with regard to the pixel pitch of an imager in object space. It is argued here that the visual system is a change detector and the appropriate application of the Nyquist Limit is with regard to the temporal frequency capability of the system.

11.8.5.4.1 Hyperacuity as a concept

The ramifications of hyperacuity has been discussed recently by Morgan[115]. Words become extremely important when discussing this subject because they are frequently used in different contexts when discussing the conceptual framework related to this concept. Historically, the subject has been discussed in the context of the retinas and photoreceptors as imaging devices. Therefore, the discussion usually relate to pixel sizes and pixel spacings. Morgan addresses this problem (page 87) without resolving it by using a reentrant definition of the term hyperacuity. His discussion and bibliography of a variety of competing fragmentary models of the concept leaves the question unresolved (page 110). All of the approaches rest on the concept of the retina and photoreceptors as imagers rather than a change detectors. As indicated in the previous paragraph, if the retina and photoreceptors are considered change detectors and the 23,000 independent signals from the photoreceptors of the foveola are processed primarily in saliency space as the first step in perception, an entirely different conceptual framework for the description of acuity results. Acuity is primarily determined by the change in signal amplitude resulting from a change in the position of an element in object space versus the edge of a photoreceptor cell in the retina. (Precise analysis requires the recognition of Snell's Law in this area.) If this change is above the threshold limit of the individual signal channel (but not necessarily above the perception threshold based on a given scene criteria), the

---

[114]Wilson, H. (1991) Psychophysical models of spatial vision and hyperactivity. *In* Regan, D. (1991) Op. Cit. Chap. 3.
[115]Morgan, M. (1991) Hyperacuity. *In* Regan, D. (1991) Op. Cit. Chap. 4.

signals from adjacent channels can be integrated in the initial perceptual process resulting in a greater effective acuity for specific patterns in object space–long thin lines, long parallel lines, tri-bar patterns etc. This increase in apparent acuity can be described as hyperacuity, but without the mystic overtones.

11.8.5.4.2 Theories of visual texture perception

Bergen has provided a review of the common conceptual framework of texture in vision[116]. He clearly differentiates visual texture perception from the analysis of surface texture as such. His statement that "we can perceive shapes defined only by texture, without the usual accompanying differences in color or brightness." is curious (underline added). He is apparently limiting this concept to relatively large areas compared to the projected size of individual photoreceptors **and** the scope of tremor. He offers no model of the visual process he is assuming. It is clear that if there is no signal in the P, Q and R channels of the model of this work, the subject perceives, interprets and cognates nothing via either the **awareness path** or the **analytical path**.

If one introduces the change detection concepts of the preceding paragraphs and recognizes that the visual system relies upon edges to determine recognizable features at the saliency level, the perception, or probably more appropriately the cognition of texture takes on a different conceptual framework. It is the higher level of cognition that allows the observer to note and quantify the variation in orientation of the crosses in the perimeter of Bergen's figure 5.1 while simultaneously quantifying the orientation of the L's in the center of the field. It should be noted that the L's in the center of the figure vary significantly in brightness as well as orientation at the resolution level of the perceptual process. The brightness changes are clearly well above the threshold level of the brightness signal channels associated with individual photoreceptors. This leads to perception on a multi-photoreceptor basis within the POS that easily recognizes the differences in orientation of the individual strokes and simple shapes present.

11.8.5.4.3 Theories of pattern recognition

As Bergen differentiated between the study of perceptual texture recognition and surface texture per se, it is important to distinguish between pattern recognition by machines and pattern recognition by biological systems. At the present time, the recognition of patterns by biological systems is far more advanced than in man-made machines. However, man is making great efforts to uncover the algorithms employed by biological systems. In general, these rules rely upon the immense parallel processing capabilities of both the midbrain (old brain) and the cortex (new brain).

Watt has presented a review of the entire field of man and machine pattern recognition as of 1991[117]. Mohn and Van Hof-van Duin have presented material on vision in infants that also bears on this overall subject[118]. Although presenting useful information, both of these must be considered archaic at the rate this field is moving. Neither presents a precise model of what they are trying to understand.

