

I. PHOTON COUNTING IN VISION

Imagine sitting quietly in a dark room, staring straight ahead. A light flashes. Do you see it? Surely if the flash is bright enough the answer is yes, but how dim can the flash be before we fail? Do we fail abruptly, so that there is a well defined threshold—lights brighter than threshold are always seen, lights dimmer than threshold are never seen—or is the transition from seeing to not seeing somehow more gradual? These questions are classical examples of “psychophysics,” studies on the relationship between perception and the physical variables in the world around us, and have a history reaching back at least into the nineteenth century.

In 1911, the physicist Lorentz was sitting in a lecture that included an estimate of the “minimum visible,” the energy of the dimmest flash of light that could be consistently and reliably perceived by human observers. But by 1911 we knew that light was composed of photons, and if the light is of well defined frequency or wavelength then the energy E of the flash is equivalent to an easily calculable number of photons n , $n = E/h\nu$. Doing this calculation, Lorentz found that just visible flashes of light correspond to roughly 100 photons incident on our eyeball. Turning to his physiologist colleague Zwaa-dermaker, Lorentz asked if much of the light incident on the cornea might get lost (scattered or absorbed) on its way through the gooey interior of the eyeball, or if the experiments could be off by as much as a factor of ten. In other words, is it possible that the real limit to human vision is the counting of single photons?

Lorentz’ suggestion really is quite astonishing. If correct, it would mean that the boundaries of our perception are set by basic laws of physics, and that we reach the limits of what is possible. Further, if the visual system really is sensitive to individual light quanta, then some of the irreducible randomness of quantum events should be evident in our perceptions of the world around us, which is a startling thought.

In this Chapter, we will see that humans (and other animals) really can detect the arrival of individual photons at the retina. Tracing through the many steps from photon arrival to perception we will see a sampling of the physics problems posed by biological systems, ranging from the dynamics of single molecules through amplification and adaptation in biochemical reaction networks, coding and computation in neural networks, all the way to learning and cognition. For photon counting some of these problems are solved, but even in this well studied case many problems are open and ripe for new theoretical and experimental work. The problem of photon counting also introduces us to methods and concepts of much broader applicability. We begin by exploring the phenomenology, aiming at the formulation of the key physics problems. By the end of the Chapter I hope to have formulated an approach to the exploration of bio-

logical systems more generally, and identified some of the larger questions that will occupy us in Chapters to come.

A. Posing the problem

One of the fundamental features of quantum mechanics is randomness. If we watch a single molecule and ask if it absorbs a photon, this is a random process, with some probability per unit time related to the light intensity. The emission of photons is also random, so that a typical light source does not deliver a fixed number of photons. Thus, when we look at a flash of light, the number of photons that will be absorbed by the sensitive cells in our retina is a random number, not because biology is noisy but because of the physics governing the interaction of light and matter. One way of testing whether we can count single photons, then, is to see if we can detect the signatures of this quantum randomness in our perceptions. This line of reasoning came to fruition in experiments by Hecht, Schlaer and Pirenne (in New York) and by van der Velden (in the Netherlands) in the early 1940s. [Need to check what was done by Barnes & Czerny, between Lorentz and 1940s]

What we think of classically as the intensity of a beam of light is proportional to the *mean* number of photons per second that arrive at some region where they can be counted.⁷ For most conventional light sources, however, the stream of photons is not regular, but completely random. Thus, in any very small window of time dt , there is a probability $r dt$ that a photon will be counted, where r is the mean counting rate or light intensity, and the events in different time windows are independent of one another. These are the defining characteristics of a “Poisson process,” which is the maximally random sequence of point events in time—if we think of the times at which photons are counted as being like the positions of particles, then the sequence of photon counts is like an ideal gas, with no correlations or “interactions” among the particles at different positions.

As explained in detail in Appendix A.1, if events occur as a Poisson process with rate r , then if we count the events over some time T , the mean number of counts will be $M = rT$, but the probability that we actually obtain a count of n will be given by the Poisson distribution,

$$P(n|M) = e^{-M} \frac{M^n}{n!}. \quad (1)$$

In our case, the mean number of photons that will be counted at the retina is proportional to the classical intensity of the light flash, $M = \alpha I$, where the constant

⁷ More precisely, we can measure the mean number of photons per second per unit area.

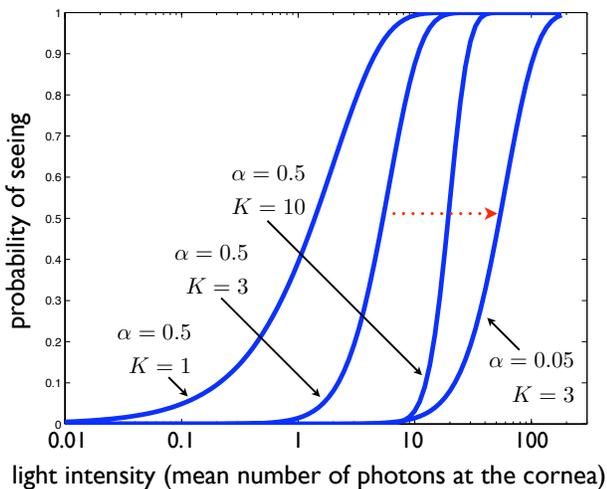


FIG. 1 Probability of seeing calculated from Eq. (2), where the intensity I is measured as the mean number of photons incident on the cornea, so that α is dimensionless. Curves are shown for different values of the threshold photon count K and the scaling factor α . Note the distinct shapes for different K , but when we change α at fixed K we just translate the curve along the the log intensity axis, as shown by the red dashed arrow.

α includes all the messy details of what happens to the light on its way through the eyeball.⁸ Thus, when we deliver the “same” flash of light again and again, the actual physical stimulus to the retina will fluctuate, and it is plausible that our perceptions will fluctuate as well.

Let’s be a bit more precise about all of this. In the simplest view, you would be willing to say “yes, I saw the flash” once you had counted K photons. Equation (1) tell us the probability of counting exactly n photons given the mean, and the mean is connected to the intensity of the flash by $M = \alpha I$. Thus we predict that there is a probability of seeing a flash of intensity I ,

$$P_{\text{see}}(I) = \sum_{n=K}^{\infty} P(n|M = \alpha I) = e^{-\alpha I} \sum_{n=K}^{\infty} \frac{(\alpha I)^n}{n!}. \quad (2)$$

So, if we sit in a dark room and watch as dim lights are flashed, we expect that our perceptions will fluctuate—sometimes we see the flash and sometimes we don’t—but there will be an orderly dependence of the *probability* of seeing on the intensity, given by Eq (2). Importantly, if

⁸ The units for light intensity are especially problematic. Today we know that we should measure the number of photons arriving per second, per unit area, but many of the units were set before this was understood. Also, if we have a broad spectrum of wavelengths, we might want to weight the contributions from different wavelengths not just by their contribution to the total energy but by their contribution to the overall appearance of brightness. Thus, some of the complications have honest origins.

we plot P_{see} vs. $\log I$, as in Fig. 1, then the *shape* of the curve depends crucially on the threshold photon count K , but changing the unknown constant α just translates the curve along the x-axis. So we have a chance to measure the threshold K by looking at the shape of the curve; more fundamentally we might say we are testing the hypothesis that the probabilistic nature of our perceptions is determined by the physics of photon counting.

Problem 1: Photon statistics, part one. There are two reasons why the arrival of photons might be described by a Poisson process. The first is a very general “law of small numbers” argument. Imagine a general point process in which events occur at times $\{t_i\}$, with some correlations among the events. Assume that these correlations die away with some correlation time, so that events i and j are independent if $|t_i - t_j| \gg \tau_c$. Explain qualitatively why, if we select events out of the original sequence at random, then if we select a sufficiently small fraction of these events the resulting sparse sequence will be approximately Poisson. What is the condition for the Poisson approximation to be a good one? What does this have to do with why, for example, the light which reaches us from an incandescent light bulb comes in a Poisson stream of photons?

Problem 2: How many sources of randomness? As noted above, the defining feature of a Poisson process is the independence of events at different times, and typical light sources generate a stream of photons whose arrival times approximate a Poisson process. But when we count these photons, we don’t catch every one. Show that if the photon arrivals are a Poisson process with rate r , and we count a fraction f these, selected at random, then the times at which events are counted will also be a Poisson process, with rate fr . Why doesn’t the random selection of events to be counted result in some “extra” variance beyond expectations for the Poisson process?

Hecht, Shlaer and Pirenne did exactly the experiment we are analyzing. Subjects (the three co-authors) sat in a dark room, and reported whether they did or did not see a dim flash of light. For each setting of the intensity, there were many trials, and responses were variable, but the subjects were forced to say yes or no, with no “maybe.” Thus, it was possible to measure at each intensity the probability that the subject would say yes, and this is plotted in Fig 2.

The first nontrivial result of these experiments is that human perception of dim light flashes really is probabilistic. No matter how hard we try, there is a range of light intensities in which our perceptions fluctuate from flash to flash of the same intensity, seeing one and missing another. Quantitatively, the plot of probability of seeing vs $\log(\text{intensity})$ is fit very well by the predictions from the Poisson statistics of photon arrivals. In particular, Hecht, Shlaer and Pirenne found a beautiful fit in the range from $K = 5$ to $K = 7$; subjects of different age had very different values for α (as must be true if light transmission through the eye gets worse with age) but

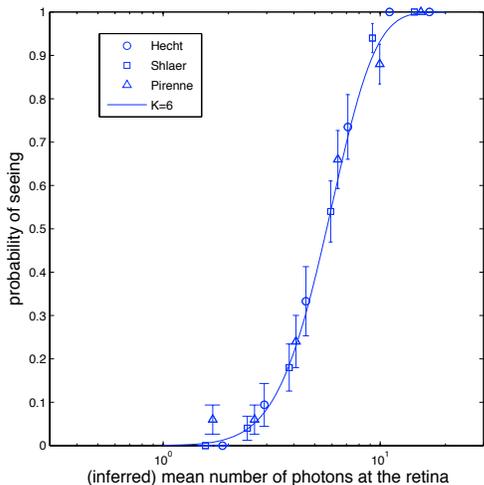


FIG. 2 Probability of seeing calculated from Eq. (2), with the threshold photon count $K = 6$, compared with experimental results from Hecht, Shlaer and Pirenne. For each observer we can find the value of α that provides the best fit, and then plot all the data on a common scale as shown here. Error bars are computed on the assumption that each trial is independent, which probably generates errors bars that are slightly too small.

similar values of K . In Fig 2 I've shown all three observers' data fit to $K = 6$, along with error bars (absent in the original paper); although one could do better by allowing each person to have a different value of K , it's not clear that this would be supported by the statistics. The different values of α , however, are quite important.

Details aside, the frequency of seeing experiment brings forward a beautiful idea: the probabilistic nature of our perceptions just reflects the physics of random photon arrivals. An absolutely crucial point is that Hecht, Shlaer and Pirenne chose stimulus conditions such that the 5 to 7 photons needed for seeing are distributed across a broad area on the retina, an area that contains hundreds of photoreceptor cells [perhaps this needs to be clearer?] Thus the probability of one receptor (rod) cell getting more than one photon is very small. The experiments on human behavior therefore indicate that individual photoreceptor cells generate reliable responses to single photons. In fact, vision begins (as we discuss in more detail soon) with the absorption of light by the visual pigment rhodopsin, and so sensitivity to single photons means that each cell is capable of responding to a single molecular event. This is a wonderful example of using macroscopic experiments to draw conclusions about single cells and their microscopic mechanisms.

Problem 3: Simulating a Poisson process. Much of what we want to know about Poisson processes can be determined analytically (see Appendix A.1). Thus if we do simulations we know

what answer we should get (!). This provides us with an opportunity to exercise our skills, even if we don't get any new answers. In particular, *doing* a simulation is never enough; you have to analyze the results, just as you analyze the results of an experiment. Now is as good a time as any to get started. If you are comfortable doing everything in a programming language like C or Fortran, that's great. On the other hand, high-level languages such as MATLAB or Mathematica have certain advantages. Here you should use MATLAB to simulate a Poisson process, and then analyze the results to be sure that you actually did what you expected to do. [Before finalizing, check on the use of free version of MATLAB, Octave.]

(a) MATLAB has a command `rand` that generates random numbers with a uniform distribution from 0 to 1. Consider a time window of length T , and divide this window into many small bins of size dt . In each bin you can use `rand` to generate a number which you can compare with a threshold—if the random number is above threshold you put an event in the bin, and you can adjust the threshold to set the average number of events in the window. You might choose $T = 10^3$ sec and arrange that the average rate of the events is $\bar{r} \sim 10$ per second; note that you should be able to relate the threshold to the mean rate \bar{r} analytically. Notice that this implements (in the limit $dt \rightarrow 0$) the definition of the Poisson process as independent point events.

(b) The next step is to check that the events you have made really do obey Poisson statistics. Start by counting events in windows of some size τ . What is the mean count? The variance? Do you have enough data to fill in the whole probability distribution $P_\tau(n)$ for counting n of events in the window? How do all of these things change as you change τ ? What if you go back and make events with a different average rate? Do your numerical results agree with the theoretical expressions? In answering this question, you could try to generate sufficiently large data sets that the agreement between theory and experiment is almost perfect, but you could also make smaller data sets and ask if the agreement is good within some estimated error bars; this will force you to think about how to put error bars on a probability distribution. [Do we need to have some more about error bars somewhere in the text?] You should also make a histogram (`hist` should help) of the times between successive events; this should be an exponential function, and you should work to get this into a form where it is a properly normalized probability density. Relate the mean rate of the events to the shape of this distribution, and check this in your data.

(c) Instead of deciding within each bin about the presence or absence of an event, use the command `rand` to choose N random times in the big window T . Examine as before the statistics of counts in windows of size $\tau \ll T$. Do you still have an approximately Poisson process? Why? Do you see connections to the statistical mechanics of ideal gases and the equivalence of ensembles?

Problem 4: Photon statistics, part two. The other reason why we might find photon arrivals to be a Poisson process comes from a very specific quantum mechanical argument about coherent states. This argument may be familiar from your quantum mechanics courses, but this is a good time to review. If you are not familiar with the description of the harmonic oscillator in terms of raising and lowering or creation and annihilation operators, try the next problem, which derives many of the same conclusions by making explicit use of wave functions.

(a.) We recall that modes of the electromagnetic field (in a free space, in a cavity, or in a laser) are described by harmonic oscillators. The Hamiltonian of a harmonic oscillator with frequency ω can be written as

$$\mathbf{H} = \hbar\omega(a^\dagger a + 1/2), \quad (3)$$

where a^\dagger and a are the creation and annihilation operators that connect states with different numbers of quanta,

$$a^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle, \quad (4)$$

$$a |n\rangle = \sqrt{n} |n-1\rangle. \quad (5)$$

There is a special family of states called coherent states, defined as

eigenstates of the annihilation operator,

$$a|\alpha\rangle = \alpha|\alpha\rangle. \quad (6)$$

If we write the coherent state as a superposition of states with different numbers of quanta,

$$|\alpha\rangle = \sum_{n=0}^{\infty} \psi_n |n\rangle, \quad (7)$$

then you can use the defining Eq (6) to give a recursion relation for the ψ_n . Solve this, and show that the probability of counting n quanta in this state is given by the Poisson distribution, that is

$$P_\alpha(n) \equiv \left| \langle n|\alpha\rangle \right|^2 = |\psi_n|^2 = e^{-M} \frac{M^n}{n!}, \quad (8)$$

where the mean number of quanta is $M = |\alpha|^2$.

(b.) The specialness of the coherent states relates to their dynamics and to their representation in position space. For the dynamics, recall that any quantum mechanical state $|\phi\rangle$ evolves in time according to

$$i\hbar \frac{d|\phi\rangle}{dt} = \mathbf{H}|\phi\rangle. \quad (9)$$

Show that if the system starts in a coherent state $|\alpha(0)\rangle$ at time $t = 0$, it remains in a coherent state for all time. Find $\alpha(t)$.

(c.) If we go back to the mechanical realization of the harmonic oscillator as a mass m hanging from a spring, the Hamiltonian is

$$\mathbf{H} = \frac{1}{2m} p^2 + \frac{m\omega^2}{2} q^2, \quad (10)$$

where p and q are the momentum and position of the mass. Remind yourself of the relationship between the creation and annihilation operators and the position and momentum operators (\hat{q}, \hat{p}). In position space, the ground state is a Gaussian wave function,

$$\langle q|0\rangle = \frac{1}{(2\pi\sigma^2)^{1/4}} \exp\left(-\frac{q^2}{4\sigma^2}\right), \quad (11)$$

where the variance of the zero point motion $\sigma^2 = \hbar/(4m\omega)$. The ground state is also a “minimum uncertainty wave packet,” so called because the variance of position and the variance of momentum have a product that is the minimum value allowed by the uncertainty principle; show that this is true. Consider the state $|\psi(q_0)\rangle$ obtained by displacing the ground state to a position q_0 ,

$$|\psi(q_0)\rangle = e^{iq_0\hat{p}}|0\rangle. \quad (12)$$

Show that this is a minimum uncertainty wave packet, and also a coherent state. Find the relationship between the coherent state parameter α and the displacement q_0 .

(d.) Put all of these steps together to show that the coherent state is a minimum uncertainty wave packet with expected values of the position and momentum that follow the classical equations of motion.

Problem 5: Photon statistics, part two, with wave functions. Work out a problem that gives the essence of the above using wave functions, without referring to α and a^\dagger .

There is a very important point in the background of this discussion. By placing results from all three observers on the same plot, and fitting with the same value of K , we are claiming that there is something reproducible, from individual to individual, about our perceptions. On the other hand, the fact that each observer has a different value for α means that there are individual differences, even in this simplest of tasks. Happily, what seems to be reproducible is something that feels

like a fundamental property of the system, the number of photons we need to count in order to be sure that we saw something. But suppose that we just plot the probability of seeing vs the (raw) intensity of the light flash. If we average across individuals with different α s, we will obtain a result which does not correspond to the theory, and this failure might even lead us to believe that the visual system does not count single photons. This shows us that (a) finding what is reproducible can be difficult, and (b) averaging across an ensemble of individuals can be *qualitatively* misleading. Here we see these conclusions in the context of human behavior, but it seems likely that similar issues arise in the behavior of single cells. The difference is that techniques for monitoring the behavior of single cells (e.g., bacteria), as opposed to averages over populations of cells, have emerged much more recently. As an example, it still is almost impossible to monitor, in real time, the metabolism of single cells, whereas simultaneous measurements on many metabolic reactions, averaged over populations of cells, have become common. We still have much to learn from these older experiments!

Problem 6: Averaging over observers. Go back to the original paper by Hecht, Shlaer and Pirenne⁹ and use their data to plot, vs. the intensity of the light flash, the probability of seeing averaged over all three observers. Does this look anything like what you find for individual observers? Can you simulate this effect, say in a larger population of subjects, by assuming that the factor α is drawn from a distribution? Explore this a bit, and see how badly misled you could be. This is not too complicated, I hope, but deliberately open ended.

Before moving on, a few more remarks about the history. [I have some concern that this is a bit colloquial, and maybe more like notes to add to the references than substance for the text. Feedback is welcome here.] It's worth noting that van der Velden's seminal paper was published in Dutch, a reminder of a time when anglophone cultural hegemony was not yet complete. Also (maybe more relevant for us), it was published in a physics journal. The physics community in the Netherlands during this period had a very active interest in problems of noise, and van der Velden's work was in this tradition. In contrast, Hecht was a distinguished contributor to understanding vision but had worked within a “photochemical” view which he would soon abandon as inconsistent with the detectability of single photons and hence single molecules of activated rhodopsin. Parallel

⁹ As will be true throughout the text, references are found at the end of the section.

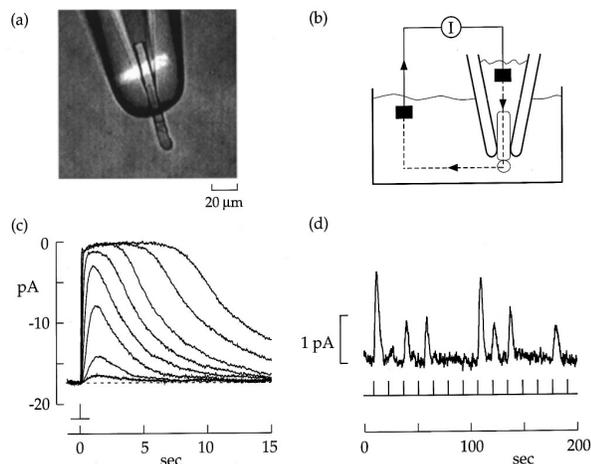


FIG. 3 (a) A single rod photoreceptor cell from a toad, in a suction pipette. Viewing is with infrared light, and the bright bar is a stimulus of 500 nm light. (b) Equivalent electrical circuit for recording the current across the cell membrane [really needs to be redrawn, with labels!]. (c) Mean current in response to light flashes of varying intensity. Smallest response is to flashes that deliver a mean ~ 4 photons, successive flashes are brighter by factors of 4. (d) Current responses to repeated dim light flashes at times indicated by the tick marks. Note the apparently distinct classes of responses to zero, one or two photons. From Rieke & Baylor (1998).

to this work, Rose and de Vries (independently) emphasized that noise due to the random arrival of photons at the retina also would limit the reliability of perception at intensities well above the point where things become barely visible. In particular, de Vries saw these issues as part of the larger problem of understanding the physical limits to biological function, and I think his perspective on the interaction of physics and biology was far ahead of its time.

It took many years before anyone could measure directly the responses of photoreceptors to single photons. It was done first in the (invertebrate) horseshoe crab [be sure to add refs to Fuortes & Yeandle; maybe show a figure?], and eventually by Baylor and coworkers in toads and then in monkeys. The complication in the lower vertebrate systems is that the cells are coupled together, so that the retina can do something like adjusting the size of pixels as a function of light intensity. This means that the nice big current generated by one cell is spread as a small voltage in many cells, so the usual method of measuring the voltage across the membrane of one cell won't work; you have to suck the cell into a pipette and collect the current, as seen in Fig 3.

Problem 7: Gigaseals. As we will see, the currents that are relevant in biological systems are on the order of picoAmps.

Although the response of rods to single photons is slow, many processes in the nervous system occur on the millisecond timescale. Show that if we want to resolve picoAmps in milliseconds, then the leakage resistance (e.g. between rod cell membrane and the pipette in Fig 3) must be $\sim 10^9$ ohm, to prevent the signal being lost in Johnson noise.

In complete darkness, there is a 'standing current' of roughly 20 pA flowing through the membrane of the rod cell's outer segment. You should keep in mind that currents in biological systems are carried not by electrons or holes, as in solids, but by ions moving through water; we will learn more about this below [be sure we do!]. In the rod cell, the standing current is carried largely by sodium ions, although there are contributions from other ions as well. This is a hint that the channels in the membrane that allow the ions to pass are not especially selective for one ion over the other. The current which flows across the membrane of course has to go somewhere, and in fact the circuit is completed within the rod cell itself, so that what flows across the outer segment of the cell is compensated by flow across the inner segment [improve the figures to show this clearly]. When the rod cell is exposed to light, the standing current is reduced, and with sufficiently bright flashes it is turned off all together.