The Chapter by Valentine in Watt illustrates the problem with the lack of a precise model[119]. In attempting to determine how the human visual system recognizes a face, he concentrates on large areas of the face that are much larger than the individual scenes imaged on the foveola. By doing this, he does not address:

+ the fundamental two-dimensional signal manipulation capability of the **analytical signal path** of the human visual system, or

+ the sequential technique used by the midbrain for analyzing and perceiving primitive shapes within the duration of a gaze, or

+ the ability of the interpretive and cognitive engines of the midbrain and cortex to assemble these individual perceived shapes into a (abstract) saliency map that can be compared with similar maps in its memory.

The first three chapters in Watt provide a variety of potentially relevant techniques associated with scaling that may play a significant role in achieving the abstract saliency map associated with a given scene. However, some of the

---

[116]Bergen, J. (1991) Theories of visual texture perception. *In* Regan, D. (1991) Op. Cit. Chap. 5.
[117]Watt, R. (1991) Pattern Recognition by Man and Machine. Vol. 14 of Vision and Visual Dysfunction, Cronly-Dillon, J. *general ed*. Boca Raton, FL: CRC Press
[118]Mohn, G. & Van Hof-van Duin, J. (1991) *In* Regan, D. (1991) Op. Cit. Chap. 7.
[119]Valentine, T. (1991) Representation and process in face recognition. In Watt, R. (1991) Op. Cit. Chap 9

concepts discussed in the first chapter, particularly on page 2, suffer from a lack of appreciation of how the visual system works. His definition of pattern as only occurring in a cognitive psychophysical context is particularly difficult to rationalize. A railroad track has a pattern whether anyone is looking at it or not. Perception of that pattern may differ from the analysis of the geometric qualities of that pattern using a ruler and protractor just as Bergen suggests the perception of texture may differ from the innate texture of a surface (**Section 11.8.5.4.2**). The volume by Watt does not develop any models of the system that is being studied. This is a critical shortcoming in a system as complex and multi-pathed as human vision. As developed in **Sections 7.5** & **15.4**, the **awareness signal path** plays a major role in determining the areas of a scene to be analyzed further by the **analytical signal path**. It is the actions of signal processing elements associated with the **awareness path** that bound the signal manipulation effort required of the similar elements of the **analytical signal path** to a much smaller domain than envisioned by Watts's numbers. Most of the "data" in a scene imaged on the retina is never treated as "information." It is overlooked, first as not representing a threat (no rapid change in the signal intensity of individual scene elements, and then as not salient to the interests of the mind at that time. The latter priority can be overridden by the **volition signal path** controlled by the anterior lobe of the cortex.

The end of the first decade of the 21st Century has led to an entirely different understanding of pattern recognition within the neural system based on the concepts of biological vision developed within the machine vision community. The hardware implementations developed within that community based on what is known as a Restricted Boltzman Machine (RBM) provides an analog of the operation of the major stage 4 engines of the occipital lobe and the pulvinar. These developments are described in detail in **Sections 15.6.5.7** and **19.10.5**. They particularly affect our understanding of the operation of the occipital lobe within the LGN/occipital couple of the **awareness signal path** and the operation of the pulvinar within the PGN/pulvinar couple of the **analytical signal path**.

11.8.5.5 Potential transfer of machine vision concepts to biological vision

Since the 1980's, the machine vison community has been attempting to emulate biological vision for their purposes. A review of their work shows it is all based on simplistic models of the visual processes and mechanisms taken primarily from introductory textbooks on vision used in first or second year university courses. The material has relied upon the concept of the retina as an 2D integrating-type imaging sensor with at least an effective frame rate of about 25 frames per second. While developed into an elaborate set of computational models, these models show little similarity to the models of biological vision based on change detector-based sensing. They offer little to no insights relative to the visual modality. The differences between machine and biological vision are reviewed in **Section 15.9**.