As in any circuit, current flow generates changes in the voltage across the cell membrane. Near the bottom of the cell [should point to better schematic, one figure with everything we need for this paragraph] there are special channels that allow calcium ions to flow into the cell in response to these voltage changes, and calcium in turn triggers the fusion of vesicles with the cell membrane. These vesicles are filled with a small molecule, a neurotransmitter, which can then diffuse across a small cleft and bind to receptors on the surface of neighboring cells; these receptors then respond (in the simplest case) by opening channels in the membrane of this second cell, allowing currents to flow. In this way, currents and voltages in one cell are converted, via a chemical intermediate, into currents and voltages in the next cell, and so on through the nervous system. The place where two cells connect in this way is called a synapse, and in the retina the rod cells form synapses onto a class of cells called bipolar cells. More about this later, but for now you should keep in mind that the electrical signal we are looking at in the rod cell is the first in a sequence of electrical signals that ultimately are transmitted along the cells in the optic nerve, connecting the eye to the brain and hence providing the input for visual perception.

Very dim flashes of light seem to produce a quantized reduction in the standing current, and the magnitude of these current pulses is roughly 1 pA, as seen in Fig 3. When we look closely at the standing current, we see that it is fluctuating, so that there is a continuous background noise of ~ 0.1 pA, so the quantal events are

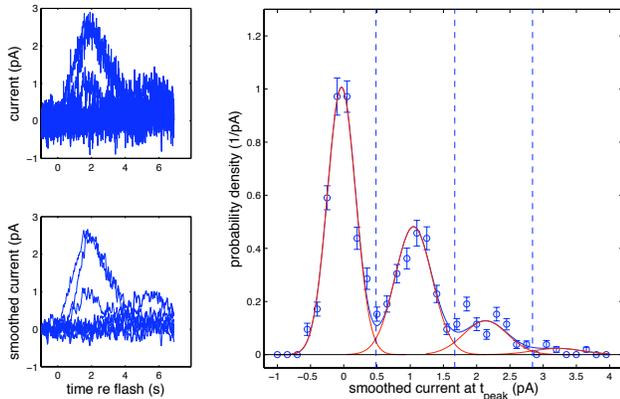


FIG. 4 A closer look at the currents in toad rods. At left, five instances in which the rod is exposed to a dim flash at $t = 0$. It looks as if two of these flashes delivered two photons (peak current ~ 2 pA), one delivered one photon (peak current ~ 1 pA), and two delivered zero. The top panel shows the raw current traces, and the bottom panel shows what happens when we smooth with a 100 ms window to remove some of the high frequency noise. At right, the distribution of smoothed currents at the moment t_{peak} when the average current peaks; the data (circles) are accumulated from 350 flashes in one cell, and the error bars indicate standard errors of the mean due to this finite sample size. Solid green line is the fit to Eq (19), composed of contributions from $n = 0, n = 1, \dots$ photon events, shown red. Dashed blue lines divide the range of observed currents into the most likely assignments to different photon counts. These data are from unpublished experiments by FM Rieke at the University of Washington; many thanks to Fred for providing the data in raw form.

easily detected. It takes a bit of work to convince yourself that these events really are the responses to single photons. Perhaps the most direct experiment is to measure the cross-section for generating the quantal events, and compare this with the absorption cross-section of the rod, showing that a little more than 2/3 of the photons which are absorbed produce current pulses. In response to steady dim light, we can observe a continuous stream of pulses, the rate of the pulses is proportional to the light intensity, and the intervals between pulses are distributed exponentially, as expected if they represent the responses to single photons (cf Section A.1).

Problem 8: Are they really single photon responses?

Work out a problem to ask what aspects of experiments in Fig 4 are the smoking gun. In particular, if one pulse were from the coincidence of two photons, how would the distribution of peak heights shift with changing flash intensity?

When you look at the currents flowing across the rod cell membrane, the statement that single photon events

are detectable above background noise seems pretty obvious, but it would be good to be careful about what we mean here. In Fig 4 we take a closer look at the currents flowing in response to dim flashes of light. These data were recorded with a very high bandwidth, so you can see a lot of high frequency noise. Nonetheless, in these five flashes, it's pretty clear that twice the cell counted zero photons, once it counted one photon (for a peak current ~ 1 pA) and twice it counted two photons; this becomes even clearer if we smooth the data to get rid of some of the noise. Still, these are anecdotes, and one would like to be more quantitative.

Even in the absence of light there are fluctuations in the current, and for simplicity let's imagine that this background noise is Gaussian with some variance σ_0^2 . The simplest way to decide whether we saw something is to look at the rod current at one moment in time, say at $t = t_{\text{peak}} \sim 2$ s after the flash, where on average the current is at its peak. Then given that no photons were counted, this current i should be drawn out of the probability distribution

$$P(i|n=0) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left[-\frac{i^2}{2\sigma_0^2}\right]. \quad (13)$$

If one photon is counted, then there should be a mean current $\langle i \rangle = i_1$, but there is still some noise. Plausibly the noise has two pieces—first, the background noise still is present, with its variance σ_0^2 , and in addition the amplitude of the single photon response itself can fluctuate; we assume that these fluctuations are also Gaussian and independent of the background, so they just add σ_1^2 to the variance. Thus we expect that, in response to one photon, the current will be drawn from the distribution

$$P(i|n=1) = \frac{1}{\sqrt{2\pi(\sigma_0^2 + \sigma_1^2)}} \exp\left[-\frac{(i - i_1)^2}{2(\sigma_0^2 + \sigma_1^2)}\right]. \quad (14)$$

If each single photon event is independent of the others, then we can generalize this to get the distribution of currents expected in response of $n = 2$ photons, [need to explain additions of variances for multiphoton responses]

$$P(i|n=2) = \frac{1}{\sqrt{2\pi(\sigma_0^2 + 2\sigma_1^2)}} \exp\left[-\frac{(i - 2i_1)^2}{2(\sigma_0^2 + 2\sigma_1^2)}\right], \quad (15)$$

and more generally n photons,

$$P(i|n) = \frac{1}{\sqrt{2\pi(\sigma_0^2 + n\sigma_1^2)}} \exp\left[-\frac{(i - ni_1)^2}{2(\sigma_0^2 + n\sigma_1^2)}\right]. \quad (16)$$

Finally, since we know that the photon count n should be drawn out of the Poisson distribution, we can write

the expected distribution of currents as

$$P(i) = \sum_{n=0}^{\infty} P(i|n)P(n) \quad (17)$$

$$= \sum_{n=0}^{\infty} P(i|n)e^{-\bar{n}} \frac{\bar{n}^n}{n!} \quad (18)$$

$$= \sum_{n=0}^{\infty} \frac{\bar{n}^n}{n!} \frac{e^{-\bar{n}}}{\sqrt{2\pi(\sigma_0^2 + n\sigma_1^2)}} \exp\left[-\frac{(i - n\mu_1)^2}{2(\sigma_0^2 + n\sigma_1^2)}\right] \quad (19)$$

In Fig 4, we see that this really gives a very good description of the distribution that we observe when we sample the currents in response to a large number of flashes.

Problem 9: Exploring the sampling problem. The data that we see in Fig 4 are not a perfect fit to our model. On the other hand, there are only 350 samples that we are using to estimate the shape of the underlying probability distribution. This is an example of a problem that you will meet many times in comparing theory and experiment; perhaps you have some experience from physics lab courses which is relevant here. We will return to these issues of sampling and fitting nearer the end of the course, when we have some more powerful mathematical tools, but for now let me encourage you to play a bit. Use the model that leads to Eq (19) to generate samples of the peak current, and then use these samples to estimate the probability distribution. For simplicity, assume that $\mu_1 = 1$, $\sigma_0 = 0.1$, $\sigma_1 = 0.2$, and $\bar{n} = 1$. Notice that since the current is continuous, you have to make bins along the current axis; smaller bins reveal more structure, but also generate noisier results because the number of counts in each bin is smaller. As you experiment with different size bins and different numbers of samples, try to develop some feeling for whether the agreement between theory and experiment in Fig 4 really is convincing.

Seeing this distribution, and especially seeing analytically how it is constructed, it is tempting to draw lines along the current axis in the ‘troughs’ of the distribution, and say that (for example) when we observe a current of less than 0.5 pA, this reflects zero photons. Is this the right way for us—or for the toad’s brain—to interpret these data? To be precise, suppose that we want to set a threshold for deciding between $n = 0$ and $n = 1$ photon. Where should we put this threshold to be sure that we get the right answer as often as possible?

Suppose we set our threshold at some current $i = \theta$. If there really were zero photons absorbed, then if by chance $i > \theta$ we will incorrectly say that there was one photon. This error has a probability

$$P(\text{say } n = 1|n = 0) = \int_{\theta}^{\infty} di P(i|n = 0). \quad (20)$$

On the other hand, if there really was one photon, but by chance the current was less than the threshold, then we’ll say 0 when we should have said 1, and this has a probability

$$P(\text{say } n = 0|n = 1) = \int_{-\infty}^{\theta} di P(i|n = 1). \quad (21)$$

There could be errors in which we confuse two photons for zero photons, but looking at Fig 4 it seems that these higher order errors are negligible. So then the total probability of making a mistake in the $n = 0$ vs. $n = 1$ decision is

$$P_{\text{error}}(\theta) = P(\text{say } n = 1|n = 0)P(n = 0) + P(\text{say } n = 0|n = 1)P(n = 1) \quad (22)$$

$$= P(n = 0) \int_{\theta}^{\infty} di P(i|n = 0) + P(n = 1) \int_{-\infty}^{\theta} di P(i|n = 1). \quad (23)$$

We can minimize the probability of error in the usual way by taking the derivative and setting the result to zero at the optimal setting of the threshold, $\theta = \theta^*$:

$$\frac{dP_{\text{error}}(\theta)}{d\theta} = P(n = 0) \frac{d}{d\theta} \int_{\theta}^{\infty} di P(i|n = 0) + P(n = 1) \frac{d}{d\theta} \int_{-\infty}^{\theta} di P(i|n = 1) \quad (24)$$

$$= P(n = 0)(-1)P(i = \theta|n = 0) + P(n = 1)P(i = \theta|n = 1); \quad (25)$$

$$\left. \frac{dP_{\text{error}}(\theta)}{d\theta} \right|_{\theta=\theta^*} = 0 \Rightarrow P(i = \theta^*|n = 0)P(n = 0) = P(i = \theta^*|n = 1)P(n = 1). \quad (26)$$

This result has a simple interpretation. Given that we have observed some current i , we can calculate the probability that n photons were detected using Bayes’ rule for

conditional probabilities:

$$P(n|i) = \frac{P(i|n)P(n)}{P(i)}. \quad (27)$$

The combination $P(i|n=0)P(n=0)$ thus is proportional to the probability that the observed current i was generated by counting $n=0$ photons, and similarly the combination $P(i|n=1)P(n=1)$ is proportional to the probability that the observed current was generated by counting $n=1$ photons. The optimal setting of the threshold, from Eq (26), is when these two probabilities are equal. Another way to say this is that for each observable current i we should compute the probability $P(n|i)$, and then our “best guess” about the photon count n is the one which maximizes this probability. This guess is best in the sense that it minimizes the total probability of errors. This is how we draw the boundaries shown by dashed lines in Fig 4 [Check details! Also introduce names for these things—maximum likelihood, maximum a posteriori probability, This is also a place to anticipate the role of prior expectations in setting thresholds!]

Problem 10: More careful discrimination. You observe some variable x (e.g., the current flowing across the rod cell membrane) that is chosen either from the probability distribution $P(x|+)$ or from the distribution $P(x|-)$. Your task is to look at a particular x and decide whether it came from the $+$ or the $-$ distribution. Rather than just setting a threshold, as in the discussion above, suppose that when you see x you assign it to the $+$ distribution with a probability $p(x)$. You might think this is a good idea since, if you’re not completely sure of the right answer, you can hedge your bets by a little bit of random guessing. Express the

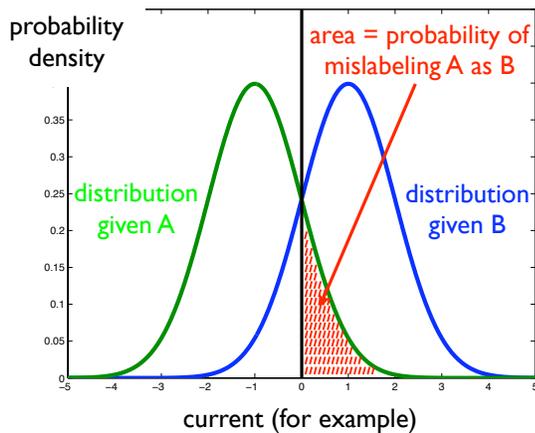


FIG. 5 Schematic of discrimination in the presence of noise. We have two possible signals, A and B, and we measure something, for example the current flowing across a cell membrane. Given either A or B, the current fluctuates. As explained in the text, the overall probability of confusing A with B is minimized if we draw a threshold at the point where the probability distributions cross, and identify all currents larger than this threshold as being B, all currents smaller than threshold as being A. Because the distributions overlap, it is not possible to avoid errors, and the area of the red shaded region counts the probability that we will misidentify A as B.

probability of a correct answer in terms of $p(x)$; this is a functional $P_{\text{correct}}[p(x)]$. Now solve the optimization problem for the function $p(x)$, maximizing P_{correct} . Show that the solution is deterministic [$p(x) = 1$ or $p(x) = 0$], so that if the goal is to be correct as often as possible you shouldn’t hesitate to make a crisp assignment even at values of x where you aren’t sure (!). Hint: Usually, you would try to maximize the P_{correct} by solving the variational equation $\delta P_{\text{correct}}/\delta p(x) = 0$. You should find that, in this case, this approach doesn’t work. What does this mean? Remember that $p(x)$ is a probability, and hence can’t take on arbitrary values.

Once we have found the decision rules that minimize the probability of error, we can ask about the error probability itself. As schematized in Fig 5, we can calculate this by integrating the relevant probability distributions on the ‘wrong sides’ of the threshold. For Fig 4, this error probability is less than three percent. Thus, under these conditions, we can look at the current flowing across the rod cell membrane and decide whether we saw $n=0, 1, 2 \dots$ photons with a precision such that we are wrong only on a few flashes out of one hundred. In fact, we might even be able to do better if instead of looking at the current at one moment in time we look at the whole trajectory of current vs. time, but to do this analysis we need a few more mathematical tools. Even without such a more sophisticated analysis, it’s clear that these cells really are acting as near perfect photon counters, at least over some range of conditions.

Problem 11: Asymptotic error probabilities. Should add a problem deriving the asymptotic probabilities of errors at high signal-to-noise ratios, including effects of prior probability.

A slight problem in our simple identification of the probability of seeing with the probability of counting K photons is that van der Velden found a threshold photon count of $K=2$, which is completely inconsistent with the $K=5-7$ found by Hecht, Shlaer and Pirenne. Barlow explained this discrepancy by noting that even when counting single photons we may have to discriminate (as in photomultipliers) against a background of dark noise.

Hecht, Shlaer and Pirenne inserted blanks in their experiments to be sure that you almost never say “I saw it” when nothing is there, which means you have to set a high threshold to discriminate against any background noise. On the other hand, van der Velden was willing to allow for some false positive responses, so his subjects could afford to set a lower threshold. Qualitatively, as shown in Fig 6, this makes sense, but to be a quantitative explanation the noise has to be at the right level.

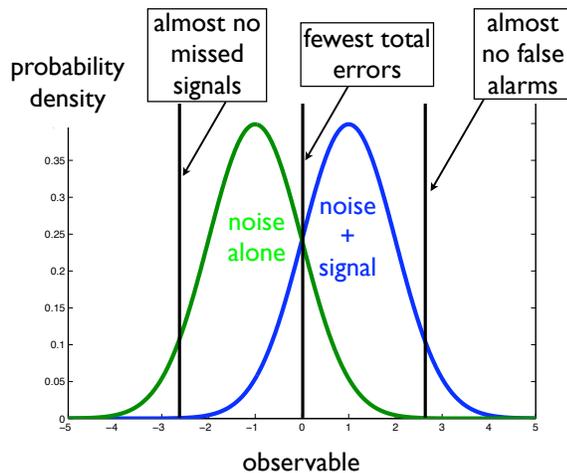


FIG. 6 Trading of errors in the presence of noise. We observe some quantity that fluctuates even in the absence of a signal. When we add the signal these fluctuations continue, but the overall distribution of the observable is shifted. If set a threshold, declaring the signal is present whenever the threshold is exceeded, then we can trade between the two kinds of errors. At low thresholds, we never miss a signal, but there will be many false alarms. At high thresholds, there are few false alarms, but we miss most of the signals too. At some intermediate setting of the threshold, the total number of errors will be minimized.

One of the key ideas in the analysis of signals and noise is “referring noise to the input,” and we will meet this concept many times in what follows [more specific pointers]. Imagine that we have a system to measure something (here, the intensity of light, but it could be anything), and it has a very small amount of noise somewhere along the path from input to output. In many systems we will also find, along the path from input to output, an amplifier that makes all of the signals larger. But the amplifier doesn’t “know” which of its inputs are signal and which are noise, so everything is amplified. Thus, a small noise near the input can become a large noise near the output, but the size of this noise at the output does not, by itself, tell us how hard it will be to detect signals at the input. What we can do is to imagine that the whole system is noiseless, and that any noise we see at the output really was injected at the input, and thus followed exactly the same path as the signals we are trying to detect. Then we can ask how big this effective input noise needs to be in order to account for the output noise.

If the qualitative picture of Fig 6 is correct, then the minimum number of photons that we need in order to say “I saw it” should be reduced if we allow the observer the option of saying “I’m pretty sure I saw it,” in effect taking control over the trade between misses and false alarms. Barlow showed that this worked, quantitatively.

In the case of counting photons, we can think of the effective input noise as being nothing more than extra

“dark” photons, also drawn from a Poisson distribution. Thus if in the relevant window of time for detecting the light flash there are an average of 10 dark photons, for example, then because the variance of the Poisson distribution is equal to the mean, there will be fluctuations on the scale of $\sqrt{10}$ counts. To be very sure that we have seen something, we need an extra K real photons, with $K \gg \sqrt{10}$. Barlow’s argument was that we could understand the need for $K \sim 6$ in the Hecht, Shaler and Pirenne experiments if indeed there were a noise source in the visual system that was equivalent to counting an extra ten photons over the window in time and area of the retina that was being stimulated. What could this noise be?

In the frequency of seeing experiments, as noted above, the flash of light illuminated roughly 500 receptor cells on the retina, and subsequent experiments showed that one could find essentially the same threshold number of photons when the flash covered many thousands of cells. Furthermore, experiments with different durations for the flash show that human observers are integrating over ~ 0.1 s in order to make their decisions about whether they saw something. Thus, the “dark noise” in the system seems to be equivalent, roughly, to 0.1 photon per receptor cell per second, or less. To place this number in perspective, it is important to note that vision begins when the pigment molecule rhodopsin absorbs light and changes its structure to trigger some sequence of events in the receptor cell. We will learn much more about the dynamics of rhodopsin and the cascade of events responsible for converting this molecular event into electrical signals that can be transmitted to the brain, but for now we should note that if rhodopsin can change its structure by absorbing a photon, there must also be some (small) probability that this same structural change or “isomerization” will happen as the result of a thermal fluctuation. If this does happen, then it will trigger a response that is identical to that triggered by a real photon. Further, such rare, thermally activated events really are Poisson processes (see Section II.A), so that thermal activation of rhodopsin would contribute exactly a “dark light” of the sort we have been trying to estimate as a background noise in the visual system. But there are roughly one billion rhodopsin molecules per receptor cell, so that a dark noise of ~ 0.1 per second per cell corresponds to a rate of once per ~ 1000 years for the spontaneous isomerization of rhodopsin.

One of the key points here is that Barlow’s explanation works only if people actually can adjust the “threshold” K in response to different situations. The realization that this is possible was part of the more general recognition that detecting a sensory signal does not involve a true threshold between (for example) seeing and not seeing. Instead, *all* sensory tasks involve a discrimination between signal and noise, and hence there are different strategies which provide different ways of trading

off among the different kinds of errors. Notice that this picture matches what we know from the physics lab.

Problem 12: Simple analysis of dark noise. Suppose that we observe events drawn out of a Poisson distribution, and we can count these events perfectly. Assume that the mean number of events has two contributions, $\bar{n} = \bar{n}_{\text{dark}} + \bar{n}_{\text{flash}}$, where $\bar{n}_{\text{flash}} = 0$ if there is no light flash and $\bar{n}_{\text{flash}} = N$ if there is a flash. As an observer, you have the right to set a criterion, so that you declare the flash to be present only if you count $n \geq K$ events. As you change K , you change the errors that you make—when K is small you often say you saw something when nothing was there, but of hardly ever miss a real flash, while at large K the situation is reversed. The conventional way of describing this is to plot the fraction of “hits” (probability that you correctly identify a real flash) against the probability of a false alarm (i.e., the probability that you say a flash is present when it isn’t), with the criterion changing along the curve. Plot this “receiver operating characteristic” for the case $\bar{n}_{\text{dark}} = 10$ and $N = 10$. Hold \bar{n}_{dark} fixed and change N to see how the curve changes. Explain which slice through this set of curves was measured by Hecht et al, and the relationship of this analysis to what we saw in Fig 2.

There are classic experiments to show that people will adjust their thresholds automatically when we change the a priori probabilities of the signal being present, as expected for optimal performance. This can be done without any explicit instructions—you don’t have to tell someone that you are changing the probabilities—and it works in all sensory modalities, not just vision. At least implicitly, then, people learn something about probabilities and adjust their criteria appropriately. Threshold adjustments also can be driven by changing the rewards for correct answers or the penalties for wrong answers. In this view, it is likely that Hecht et al. drove their observers to high thresholds by having a large effective penalty for false positive detections. Although it’s not a huge literature, people have since manipulated these penalties and rewards in frequency of seeing experiments, with the expected results. Perhaps more dramatically, modern quantum optics techniques have been used to manipulate the statistics of photon arrivals at the retina, so that the tradeoffs among the different kinds of errors are changed ... again with the expected results.¹⁰

Not only did Baylor and coworkers detect the single photon responses from toad photoreceptor cells, they also found that single receptor cells in the dark show spontaneous photon-like events roughly at the right rate to be the source of dark noise identified by Barlow. If you

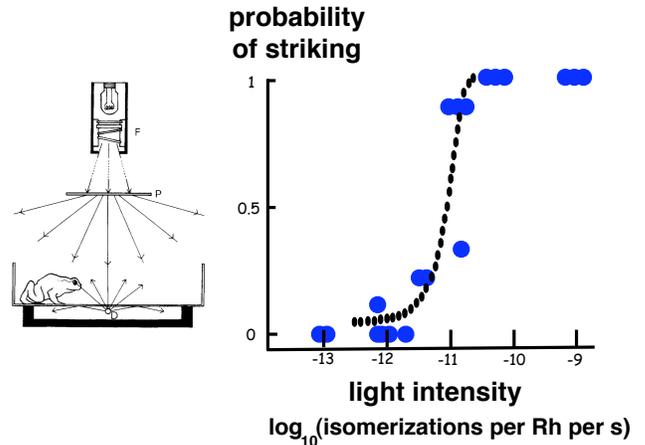


FIG. 7 [fill in the caption] From Aho et al (1988).

look closely you can find one of these spontaneous events in the earlier illustration of the rod cell responses to dim flashes, Fig 3. Just to be clear, Barlow identified a *maximum* dark noise level; anything higher and the observed reliable detection is impossible. The fact that the real rod cells have essentially this level of dark noise means that the visual system is operating near the limits of reliability set by thermal noise in the input. It would be nice to give a more direct test of this idea.

In the lab we often lower the noise level of photodetectors by cooling them. This should work in vision too, since one can verify that the rate of spontaneous photon-like events in the rod cell current is strongly temperature dependent, increasing by a factor of roughly four for every ten degree increase in temperature. Changing temperature isn’t so easy in humans, but it does work with cold blooded animals like frogs and toads. To set the stage, it is worth noting that one species of toad in particular (*Bufo bufo*) manages to catch its prey under conditions so dark that human observers cannot see the toad, much less the prey [add the reference!]. So, Aho et al. convinced toads to strike with their tongues at small worm-like objects illuminated by very dim lights, and measured how the threshold for reliable striking varied with temperature, as shown in Fig 7. Because one can actually make measurements on the retina itself, it is possible to calibrate light intensities as the rate at which rhodopsin molecules are absorbing photons and isomerizing, and the toad’s responses are almost deterministic once this rate is $r \sim 10^{-11} \text{ s}^{-1}$ in experiments at 15°C , and responses are detectable at intensities a factor of three to five below this. For comparison, the rate of thermal isomerizations at this temperature is $\sim 5 \times 10^{-12} \text{ s}^{-1}$.

If the dark noise consists of rhodopsin molecules spontaneously isomerizing at a rate r_d , then the mean number of dark events will be $n_d = r_d T N_r N_c$, where $T \sim 1 \text{ s}$ is the

¹⁰ It is perhaps too much to go through all of these results here, beautiful as they are. To explore, see the references at the end of the section.

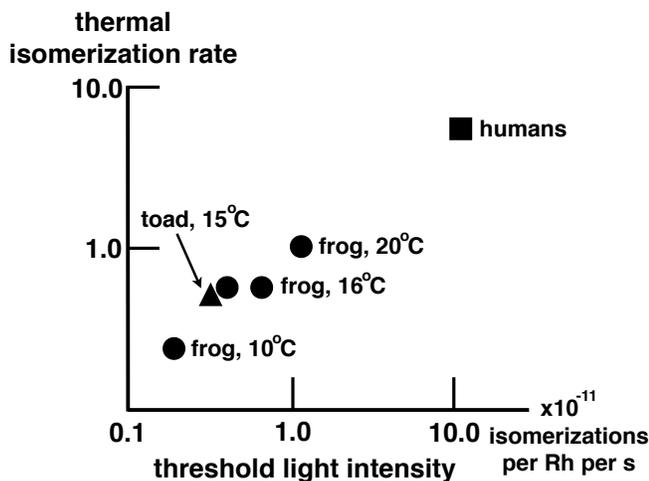


FIG. 8 [fill in the caption] From Aho et al (1987, 1988).

relevant integration time for the decision, $N_r \sim 3 \times 10^9$ is the number of rhodopsin molecules per cell in this retina, and $N_c \sim 4,500$ is the number of receptor cells that are illuminated by the image of the worm-like object. Similarly, the mean number of real events is $n = rTN_rN_c$, and reliable detection requires $n > \sqrt{n_d}$, or

$$r > \sqrt{\frac{r_d}{TN_rN_c}} \sim 6 \times 10^{-13} \text{ s}^{-1}. \quad (28)$$

Thus, if the toad knows exactly which part of the retina it should be looking at, then it should reach a signal-to-noise ratio of one at light intensities a factor of ten below the nominal dark noise level. But there is no way to be sure where to look before the target appears, and the toad probably needs a rather higher signal-to-noise ratio before it is willing to strike. Thus it is plausible that the threshold light intensities in this experiment should be comparable to the dark noise level, as observed.

One can do an experiment very similar to the one with toads using human subjects (who say yes or no, rather than sticking out their tongues), asking for a response to small targets illuminated by steady, dim lights. Frogs will spontaneously jump at dimly illuminated patch of the ceiling, in an attempt to escape from an otherwise dark box. Combining these experiments, with the frogs held at temperatures from 10 to 20 °C, one can span a range of almost two orders of magnitude in the thermal isomerization rate of rhodopsin. It's not clear whether individual organisms hold their integration times T fixed as temperature is varied, or if the experiments on different organisms correspond to asking for integration over a similar total number of rhodopsin molecules (N_rN_c). Nonetheless, it is satisfying to see, in Fig 8, that the “threshold” light intensity, where response occur 50% of the time, is varying systematically with the dark noise level. It is certainly true that operating at lower temperatures allows

the detection of dimmer lights, or equivalently more reliable detection of the same light intensity,¹¹ as expected if the dominant noise source was thermal in origin. These experiments support the hypothesis that visual processing in dim lights really is limited by input noise and not by any inefficiencies of the brain.

Problem 13: Getting a feel for the brain's problem. Let's go back to Problem 3, where you simulated a Poisson process.

(a) If you use the strategy of making small bins $\Delta\tau$ and testing a random number in each bin against a threshold, then it should be no problem to generalize this to the case where the threshold is different at different times, so you are simulating a Poisson process in which the rate is varying as a function of time. As an example, consider a two second interval in which the counting rate has some background (like the dark noise in rods) value r_{dark} except in a 100 msec window where the rate is higher, say $r = r_{\text{dark}} + r_{\text{signal}}$. Remember that for one rod cell, r_{dark} is $\sim 0.02 \text{ sec}^{-1}$, while humans can see flashes which have $r_{\text{signal}} \sim 0.01 \text{ sec}^{-1}$ if they can integrate over 1000 rods. Try to simulate events in this parameter range and actually look at examples, perhaps plotted with x's to show you where the events occur on a single trial.

(b) Can you tell the difference between a trial where you have $r_{\text{signal}} = 0.01 \text{ sec}^{-1}$ and one in which $r_{\text{signal}} = 0$? Does it matter whether you know when to expect the extra events? In effect these plots give a picture of the problem that the brain has to solve in the Hecht-Shaler-Pirenne experiment, or at least an approximate picture.

(c) Sitting in a dark room to repeat the HSP experiment would take a long time, but maybe you can go from your simulations here to design a psychophysical experiment simple enough that you can do it on one another. Can you measure the reliability of discrimination between the different patterns of x's that correspond to the signal being present or absent? Do you see an effect of “knowing when to look”? Do people seem to get better with practice? Can you calculate the theoretical limit to how well one can do this task? Do people get anywhere near this limit? This is an open ended problem.

Problem 14: A better analysis? Go back to the original paper by Aho et (1988) and see if you can give a more compelling comparison between thresholds and spontaneous isomerization rates. From Eq (28), we expect that the light intensity required for some criterion level of reliability scales as the square root of the dark noise level, but also depends on the total number of rhodospin molecules over which the subject must integrate. Can you estimate this quantity for the experiments on frogs and humans? Does this lead to an improved version of Fig 8? Again, this is an open ended problem.

The dominant role of spontaneous isomerization as a source of dark noise leads to a wonderfully counterintuitive result, namely that the photoreceptor which is

¹¹ The sign of this prediction is important. If we were looking for more reliable behaviors at higher temperatures, there could be many reasons for this, such as quicker responses of the muscles. Instead, the prediction is that we should see more reliable behavior as you cool down—all the way down to the temperature where behavior stops—and this is what is observed.

designed to maximize the signal-to-noise ratio for detection of dim lights will allow a significant number of photons to pass by undetected. Consider a rod photoreceptor cell of length ℓ , with concentration C of rhodopsin; let the absorption cross section of rhodopsin be σ . [Do I need to explain the definition of cross sections, and/or the derivation of Beer's law?] As a photon passes along the length of rod, the probability that it will be absorbed (and, presumably, counted) is $p = 1 - \exp(-C\sigma\ell)$, suggesting that we should make C or ℓ larger in order to capture more of the photons. But, as we increase C or ℓ , we are increasing the number of rhodopsin molecules, $N_{\text{rh}} = CA\ell$, with A the area of the cell, so we also increase the rate of dark noise events, which occurs at a rate r_{dark} per molecule.

If we integrate over a time τ , we will see a mean number of dark events (spontaneous isomerizations) $\bar{n}_{\text{dark}} = r_{\text{dark}}\tau N_{\text{rh}}$. The actual number will fluctuate, with a standard deviation $\delta n = \sqrt{\bar{n}_{\text{dark}}}$. On the other hand, if n_{flash} photons are incident on the cell, the mean number counted will be $\bar{n}_{\text{count}} = n_{\text{flash}}p$. Putting these factors together we can define a signal-to-noise ratio

$$SNR \equiv \frac{\bar{n}_{\text{count}}}{\delta n} = n_{\text{flash}} \frac{[1 - \exp(-C\sigma\ell)]}{\sqrt{CA\ell r_{\text{dark}}\tau}}. \quad (29)$$

The absorption cross section σ and the spontaneous isomerization rate r_{dark} are properties of the rhodopsin molecule, but as the rod cell assembles itself, it can adjust both its length ℓ and the concentration C of rhodopsin; in fact these enter together, as the product $C\ell$. When $C\ell$ is larger, photons are captured more efficiently and this leads to an increase in the numerator, but there also are more rhodopsin molecules and hence more dark noise, which leads to an increase in the denominator. Viewed as a function of $C\ell$, the signal-to-noise ratio has a maximum at which these competing effects balance; working out the numbers one finds that the maximum is reached when $C\ell \sim 1.26/\sigma$, and we note that all the other parameters have dropped out. In particular, this means that the probability of an incident photon *not* being absorbed is $1 - p = \exp(-C\sigma\ell) \sim e^{-1.26} \sim 0.28$. Thus, to maximize the signal-to-noise ratio for detecting dim flashes of light, nearly 30% of photons should pass through the rod without being absorbed (!). **Say something about how this compares with experiment!**

Problem 15: Escape from the tradeoff. Derive for yourself the numerical factor $(C\ell)_{\text{opt}} \sim 1.26/\sigma$. Can you see any way to design an eye which gets around this tradeoff between more efficient counting and extra dark noise? Hint: Think about what you see looking into a cat's eyes at night.

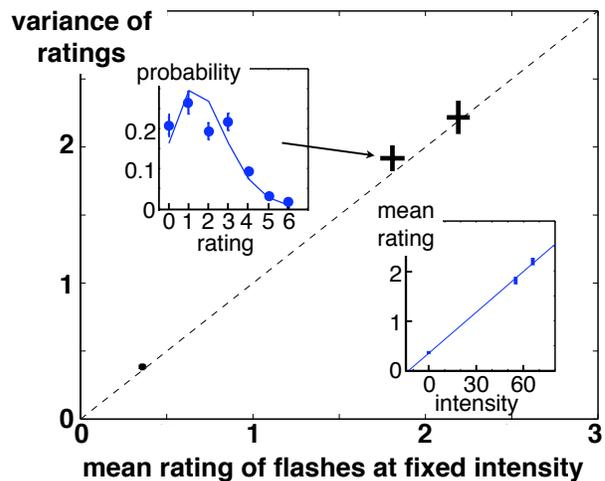


FIG. 9 Results of experiments in which observers are asked to rate the intensity of dim flashes, including blanks, on a scale from 0 to 6. Main figure shows that the variance of the ratings at fixed intensity is equal to the mean, as expected if the ratings are Poisson distributed. Insets show that the full distribution is approximately Poisson (upper) and that the mean rating is linearly related to the flash intensity, measured here as the mean number of photons delivered to the cornea. From Sakitt (1972).

If this is all correct, it should be possible to coax human subjects into giving responses that reflect the counting of individual photons, rather than just the summation of multiple counts up to some threshold of confidence or reliability. Suppose we ask observers not to say yes or no, but rather to rate the apparent intensity of the flash, say on a scale from 0 to 7. Remarkably, as shown in Fig 9, in response to very dim flashes interspersed with blanks, at least some observers will generate rating that, given the intensity, are approximately Poisson distributed: the variance of the ratings is essentially equal to the mean, and even the full distribution of ratings over hundreds of trials is close to Poisson. Further, the mean rating is linearly related to the light intensity, with an offset that agrees with other measurements of the dark noise level. Thus, the observers behaves exactly as if she can give a rating that is equal to the number of photons counted. This astonishing result would be almost too good to be true were it not that some observers deviate from this ideal behavior—they starting counting at two or three, but otherwise follow all the same rules.

While the phenomena of photon counting are very beautiful, one might worry that this represents just a very small corner of vision. Does the visual system continue to count photons reliably even when it's not completely dark outside? To answer this let's look at vision in a rather different animal, as in Fig 10. When you look down on the head of a fly, you see—almost to the exclusion of anything else—the large compound eyes. Each little hexagon that you see on the fly's head is a sepa-

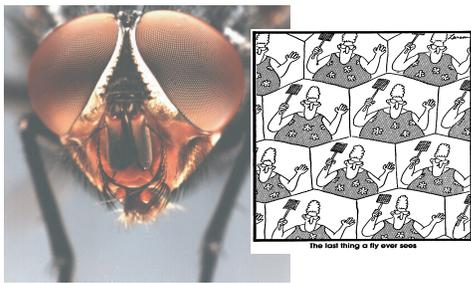


FIG. 10 The fly’s eye(s). At left a photograph taken by H Leertouwer at the Rijksuniversiteit Groningen, showing (even in this poor reproduction) the hexagonal lattice of lenses in the compound eye. This is the blowfly *Calliphora vicina*. At right, a schematic of what a fly might see, due to Gary Larson. The schematic is incorrect because each lens actually looks in a different direction, so that whole eye (like ours) only has one image of the visual world. In our eye the “pixelation” of the image is enforced by the much less regular lattice of receptors on the retina; in the fly pixelation occurs already with the lenses.

rate lens, and in large flies there are $\sim 5,000$ lenses in each eye, with approximately 1 receptor cell behind each lens, and roughly 100 brain cells per lens devoted to the processing of visual information. The lens focuses light on the receptor, which is small enough to act as an optical waveguide. Each receptor sees only a small portion of the world, just as in our eyes; one difference between flies and us is that diffraction is much more significant for organisms with compound eyes—because the lenses are so small, flies have an angular resolution of about 1° , while we do about $100\times$ better. [Add figure to emphasize similarity of two eye types.]

The last paragraph was a little sloppy (“approximately one receptor cell”), so let’s try to be more precise. For flies there actually are eight receptors behind each lens. Two provide sensitivity to polarization and some color vision, which we will ignore here. The other six receptors look out through the same lens in different directions, but as one moves to neighboring lenses one finds that there is one cell under each of six neighboring lenses which looks in the same direction. Thus these six cells are equivalent to one cell with six times larger photon capture cross section, and the signals from these cells are collected and summed in the first processing stage (the lamina); one can even see the expected six fold improvement in signal to noise ratio, in experiments we’ll describe shortly.¹²

Because diffraction is such a serious limitation, one might expect that there would be fairly strong selection

¹² Talk about the developmental biology issues raised by these observations, and the role of the photoreceptors as a model system in developmental decision making. For example, Lubensky et al (2011). Not sure where to put this, though.

for eyes that make the most of the opportunities within these constraints. Indeed, there is a beautiful literature on optimization principles for the design of the compound eye; the topic even makes an appearance in Feynman’s undergraduate physics lectures. Roughly speaking (Fig 11), we can think of the fly’s head as being a sphere of radius R , and imagine that the lens are pixels of linear dimension d on the surface. Then the geometry determines an angular resolution (in radians) of $\delta\phi_{\text{geo}} \sim d/R$; resolution gets better if d gets smaller. On the other hand, diffraction through an aperture of size d creates a blur of angular width $\delta\phi_{\text{diff}} \sim \lambda/d$, where $\lambda \sim 500$ nm is the wavelength of the light we are trying to image; this limit of course improves as the aperture size d gets larger. Although one could try to give a more detailed theory, it seems clear that the optimum is reached when the two different limits are about equal, corresponding to an optimal pixel size

$$d_* \sim \sqrt{\lambda R}. \quad (30)$$

This is the calculation in the Feynman lectures, and Feynman notes that it gives the right answer within 10% in the case of a honey bee.

A decade before Feynman’s lectures, Barlow had derived the same formula and went into the drawers of the natural history museum in Cambridge to find a variety of insects with varying head sizes, and he verified that the pixel size really does scale with the square root of the head radius, as shown in Fig 12. I think this work should be more widely appreciated, and it has several features we might like to emulate. First, it explicitly brings measurements on many species together in a quantitative way. Second, the fact that multiple species can put

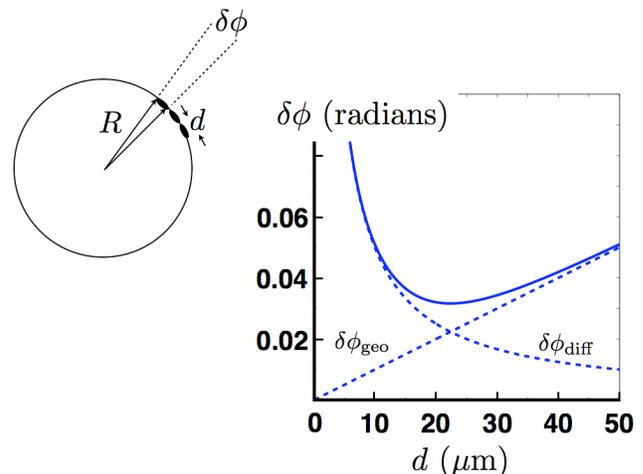


FIG. 11 At left, a schematic of the compound eye, with lenses of width d on the surface of a spherical eye with radius R . At right, the angular resolution of the eye as a function of the lens size, showing the geometric ($\delta\phi_{\text{geo}} \sim d/R$) and diffraction ($\delta\phi_{\text{diff}} \sim \lambda/d$) contributions in dashed lines; the full resolution in solid lines.

onto the same graph is not a phenomenological statement about, for example, scaling of one body part relative to another, but rather is based on a clearly stated physical principle. Finally, and most importantly for our later discussion in this course, Barlow makes an important transition: rather than just asking whether a biological system approaches the physical limits to performance, he assumes that the physical limits are reached and uses this hypothesis to predict something else about the structure of the system. This is, to be sure, a simple example, but an early and interesting example nonetheless.¹³

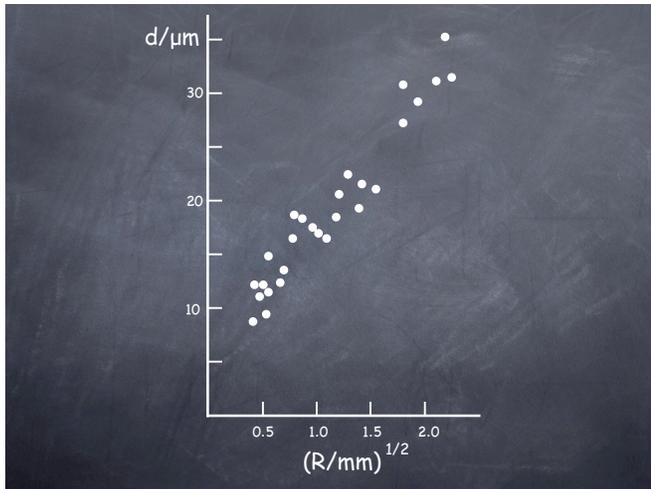


FIG. 12 The size of lenses in compound eyes as a function of head size, across many species of insect. From Barlow (1952).

[Should also point back to Mallock!]

Pushing toward diffraction-limited optics can't be the whole story, since at low light levels having lots of small pixels isn't much good—so few photons are captured in each pixel that one has a dramatic loss of intensity resolution. There must be some tradeoff between spatial resolution and intensity resolution, and the precise form of this tradeoff will depend on the statistical structure of the input images (if you are looking at clouds it will be different than looking at tree branches). The difficult question is how to quantify the relative worth of extra resolution in space vs intensity, and it has been suggested

¹³ This example also raises an interesting question. In Fig 12, each species of insect is represented by a single point. But not all members of the same species are the same size, as you must have noticed. Is the relationship between R and d that optimizes function preserved across the natural sizes variations among individuals? Does it matter whether the size differences are generated by environmental or genetic factors? This is a question about the reproducibility of spatial structures in development, a question we will come back to (albeit in simpler forms) in Section III.C. It would be good, though, if someone just *measured* the variations in eye dimensions across many individuals!

that the right way to do this is to count bits—design the eye not to maximize resolution, but rather to maximize the information that can be captured about the input image. This approach was a semi-quantitative success, showing how insects that fly late at night or with very high speeds (leading to blurring by photoreceptors with finite time resolution) should have less than diffraction limited spatial resolving power. I still think there are open questions here, however.

Coming back to the question of photon counting, one can record the voltage signals in the photoreceptor cells and detect single photon responses, as in vertebrates. If we want to see what happens at higher counting rates, we have to be sure that we have the receptor cells in a state where they don't "run down" too much because the increased activity. In particular, the rhodopsin molecule itself has to be recycled after it absorbs a photon. In animals with backbones, this actually happens not within the photoreceptor, but in conjunction with other cells that form the pigment epithelium. In contrast, in invertebrates the "resetting" of the rhodopsin molecule occurs within the receptor cell and can even be driven by absorption of additional long wavelength photons. Thus, if you want to do experiments at high photon flux on isolated vertebrate photoreceptors, there is a real problem of running out of functional rhodospin, but this doesn't happen in the fly's eye. Also, the geometry of the fly's eye makes it easier to do stable intracellular measurements without too much dissection.

To set the stage for experiments at higher counting rates, consider a simple model in which each photon arriving at time t_i produces a pulse $V_0(t - t_i)$, and these pulses just add to give the voltage [maybe there should be a sketch showing the summation of pulses to give the total voltage]

$$V(t) = V_{\text{DC}} + \sum_i V_0(t - t_i), \quad (31)$$

where V_{DC} is the constant voltage that one observes across the cell membrane in the absence of light. In Section A.1, we can find the distribution of the arrival times $\{t_i\}$ on the hypothesis that the photons arrive as a Poisson process with a time dependent rate $r(t)$; from Eq (A13) we have

$$P[\{t_i\}|r(t)] = \exp\left[-\int_0^T d\tau r(\tau)\right] \frac{1}{N!} r(t_1)r(t_2)\cdots r(t_N), \quad (32)$$

where $r(t)$ is the rate of photon arrivals—the light intensity in appropriate units. To compute the average voltage response to a given time dependent light intensity, we have to do a straightforward if tedious calculation:

$$\left\langle \sum_i V_0(t - t_i) \right\rangle = \sum_{N=0}^{\infty} \int_0^T d^N t_i P[\{t_i\}|r(t)] \sum_i V_0(t - t_i). \quad (33)$$

This looks a terrible mess. Actually, it's not so bad, and one can proceed systematically to do all of the integrals. Once you have had some practice, this isn't too difficult, but the first time through it is a bit painful, so I'll push the details off into Section A.1, along with all the other details about Poisson processes. When the dust settles [leading up to Eq (A64)], the voltage responds linearly to the light intensity,

$$\langle V(t) \rangle = V_{\text{DC}} + \int_{-\infty}^{\infty} dt' V_0(t-t') r(t'). \quad (34)$$

In particular, if we have some background photon counting rate \bar{r} that undergoes fractional modulations $C(t)$, so that

$$r(t) = \bar{r}[1 + C(t)], \quad (35)$$

then there is a linear response of the voltage to the contrast C ,

$$\langle \Delta V(t) \rangle = \bar{r} \int_{-\infty}^{\infty} dt' V_0(t-t') C(t'). \quad (36)$$

Recall that such integral relationships (convolutions) simplify when we use the Fourier transform. For a function of time $f(t)$ we will define the Fourier transform with the conventions

$$\tilde{f}(\omega) = \int_{-\infty}^{\infty} dt e^{+i\omega t} f(t), \quad (37)$$

$$f(t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega t} \tilde{f}(\omega). \quad (38)$$

Then, for two functions of time $f(t)$ and $g(t)$, we have

$$\int_{-\infty}^{\infty} dt e^{+i\omega t} \left[\int_{-\infty}^{\infty} dt' f(t-t') g(t') \right] = \tilde{f}(\omega) \tilde{g}(\omega). \quad (39)$$

Problem 16: Convolutions. Verify the “convolution theorem” in Eq (39). If you need some reminders, see, for example, Lighthill (1958).