TABLE OF CONTENTS 4/30/17

11 Introduction to modeling of biological phenomena 1
11.1 Background 1
11.1.1 Clarification of concepts 2
11.1.1.1 Models and Diagrams to aid understanding 2
11.1.1.1.1 Domains of signal manipulation 3
11.1.1.1.2 Bayesian (or other statistical) mathematics are not required in vision 4
11.1.1.2 Proper assignment of signal sources 4
11.1.1.2.1 Proper assignment of dependent variables 4
11.1.1.3 Adequately quantifying stimuli 4
11.1.1.4 The grounding plan 6
11.1.1.4.1 The detailed ground plan of the retina 6
11.1.1.4.2 The large scale ground plan of the visual system 8
11.1.1.5 Inappropriateness of Ohm's Law in vision 10
11.1.1.6 Inevitable results of the non-linearity of animal tissue 10
11.1.1.7 Availability of Complete Photoexcitation/De-excitation Equation 11
11.1.1.8 Obsolescence of the Michaelis Equation and the Logistic Equation 11
11.1.1.9 Filtering of the P/D Equation by proximal circuits 12
11.1.1.10 Elucidation of additional ERG waveforms 12
11.1.1.11 Exceptions to the Principle of Univariance 12
11.1.1.11.1 Failure of Univariance with respect to luminousity 13
11.1.1.11.2 Failure of Univariance with respect to chrominance 14
11.1.1.12 Emergence of the Biological Transistor 14
11.1.1.12.1 The Activas within the Outer Segment 15
11.1.1.13 Ramifications of the Biological Transistor 16
11.1.1.14 Rationalizing the confusion over polarity 16
11.1.1.15 Important consequences of the adaptation mechanism 17
11.1.2 Terminology and conventions 17
11.1.2.1 Responses 17
11.1.2.1.1 Mathematical manipulation of the responses 19
11.1.2.1 Polarity 20
11.1.2.1.1 Polarization of radiation 20
11.1.2.1.2 Polarization in morphology 20
11.1.2.1.3 Potential and Polarization 21
11.1.2.2 Electrical fields 24
11.1.2.3 Signal transfer between neurons 24
11.1.2.4 Signal control 25
11.1.2.5 Signal measurement (Reserved) 25
11.1.2.5.1 Voltage polarization (Reserved) 25
11.1.2.5.2 Signal isolation 25
11.1.2.5.4 Decoupling 27
11.1.2.5.5 Adaptation 27
11.1.2.6 Signal classification 27
11.1.2.6.1 "ON" and "OFF" in psychophysics 28
11.1.2.7 Resonance 28
11.1.2.8 Bandwidth and risetime 28
11.1.2.8.1 Spatial versus temporal frequency response 28
11.1.2.8.2 Individual circuits 28
11.1.2.8.3 Types of temporal frequency filters in vision 29
11.1.2.8.4 Scene information modulation by saccadic motion-TREMOR . . . 30
11.1.2.8.5 Cascaded Circuits 31
11.1.2.9 Delay 31
11.1.2.9.1 Transport delay as a tradeoff 32
11.1.2.10 Feedback 32
11.1.2.11 Charge transport 32
11.1.2.12 Reversible Thermodynamics, the non-dissipative resistor 33
11.1.2.13 Inhibition versus differencing 34
11.1.2.14 Use of the term "Neural Networks" 34