Armed with Eq (39), we can write the response of the photoreceptor in the frequency domain,

$$\langle \Delta \tilde{V}(\omega) \rangle = \bar{r} \tilde{V}_0(\omega) \tilde{C}(\omega), \quad (40)$$

so that there is a transfer function, analogous to impedance relating current and voltage in an electrical circuit,

$$\tilde{T}(\omega) \equiv \frac{\langle \Delta \tilde{V}(\omega) \rangle}{\tilde{C}(\omega)} = \bar{r} \tilde{V}_0(\omega). \quad (41)$$

Recall that this transfer function is a complex number at every frequency, so it has an amplitude and a phase,

$$\tilde{T}(\omega) = |\tilde{T}(\omega)| e^{i\phi_T(\omega)}. \quad (42)$$

The units of \tilde{T} are simply voltage per contrast. The interpretation is that if we generate a time varying contrast $C(t) = C \cos(\omega t)$, then the voltage will also vary at frequency ω ,

$$\langle \Delta V(t) \rangle = |\tilde{T}(\omega)| C \cos[\omega t - \phi_T(\omega)]. \quad (43)$$

[Should we have one extra problem to verify this last equation? Or is it obvious?]

If every photon generates a voltage pulse $V_0(t)$, but the photons arrive at random, then the voltage must fluctuate. To characterize these fluctuations, we'll use some of the general apparatus of correlation functions and power spectra. A review of these ideas is given in Appendix A.2.

We want to analyze the fluctuations of the voltage around its mean, which we will call $\delta V(t)$. By definition, the mean of this fluctuation is zero, $\langle \delta V(t) \rangle = 0$. There is a nonzero variance, $\langle [\delta V(t)]^2 \rangle$, but to give a full description we need to describe the covariance between fluctuations at different times, $\langle \delta V(t) \delta V(t') \rangle$. Importantly, we are interested in systems that have no internal clock, so this covariance or correlation can't depend separately on t and t' , only on the difference. More formally, if *we* shift *our* clock by a time τ , this can't matter, so we must have

$$\langle \delta V(t) \delta V(t') \rangle = \langle \delta V(t + \tau) \delta V(t' + \tau) \rangle; \quad (44)$$

this is possible only if

$$\langle \delta V(t) \delta V(t') \rangle = C_V(t - t'), \quad (45)$$

where $C_V(t)$ is the “correlation function of V .” Thus, invariance under time translations restricts the form of the covariance. Another way of expressing time translation invariance in the description of random functions is to say that any particular wiggle in plotting the function is equally likely to occur at any time. This property also is called “stationarity,” and we say that fluctuations that have this property are stationary fluctuations.

In Fourier space, the consequence of invariance under time translations can be stated more simply—if we compute the covariance between two frequency components, we find

$$\langle \delta \tilde{V}(\omega_1) \delta \tilde{V}(\omega_2) \rangle = 2\pi \delta(\omega_1 + \omega_2) S_V(\omega_1), \quad (46)$$

where $S_V(\omega)$ is called the power spectrum (or power spectral density) of the voltage V . Remembering that $\delta \tilde{V}(\omega)$ is a complex number, it might be more natural to write this as

$$\langle \delta \tilde{V}(\omega_1) \delta \tilde{V}^*(\omega_2) \rangle = 2\pi \delta(\omega_1 - \omega_2) S_V(\omega_1). \quad (47)$$

Time translation invariance thus implies that fluctuations at different frequencies are independent.¹⁴ This makes sense, since if (for example) fluctuations at 2 Hz and 3 Hz were correlated, we could form beats between these components and generate a clock that ticks every second. Finally, the Wiener–Khinchine theorem states that the power spectrum and the correlation function are a Fourier transform pair,

$$S_V(\omega) = \int d\tau e^{+i\omega\tau} C_V(\tau), \quad (48)$$

$$C_V(\tau) = \int \frac{d\omega}{2\pi} e^{-i\omega\tau} S_V(\omega). \quad (49)$$

Notice that

$$\langle [\Delta V(t)]^2 \rangle \equiv C_V(0) = \int \frac{d\omega}{2\pi} S_V(\omega). \quad (50)$$

Thus we can think of each frequency component as having a variance $\sim S_V(\omega)$, and by summing these components we obtain the total variance.

Problem 17: More on stationarity. Consider some fluctuating variable $x(t)$ that depends on time, with $\langle x(t) \rangle = 0$. Show that, because of time translation invariance, higher order correlations among Fourier components are constrained:

$$\langle \tilde{x}(\omega_1) \tilde{x}^*(\omega_2) \tilde{x}^*(\omega_3) \rangle \propto 2\pi \delta(\omega_1 - \omega_2 - \omega_3) \quad (51)$$

$$\langle \tilde{x}(\omega_1) \tilde{x}(\omega_2) \tilde{x}^*(\omega_3) \tilde{x}^*(\omega_4) \rangle \propto 2\pi \delta(\omega_1 + \omega_2 - \omega_3 - \omega_4). \quad (52)$$

If you think of \tilde{x}^* (or \tilde{x}) as being analogous to the operators for creation (or annihilation) of particles, explain how these relations are related to conservation of energy for scattering in quantum systems.

Problem 18: Brownian motion in a harmonic potential. [The harmonic oscillator gets used more than once, of course; check for redundancy among problems in different sections!] Consider a particle of mass m hanging from a spring of stiffness κ , surrounded through a fluid. The effect of the fluid is, on average, to generate a drag force, and in addition there is a ‘Langevin force’ that describes the random collisions of the fluid molecules with the particle, resulting in Brownian motion. The equation of motion is

$$m \frac{d^2 x(t)}{dt^2} + \gamma \frac{dx(t)}{dt} + \kappa x(t) = \eta(t), \quad (53)$$

where γ is the drag coefficient and $\eta(t)$ is the Langevin force. A standard result of statistical mechanics is that the correlation function of the Langevin force is

$$\langle \eta(t) \eta(t') \rangle = 2\gamma k_B T \delta(t - t'), \quad (54)$$

where T is the absolute temperature and $k_B = 1.36 \times 10^{-23}$ J/K is Boltzmann’s constant.

(a.) Show that the power spectrum of the Langevin force is $S_\eta(\omega) = 2\gamma k_B T$, independent of frequency. Fluctuations with such a constant spectrum are called ‘white noise.’

(b.) Fourier transform Eq (53) and solve, showing how $\tilde{x}(\omega)$ is related to $\tilde{\eta}(\omega)$. Use this result to find an expression for the power spectrum of fluctuations in x , $S_x(\omega)$.

(c.) Integrate the power spectrum $S_x(\omega)$ to find the total variance in x . Verify that your result agrees with the equipartition theorem,

$$\left\langle \frac{1}{2} \kappa x^2 \right\rangle = \frac{1}{2} k_B T. \quad (55)$$

Hint: The integral over ω can be done by closing a contour in the complex plane.

(d.) Show that the power spectrum of the velocity, $S_v(\omega)$, is related to the power spectrum of position through

$$S_v(\omega) = \omega^2 S_x(\omega). \quad (56)$$

Using this result, verify the other prediction of the equipartition theorem for this system,

$$\left\langle \frac{1}{2} m v^2 \right\rangle = \frac{1}{2} k_B T. \quad (57)$$

Now we have a language for describing the signals and noise in the receptor cell voltage, by going to the frequency domain. What does this have to do with counting photons? The key point is that we can do a calculation similar to the derivation of Eq (40) for $\langle \Delta V(t) \rangle$ to show that, at $C = 0$, the voltage will undergo fluctuations—responding to the random arrival of photons—with power spectrum

$$N_V(\omega) = \bar{r} |V_0(\omega)|^2. \quad (58)$$

We call this N_V because it is noise. The noise has a spectrum shaped by the pulses V_0 , and the magnitude is determined by the photon counting rate; again see Appendix A.1 for details.

Notice that both the transfer function and noise spectrum depend on the details of $V_0(t)$. In particular, because this pulse has finite width in time, the transfer function gets smaller at higher frequencies. Thus if you watch a flickering light, the strength of the signal transmitted by your photoreceptor cells will decrease with increasing frequency.

The crucial point is that, for an ideal photon counter, although higher frequency signals are attenuated the signal-to-noise ratio actually doesn’t depend on frequency. Thus if we form the ratio

$$\frac{|\tilde{T}(\omega)|^2}{N_V(\omega)} = \frac{|\bar{r} \tilde{V}_0(\omega)|^2}{\bar{r} |V_0(\omega)|^2} = \bar{r}, \quad (59)$$

we just recover the photon counting rate, independent of details. Since this is proportional to the signal-to-noise ratio for detecting contrast modulations $\tilde{C}(\omega)$, we expect that real photodetectors will give less than this ideal value. [Should be able to make a crisper statement here—is it a theorem? Prove it, or give the proof as a problem?]

¹⁴ Caution: this is true only at second order; it is possible for different frequencies to be correlated when we evaluate products of three or more terms. See the next problem for an example.

Problem 19: Frequency vs counting rate. [Need to give more guidance through this problem! Step by step ...] If we are counting photons at an average rate \bar{r} , you might think that it is easier to detect variations in light intensity at a frequency $\omega \ll \bar{r}$ than at higher frequencies, $\omega \gg \bar{r}$; after all, in the high frequency case, the light changes from bright to dim and back even before (on average) a single photon has been counted. But Eq (59) states that the signal-to-noise ratio for detecting contrast in an ideal photon counter is independent of frequency, counter to this intuition. Can you produce a simple simulation to verify the predictions of Eq (59)? As a hint, you should think about observing the photon arrivals over a time T such that $\bar{r}T \gg 1$. Also, if you are looking for light intensity variations of the form $r(t) = \bar{r}[1 + C \cos(\omega t)]$, you should process the photon arrival times $\{t_i\}$ to form a signal $s = \sum_i \cos(\omega t_i)$.

So now we have a way of testing the photoreceptors: Measure the transfer function $\tilde{T}(\omega)$ and the noise spectrum $N_V(\omega)$, form the ratio $|\tilde{T}(\omega)|^2/N_V(\omega)$, and compare this with the actual photon counting rate \bar{r} . This was done for the fly photoreceptors, with the results shown in Fig 13. It's interesting to look back at the original papers and understand how they calibrated the measurement of \bar{r} (I'll leave this as an exercise for you!). [This account of the experiments is too glib. I will go back to expand and clarify. Rob has also offered new versions of the figures.]

What we see in Fig 13 is that, over some range of frequencies, the performance of the fly photoreceptors is close to the level expected for an ideal photon counter. It's interesting to see how this evolves as we change the mean light intensity, as shown in Fig 14. The performance of the receptors tracks the physical optimum up

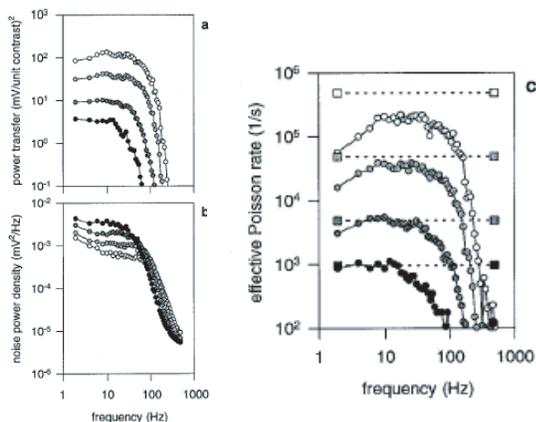


FIG. 13 Signal and noise in fly photoreceptors, with experiments at four different mean light intensities, from de Ruyter van Steveninck & Laughlin (1996b). (a) Transfer function $|\tilde{T}(\omega)|^2$ from contrast to voltage. (b) Power spectrum of voltage noise, $N_V(\omega)$. (c) The ratio $|\tilde{T}(\omega)|^2/N_V(\omega)$, which would equal the photon counting rate if the system were ideal; dashed lines show the actual counting rates.

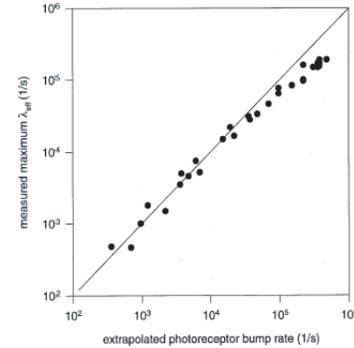


FIG. 14 Performance of fly photoreceptors vs light intensity. [Should redraw this, and label with consistent notation.] Having measured the quantity $\lambda_{\text{eff}} = |\tilde{T}(\omega)|^2/N_V(\omega)$, as in Fig 13, we plot the maximum value (typically at relatively low frequencies) vs the actual photon counting rate \bar{r} . We see that, over an enormous dynamic range, the signal-to-noise ratio tracks the value expected for an ideal photon counter.

to counting rates of $\bar{r} \sim 10^5$ photons/s. Since the integration time of the receptors is ~ 10 ms, this means that the cell can count, almost perfectly, up to about 1000.

An important point about these results is that they wouldn't work if the simple model were literally true. At low photon counting rates \bar{r} , the pulse V_0 has an amplitude of several millivolts, as you can work out from panel (a) in Fig 13. If we count $\sim 10^3$ events, this should produce a signal of several Volts, which is absolutely impossible in a real cell! What happens is that the system has an automatic gain control which reduces the size of the pulse V_0 as the light intensity is increased. Remarkably, this gain control or *adaptation* occurs while preserving (indeed, enabling) nearly ideal photon counting. Thus as the lights go up, the response to each photon become smaller (and, if you look closely, faster), but no less reliable.

Problem 20: Looking at the data. Explain how the data in Fig 13 provide evidence for the adaption of the pulse V_0 with changes in the mean light intensity.

[This seems a little brief! Maybe there should be a summary of what has happened, what we conclude ... also explain where the loose ends remain vs where things are solid.] These observations on the ability of the visual system to count single photons—down to the limit set by thermal noise in rhodopsin and up to counting rates of $\sim 10^5 \text{ s}^{-1}$ —raise questions at several different levels:

1. At the level of single molecules, we will see that the performance of the visual system depends crucially on the dynamics of rhodopsin itself. In particular, the structural response of the molecule to photon absorption is astonishingly fast, while the dark noise level means that the rate of spontaneous structural changes is extremely slow.

2. At the level of single cells, there are challenges in understanding how a network of biochemical reactions converts the structural changes of single rhodopsin molecules into macroscopic electrical currents across the rod cell membrane.

3. At the level of the retina as a whole, we would like to understand how these signals are integrated without being lost into the inevitable background of noise. Also at the level of the retina, we need to understand how single photon signals are encoded into the stereotyped pulses that are the universal language of the brain.

4. At the level of the whole organism, there are issues about how the brain learns to make the discriminations that are required for optimal performance.

In the next sections we'll look at these questions, in order.

It is a pleasure to read classic papers, and surely Hecht et al (1942) and van der Velden (1944) are classics, as is the discussion of dark noise by Barlow (1956). The pre-history of the subject, including the story about Lorentz, is covered by Bouman (1961). The general idea that our perceptual "thresholds" are really thresholds for discrimination against background noise with some criterion level of reliability made its way into quantitative psychophysical experiments in the 1950s and 60s, and this is now (happily) a standard part of experimental psychology; the canonical treatment is by Green and Swets (1966). The origins of these ideas are an interesting mix of physics and psychology, developed largely for radar in during World War II, and a summary of this early work is in the MIT Radar Lab series (Lawson & Uhlenbeck 1950). Another nice mix of physics and psychology is the revisiting of the original photon counting experiments using light sources with non-Poisson statistics (Teich et al 1982). The idea that random arrival of photons could limit our visual perception beyond the "just visible" was explored, early on, by de Vries (1943) and Rose (1948). Some of the early work by de Vries and coworkers on the physics of the sense organs (not just vision) is described in a lovely review (de Vries 1956). As a sociological note, de Vries was an experimental physicist with very broad interests, from biophysics to radiocarbon dating; for a short biography see de Waard (1960).

Barlow 1956: Retinal noise and absolute threshold. HB Barlow, *J Opt Soc Am* **46**, 634–639 (1956).

Bouman 1961: History and present status of quantum theory in vision. MA Bouman, in *Sensory Communication*, W Rosenblith, ed, pp 377–401 (MIT Press, Cambridge, 1961).

Green & Swets 1966: *Signal Detection Theory and Psychophysics*. DM Green, & JA Swets (Wiley, New York, 1966).

Hecht et al 1942: Energy, quanta and vision. S Hecht, S Schlaer & MH Pirenne, *J Gen Physiol* **25**, 819–840 (1942).

Lawson & Uhlenbeck 1950: *Threshold Signals*. MIT Radiation Laboratory Series vol 24 JL Lawson & GE Uhlenbeck (McGraw-Hill, New York, 1950).

Rose 1948: The sensitivity performance of the human eye on an absolute scale. A Rose, *J Opt Soc Am* **38**, 196–208 (1948).

Teich et al 1982: Multiplication noise in the human visual system at threshold. III: The role of non-Poisson quantum fluctuations. MC Teich, PR Prucnal, G Vannucci, ME Breton & WJ McGill, *Biol Cybern* **44**, 157–165 (1982).

van der Velden 1944: Over het aantal lichtquanta dat nodig is voor een lichtprikkel bij het menselijk oog. HA van der Velden, *Physica* **11**, 179–189 (1944).

de Vries 1943: The quantum character of light and its bearing upon threshold of vision, the differential sensitivity and visual acuity of the eye. Hl de Vries, *Physica* **10**, 553–564 (1943).

de Vries 1956: Physical aspects of the sense organs. Hl de Vries, *Prog Biophys Biophys Chem* **6**, 207–264 (1956).

de Waard 1960: Hessel de Vries, physicist and biophysicist. H de Waard, *Science* **131**, 1720–1721 (1960).

Single photon responses in receptor cells of the horseshoe crab were reported by Fuortes and Yeandle (1964). The series of papers from Baylor and co-workers on single photon responses in vertebrate rod cells, first from toads and then from monkeys, again are classics, well worth reading today, not least as examples of how to do quantitative experiments on biological systems. Aho, Donner, Reuter and co-workers have made a major effort to connect measurements on rod cells and ganglion cells with the behavior of the whole organism, using the toad as an example; among their results are the temperature dependence of dark noise (Fig 8), and the latency/anticipation results in Section I.D. The remarkable experiments showing that people really can count every photon are by Sakitt (1972). We will learn more about currents and voltages in cells very soon, but for background I have always liked Aidley's text, now in multiple editions; as is often the case, the earlier editions can be clearer and more compact.

Aidley 1998: *The Physiology of Excitable Cells, 4th Edition* DJ Aidley (Cambridge University Press, Cambridge, 1998).

Aho et al 1987: Retinal noise, the performance of retinal ganglion cells, and visual sensitivity in the dark-adapted frog. A-C Aho, K Donner, C Hydén, T Reuter & OY Orlov, *J Opt Soc Am A* **4**, 2321–2329 (1987).

Aho et al 1988: Low retinal noise in animals with low body temperature allows high visual sensitivity. A-C Aho, K Donner, C Hydén, LO Larsen & T Reuter, *Nature* **334**, 348–350 (1988).

Aho et al 1993: Visual performance of the toad (*Bufo bufo*) at low light levels: retinal ganglion cell responses and prey-catching accuracy. A-C Aho, K Donner, S Helenius, LO Larsen & T Reuter, *J Comp Physiol A* **172**, 671–682 (1993).

Baylor et al 1979a: The membrane current of single rod outer segments. DA Baylor, TD Lamb & K-W Yau, *J Physiol (Lond)* **288**, 589–611 (1979).

Baylor et al 1979b: Rod responses to single photons. DA Baylor, TD Lamb & K-W Yau, *J Physiol (Lond)* **288**, 613–634 (1979).

Baylor et al 1980: Two components of electrical dark noise in toad retinal rod outer segments. DA Baylor, G Matthews & K-W Yau, *J Physiol (Lond)* **309**, 591–621 (1980).

Baylor et al 1984: The photocurrent, noise and spectral sensitivity of rods of the monkey *Macaca fascicularis*. DA Baylor, BJ Nunn & JF Schnapf, *J Physiol (Lond)* **357**, 575–607 (1984).

Fuortes & Yeandle 1964: Probability of occurrence of discrete potential waves in the eye of *Limulus*. MGF Fuortes & S Yeandle, *J Gen Physiol* **47**, 443–463 (1964).

Sakitt 1972: Counting every quantum. B Sakitt, *J Physiol* **223**, 131–150 (1972).

For the discussion of compound eyes, useful background is contained in Stavenga and Hardie (1989), and in the beautiful compilation of insect brain anatomy by Strausfeld (1976), although this is hard to find; as an alternative there is an online atlas, <http://flybrain.neurobio.arizona.edu/Flybrain/html/>. There is also the more recent Land & Nilsson (2002). Evidently Larson (2003) is an imperfect guide to these matters. Everyone should have a copy of the Feynman lectures (Feynman et al 1963), and check the chapters on vision. The early work by Barlow (1952) deserves more appreciation, as noted in the main text, and the realization that diffraction must be important for insect eyes goes back to Mallock (1894). For a gentle introduction to the wider set of ideas about scaling relations between different body parts, see McMahon & Bonner (1983). The experiments on signal-to-noise ratio in fly photoreceptors are by de Ruyter van Steveninck and Laughlin (1996a, 1996b). For a review of relevant ideas in Fourier analysis and related matters, see Appendix A.2 and Lighthill (1958). You should come back to the ideas of Snyder et al (Snyder 1977, Snyder et al 1977) near the end of the book, after we have covered some of the basics of information theory.

Barlow 1952: The size of ommatidia in apposition eyes. HB Barlow, *J Exp Biol* **29**, 667–674 (1952).

Feynman et al 1963: *The Feynman Lectures on Physics*. RP Feynman, RB Leighton & M Sands (Addison-Wesley, Reading, 1963).

Larson 2003: *The Complete Far Side*. G Larson (Andrews McNeel Publishing, Kansas City, 2003).

Land & Nilsson 2002: *Animal Eyes* MF Land & D-E Nilsson (Oxford University Press, Oxford, 2002).

Lighthill 1958: *Introduction to Fourier Analysis and Generalized Functions*. MJ Lighthill (Cambridge University Press, Cambridge, 1958)

Mallock 1894: Insect sight and the defining power of composite eyes. A Mallock, *Proc R Soc Lond* **55**, 85–90 (1894).

McMahon & Bonner 1983: *On Size and Life*. TA McMahon & JT Bonner (WH Freeman, New York, 1983).

de Ruyter van Steveninck & Laughlin 1996a: The rate of information transfer at graded-potential synapses. RR de Ruyter van Steveninck & SB Laughlin, *Nature* **379**, 642–645 (1996).

de Ruyter van Steveninck & Laughlin 1996b: Light adaptation and reliability in blowfly photoreceptors. R de Ruyter van Steveninck & SB Laughlin, *Int. J. Neural Syst.* **7**, 437–444 (1996)

Snyder 1977: Acuity of compound eyes: Physical limitations and design. AW Snyder, *J Comp Physiol* **116**, 161–182 (1977).

Snyder et al 1977: Information capacity of compound eyes. AW Snyder, DS Stavenga & SB Laughlin, *J Comp Physiol* **116**, 183–207 (1977).

Stavenga & Hardie 1989: *Facets of Vision*. DG Stavenga & RC Hardie, eds (Springer-Verlag, Berlin, 1989).

Strausfeld 1976: *Atlas of an Insect Brain*. N Strausfeld (Springer-Verlag, Berlin, 1976).

Finally, a few reviews that place the results on photon counting into a broader context.

Barlow 1981: Critical limiting factors in the design of the eye and visual cortex. HB Barlow, *Proc R Soc Lond Ser B* **212**, 1–34 (1981).

Bialek 1987: Physical limits to sensation and perception. W Bialek, *Ann Rev Biophys Biophys Chem* **16**, 455–478 (1987).

Rieke & Baylor 1998: Single photon detection by rod cells of the retina. F Rieke & DA Baylor, *Rev Mod Phys* **70**, 1027–1036 (1998).

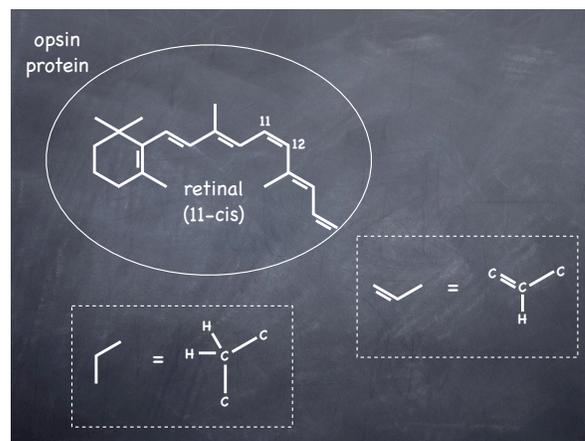


FIG. 15 Schematic structure of rhodopsin, showing the organic pigment retinal nestled in a pocket formed by the surrounding opsin protein. This conformation of the retinal is called 11-cis, since there is a rotation around the bond between carbons numbered 11 and 12 (starting at the lower right in the ring). Insets illustrate the conventions in such chemical structures, with carbons at nodes of the skeleton, and hydrogens not shown, but sufficient to make sure that each carbon forms four bonds.

B. Single molecule dynamics

To a remarkable extent, our ability to see in the dark is limited by the properties of rhodopsin itself, essentially because everything else works so well. Rhodopsin consists of a medium sized organic pigment, retinal, enveloped by a large protein, opsin (cf Fig 15). The primary photo-induced reaction is isomerization of the retinal, which ultimately couples to structural changes in the protein. The effort to understand the dynamics of these processes goes back to Wald's isolation of retinal (a vitamin A derivative) in the 1930s, his discovery of the isomerization, and the identification of numerous states through which the molecule cycles. The field was given a big boost by the discovery that there are bacterial rhodopsins, some of which serve a sensory function while others are energy transducing molecules, using the energy of the absorbed photon to pump protons across the cell membrane; the resulting difference in electrochemical potential for protons is a universal intermediate in cellular energy conversion, not just in bacteria but in us as well. [Maybe a pointer to channel rhodopsins would be good here too.]

By now we know much more than Wald did about the structure of the rhodopsin molecule [need to point to a better figure, more details].

While there are many remarkable features of the rhodopsin molecule, we would like to understand those particular features that contribute to the reliability of photon counting. First among these is the very low spontaneous isomerization rate, roughly once per thousand years. As we have seen, these photon-like events provide

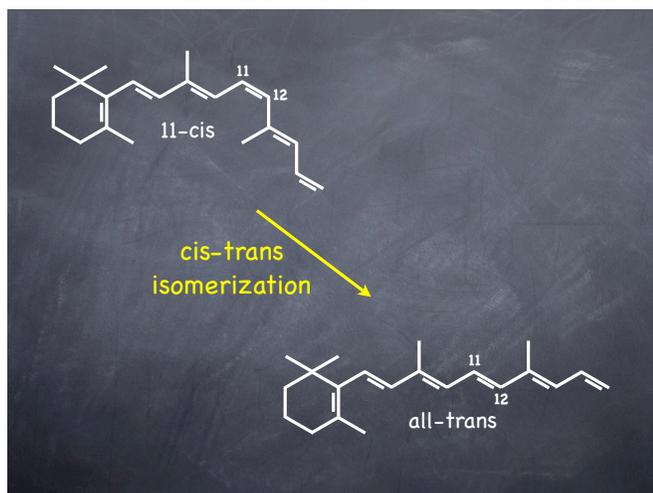


FIG. 16 Isomerization of retinal, the primary event at the start of vision. The π -bonds among the carbons favor planar structures, but there are still alternative conformations. The 11-cis conformation is the ground state of rhodopsin, and after photon absorption the molecule converts to the all-trans configuration. These different structures have different absorption spectra, as well as other, more subtle differences. Thus we can monitor the progress of the transition 11-cis \rightarrow all-trans essentially by watching the molecule change color, albeit only slightly. [Show the spectra!]

the dominant noise source that limits our ability to see in the dark, so there is a clear advantage to having the lowest possible rate. When we look at the molecules themselves, purified from the retina, we can “see” the isomerization reaction because the initial 11-cis state and the final all-trans states (see Fig 16) have different absorption spectra [add this to the figure]. For rhodopsin itself, the spontaneous isomerization rate is too slow to observe in a bulk experiment. If we isolate the pigment retinal, however, we find that it has a spontaneous isomerization rate of $\sim 1/\text{yr}$, so that a bottle of 11-cis retinal is quite stable, but the decay to all-trans is observable.

How can we understand that rhodopsin has a spontaneous isomerization rate 1000 \times less than that of retinal? The spontaneous isomerization is thermally activated, and has a large “activation energy” as estimated from the temperature dependence of the dark noise.¹⁵ It seems reasonable that placing the retinal molecule into the pocket formed by the protein opsin would raise the activation energy, essentially because parts of the protein need to be pushed out of the way in order for the retinal to rotate and isomerize. Although this sounds plausible, it’s probably wrong. If we write the

¹⁵ I am assuming here that the ideas of activation energy and Arrhenius behavior of chemical reaction rates are familiar. For more on this, see Section II.A.

dark isomerization rate as $r = Ae^{-E_{\text{act}}/k_B T}$, retinal and rhodopsin have the same value of the activation energy $E_{\text{act}} = 21.9 \pm 1.6 \text{ kcal/mole}$ [this is from measurements on rods; give the number in useful units! maybe footnote about difficulties of units] within experimental error, but different values of the prefactor A . If we look at photoreceptor cells that are used for daytime vision—the cones, which also provide us with sensitivity to colors, as discussed below [check where this gets done!]¹⁶—the dark noise level is higher (presumably single photon counting is unnecessary in bright light), but again this is a difference in the prefactor, not in the activation energy. As we will see when we discuss the theory of reaction rates in Section II.A, understanding prefactors is much harder than understanding activation energies, and I think we don’t really have a compelling theoretical picture that explains the difference between retinal and rhodopsin. [Fred Rieke gave me some pointers I have to chase down before deciding on that last sentence!]

The isolated retinal pigment isomerization at a rate that is faster than rhodopsin. On the other hand, if we excite the isolated retinal with a very short pulse of light, and follow the resulting changes in absorption spectrum, these photo-induced dynamics are not especially fast, with isomerization occurring at a rate $\sim 10^9 \text{ s}^{-1}$. Although this is fast compared to the reactions that we can see directly, it is actually so slow that it is comparable to the rate at which the molecule will re-emit the photon. We recall from quantum mechanics that the spontaneous emission rates from electronic excited states are constrained by sum rules if they are dipole-allowed. This means that emission lifetimes for visible photons are order 1 nanosecond for almost all of the simple cases. In a big molecule, there can be some re-arrangement of the molecular structure before the photon is emitted (see the discussion below), and this results in the emitted or fluorescent photon being of longer wavelength. Nonetheless, the natural time scale is nanoseconds, and the isomerization of retinal is not fast enough to prevent fluorescence and truly capture the energy of the photon with high probability.

Problem 21: Why nanoseconds? Explain why spontaneous emission of visible photons typically occurs with a rate $\sim 10^9 \text{ s}^{-1}$. [Need to explain where to start!]

Now fluorescence is a disaster for visual pigment—not only don’t you get to count the photon where it was absorbed, it might get counted somewhere else, blurring the image. In fact rhodopsin does not fluoresce: The quantum yield or branching ratio for fluorescence is $\sim 10^{-5}$.

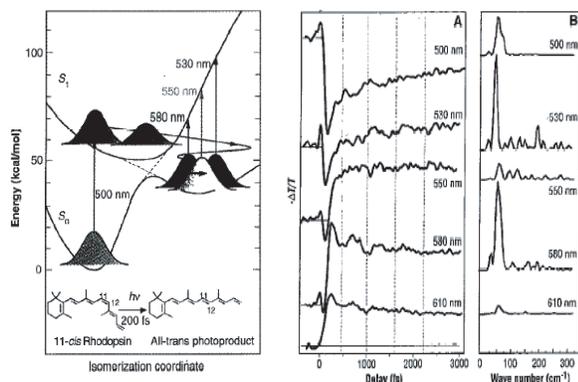


FIG. 17 [This needs to be redrawn; maybe two figures to make different points? Convert all the units once and for all?] Femtosecond dynamics of rhodopsin, from Wang et al (1994). At left, schematic potential energy surfaces in the electronic ground and excited states. At right, panel (A) shows transient absorption spectra following a 35 fs pulse of 500 nm light. Panel (B) shows the magnitude of the Fourier transform of the time dependent absorption at each of several wavelengths, illustrating the oscillations expected if the vibrational dynamics is coherent. You might like to convert the kcal/mol and cm^{-1} into more conventional physical units!

If we imagine the molecule sitting in the excited state, transitioning to the ground state via fluorescence at a rate $\sim 10^9 \text{ s}^{-1}$, then to have a branching ratio of 10^{-5} the competing process must have a rate of $\sim 10^{14} \text{ s}^{-1}$. Thus, the rhodopsin molecule must leave the excited state by some process on a time scale of ~ 10 femtoseconds, which is extraordinarily fast. Indeed, for many years, every time people built faster pulsed lasers, they went back to rhodopsin to look at the initial events, culminating in the direct demonstration of femtosecond isomerization, making this one of the fastest molecular events ever observed.

The 11-cis and all trans configurations of retinal have different absorption spectra, and this is why we can observe the events following photon absorption as an evolution of the spectrum. The basic design of such experiments is to excite the molecules with a brief pulse of light, elevating them into the excited state, and then probe with another brief pulse after some delay. In the simplest version, one repeats the experiment many times with different choices of the delay and the energy or wavelength of the probe pulse. An example of the results from such an experiment are shown in Fig 17. The first thing to notice is that the absorption at a wavelength of 550 nm, characteristic of the all-trans structure, rises very quickly after the pulse which excites the system, certainly within tens of femtoseconds. In fact this experiment reveals all sorts of interesting structure, to which we will return below.

The combination of faster photon induced isomerization and slower thermal isomerization means that the

protein opsin acts as an electronic state selective catalyst: ground state reactions are inhibited, excited state reactions accelerated, each by orders of magnitude. It is fair to say that if these state dependent changes in reaction rate did not occur—that is, if the properties of rhodopsin were those of retinal—then we simply would not be able to see in the dark of night.

Problem 22: What would vision be like if ... ? Imagine that the spontaneous isomerization rate and quantum yield for photoisomerization in rhodopsin were equal to those in retinal. Estimate, quantitatively, what this would mean for our ability to see at night. [we should try to connect with real intensities at dusk etc.]

In order to make sense out of all of this, and get started in understanding how rhodopsin achieves its function, we need to understand something about electronic transitions in large molecules, as opposed to the case of atoms that we all learned about in our quantum mechanics classes. The absorption of a photon by an atom involves a transition between two electronic states, and this is also true for a large molecule. But for the atom the absorption line is very narrow, while for big molecules it is very broad. For rhodopsin, there is a nice way of measuring the absorption spectrum over a very large dynamic range, and this is to use the rod cell as a sensor. Instead of asking how much light is absorbed, we can try assuming¹⁶ that all absorbed photons have a constant probability of generating a pulse of current at the rod's output, and so we can adjust the light intensity at each wavelength to produce the same current. If the absorption is stronger, we need less light, and conversely more light if the absorption is weaker. The results of such an experiment are shown in Fig 18. It is beautiful that in this way one can follow the long wavelength tail of the spectrum down to cross-sections that are $\sim 10^{-5}$ of the peak. More qualitatively, we see that the width of the spectrum, say at half maximum, is roughly 20% of the peak photon energy, which is enormous in contrast with atomic absorption lines.

As an aside, the fact that one can follow the sensitivity of the photoreceptor cell deep into the long wavelength tail opens the possibility of asking a very different question about the function of these cells (and all cells). We recall that every cell in our bodies has the same genetic

¹⁶ This assumption can also be checked. It's true, but I think there have not been very careful measurements in the long wavelength tail, where something interesting might happen.

material, and hence the instructions for making all possible proteins. In particular, all photoreceptor cells have the ability to make all visual pigments. But the different classes of receptors—rods and the three kinds of cones—make different pigments, corresponding to different proteins surrounding more or less the same retinal molecule, and the resulting differences in absorption spectra provide the basis for color vision. If a single cone couldn't reliably turn on the expression of one rhodopsin gene, and turn off all of the others, then the retina wouldn't be able to generate a mix of spectral sensitivities, and we wouldn't see colors. But how “off” is “off”?

In a macaque monkey (not so different from us in these matters), “red” cones have their peak sensitivity at a wavelength ~ 570 nm, but at this wavelength the “blue” cones have sensitivities that are $\sim 10^5 \times$ reduced relative to their own peak. Since the peak absorption cross-sections are comparable, this tells us that the relative concentration of red pigments in the blue cones must be less than 10^{-5} . That is, the cell makes at least 10^5 times as much of the correct protein as it does of the incorrect proteins, which I always thought was pretty impressive.¹⁷

Returning to the absorption spectrum itself, we realize that a full treatment would describe molecules by doing

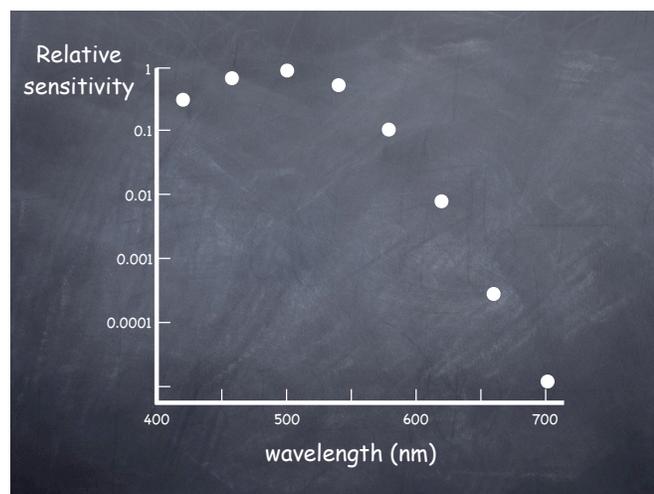


FIG. 18 Sensitivity of the rod photoreceptor as a function of wavelength. This is measured, as explained in the text, by adjusting the intensity of light to give a criterion output, so that very low sensitivity corresponds to shining a bright light, rather than measuring a small output. Redrawn from Baylor et al (1979a).

¹⁷ Many thanks to Denis Baylor for reminding me of this argument. Since there are $\sim 10^9$ rhodopsins in one cell, errors of even one part in 10^5 would mean that there are thousands of “wrong” molecules floating around. I wonder if this is true, or if the true errors are even smaller. [Apparently there is evidence that some cones are less precise about what defines “off;” should check this!]

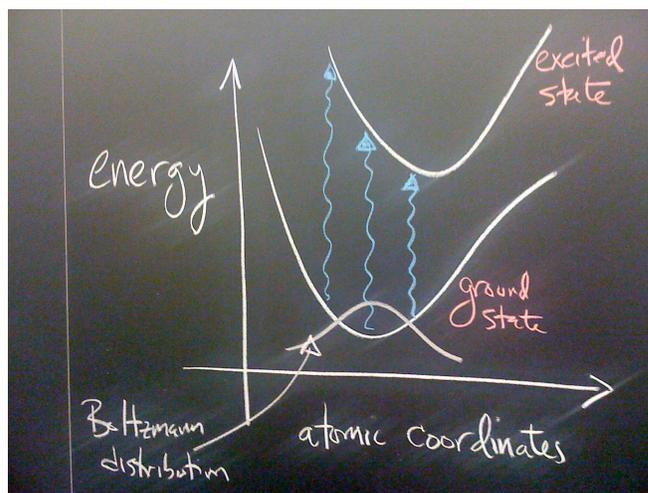


FIG. 19 Schematic of the electronic states in a large molecule, highlighting their coupling to motion of the nuclei. The sketch shows two states, with photon absorption (in blue) driving transitions between them. If we think in semi-classical terms, as explained in the text, then these transitions are ‘too fast’ for the atoms to move, and hence are vertical on such plots (the Franck–Condon approximation). Because the atomic coordinates fluctuate, as indicated by the Boltzmann distribution, the energy of the photon required to drive the transition also fluctuates, and this broadens the absorption spectrum.

the quantum mechanics of a combined system of electrons and nuclei. But the nuclei are very much heavier than the electrons, and hence move more slowly. More rigorously, the large ratio of masses means that we can think of solving the quantum mechanics of the electrons with the nuclei in fixed position, and then for each such atomic configuration the energy of the electrons contributes to the potential energy; as the nuclei move in this potential (whether classically or quantum mechanically) the electrons follow adiabatically.¹⁸ This is the Born–Oppenheimer approximation, which is at the heart of all attempts to understand molecular dynamics.¹⁹

Figure 19 shows the energy of two different electronic states, plotted schematically against (one of the) atomic coordinates. In the ground state, we know that there is some arrangement of the atoms that minimizes the en-

¹⁸ Because the electrons (mostly) follow the nuclei, I will use “nuclei” and “atoms” interchangeably in what follows.

¹⁹ I assume that most readers know something about the Born–Oppenheimer approximation, since it is a pretty classical subject. It is also one of the first adiabatic approximations in quantum mechanics. It took many years to realize that some very interesting things can happen in the adiabatic limit, notably the appearance of non-trivial phase factors in the adiabatic evolution of wave functions. Some of these ‘complications’ (to use a word from one of original papers) were actually discovered in the context of the Born–Oppenheimer approximation itself, but now we know that this circle of ideas is much bigger, extending out to quantum optics and quite exotic field theories.

ergy, and that in the neighborhood of this minimum the potential surface must look roughly like that of a system of Hookean springs. Once we lift the electrons into the first excited state, there is again some configuration of the atoms that minimizes the energy (unless absorbing one photon is enough to break the molecule apart!), but unless there is some symmetry this equilibrium configuration will be different than in the ground state, and the stiffness of the spring holding the molecule in this equilibrium configuration also will be different. Hence in Fig 19, the energy surfaces for the ground and excited states are shown displaced and with different curvatures.

It is important to realize that sketches such as that in Fig 19 are approximations in many senses. Most importantly, this sketch involves only one coordinate. You may be familiar with a similar idea in the context of chemical reactions, where out of all the atoms that move during the reaction we focus on one “reaction coordinate” that forms a path from the reactants to products; for more about this see Section II.A. One view is that this is just a convenience—we can’t draw in many dimensions, so we just draw one, and interpret the figure cautiously. Another view is that the dynamics *are* effectively one dimensional, either because there is a separation of time scales, or because we can change coordinates to isolate, for example, a single coordinate that couples to the difference in energy between the ground and excited electronic states. The cost of this reduction in dimensionality might be a more complex dynamics along this one dimension, for example with a “viscosity” that is strongly frequency dependent, which again means that we need to be cautious in interpreting the picture that we draw. In what follows I’ll start by being relatively informal, and try to become more precise as we go along.

In the limit that the atoms are infinitely heavy, they don’t move appreciably during the time required for an electronic transition. On the other hand, the positions of the atoms still have to come out of the Boltzmann distribution, since the molecule is in equilibrium with its environment at temperature T . In this limit, we can think of transitions between electronic states as occurring without atomic motion, corresponding to vertical lines on the schematic in Fig 19. If the photon happens to arrive when the atomic configuration is a bit to the left of the equilibrium point, then as drawn the photon energy needs to be larger in order to drive the transition; if the configuration is a bit to the right, then the photon en-

ergy is smaller. In this way, the Boltzmann distribution of atomic positions is translated into a broadening of the absorption line. In particular, the transition can occur with a photon that has very little energy if we happen to catch a molecule in the rightward tail of the Boltzmann distribution: the electronic transition can be made up partly from the energy of the photon and partly from energy that is “borrowed” from the thermal bath. As a result, the absorption spectrum should have a tail at long wavelengths, and this tail will be strongly temperature dependent, and this is observed in rhodopsin and other large molecules. Since our perception of color depends on the relative absorption of light by rhodopsins with different spectra, this means that there must be wavelengths such that the apparent color of the light will depend on temperature [need a pointer and refs for this .. maybe tell the story of de Vries and the hot tub?]