11.1.3 Methods and capabilities of available tools 35
11.1.3.1 Traffic analysis as a major adjunct to anatomy 35
11.1.3.2 Spatial and temporal resolution of available tools 36
11.1.4 Architecture of the neural system 37
11.1.4.1 The basic neural unit 39
11.1.4.2 A hierarchy of the neural system architecture 40
11.1.4.2.1 A top level block diagram of the complete neural system 41
11.1.4.2.2 The hierarchy within the CNS 43
11.1.4.3 Consequences of introducing memory in the CNS 45
11.1.4.3.1 Anachronistic consequences of memory in the CNS 45
11.1.4.4 Recent fMRI data confirming the distribution of the feature extraction engines of the brain 46
11.1.4.5 Chemical versus electrical foundation of the architecture 46
11.1.4.6 Electrotonic versus pulse signaling 46
11.1.5 Categorization of neurons by stage and function 47
11.1.5.1 Notation for the ganglion neurons 51
11.1.6 Notation of principal waveforms of electrophysiology 52
11.1.6.1 Component waveforms of the composite ERG and LERG 52
11.1.6.1.1 The Class A waveform–of the P/D process 52
11.1.6.1.2 The Class B waveform–at the inner segment, IS 53
11.1.6.1.3 The Class C waveform–at the photoreceptor cell 53
11.1.6.1.4 The Class D waveform–at the photoreceptor pedicle 53
11.1.6.1.5 The Class E waveform–at the 1st matrix 53
11.1.6.1.6 The Class F waveform–at 2nd matrix 54
11.1.6.1.7 The Class G waveform–at ganglion output 55
11.1.6.1.8 The Class H waveform–at the stellate cell 55
11.1.6.1.9 The Class J waveform–following the stellate cell 55
11.1.6.1.10 A complete ERG as an example 55
11.1.6.1.11 Comments applying to the above waveforms 56
11.1.6.2 Features of the composite waveforms of the ERG 57
11.1.6.2.1 The a-wave 59
11.1.6.2.2 The b-wave 59
11.1.6.2.3 The c-wave 59
11.1.6.2.4 The d-wave 60
11.1.6.2.5 The e-wave 61
11.1.6.2.6 The m-wave 61
11.1.6.2.7 The early and late receptor potentials 61
11.1.6.2.8 The OP or oscillatory potential feature 61
11.1.2.5.9 The PNR–proximal negative response 62
11.1.6.3 Features of the multi-focal ERG 63
11.1.7 Graphical presentations 64
11.1.7.1 Obsolescence of the spectral sensitivity nomograph 64
11.1.7.2 Graphs from semiconductor physics 64
11.1.7.3 Multiple quadrant graphs 65
11.1.7.4 Graphs from photography 66
11.1.8 Data rates versus information rates 67
11.1.8.1 The data capacity of the visual system 68
11.1.8.2 The information capacity of the visual system 69
11.1.8.2.1 The problem of calculating information content in the analysis channel 69
11.1.8.2.2 The problem of calculating information content in the alarm channels 69
11.1.8.3 Calculating the data transfer rate of humans 69
11.1.8.4 Calculating the information transfer rate of humans 71
11.2 Description of the circuits of visual signal processing 71
11.2.1 Nature of the active devices within a neuron 72
11.2.2 Neurophysiological-reductionistic models 73
11.2.2.1 The single-neuron theories 73
11.3 Baseline electrical analogs for vision 73
11.4 Top level power distribution/consumption 76
11.5 Principle Signal Sources 78

11.6 Functional and tabular models of vision 78
11.6.1 Background 79
11.6.2 The Global Model 80
11.6.2.1 The Signaling environment of the Global Model 80
11.6.3 The Functional Block Diagram 83
11.6.3.1 Other signaling channels 87
11.6.3.1.1 The appearance channels 87
11.6.3.1.2 The polarization channel(s) in *Arthropoda* & *Mollusca* 87
11.6.3.1.3 The polarization channel(s) in *Chordata* 88
11.6.3.2 The relative complexity of the functional block diagram 88
11.6.3.2.1 Possible diversity encoding schemes 89
11.6.4 The Top Level Schematic 90
11.6.4.1 The Motor System and Platform 90
11.6.4.2 The Cortical System 91
11.6.4.3 The Vascular System 91
11.6.4.3.1 The chromophore regeneration subsystem 92
11.6.4.3.2 The electrical power supply subsystem 92
11.6.4.4 The Signaling System 92
11.6.4.4.1 Luminance signaling 95
11.6.4.4.2 Chromatic signal separation 95
11.6.4.4.3 Polarization signaling 96
11.6.4.4.4 Appearance signaling 96
11.7 Circuit diagrams of vision 96
11.7.1 Background 97
11.7.1.1 The grounding plan 97
11.7.1.2 Effects of temperature 97
11.7.1.3 Large versus small signal conditions 97
11.7.2 Large signal circuit diagrams 98
11.7.2.1 Discussion of the large signal circuit diagram 99
11.7.2.2 Gross dynamics of visual performance 104
11.7.2.2.1 The photopic regime 106
11.7.2.2.2 The mesotopic regime 107
11.7.2.2.3 The hypertopic regime 107
11.7.2.2.4 The scotopic regime 108
11.7.2.2.5 A hybrid regime 108
11.7.3 Medium signal circuit diagrams 108
11.7.4 Small signal circuit diagrams 111
11.7.5 Noise model circuit diagrams 112
11.7.5.1 Defining the model 113
11.7.5.1.1 The “RMS noise” model 114
11.7.5.1.2 The band-limited noise model 115
11.7.5.1.3 Noise and the signal processing stage 116
11.7.5.1.4 Noise at the dendritic input to the PC’s of the outer segment EMPTY 116
11.7.5.2 The spectral and noise experiments of Fyhrquist et al. 116
11.8 Comparison with other models 117
11.8.1 Morphological descriptions 117
11.8.2 Sub-stages of the Overall Block Diagram 119
11.8.2.1 Monophasic signaling channels 119
11.8.2.2 Biphasic signaling channels 119
11.8.2.3 Special foveola channels 120
11.8.3 Other overall block diagrams 120
11.8.4 Other circuit diagrams 123
11.8.5 Other conceptual descriptions 124
11.8.5.1 Detection theories 124
11.8.5.2 Chromatic theories 124
11.8.5.2.1 Trilateral color theories 124
11.8.5.2.2 Opponent color theories 125
11.8.5.3 The glutamate and Calcium cascades 125
11.8.5.3.1 The putative glutamate cascade 125
11.8.5.3.2 The Calcium cascade EMPTY 126
11.8.5.4 Performance based psychophysical models 126