Concretely, if we imagine that the potential surfaces are perfect Hookean springs, but with displaced equilibrium positions, then we can relate the width of the spectrum directly to the magnitude of this displacement. In the ground state we have the potential

$$V_g(q) = \frac{1}{2}\kappa q^2, \quad (60)$$

and in the excited state we have

$$V_e(q) = \epsilon + \frac{1}{2}\kappa(q - \Delta)^2, \quad (61)$$

where ϵ is the minimum energy difference between the two electronic states and Δ is the shift in the equilibrium position, as indicated in Fig 20. With q fixed, the condition for absorbing a photon is that the energy $\hbar\Omega$ match the difference in electronic energies,

$$\hbar\Omega = V_e(q) - V_g(q) = \epsilon + \frac{1}{2}\kappa\Delta^2 - \kappa\Delta q. \quad (62)$$

The probability distribution of q when molecules are in the ground state is given by

$$P(q) = \frac{1}{Z} \exp\left[-\frac{V_g(q)}{k_B T}\right] = \frac{1}{\sqrt{2\pi\kappa k_B T/\kappa}} \exp\left[-\frac{\kappa q^2}{2k_B T}\right], \quad (63)$$

so we expect the cross-section for absorbing a photon of frequency Ω to have the form

$$\sigma(\Omega) \propto \int dq P(q) \delta\left[\hbar\Omega - \left(\epsilon + \frac{1}{2}\kappa\Delta^2 - \kappa\Delta q\right)\right] \quad (64)$$

$$\propto \int dq \exp\left[-\frac{\kappa q^2}{2k_B T}\right] \delta\left[\hbar\Omega - \left(\epsilon + \frac{1}{2}\kappa\Delta^2 - \kappa\Delta q\right)\right] \quad (65)$$

$$\propto \exp\left[-\frac{(\hbar\Omega - \hbar\Omega_{\text{peak}})^2}{4\lambda k_B T}\right], \quad (66)$$

where the peak of the absorption is at

$$\hbar\Omega_{\text{peak}} = \epsilon + \lambda, \quad (67)$$

and

$$\lambda = \frac{1}{2}\kappa\Delta^2 \quad (68)$$

is the energy required to distort the molecule into the equilibrium configuration of the excited state if we stay in the ground state.

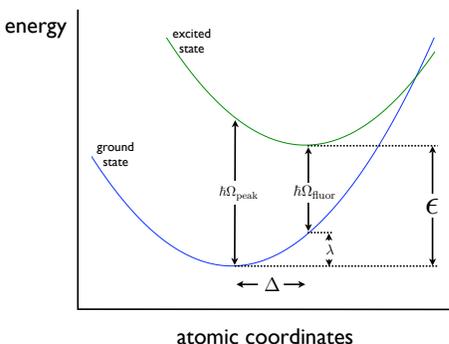


FIG. 20 The potential surfaces of Fig 19, redrawn in the special case where they are parabolic. Then, as in Eqs (60) through (68), there are just a few key parameters that determine the shape of the absorption spectrum and also the fluorescence emission. **Redraw figure to show that $\hbar\Omega_{\text{peak}} = \epsilon + \lambda$; ref to Eq (67).**

The energy λ is known, in different contexts, as the reorganization energy or the Stokes shift. If the molecule stays in the excited state for a long time, the distribution of coordinates will re-equilibrate to the Boltzmann distribution appropriate to $V_e(q)$, so that the most likely coordinate becomes $q = \Delta$. At this coordinate, if the molecule returns to the ground state by emitting a photon—fluorescence—the energy of this photon will be $\hbar\Omega_{\text{fluor}} = \epsilon - \lambda$. Thus the peak fluorescence is at lower energies, or red shifted from the absorption peak by an amount 2λ , as one can read off from Fig 20. This connects the width of the absorption band to the red shift that occurs in fluorescence, and for many molecules this prediction is correct, quantitatively, giving us confidence in the basic picture. **[I wonder if all of this needs more figures in order to be clear?]**

In the case of rhodopsin, the peak absorption is at a wavelength of 500 nm or an energy of $\hbar\Omega_{\text{peak}} = 2.5$ eV. The width of the spectrum is described roughly by a Gaussian with a standard deviation of $\sim 10\%$ of the peak energy, so that $2\lambda k_B T \sim (0.25 \text{ eV})^2$, or $\lambda \sim 1.25$ eV. Surely we can't take this seriously, since this reorganization energy is enormous, and would distort the molecule well beyond the point where we could describe the potential surfaces by Hookean springs. Amusingly, if we took

this result literally, the peak fluorescence would be at zero energy (!). Probably the correct conclusion is that there is a tremendously strong coupling between excitation of the electrons and motion of the atoms, and presumably this is related to the fact that photon absorption leads to very rapid structural changes.

Before proceeding, it would be nice to do an honest calculation that reproduces the intuition of Figs 19 and 20, and this is done in Section A.3. The results of the calculation show, in more detail, how the coupling of electronic states to the vibrational motion of the molecule can shape the absorption spectrum. If there is just one lightly damped vibrational mode, then the single sharp absorption line which we expect from atomic physics becomes a sequence of lines, corresponding to changing electronic state and exciting one, two, three, ... or more vibrational quanta. If there are many modes, and these modes are damped by interaction with other degrees of freedom, these “vibronic” lines merge into a smooth spectrum which we can calculate in a semi-classical approximation.

The coupling of electronic transitions to vibrational motion generates the phenomenon of Raman scattering—a photon is inelastically scattered, making a virtual transition to the electronically excited state and dropping back down to the ground state, leaving behind a vibrational quantum **[add a figure illustrating Raman scattering]**. The energy shifts of the scattered photons allow us to read off, directly, the frequencies of the relevant vibrational modes. With a bit more sophistication, we can connect the strength of the different lines to the coupling constants (e.g., the displacements Δ_i along each mode, generalizing the discussion above) that characterize the interactions between electronic and vibrational degrees of freedom. If everything works, it should be possible to reconstruct the absorption spectrum from these estimates of frequencies and couplings. This whole program has been carried through for Rhodospin. Importantly, in order to get everything right, one has to include motions which are effectively unstable in the excited state, presumably corresponding to the torsional motions that lead to cis–trans isomerization. **[This is all a little quick. On the other hand, there is a huge amount of detail here that might take us away from the goal. Advice is welcome!]**

Problem 23: Raman scattering. Take the students through a simple calculation of Raman scattering ...

If we try to synthesize all of these ideas into a single schematic, we might get something like Fig 21. If we take

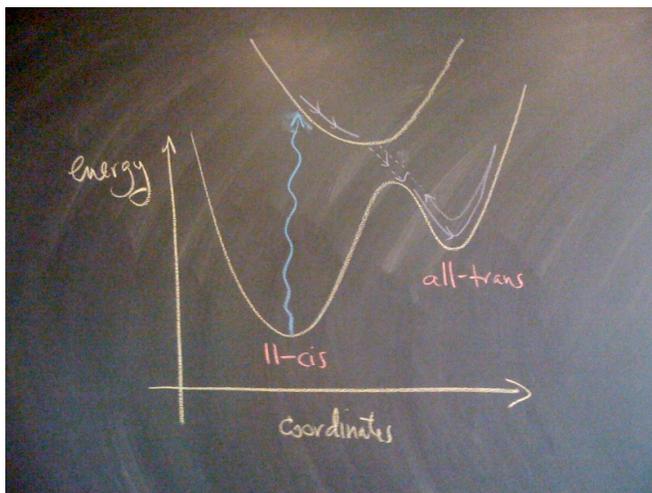


FIG. 21 Schematic model of the energy surfaces in Rhodopsin. The ground state has minima at both the 11-cis and the all-trans structures. A single excited state sits above this surface. At some intermediate structure, the surfaces come very close. At this point, the Born–Oppenheimer approximation breaks down, and there will be some mixing between the two states. A molecule lifted into the excited state by absorbing a photon slides down the upper surface, and can pass non-adiabatically into the potential well whose minimum is at all-trans.

this picture seriously, then after exciting the molecule with a pulse of light, we should see the disappearance of the absorption band associated with the 11-cis structure, the gradual appearance of the absorption from the all-trans state, and with a little luck, stimulated emission while the excited state is occupied. All of this is seen. Looking closely (e.g., at Fig 17), however, one sees that spectra are oscillating in time. Rather than sliding irreversibly down the potential surfaces toward their minima, the atomic structure oscillates. More remarkably, detailed analysis of the time evolution of the spectra demonstrates that there is coherent quantum mechanical mixing among the relevant electronic and vibrational states.

Our usual picture of molecules and their transitions comes from chemical kinetics: there are reaction rates, which represent the probability per unit time for the molecule to make transitions among states which are distinguishable by some large scale rearrangement; these transitions are cleanly separated from the time scales for molecules to come to equilibrium in each state. The initial isomerization event in rhodopsin is so fast that this approximation certainly breaks down. More profoundly, the time scale of the isomerization is so fast that it competes with the processes that destroy quantum mechanical coherence among the relevant electronic and vibrational states. The whole notion of an irreversible transition from one state to another necessitates the loss of coherence between these states (recall Schrödinger’s cat),

and so in this sense the isomerization is proceeding as rapidly as possible.

At this point what we would like to do is an honest, if simplified calculation that generates the schematic in Fig 21 and explains how the dynamics on these surfaces can be so fast. As far as I know, there is no clear answer to this challenge, although there are many detailed simulations, in the quantum chemical style, that probably capture elements of the truth.[\[it would be nice to be a little more explicit here!\]](#) The central ingredient is the special nature of the π bonds along the retinal. In the ground state, electron hopping between neighboring p_z orbitals lowers the energy of the system, and this effect is maximized in planar structures where the orbitals are all in the same orientation. But this lowering of the energy depends on the character of the electron wave functions—in the simplest case of bonding between two atoms, the symmetric state (the ‘bonding orbital’) has lower energy in proportion to the hopping matrix element, while the anti-symmetric state (‘anti-bonding orbital’) has higher energy, again in proportion to the matrix element. Thus, if we excite the electrons, it is plausible that the energy of the excited state could be reduced by structural changes that reduce the hopping between neighboring carbons, which happens if the molecule rotates to become non-planar. In this way we can understand why there is a force for rotation in the excited state, and why there is another local minimum in the ground state at the 11-cis structure.

Problem 24: Energy levels in conjugated molecules. The simplest model for a conjugated molecule is that the electrons which form the π orbitals can sit on each carbon atom with some energy that we can set to zero, and they can hop from one atom to its neighbors. Note that there is one relevant electron per carbon atom. If we write the Hamiltonian for the electrons as a matrix, then for a ring of six carbons (benzene) we have

$$\mathbf{H}_6 = \begin{pmatrix} 0 & -t & 0 & 0 & 0 & -t \\ -t & 0 & -t & 0 & 0 & 0 \\ 0 & -t & 0 & -t & 0 & 0 \\ 0 & 0 & -t & 0 & -t & 0 \\ 0 & 0 & 0 & -t & 0 & -t \\ -t & 0 & 0 & 0 & -t & 0 \end{pmatrix}, \quad (69)$$

where the “hopping matrix element” $-t$ is negative because the electrons can lower their energy by being shared among neighboring atoms—this is the essence of chemical bonding! Models like this are called *tight binding* models in the condensed matter physics literature and *Hückel* models in the chemical literature. Notice that they leave out any direct interactions among the electrons. This problem is about solving Schrödinger’s equation, $\mathbf{H}\psi = E\psi$, to find the energy eigenstates and the corresponding energy levels. Notice that for the case of benzene if we write the wave function ψ in terms of its six components (one for each carbon atom) then

Schrödinger's equation becomes

$$-t(\psi_2 + \psi_6) = E\psi_1 \quad (70)$$

$$-t(\psi_1 + \psi_3) = E\psi_2 \quad (71)$$

$$-t(\psi_2 + \psi_4) = E\psi_3 \quad (72)$$

$$-t(\psi_3 + \psi_5) = E\psi_4 \quad (73)$$

$$-t(\psi_4 + \psi_6) = E\psi_5 \quad (74)$$

$$-t(\psi_5 + \psi_1) = E\psi_6. \quad (75)$$

(a.) Considering first the case of benzene, show that solutions to the Schrödinger equation are of the form $\psi_n \propto \exp(ikn)$. What are the allowed values of the “momentum” k ? Generalize to an arbitrary N -membered ring.

(b.) What are the energies corresponding to the states labeled by k ? Because of the Pauli principle, the ground state of the molecule

is constructed by putting the electrons two-by-two (spin up and spin down) into the lowest energy states; thus the ground state of benzene has two electrons in each of the lowest three states. What is the ground state energy of benzene? What about for an arbitrary N -membered ring (with N even)? Can you explain why benzene is especially stable?

(c.) Suppose that the bonds between carbon atoms stretch and compress a bit, so that they become alternating single and double bonds rather than all being equivalent. To first order, if the bond stretches by an amount u then the hopping matrix element should go down (the electron has farther to hop), so we write $t \rightarrow t - \alpha u$; conversely, if the bond compresses, so that u is negative, the hopping matrix element gets larger. If we have alternating long and short (single and double) bonds, then the Hamiltonian for an six membered ring would be

$$\mathbf{H}_6(u) = \begin{pmatrix} 0 & -t + \alpha u & 0 & 0 & 0 & -t - \alpha u \\ -t + \alpha u & 0 & -t - \alpha u & 0 & 0 & 0 \\ 0 & -t - \alpha u & 0 & -t + \alpha u & 0 & 0 \\ 0 & 0 & -t + \alpha u & 0 & -t - \alpha u & 0 \\ 0 & 0 & 0 & -t - \alpha u & 0 & -t + \alpha u \\ -t - \alpha u & 0 & 0 & 0 & -t + \alpha u & 0 \end{pmatrix}. \quad (76)$$

Find the ground state energy of the electrons as a function of u , and generalize to the case of N -membered rings. Does the “dimerization” of the system ($u \neq 0$) raise or lower the energy of the electrons? Note that if your analytic skills (or patience!) give out, this is a relatively simple numerical problem; feel free to use the computer, but be careful to explain what units you are using when you plot your results.

(d.) In order to have bonds alternately stretched and compressed by an amount u , we need an energy $\frac{1}{2}\kappa u^2$ in each bond, where κ is the stiffness contributed by all the other electrons that we're not keeping track of explicitly. Consider parameter values $t = 2.5$ eV, $\alpha = 4.1$ eV/Å, and $\kappa = 21$ eV/Å². Should benzene have alternating single and double bonds ($u \neq 0$) or should all bonds be equivalent ($u = 0$)?

(e.) Peierls' theorem about one-dimensional electron systems predicts that, for N -carbon rings with N large, the minimum total energy will be at some non-zero u_* . Verify that this is true in this case, and estimate u_* . How large does N have to be before it's “large”? What do you expect for retinal?

I could try to do a full calculation here that puts flesh on the outline in the previous paragraph, using the tools from the problem above. But there still is a problem even if this works ...

Suppose that we succeed, and have a semi-quantitative theory of the excited state dynamics of rhodopsin, enough to understand why the quantum yield of fluorescence is so low, and what role is played by quantum coherence. We would then have to check that the barrier between the 11-cis and the all-trans structures in Fig 21 comes out to have the right height to explain the activation energy for spontaneous isomerization. But then how do we account for the anomalously low prefactor in

this rate, which is where, as discussed above, the protein acts to suppress dark noise? If there is something special about the situation in the environment of the protein which makes possible the ultrafast, coherent dynamics in the excited state, why does this special environment generate almost the same barrier as for isolated retinal?

It is clear that the ingredients for understanding the dynamics of rhodopsin—and hence for understanding why we can see into the darkest times of night—involve quantum mechanical ideas more related to condensed matter physics than to conventional biochemistry, a remarkably long distance from the psychology experiments on human subjects that we started with. While Lorentz could imagine that people count single quanta, surely he couldn't have imagined that he first steps of this process are coherent. While these are the ingredients, it is clear that we don't have them put together in quite the right way yet.

If rhodopsin were the only example of this “almost coherent chemistry” that would be good enough, but in fact the other large class of photon induced events in biological systems—photosynthesis—also proceed so rapidly as to compete with loss of coherence, and the crucial events again seem to happen (if you'll pardon the partisanship) while everything is still in the domain of physics and not conventional chemistry. Again there are beautiful experiments that present a number of theoretical challenges.²⁰

²⁰ As usual, a guide is found in the references at the end of this section.

Why biology pushes to these extremes is a good question. How it manages to do all this with big floppy molecules in water at roughly room temperature also is a great question.

To get some of the early history of work on the visual pigments, one can do worse than to read Wald's Nobel lecture (Wald 1972). Wald himself (along with his wife and collaborator, Ruth Hubbard) was quite an interesting fellow, much involved in politics; to connect with the previous section, his PhD adviser was Selig Hecht. [\[need more about dark noise and temperature dependence?\]](#) For a measurement of dark noise in cones, see Sampath & Baylor (2002). The remarkable result that the quantum yield of fluorescence in rhodopsin is $\sim 10^{-5}$ is due to Doukas et al (1984); it's worth noting that measuring this small quantum yield was possible at a time when one could not directly observe the ultrafast processes that are responsible for making the branching ratio this small. Direct measurements were finally made by Mathies et al (1988), Schoenlein et al (1991), and Wang et al (1994), the last paper making clear that the initial events are quantum mechanically coherent. A detailed analysis of the Raman spectra of Rhodopsin has been done by Loppnow & Mathies (1988).

Doukas et al 1984: Fluorescence quantum yield of visual pigments: Evidence for subpicosecond isomerization rates. AG Doukas, MR Junnarkar, RR Alfano, RH Callender, T Kakitani & B Honig, *Proc Nat Acad Sci (USA)* **81**, 4790–4794 (1984).

Loppnow & Mathies 1988: Excited-state structure and isomerization dynamics of the retinal chromophore in rhodopsin from resonance Raman intensities GR Loppnow & RA Mathies, *Biophys J* **54**, 35–43 (1988).

Mathies et al 1988: Direct observation of the femtosecond excited-state cis–trans isomerization in bacteriorhodopsin. RA Mathies, CH Brito Cruz, WT Pollard & CV Shank, *Science* **240**, 777–779 (1988).

Sampath & Baylor 2002: Molecular mechanisms of spontaneous pigment activation in retinal cones. AP Sampath & DA Baylor, *Biophys J* **83**, 184–193 (2002).

Schoenlein et al 1991: The first step in vision: Femtosecond isomerization of rhodopsin. RW Schoenlein, LA Peteanu, RA Mathies & CV Shank, *Science* **254**, 412–415 (1991).

Wald 1972: The molecular basis of visual excitation. G Wald, in *Nobel Lectures: Physiology or Medicine 1963–1970* (Elsevier, Amsterdam, 1972). Also available at <http://nobelprize.org>.

Wang et al 1994: Vibrationally coherent photochemistry in the femtosecond primary event of vision. Q Wang, RW Schoenlein, LA Peteanu, RA Mathies & CV Shank, *Science* **266**, 422–424 (1994).

The Born–Oppenheimer approximation is discussed in almost all quantum mechanics textbooks. For a collection of the key papers, with commentary, on the rich phenomena that can emerge in such adiabatic approximations, see Shapere & Wilczek (1989). Models for coupling of electron hopping to bond stretching (as in the last problem) were explored by Su, Schrieffer and Heeger in relation to polyacetylene. Importantly, these models predict that the excitations (e.g., upon photon absorption) are not just electrons and holes in the usual ladder of molecular orbitals, but that there are localized, mobile objects with unusual quantum numbers. These mobile objects can be generated by doping, which is the basis for conductivity in these quasi–one dimensional materials. The original work in Su et al (1980); a good review is Heeger et al (1988). Many people must have realized that the dynamical models being used by condensed matter physicists for (ideally) infinite chains

might also have something to say about finite chains. For ideas in this direction, including some specifically relevant to Rhodopsin, see Bialek et al (1987), Vos et al (1996), and Aalberts et al (2000).

Aalberts et al 2000: Quantum coherent dynamics of molecules: A simple scenario for ultrafast photoisomerization. DP Aalberts, MSL du Croo de Jongh, BF Gerke & W van Saarloos, *Phys Rev A* **61**, 040701 (2000).

Heeger et al 1988: Solitons in conducting polymers. AJ Heeger, S Kivelson, JR Schrieffer & W–P Su, *Rev Mod Phys* **60**, 781–850 (1988).

Bialek et al 1987: Simple models for the dynamics of biomolecules: How far can we go?. W Bialek, RF Goldstein & S Kivelson, in *Structure, Dynamics and Function of Biomolecules: The First EBSA Workshop*, A Ehrenberg, R Rigler, A Graslund & LJ Nilsson, eds, pp 65–69 (Springer–Verlag, Berlin, 1987).

Shapere & Wilczek 1989: *Geometric Phases in Physics* A Shapere and F Wilczek (World Scientific, Singapore, 1989)

Su et al 1980: Soliton excitations in polyacetylene. W–P Su, JR Schrieffer & AJ Heeger, *Phys Rev B* **22**, 2099–2111 (1980).

Vos et al 1996: Su–Schrieffer–Heeger model applied to chains of finite length. FLJ Vos, DP Aalberts & W van Saarloos, *Phys Rev B* **53**, 14922–14928 (1996).

Going beyond the case of rhodopsin, you may want to explore the role of quantum coherence in the initial events of photosynthesis; for an introduction see Fleming & van Grondelle (1994). The first experiments focused on photo–induced electron transfer, and looked at systems that had been genetically modified so that the electron, once excited, had no place to go (Vos et al 1991, Vos et al 1993); this made it possible to see the coherent vibrational motion of the molecule more clearly in spectroscopic experiments. Subsequent experiments used more intact systems, but looked first at low temperatures (Vos et al 1994a) and finally at room temperature (Vos et al 1994b). Eventually it was even possible to show that photo–triggering of electron transfer in other systems could reveal coherent vibrational motions (Liebl et al 1999). More or less at the same time as the original Vos et al experiments, my colleagues and I made the argument that photo–induced electron transfer rates in the initial events of photosynthesis would be maximized if the system were poised on the threshold of revealing coherent effects; maybe (although there were uncertainties about all the parameters) one could even strengthen this argument to claim that the observed rates were possible only in this regime (Skourtis et al 1992). Most recently, it has been discovered that when energy is trapped in the “antenna pigments” of photosynthetic systems, the migration of energy toward the reaction center (where the electron transfer occurs) is coherent, and it has been suggested that this allows for a more efficient exploration of space, finding the target faster than is possible in diffusive motion (Engel et al 2007). [\[Decide what to say about the large follow up literature!\]](#)

Engel et al 2007: Evidence for wavelike energy transfer through quantum coherence in photosynthetic systems. GS Engel, TR Calhoun, EL Read, T–K Ahn, T Mančal, Y–C Cheng, RE Blankenship & GR Fleming, *Nature* **446**, 782–786 (2007).

Fleming & van Grondelle 1994: The primary steps of photosynthesis. GR Fleming & R van Grondelle, *Physics Today* pp 48–55, February 1994.

Liebl et al 1999: Coherent reaction dynamics in a bacterial cytochrome c oxidase. U Liebl, G Lipowski, M Négrerie, JC Lambry, JL Martin & MH Vos, *Nature* **401**, 181–184 (1999).

Skourtis et al 1992: A new look at the primary charge separation in bacterial photosynthesis. SS Skourtis, AJR DaSilva, W Bialek & JN Onuchic, *J Phys Chem* **96**, 8034–8041 (1992).

Vos et al 1991: Direct observation of vibrational coherence in bacterial reaction centers using femtosecond absorption

spectroscopy. MH Vos, JC Lambry, SJ Robles, DC Youvan, J Breton & JL Martin, *Proc Nat'l Acad Sci (USA)* **88**, 8885–8889 (1991).

Vos et al 1993: Visualization of coherent nuclear motion in a membrane protein by femtosecond spectroscopy. MH Vos, F Rappaport, JC Lambry, J Breton & JL Martin, *Nature* **363**, 320–325 (1993).

Vos et al 1994a: Coherent dynamics during the primary electron transfer reaction in membrane-bound reaction centers of *Rhodobacter sphaeroides*. MH Vos, MR Jones, CN Hunter, J Breton, JC Lambry & JL Martin, *Biochemistry* **33**, 6750–6757 (1994).

Vos et al 1994b: Coherent nuclear dynamics at room temperature in bacterial reaction centers. MH Vos, MR Jones, CN Hunter, J Breton, JC Lambry & JL Martin, *Proc Nat'l Acad Sci (USA)* **91**, 12701–12705 (1994).

C. Dynamics of biochemical networks

Section still needs editing, as of September 18, 2011. The material here seems to have accreted during the early versions of the course, and much time is spent on things which we now know aren't productive On the other hand, I would like to say more about, for example, Sengupta et al (2000) on SNR in cascades and gain-bandwidth, as well as returning to the problem of transduction in invertebrates, e.g. theoretical work from Shraiman, Ranganathan et al.. So, I'd like make a more thorough overhaul here!

We have known for a long time that light is absorbed by rhodopsin, and that light absorption leads to an electrical response which is detectable as a modulation in the current flowing across the photoreceptor cell membrane. It is only relatively recently that we have come to understand the mechanisms which link these two events. The nature of the link is qualitatively different in different classes of organisms. For vertebrates, including us, the situation is as schematized in Fig 22. [it would be nice to come back and talk about invertebrates too]

In outline, what happens is that the excited rhodopsin changes its structure, arriving after several steps in a state where it can act as a catalyst to change the structure of another protein called transducin (T). The activated transducin in turn activates a catalyst called phosphodiesterase (PDE), which breaks down cyclic guanosine monophosphate (cGMP). Finally, cGMP binds to channels in the cell membrane and opens the channels, allowing current to flow (mostly carried by Na^+ ions); breaking down the cGMP thus decreases the number of open channels and decreases the current. [This discussion needs to refer to a schematic of the rod cell. Where is this? Earlier? Here?]

In a photomultiplier, photon absorption results in the ejection of a primary photoelectron, and then the large

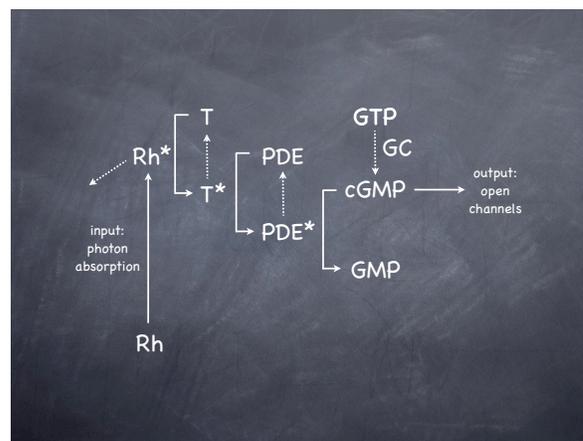


FIG. 22 The cascade leading from photon absorption to ionic current flow in rod photoreceptors. Solid lines indicate ‘forward’ steps that generate gain; dashed lines are the ‘backward’ steps that shut off the process. **T** is the transducin molecule, a member of the broad class of G-proteins that couple receptors to enzymes. **PDE** is the enzyme phosphodiesterase, named for the particular bond that it cuts when it degrades cyclic guanosine monophosphate (**cGMP**) into **GMP**. **GC** is the guanylate cyclase that synthesizes **cGMP** from guanosine triphosphate, **GTP**.

electric field accelerates this electron so that when it hits the next metal plate it ejects many electrons, and the process repeats until at the output the number of electrons is sufficiently large that it constitutes a macroscopic current. Thus the photomultiplier really is an electron multiplier. In the same way, the photoreceptor acts as a molecule multiplier, so that for one excited rhodopsin molecule there are many cGMP molecules degraded at the output of the “enzymatic cascade.”

There are lots of interesting questions about how the molecule multiplication actually works in rod photoreceptors. These questions are made more interesting by the fact that this general scheme is ubiquitous in biological systems. [need a schematic about G-protein coupled receptors!] Rhodopsin is a member of a family of proteins which share common structural features (seven alpha helices that span the membrane in which the protein is embedded) and act as receptors, usually activated by the binding of small molecules such as hormones or odorants rather than light. Proteins in this family interact with proteins from another family, the G proteins, of which transducin is an example, and the result of such interactions typically is the activation of yet another enzyme, often one which synthesizes or degrades a cyclic nucleotide. Cyclic nucleotides in turn are common intracellular messengers, not just opening ion channels but also activating or inhibiting a variety of enzymes. This universality of components means that understanding the mechanisms of photon counting in rod cells is not just a curiosity for physicists, but a place where we can provide a model for understanding an enormous range of biological processes.

In order to get started, we need to know a little bit about ion channels, which form the output of the system. We will see that even the simplest, order-of-magnitude properties of channels raise a question about the observed behavior of the rod cells.

Recall that the brain contains no metallic or semiconductor components. Signals can still be carried by electrical currents and voltages, but now currents consist of ions, such as potassium or sodium, flowing through water or through specialized conducting pores. These pores, or channels, are large molecules (proteins) embedded in the cell membrane, and can thus respond to the electric field or voltage across the membrane as well as to the binding of small molecules. The coupled dynamics of channels and voltage turns each cell into a potentially complex nonlinear dynamical system.

Imagine a spherical molecule or ion of radius a ; a typical value for this radius is 0.3 nm. From Stokes' formula we know that if this ion moves through the water at velocity v it will experience a drag force $F = \gamma v$, with the drag coefficient $\gamma = 6\pi\eta a$, where η is the viscosity; for water $\eta = 0.01$ poise, the cgs unit poise = gm/(cm · s). The inverse of the drag coefficient is called the mobility, $\mu = 1/\gamma$, and the diffusion constant of a particle is related to the mobility and the absolute temperature by the Einstein relation or fluctuation dissipation theorem, $D = k_B T \mu$, with k_B being Boltzmann's constant and T the absolute temperature. Since life operates in a narrow range of absolute temperatures, it is useful to remember that at room temperature (25°C), $k_B T \sim 4 \times 10^{-21}$ J $\sim 1/40$ eV. So let's write the diffusion constant in terms of the other quantities, and then evaluate the order of magnitude:

$$D = k_B T \mu = k_B T \cdot \frac{1}{\gamma} = \frac{k_B T}{6\pi\eta a} \quad (77)$$

$$= \frac{[4 \times 10^{-21} \text{ J}]}{6\pi \cdot [0.01 \text{ gm}/(\text{cm} \cdot \text{s})] \cdot [0.3 \times 10^{-9} \text{ m}]} \quad (78)$$

$$\sim 2 \times 10^{-9} \text{ m}^2/\text{s} = 2 \mu\text{m}^2/\text{ms}. \quad (79)$$

Ions and small molecules diffuse freely through water, but cells are surrounded by a membrane that functions as a barrier to diffusion. In particular, these membranes are composed of lipids, which are nonpolar, and therefore cannot screen the charge of an ion that tries to pass through the membrane. The water, of course, is polar and does screen the charge, so pulling an ion out of the water and pushing it through the membrane would require surmounting a large electrostatic energy barrier. This barrier means that the membrane provides an enormous resistance to current flow between the inside and the outside of the cell. If this were the whole story there would be no electrical signaling in biology. In fact, cells construct specific pores or channels through which ions can pass, and by regulating the state of these channels the cell can control the flow of electric current across the membrane. [\[need a sketch that goes with this discussion\]](#)

Ion channels are themselves molecules, but very large ones—they are proteins composed of several thousand atoms in very complex arrangements. Let's try, however, to ask a simple question: If we open a pore in the cell membrane, how quickly can ions pass through? More precisely, since the ions carry current and will move in response to a voltage difference across the membrane, how large is the current in response to a given voltage?

Imagine that one ion channel serves, in effect, as a hole in the membrane. Let us pretend that ion flow through this hole is essentially the same as through water. The electrical current that flows through the channel is

$$J = q_{\text{ion}} \cdot [\text{ionic flux}] \cdot [\text{channel area}], \quad (80)$$

where q_{ion} is the charge on one ion, and we recall that 'flux' measures the rate at which particles cross a unit area, so that

$$\text{ionic flux} = \frac{\text{ions}}{\text{cm}^2 \text{s}} = \frac{\text{ions}}{\text{cm}^3} \cdot \frac{\text{cm}}{\text{s}} \quad (81)$$

$$= [\text{ionic concentration}] \cdot [\text{velocity of one ion}] \\ = cv. \quad (82)$$

Major current carriers such as sodium and potassium are at concentrations of $c \sim 100$ mM, or $c \sim 6 \times 10^{19}$ ions/cm³.

The next problem is to compute the typical velocity of one ion. We are interested in a current, so this is not the velocity of random Brownian motion but rather the average of that component of the velocity directed along the electric field. In a viscous medium, the average velocity is related to the applied force through the mobility, or the inverse of the drag coefficient as above. The force on an ion is in turn equal to the electric field times the ionic charge, and the electric field is (roughly) the voltage difference V across the membrane divided by the thickness ℓ of the membrane:

$$v = \mu F = \mu q_{\text{ion}} E \sim \mu q_{\text{ion}} \frac{V}{\ell} = \frac{D}{k_B T} q_{\text{ion}} \frac{V}{\ell}. \quad (83)$$

Putting the various factors together we find the current

$$J = q_{\text{ion}} \cdot [\text{ionic flux}] \cdot [\text{channel area}] \\ = q_{\text{ion}} \cdot [cv] \cdot [\pi d^2/4] \quad (84)$$

$$= q_{\text{ion}} \cdot \left[c \cdot \frac{D}{\ell} \cdot \frac{q_{\text{ion}} V}{k_B T} \right] \cdot \frac{\pi d^2}{4} \quad (85)$$

$$= \frac{\pi}{4} q_{\text{ion}} \cdot \frac{cd^2 D}{\ell} \cdot \frac{q_{\text{ion}} V}{k_B T}, \quad (86)$$

where the channel has a diameter d . If we assume that the ion carries one electronic charge, as does sodium, potassium, or chloride, then $q_{\text{ion}} = 1.6 \times 10^{-19}$ C and $q_{\text{ion}} V / (k_B T) = V / (25 \text{ mV})$. Typical values for the channel diameter should be comparable to the diameter of a single ion, $d \sim 0.3$ nm, and the thickness of the membrane is $\ell \sim 5$ nm. Thus

$$\begin{aligned}
J &= \frac{\pi}{4} q_{\text{ion}} \cdot \frac{cd^2D}{\ell} \cdot \frac{q_{\text{ion}}V}{k_B T} \\
&= \frac{\pi}{4} (1.6 \times 10^{-19} \text{ C}) \cdot \frac{(6 \times 10^{19} \text{ cm}^{-3})(3 \times 10^{-8} \text{ cm})^2(10^{-5} \text{ cm}^2/\text{s})}{50 \times 10^{-8} \text{ cm}} \cdot \frac{V}{25 \text{ mV}}
\end{aligned} \tag{87}$$

$$\sim 2 \times 10^{-14} \cdot \frac{V}{\text{mV}} \text{ C/s} \sim 2 \times 10^{-11} \frac{V}{\text{Volts}} \text{ Amperes}, \tag{88}$$

or

$$J = gV \tag{89}$$

$$g \sim 2 \times 10^{-11} \text{ Amperes/Volt} = 20 \text{ picoSiemens.} \tag{90}$$

So our order of magnitude argument leads us to predict that the conductance of an open channel is roughly 20 pS.²¹ With a voltage difference across the membrane of ~ 50 mV, we thus expect that opening a single channel will cause ~ 1 picoAmp of current to flow. Although incredibly oversimplified, this is basically the right answer, as verified in experiments where one actually measures the currents flowing through single channel molecules.

The first problem in understanding the enzymatic cascade in rods is accessible just from these back of the envelope arguments. When we look at the total change in current that results from a single photon arrival, it is also ~ 1 pA. But if this were just the effect of (closing) one channel, we'd see "square edges" in the current trace as the single channels opened or closed. It would also be a little weird to have sophisticated (and expensive!) mechanisms for generating macroscopic changes in cGMP concentration only to have this act once again on a single molecule—if we have a single molecule input and a single molecule output, it really isn't clear why we would need an amplifier. What's going on?

The answer turns out to be that these channels flicker very rapidly between their open and closed states, so that on the relatively slow time scale of the rod response one sees essentially a graded current proportional to the probability of the channel being open. Thus the population of channels in the rod cell membrane produces a current that depends continuously on the concentration of cGMP. Alternatively, the noise variance that is associated with the random binary variable open/closed has been spread over a very broad bandwidth, so that in the frequency range of interest (recall that the single photon response is on a time scale of ~ 1 s) the noise is much reduced. This idea is made precise in the following problem, which you can think of as an introduction to the

analysis of noise in "chemical" systems where molecules fluctuate among multiple states.

Problem 25: Flickering channels. Imagine a channel that has two states, open and closed. There is a rate k_{open} at which the molecule makes transitions from the closed state to the open state, and conversely there is a rate k_{close} at which the open channels transition into the closed state. If we write the number of open channels as n_{open} , and similarly for the number of closed channels, this means that the deterministic kinetic equations are

$$\frac{dn_{\text{open}}}{dt} = k_{\text{open}}n_{\text{closed}} - k_{\text{close}}n_{\text{open}} \tag{91}$$

$$\frac{dn_{\text{close}}}{dt} = k_{\text{close}}n_{\text{open}} - k_{\text{open}}n_{\text{close}}, \tag{92}$$

or, since $n_{\text{open}} + n_{\text{closed}} = N$, the total number of channels,

$$\frac{dn_{\text{open}}}{dt} = k_{\text{open}}(N - n_{\text{open}}) - k_{\text{close}}n_{\text{open}} \tag{93}$$

$$= -(k_{\text{open}} + k_{\text{close}})n_{\text{open}} + k_{\text{open}}N. \tag{94}$$

For a single channel molecule, these kinetic equations should be interpreted as saying that an open channel has a probability $k_{\text{close}}dt$ of making a transition to the closed state within a small time dt , and conversely a closed channel has a probability $k_{\text{open}}dt$ of making a transition to the open state. We will give a fuller account of noise in chemical systems in the next Chapter, but for now you should explore this simplest of examples.

(a.) If we have a finite number of channels, then really the number of channels which make the transition from the closed state to the open state in a small window dt is a random number. What is the mean number of these closed \rightarrow open transitions? What is the mean number of open \rightarrow closed transitions? Use your results to show that macroscopic kinetic equations such as Eqs (91) and (92) should be understood as equations for the mean numbers of open and closed channels,

$$\frac{d\langle n_{\text{open}} \rangle}{dt} = k_{\text{open}}\langle n_{\text{closed}} \rangle - k_{\text{close}}\langle n_{\text{open}} \rangle \tag{95}$$

$$\frac{d\langle n_{\text{close}} \rangle}{dt} = k_{\text{close}}\langle n_{\text{open}} \rangle - k_{\text{open}}\langle n_{\text{close}} \rangle. \tag{96}$$

(b.) Assuming that all the channels make their transitions independently, what is the variance in the number of closed \rightarrow open transitions in the small window dt ? In the number of open \rightarrow closed transitions? Are these fluctuations in the number of transitions independent of one another?

(c.) Show that your results in [b] can be summarized by saying that the change in the number of open channels during the time dt obeys an equation

$$n_{\text{open}}(t+dt) - n_{\text{open}}(t) = dt[k_{\text{open}}n_{\text{closed}} - k_{\text{close}}n_{\text{open}}] + \eta(t), \tag{97}$$

where $\eta(t)$ is a random number that has zero mean and a variance

$$\langle \eta^2(t) \rangle = dt[k_{\text{open}}n_{\text{closed}} + k_{\text{close}}n_{\text{open}}]. \tag{98}$$

²¹ Siemens are the units of conductance, which are inverse to units of resistance, ohms. In the old days, this inverse of resistance had the rather cute unit 'mho' (pronounced 'moe,' like the Stooge).

Explain why the values of $\eta(t)$ and $\eta(t')$ are independent if $t \neq t'$.

(d.) This discussion should remind you of the description of Brownian motion by a Langevin equation, in which the deterministic dynamics are supplemented by a random force that describes molecular collisions. In this spirit, show that, in the limit $dt \rightarrow 0$, you can rewrite your results in [c] to give a Langevin equation for the number of open channels,

$$\frac{dn_{\text{open}}}{dt} = -(k_{\text{open}} + k_{\text{close}})n_{\text{open}} + k_{\text{open}}N + \zeta(t), \quad (99)$$

where

$$\langle \zeta(t)\zeta(t') \rangle = \delta(t - t')[k_{\text{open}}n_{\text{closed}} + k_{\text{close}}n_{\text{open}}]. \quad (100)$$

In particular, if the noise is small, show that $n_{\text{open}} = \langle n_{\text{open}} \rangle + \delta n_{\text{open}}$, where

$$\frac{d\delta n_{\text{open}}}{dt} = -(k_{\text{open}} + k_{\text{close}})\delta n_{\text{open}} + \zeta_s(t), \quad (101)$$

$$\langle \zeta_s(t)\zeta_s(t') \rangle = 2k_{\text{open}}\langle n_{\text{closed}} \rangle. \quad (102)$$

(e.) Solve Eq (101) to show that

$$\langle \delta n_{\text{open}}^2 \rangle = Np_{\text{open}}(1 - p_{\text{open}}) \quad (103)$$

$$\langle \delta n_{\text{open}}(t)\delta n_{\text{open}}(t') \rangle = \langle \delta n_{\text{open}}^2 \rangle \exp\left[-\frac{|t - t'|}{\tau_c}\right], \quad (104)$$

where the probability of a channel being open is $p_{\text{open}} = k_{\text{open}}/(k_{\text{open}} + k_{\text{close}})$, and the correlation time $\tau_c = 1/(k_{\text{open}} + k_{\text{close}})$. Explain how the result for the variance $\langle \delta n_{\text{open}}^2 \rangle$ could be derived more directly.

(f.) Give a critical discussion of the approximations involved in writing down these Langevin equations. In particular, in the case of Brownian motion of a particle subject to ordinary viscous drag, the Langevin force has a Gaussian distribution. Is that true here?

Problem 26: Averaging out the noise. Consider a random variable such as n_{open} in the previous problem, for which the noise has exponentially decaying correlations, as in Eq (104). Imagine that we average over a window of duration τ_{avg} , to form a new variable

$$z(t) = \frac{1}{\tau_{\text{avg}}} \int_0^{\tau_{\text{avg}}} d\tau \delta n_{\text{open}}(t - \tau). \quad (105)$$

Show that, for $\tau_{\text{avg}} \gg \tau_c$, the variance of z is smaller than the variance of δn_{open} by a factor of τ_{avg}/τ_c . Give some intuition for why this is true (e.g., how many statistically independent samples of n_{open} will you see during the averaging time?). What happens if your averaging time is shorter?

I think this is a fascinating example, because evolution has selected for very fast channels to be present in a cell that signals very slowly! Our genome (as well as those of many other animals) codes for hundreds if not thousands of different types of channels once one includes the possibility of alternative splicing. These different channels differ, among other things, in their kinetics. In the fly retina, for example, the dynamics of visual inputs looking straight ahead are very different from those looking to the side, and in fact the receptor cells that look in these different directions have different kinds of channels—the faster channels to respond to the more rapidly varying signals. [I am not sure that the last statement is correct, and need to check the references; what certainly is true is that insects with different lifestyles (e.g., acrobats vs. slow fliers) use different potassium channels ...] In the

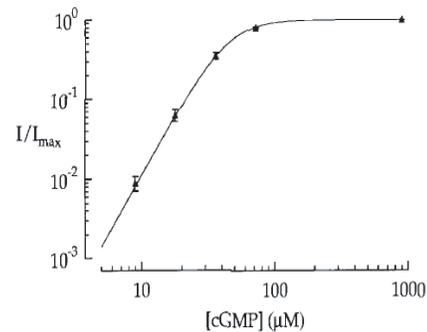


FIG. 23 Current through the rod cell membrane as a function of the cyclic GMP concentration. The fit is to Eq (106), with $n = 2.9 \pm 0.1$ and $G_{1/2} = 45 \pm 4 \mu\text{M}$. From Rieke & Baylor (1996).

vertebrate rod, signals are very slow but the channels are fast, and this makes sense only if the goal is to suppress the noise.

Having understood a bit about the channels, let's take one step back and see how these channels respond to cyclic GMP. Experimentally, with the rod outer segment sucked into the pipette for measuring current, one can break off the bottom of the cell and make contact with its interior, so that concentrations of small molecules inside the cell will equilibrate with concentrations in the surrounding solution. Since the cell makes cGMP from GTP, if we remove GTP from the solution then there is no source other than the one that we provide, and now we can map current vs concentration. The results of such an experiment are shown in Fig 23. We see that the current I depends on the cGMP concentration G as

$$I = I_{\text{max}} \frac{G^n}{G^n + G_{1/2}^n}, \quad (106)$$

with $n \approx 3$. This suggests strongly that the channel opens when three molecules of cGMP bind to it. This is an example of “cooperativity” or “allostery,” which is a very important theme in biochemical signaling and regulation. It's a little off to the side of our discussion here, however, so see Appendix A.4.

Let's try to write a more explicit model for the dynamics of amplification in the rod cell, working back from the channels. We have Eq (106), which tells us how the current I depends on G , the concentration of cyclic GMP. The dynamics of G has two terms, synthesis and degradation:

$$\frac{dG}{dt} = \gamma - PDE^*G, \quad (107)$$

where γ denotes the rate of synthesis by the guanylate cyclase (GC, cf Fig 22), and PDE^* measures the activity of the active phosphodiesterase. It turns out that there

is a feedback mechanism in the rod, where calcium enters through the open channels (as part of the current), and then calcium binding inhibits the activity of the guanylate cyclase. We can summarize these effects, measured in several experiments, by writing

$$\gamma = \frac{\gamma_{\max}}{1 + (Ca/K_{\text{gc}})^2} \approx \alpha Ca^{-2}, \quad (108)$$

where the last approximation is valid so long as the typical calcium concentration Ca is much larger than the binding constant $K_{\text{gc}} \sim 100$ nM, which seems to be true; the fact that the dependence is on the square of the calcium concentration presumably means that two Ca^{++} ions bind to inhibit the cyclase (see again the discussion of cooperativity in Appendix A.4). Since calcium enters the cell as a fraction of the current flowing through the open channels, and presumably is pumped back out by other mechanisms, we can write

$$\frac{dCa}{dt} = fI(G) - \beta Ca, \quad (109)$$

where f is the fraction of the current carried by calcium and $1/\beta$ is the lifetime of calcium before it is pumped out. These equations tell how the cyclic GMP concentration, and hence the current, will respond to changes in the activity of the phosphodiesterase, thus describing the last steps of the amplification cascade.