11.8.5.4.1 Hyperacuity as a concept 126
11.8.5.4.2 Theories of visual texture perception 127
11.8.5.4.3 Theories of pattern recognition 127
11.8.5.5 Potential transfer of machine vision concepts to biological vision 128

Figure 11.1.1-1 The ground plan within the retina. 7
Figure 11.1.1-2 The ground planes and bridges pertinent to the Visual system 9
Figure 11.1.2-1 Correlation between the functional and morphological neuron 21
Figure 11.1.2-2 Fundamental voltage regimes in vision. 23
Figure 11.1.2-3 ERG lead placement relative to the topography of the retina. 26
Figure 11.1.3-1 The spatial and temporal resolution range available with current techniques 37
Figure 11.1.4-1 The spheres of operation in the higher neural systems 38
Figure 11.1.4-2 The hierarchy and signaling plan of neural pathways. 41
Figure 11.1.4-3 Top level block diagram of the neural system 42
Figure 11.1.5-1 A tabulation of neuron classes and their common names 48
Figure 11.1.5-2 A tabulation of neuron classes and their common names 50
Figure 11.1.5-3 Correlation of ganglion cell names 51
Figure 11.1.6-1 The principal signal waveforms of the photoreceptor cell 52
Figure 11.1.6-2 Time and spatial diversity processing in the 2nd lateral matrix. 54
Figure 11.1.6-3 COLOR The ERG as a composite 56
Figure 11.1.6-4 Definition of ERG -waves in terms of their constituent waveforms. 58
Figure 11.1.6-5 The multi-focal ERG scheme and response 63
Figure 11.1.7-1 CR Typical common base output characteristic of a P-N-P semiconductor 65
Figure 11.1.7-2 The multi-quadrant graphical presentation 66
Figure 11.1.7-3 The reciprocity graph of photography. 67
Figure 11.1.8-1 Summary of the signal propagation velocities in neural systems. 70
Figure 11.3.1-1 Fundamental transistor circuits 75
Figure 11.6.2-1 Global model of the vision process in animals 81
Figure 11.6.3-1 Functional block diagram of vision. 85
Figure 11.6.3-2 A functional block diagram illustrating the gross convergence 89
Figure 11.6.4-1 Top level Schematic of the Visual System in *Chordata*. See text for details 90
Figure 11.6.4-2 Overall signal processing architecture of the tetrachromatic visual system 94
Figure 11.7.2-1 The large signal circuit diagram of the fundamental signal paths 102
Figure 11.7.2-2 Generic decoding circuit of the neurological system EDIT 103
Figure 11.7.2-3 Incremental threshold of a human subject versus a chromatic background 106
Figure 11.7.3-1 The medium signal equivalent circuit for the visual process 110
Figure 11.7.3-2 The small signal equivalent circuit for the visual process 112
Figure 11.7.3-3 High level noise model of vison 114
Figure 11.8.1-1 The Overall Block Diagram of Vision at several layers of detail 118
Figure 11.8.1-2 Primate visual pathways shown semi-schematically by Dacey. 123

(Inactive) AUTHOR INDEX
(Have not learned to use two separate concordances yet)
(Now must use two separate runs and freeze content of list after each run)

Brezold 67, 83
Brucke 67, 83
Carnot 36, 83
Davson 2, 83
Hubel 63, 73, 74, 77
Noback 63
Poynting 49, 83
Purkinje 66, 83
Ranvier 9, 10, 15, 20, 21, 43, 83
Valois 76, 83
Wald 67, 78
Yarbus 74, 83

(Active) SUBJECT INDEX (using advanced indexing option)

1st lateral processing matrix 59, 90, 91, 95
2nd lateral processing matrix 61, 63, 90, 96
99% 64, 122
action potential 7, 16, 20, 23, 24, 26, 40, 41, 55, 56, 69, 70, 72, 80, 84
Activa 1, 7, 15, 20-23, 25, 27, 39, 40, 43, 46, 49, 51-53, 64, 65, 69, 70, 73-78, 82, 97-100, 104
adaptation 4, 7, 10, 17, 20, 22, 27, 29, 30, 47, 49, 53, 59-62, 66, 67, 82, 83, 92, 94, 97-100, 104-108, 113-116, 120, 121, 126
adaptation amplifier 47, 49, 53, 60, 66, 67, 83, 94, 98-100, 106, 107, 115, 116
afterimage 121
alarm mode 71
amplification 6, 15, 16, 22, 27, 39, 47, 49, 51, 52, 67, 80, 82, 99, 100, 107, 114, 116, 124, 125
arborization 20, 21, 96, 103, 104, 111, 122
area 7 . . . . . 45, 82, 84, 96, 121
attention 52, 56, 68, 71, 72, 119
auditory nerve 88
axoplasm 7, 17, 18, 21-25, 64, 65, 72, 77
a-wave 12, 18, 52, 55, 56, 58-62, 77, 78
baseball 71
bat 71
Bayesian 4
bifurcation 44
bilayer 33
bioelectrochemistry 12
bistratified 122
Black Body 104, 108
bleaching 108
bouton 21
broadband 124
b-wave 12, 18, 55, 56, 58-62, 77, 78
cerebellum 42, 44, 45, 55, 82, 83, 101
cerebrum 42, 44-46, 82, 117
CIE 19, 20, 92, 93
class A 29, 52, 56, 59, 114
class B 29, 52, 53, 56, 60, 66
coaxial cable 69
colliculus 88, 91, 95, 103, 117, 119, 122
commissure 68
computation 45
computational 34, 45, 69, 71, 79, 113, 128

conexus 20, 21
consciousness 88
continuum 20, 64
cross section 117
cross-section 14, 117
dark adaptation 4, 107
data base 64
database 64
decoder 40, 114
dendrolemma 53
diencephalon 44, 122
diode 7, 10, 34, 64, 65, 72, 97, 99, 108-111, 123
dynamic range 22, 27, 66, 97, 98, 105, 106, 109
eccentric cell 4, 16
electrostenolytic process 23, 53, 70, 76, 126
encoder 40
entropy 71
ERG 9, 12, 17, 18, 25, 26, 52, 55-63, 69
evoked potential 35
evoked potentials 17, 18
evolution 16, 44, 89, 117
expanded 7, 12, 35, 62, 83, 88, 104, 111, 115, 122
exposine 107
external feedback 25, 48, 50, 73, 106
feedback 3, 25, 28, 32, 34, 48, 50, 64, 73, 83, 106, 107, 114, 119
fMRI 35, 44, 46
free running 55, 100, 104
free-running 47, 49, 111
GABA 46, 104
ganglion neuron 47, 49, 122
glutamate 91, 125
group velocity 69, 70
half-amplitude 116
hole 24, 32, 33
hole transport 24
homogeneous 8, 10, 20
hydronium 25, 99
hyperacuity 126, 127
intelligence 35, 39
internal feedback 32, 64, 106, 107, 114
inverting 22, 54, 56, 59, 100, 111
Ishihara 124
Kirchoff’s Laws 10
latency 59, 60
lateral geniculate 44, 122
lgn/occipital 128
light adaptation 99, 105, 107, 120, 121
Limulus 4, 16, 47, 49, 61, 100
liquid-crystalline 72
local circuit 24
long term memory 45, 46, 79
machine vision 128
mesencephalon 45
mesotopic 107, 114
microcircuitry 117, 125
midbrain 3, 32, 41-43, 45, 67-69, 127
modulation 28, 30, 97
monopulse 23, 55, 70
monopulse oscillator 23, 55
morphogenesis 38
MRI 35, 44, 46
multi-dimensional 78

multi-focal ERG 63
myelin 16
myelinated 32
Myelination 32
neurite 1, 20, 24, 72, 104
neurites 20, 21, 43, 47, 49, 104
Node of Ranvier 7, 16, 21, 25, 46, 49, 51, 55, 70
noise 4, 16, 27, 57, 68, 73, 105, 112-116, 120
non-inverting 22, 54, 56, 100, 111
Nyquist 126
n-type 33, 74
Ohm's Law 10
P/D equation 11, 12, 30, 52, 59, 66
pain 108
parametric 65
parietal lobe 45, 82, 83, 122
parvocellular 52, 82, 84, 91, 96, 121, 122
perigeniculate 122, 123
perigeniculate nucleus 122, 123
PET 35
pgn/pulvinar 89, 128
phase velocity 69, 70
pnp 64, 75
poda 16, 107
podites 20
poditic 7, 20, 22, 23, 54, 100, 104
POS 44, 82, 83, 89, 116, 122, 127
Pretectal 117, 122, 123
Pretectum 43, 45, 52, 82, 83, 85, 87-91, 95, 96, 103, 117, 119-121
probabilistic 120
propagation velocity 32, 70
protocol 62
pulse-to-pulse 104, 111, 124
pulvinar 45, 84, 89, 120, 122, 128
Pulvinar pathway 45, 84, 120
p-type 33, 74
quantum-mechanical 29, 33, 52
RBM 128
reading 3, 71
rectifier 111
reflex arc 39-42, 44
resonance 28, 46
retinitis pigmentosa 59
rhabdom 16
rhabdomere 101
saliency map 122, 127
segregation 121
servo loop 95
servomechanism 42, 79, 82, 91, 120, 122
signal-to-noise 27
signal-to-noise ratio 27
simple neurons 124
sleep 45
square law 13, 107, 114
stage 0 . . . . . 43, 83
stage 1 . . . . . 12, 43, 47, 49, 52, 53, 80, 83, 118
stage 1A 82
stage 1B 82
stage 2 . . . . . 12, 43, 47, 49, 53, 70, 80-83, 117, 121
stage 3 . . . . . 43, 46, 47, 49, 51, 52, 55, 70, 80, 83, 84, 114, 118, 121, 122
stage 3A 82

stage 3B 82
stage 4 . . . . . 43, 55, 83, 84, 118, 121, 128
stage 5 . . . . . 43, 83
stage 5c 83
stage 6 43
stage C 80
stellate 40, 41, 43, 55, 103, 105, 114
stellate cell 40, 43, 55, 114
stress 38, 63, 66, 104, 110
superior colliculus 88, 91, 103, 117, 122
synapse 22, 25, 36, 39, 40, 46, 54, 72, 100, 104, 109-111, 116
thalamic reticular nucleus 122
thalamus 70, 122
threshold 13, 14, 16, 23, 28, 55, 85, 97, 99, 100, 104-106, 111, 114, 126, 127
topography 25-27, 63
topology 2, 3, 29, 35, 48, 50, 97, 98, 124
transduction 56, 60, 73, 82, 83, 113, 114, 124
transistor action 15
translation 13, 14, 78, 82, 124
trans- 60, 61
tremor 3, 4, 17, 30, 31, 38, 78, 119, 126, 127
VEP 17-19, 35, 46
vestibular system 45, 48, 50
visual cortex 45, 46, 84, 95, 96, 104, 119, 120, 122
Weber's Law 2, 19, 105
white matter 122
xxx 3, 7, 24, 49, 51, 66, 68, 84, 107, 110, 114, 122
[xxx 89, 126

(Inactive) DEFINITIONS INDEX (Use individual marks)

Principle of Univariance
Retinal illuminance
Transport delay
net photoreceptor