It is convenient to express the response of G to PDE^* in the limit that the response is linear, which we expect is right when only small numbers of photons are being counted. This linearization gives us

$$\delta\dot{G} = \frac{\partial\gamma}{\partial Ca} \delta Ca - PDE_0^* \delta G - G_0 \delta PDE^* \quad (110)$$

$$\delta\dot{Ca} = fI'(G_0) \delta G - \beta \delta Ca, \quad (111)$$

where the subscript 0 denotes the values in the dark. We can solve these equations by passing to Fourier space, where

$$\delta\tilde{G}(\omega) = \int dt e^{+i\omega t} \delta G(t), \quad (112)$$

and similarly for the other variables. As usual, this reduces the linear differential equations to linear algebraic equations, and when the dust settles we find

$$\frac{\delta\tilde{G}(\omega)}{\delta\tilde{PDE}^*(\omega)} = \frac{-G_0(-i\omega + \beta)}{(-i\omega + PDE_0^*)(-i\omega + \beta) + A}, \quad (113)$$

$$A = 2\gamma_0 f I'(G_0) / Ca_0. \quad (114)$$

Already this looks like lots of parameters, so we should see how we can simplify, or else measure some of the parameters directly.

First, one finds experimentally that the cyclic GMP concentration is in the regime where $I \propto G^3$, that is

$G \ll G_{1/2}$. This means that we can express the response more compactly as a fractional change in current

$$\delta\tilde{I}(\omega) = 3I_0 \frac{-i\omega + \beta}{(-i\omega + PDE_0^*)(-i\omega + \beta) + A} \cdot \delta\tilde{PDE}^*(\omega), \quad (115)$$

where $A = 6\beta PDE_0^*$.

Problem 27: Dynamics of cGMP. Fill in all the steps leading to Eq (115).

In the same experiment where one measures the response of the channels to cGMP, one can suddenly bring the cGMP concentration of the outside solution to zero, and then the internal cGMP concentration (which we can read off from the current, after the first experiment) will fall due both to diffusion out of the cell and to any PDE which is active in the dark; one can also poison the PDE with a drug (IBMX), separating the two components. In this way one can measure $PDE_0^* = 0.1 \pm 0.02 \text{ s}^{-1}$. To measure β , you need to know that the dominant mechanism for pumping calcium out of the cell actually generates an electrical current across the membrane.²² With this knowledge, if we turn on a bright light and close all the cGMP-sensitive channels, there is no path for calcium to enter the rod outer segment, but we still see a small current as it is pumped out. This current decays with a rate $\beta \sim 2 \text{ s}^{-1}$. Thus, although this model—even for part of the process!—looks complicated, there are many independent experiments one can do to measure the relevant parameters.

In fact, the analysis of the dynamics of cGMP and calcium leads us to the point where we can more or less invert these dynamics, turning the dynamics of the current back into the dynamics of the PDE^* . An interesting application of this idea is to try and understand the continuous background noise that occurs in the dark. As we saw, there is a big source of noise in the dark that comes from spontaneous isomerization of rhodopsin. But there is also a smaller, continuous rumbling, with an amplitude $\delta I_{\text{rms}} \sim 0.1$ pA. This isn't the intrinsically random opening and closing of the channels, since we have seen that this happens very fast and thus contributes very little to the noise at reasonable frequencies. It must thus reflect

²² This needn't be true. First, there are mechanisms which exchange ions on different sides of the membrane, maintaining electrical neutrality. Second, it could be that the dominant pump sends calcium into storage spaces inside the cell, so no ions cross the cell membrane.

responses of the channels to fluctuations in the concentration of cGMP. Since this concentration is determined by a balance between synthesis and degradation, one should check whether one of these processes is dominating the noise.

The rate at which cGMP is synthesized is modulated by calcium, but we can prevent the calcium concentration from changing by using buffers, either injected into the cell or in the surrounding solution when the cell is broken open. If the calcium concentration were itself fluctuating, and these fluctuations generated noise in the synthesis of cGMP, buffering the calcium concentration should lower the continuous background noise; instead the noise goes up. On the other hand, if we poison the phosphodiesterase with IBMX, and allow synthesis to compete with diffusion out of a broken cell, the noise drops dramatically. [At this point things get a little vague .. go back and do better!] These, and other experiments as well, indicate that the dominant source of the continuous dark noise is fluctuations in the number of active phosphodiesterase molecules. Alternatively, one can say that the noise arises from ‘spontaneous’ activation of the PDE, absent any input from activated rhodopsin.

[Need to be sure we have control over the math here .. maybe connect back to problem about ion channels? Also connect to Appendix A.2. Review before giving results. Get all the number right, too!] If the activation of PDE in the dark is rare, then we expect that the variance in the number of active molecules will be equal to the mean, and the fluctuations in activity should have a correlation time equal to the lifetime of the activated state. If a is the activity of a single enzyme—that is, the factor that converts the number of active enzymes into the rate at which cGMP is degraded—then we have

$$\langle \delta PDE^*(t) \delta PDE^*(t') \rangle = a PDE_0^* e^{-|t-t'|/\tau_c}, \quad (116)$$

where τ_c is the lifetime of the active state. Putting this together with Eq (115), we can generate a prediction for the power spectrum of fluctuations in the current. Importantly, the only unknown parameters are a , which sets the over scale of the fluctuations, and τ_c , which shapes the spectrum. Fitting to the observed spectra, one finds $a = 1.6 \times 10^{-5} \text{ s}^{-1}$ and $\tau_c = 0.56 \text{ s}$. Thus, a single active phosphodiesterase causes the cGMP concentration to decrease at a rate $aG_0 \sim 2 \times 10^{-4} \mu\text{M/s}$, and this lasts for roughly half a second; with a volume of $\sim 10^{-12} \text{ l}$, this means that one PDE^* destroys ~ 60 molecules of cGMP.

Knowing how changes in concentration change the current, and how much one PDE^* can reduce the cGMP concentration, we can calculate that a single photon must activate at least 2000 phosphodiesterase molecules. More concretely, a single activated rhodopsin must trigger the activation of at least 2000 PDE^* . In order for this to happen, the activated rhodopsin has to diffuse in the disk membrane [did we actually discuss the geometry of the disk etc? check!] during its lifetime; certainly the

number of molecules that it can activate is limited by the number of molecules that it can encounter via diffusion. With measured diffusion constants and a lifetime of roughly one second (after this, the whole response starts to shut off), this seems possible, but not with much to spare. Thus, it seems likely that the gain in the first part of the amplifier is limited by the density of molecules and the physics of diffusion. [Need estimates of diffusion constant here .. either explain, or give problem, about diffusion limit to this reaction.]

[I think that before going on to discuss reproducibility we want to say a bit more about gain .. look at Detwiler et al (2000) regarding the design of G protein elements, since this would also give an excuse to discuss some more about these ... Then check segue.] So, given this dissection of the amplifier, what is it that we really want to know? Understanding gain—how you get many molecules out for only one molecule at the input—isn’t so hard, basically because catalysis rates are high, close to the diffusion limit. One might want to understand the system’s choice of other parameters, but is there really a conceptual problem here?

Perhaps the most surprising aspect of the single photon response in rods is its reproducibility. If we look at the responses to dim light flashes and isolate those responses that correspond to a single photon (you have already done a problem to assess how easy or hard this is!), one finds that the amplitude of the response fluctuates by only $\sim 15 - 20\%$; see, for example, Fig. 24. To understand *why* this is surprising we have to think about chemistry at the level of single molecules, specifically the chemical reactions catalyzed by the single activated molecule of rhodopsin.

[This discussion need to point back to the problem about ion channels.] When we write that there is a rate k for a chemical reaction, what we mean is that for one molecule there is a *probability per unit time* k that the reaction will occur—this should be familiar from the case of radioactive decay. Thus when one molecule of rhodopsin is activated at time $t = 0$, if we imagine that de-activation is a simple chemical reaction then the probability that the molecule is still active at time t obeys the usual kinetic equation

$$\frac{dp(t)}{dt} = -kp(t); \quad (117)$$

of course if there are N total molecules then $Np(t) = n(t)$ is the expected number of molecules still in the active state. Thus, $p(t) = \exp(-kt)$. The probability density $P(t)$ that the molecule is active for exactly a time t is the probability that the molecule is still active at t times the probability per unit time of de-activation, so

$$P(t) = kp(t) = k \exp(-kt). \quad (118)$$

This may seem pedantic, but it’s important to be clear—and we’ll see that far from being obvious there must be

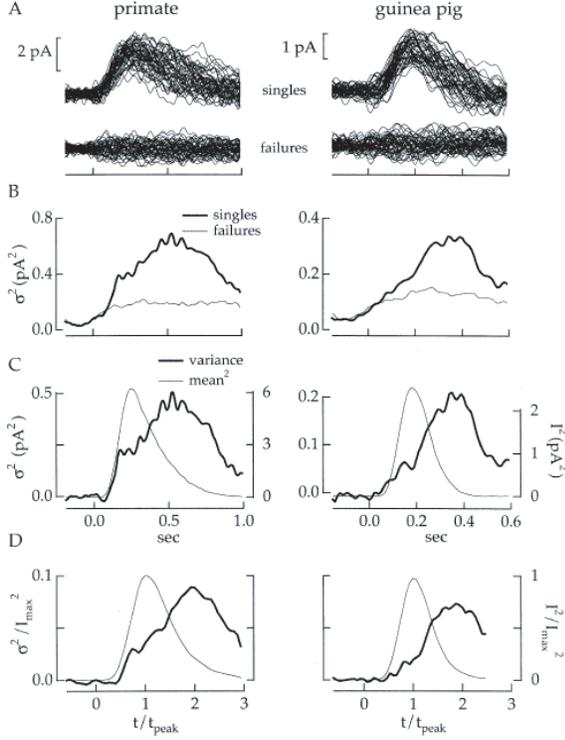


FIG. 24 Reproducibility of the single photon response, from Field & Rieke (2002b). (A) Examples of single photon responses and failures from single mammalian rods. (B) Variances of the responses in (A). (C) Variance and square of mean response to one photon; variance in the response is defined as the difference in variance between responses and failures. Finally (D) shows the mean of results as in (C) from eight primate rods and nine guinea pig rods; scales are normalized for each cell by the peak mean response and the time to peak. We see that at the peak response the relative variance is ~ 0.025 , so the root-mean-square fluctuations are ~ 0.15 .

something wrong with this simple picture.

Given the probability density $P(t)$, we can calculate the mean and variance of the time spent in the active state:

$$\langle t \rangle \equiv \int_0^{\infty} dt P(t) t \quad (119)$$

$$= k \int_0^{\infty} \exp(-kt) t = 1/k; \quad (120)$$

$$\langle (\delta t)^2 \rangle \equiv \int_0^{\infty} dt P(t) t^2 - \langle t \rangle^2 \quad (121)$$

$$= k \int_0^{\infty} dt \exp(-kt) t^2 - 1/k^2 \quad (122)$$

$$= 2/k^2 - 1/k^2 = 1/k^2. \quad (123)$$

Thus we find that

$$\delta t_{\text{rms}} \equiv \sqrt{\langle (\delta t)^2 \rangle} = 1/k = \langle t \rangle, \quad (124)$$

so that the root-mean-square fluctuations in the lifetime are equal to the mean.

How does this relate to the reproducibility of the single photon response? The photoreceptor works by having the active rhodopsin molecule act as a catalyst, activating transducin molecules. If the catalysis proceeds at some constant rate (presumably set by the time required for rhodopsin and transducin to find each by diffusion in the membrane), then the number of activated transducins is proportional to the time that rhodopsin spends in the active state—and hence we would expect that the number of active transducin molecules has root-mean-square fluctuations equal to the mean number. If the subsequent events in the enzymatic cascade again have outputs proportional to their input number of molecules, this variability will not be reduced, and the final output (the change in cGMP concentration) will again have relative fluctuations of order one, much larger than the observed 15–20%. This is a factor of 25 or 40 error in variance; we can't even claim to have an order of magnitude understanding of the reproducibility. I'd like to give an idea of the different possible solutions that people have considered, focusing on very simple versions of these ideas that we can explore analytically. At the end, we'll look at the state of the relevant experiments.

One possibility is that although the lifetime of activated rhodopsin might fluctuate, the number of molecules at the output of the cascade fluctuates less because of saturation [point to sketch of discs]. For example, if each rhodopsin has access only to a limited pool of transducin molecules, a reasonable fraction of rhodopsins might remain active long enough to hit all the molecules in the pool. The simplest version of this idea is as follows. Let the total number of transducins in the pool be N_{pool} , and let the number of activated transducins be n_T . When the rhodopsin is active, it catalyzes the conversion of inactive transducins (of which there are $N_{\text{pool}} - n_T$) into the active form at a rate r , so that (neglecting the discreteness of the molecules)

$$\frac{dn_T}{dt} = r(N_{\text{pool}} - n_T). \quad (125)$$

If the rhodopsin molecule is active for a time t then this catalysis runs for a time t and the number of activated transducins will be

$$n_T(t) = N_{\text{pool}}[1 - \exp(-rt)]. \quad (126)$$

For small t the variations in t are converted into proportionately large variations in n_T , but for large t the saturation essentially cuts off this variation.

To be more precise, recall that we can find the distribution of n_T by using the identity

$$P(n_T)dn_T = P(t)dt, \quad (127)$$

which applies whenever we have two variables that are related by a deterministic, invertible transformation. From

Eq (126) we have

$$t = -\frac{1}{r} \ln(1 - n_T/N_{\text{pool}}), \quad (128)$$

and so, going through the steps explicitly:

$$P(n_T) = P(t) \left| \frac{dn_T}{dt} \right|^{-1} \quad (129)$$

$$= k \exp(-kt) \frac{1}{r(N_{\text{pool}} - n_T)} \quad (130)$$

$$= \frac{k}{r} \exp \left[\left(\frac{k}{r} \right) \ln(1 - n_T/N_{\text{pool}}) \right] \frac{1}{(N_{\text{pool}} - n_T)} \quad (131)$$

$$= \frac{k}{rN_{\text{pool}}} \left(1 - \frac{n_T}{N_{\text{pool}}} \right)^{k/r-1}. \quad (132)$$

[Maybe a plot to show this?] When the activation rate r is small, n_T always stays much less than N_{pool} and the power law can be approximated as an exponential. When r is large, however, the probability distribution grows a power law singularity at N_{pool} ; for r finite this singularity is integrable but as $r \rightarrow \infty$ it approaches a log divergence, which means that essentially all of the weight will be concentrated at N_{pool} . In particular, the relative variance of n_T vanishes as r becomes large, as promised.

This discussion has assumed that the limited number of target molecules is set, perhaps by some fixed structural domain. Depending on details, it is possible for such a limit to arise dynamically, as a competition between diffusion and chemical reactions. In invertebrate photoreceptors, such as the flies we have met in our discussion above, there is actually a positive feedback loop in the amplifier which serves to ensure that each structural domain (which are more obvious in the fly receptor cells) ‘fires’ a saturated, stereotyped pulse in response to each photon.

[Make a sketch of the different models—either one big figure, or separate ones for each model.]

The next class of models are those that use feedback. The idea, again, is simple: If the output of the cascade is variable because the rhodopsin molecule doesn’t “know” when to de-activate, why not link the de-activation to the output of the cascade? Roughly speaking, count the molecules at the output and shut the rhodopsin molecule off when we reach some fixed count. Again let’s try the simplest version of this. When rhodopsin is active it catalyzes the formation of some molecule (which might not actually be the transducin molecule itself) at rate r , and let the number of these output molecules be x so that we simply have

$$\frac{dx}{dt} = r, \quad (133)$$

or $x = rt$. Let’s have the rate of deactivation of rhodopsin depend on x , so that instead of Eq (117) we have

$$\frac{dp(t)}{dt} = k[x(t)]p(t). \quad (134)$$

For example, if deactivation is triggered by the cooperative binding of m x molecules (as in the discussion of cGMP-gated channels), we expect that

$$k[x] = k_{\text{max}} \frac{x^m}{x_0^m + x^m}. \quad (135)$$

We can solve Eq (134) and then recover the probability density for rhodopsin lifetime as before,

$$p(t) = \exp \left(- \int_0^t d\tau k[x(\tau)] \right) \quad (136)$$

$$P(t) = k[x(t)] \exp \left(- \int_0^t d\tau k[x(\tau)] \right). \quad (137)$$

Again we can push through the steps:

$$P(t) = k[x(t)] \exp \left(- \int_0^t d\tau k[x(\tau)] \right) = k_{\text{max}} \frac{x^m(t)}{x_0^m + x^m(t)} \exp \left(-k_{\text{max}} \int_0^t d\tau \frac{x^m(\tau)}{x_0^m + x^m(\tau)} \right) \quad (138)$$

$$\approx k_{\text{max}} \left(\frac{t}{t_0} \right)^m \exp \left[-\frac{k_{\text{max}} t_0}{m+1} \left(\frac{t}{t_0} \right)^{m+1} \right], \quad (139)$$

where in the last step we identify $t_0 = x_0/r$ and assume that $t \ll t_0$.

To get a better feel for the probability distribution in

Eq (139) it is useful to rewrite it as

$$P(t) \approx k_{\text{max}} \exp[-G(t)] \quad (140)$$

$$G(t) = -m \ln \left(\frac{t}{t_0} \right) + \frac{k_{\text{max}} t_0}{m+1} \left(\frac{t}{t_0} \right)^{m+1} \quad (141)$$

We can find the most likely value of the lifetime, \bar{t} , by minimizing G , which of course means that the derivative must be set to zero:

$$G'(t) = -\frac{m}{t} + k_{\max}t_0 \cdot \frac{1}{t} \left(\frac{t}{t_0}\right)^{m+1} \quad (142)$$

$$G'(t = \bar{t}) = 0 \Rightarrow k_{\max}t_0 \cdot \frac{1}{\bar{t}} \left(\frac{\bar{t}}{t_0}\right)^{m+1} = \frac{m}{\bar{t}} \quad (143)$$

$$\frac{\bar{t}}{t_0} = \left(\frac{m}{k_{\max}t_0}\right)^{1/m} \quad (144)$$

In particular we see that for sufficiently large k_{\max} we will have $\bar{t} \ll t_0$, consistent with the approximation above. What we really want to know is how sharp the distribution is in the neighborhood of \bar{t} , so we will try a series expansion of $G(t)$:

$$P(t) \approx k_{\max} \exp \left[-G(\bar{t}) - \frac{1}{2}G''(\bar{t})(t - \bar{t})^2 - \dots \right] \quad (145)$$

$$G''(t) = \frac{m}{t^2} + (k_{\max}t_0)m \cdot \frac{1}{t^2} \left(\frac{t}{t_0}\right)^{m+1} \approx \frac{m}{\bar{t}^2}, \quad (146)$$

where again in the last step we assume $\bar{t} \ll t_0$. Thus we see that the distribution of lifetimes is, at least near its peak,

$$P(t) \approx P(\bar{t}) \exp \left[-\frac{m}{2\bar{t}^2}(t - \bar{t})^2 - \dots \right]. \quad (147)$$

This of course is a Gaussian with variance

$$\langle(\delta t)^2\rangle = \frac{1}{m} \cdot \bar{t}^2, \quad (148)$$

so the relative variance is $1/m$ as opposed to 1 in the original exponential distribution.

A concrete realization of the feedback idea can be built around the fact that the current flowing into the rod includes calcium ions, and the resulting changes in calcium concentration can regulate protein kinases—proteins which in turn catalyze the attachment of phosphate groups to other proteins—and rhodopsin shut off is known to be associated with phosphorylation at multiple sites. Calcium activation of kinases typically is cooperative, so $m \sim 4$ in the model above is plausible. Notice that in the saturation model the distribution of lifetimes remains broad and the response to these variations is truncated; in the feedback model the distribution of lifetimes itself is sharpened.

A third possible model involves multiple steps in rhodopsin de-activation. Let us imagine that rhodopsin starts in one state and makes a transition to state 2, then from state 2 to state three, and so on for K states, and then it is the transition from state K to $K + 1$ that actually corresponds to de-activation. Thus there are K active states and if the time spent in each state is t_i then the total time spent in activated states is

$$t = \sum_{i=1}^K t_i. \quad (149)$$

Clearly the mean value of t is just the sum of the means of each t_i , and if the transitions are independent (again, this is what you mean when you write the chemical kinetics with the arrows and rate constants) then the variance of t will also be the sum of the variances of the individual t_i ,

$$\langle t \rangle = \sum_{i=1}^K \langle t_i \rangle \quad (150)$$

$$\langle(\delta t)^2\rangle = \sum_{i=1}^K \langle(\delta t_i)^2\rangle. \quad (151)$$

We recall from above that for each single step, $\langle(\delta t_i)^2\rangle = \langle t_i \rangle^2$. If the multiple steps occur at approximately equal rates, we can write

$$\langle t \rangle = \sum_{i=1}^K \langle t_i \rangle \approx K \langle t_1 \rangle \quad (152)$$

$$\langle(\delta t)^2\rangle = \sum_{i=1}^K \langle(\delta t_i)^2\rangle = \sum_{i=1}^K \langle t_i \rangle^2 \approx K \langle t_1 \rangle^2 \quad (153)$$

$$\frac{\langle(\delta t)^2\rangle}{\langle t \rangle^2} \approx \frac{K \langle t_1 \rangle^2}{(K \langle t_1 \rangle)^2} = \frac{1}{K}. \quad (154)$$

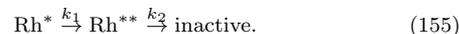
Thus the relative variance declines as one over the number of steps, and the relative standard deviation declines as one over the square root of the number of steps. This is an example of how averaging K independent events causes a $1/\sqrt{K}$ reduction in the noise level.

The good news is that allowing de-activation to proceed via multiple steps can reduce the variance in the lifetime of activated rhodopsin. Again our attention is drawn to the fact that rhodopsin shut off involves phosphorylation of the protein at multiple sites. The bad news is that to have a relative standard deviation of $\sim 20\%$ would require 25 steps.

It should be clear that a multistep scenario works only if the steps are irreversible. If there are significant “backward” rates then progress through the multiple states becomes more like a random walk, with an accompanying increase in variance. Thus each of the (many) steps involved in rhodopsin shut off must involve dissipation of a few $k_B T$ of energy to drive the whole process forward.

Problem 28: Getting the most out of multiple steps.

Consider the possibility that Rhodopsin leaves its active state through a two step process. To fix the notation, let's say that the first step occurs with a rate k_1 and the second occurs with rate k_2 :



Assume that we are looking at one molecule, and at time $t = 0$ this molecule is in state Rh^* .

(a) Write out and solve the differential equations for the time dependent probability of being in each of the three states.

(b) Use your results in [a] to calculate the probability distribution for the time at which the molecule *enters* the inactive state. This is the distribution of “lifetimes” for the two active states. Compute the mean and variance of this lifetime as a function of the parameters k_1 and k_2 .

(c) Is there a simple, intuitive argument that allows you to write down the mean and variance of the lifetime, without solving any differential equations? Can you generalize this to a scheme in which inactivation involves N steps rather than two?

(d) Given some desired mean lifetime, is there a way of adjusting the parameters k_1 and k_2 (or, more generally, $k_1, k_2 \dots, k_N$) to minimize the variance?

(e) Suppose that there is a back reaction $\text{Rh}^{**} \xrightarrow{k_{-1}} \text{Rh}^*$. Discuss what this does to the distribution of lifetimes. In particular, what happens if the rate k_{-1} is very fast? Note that “discuss” is deliberately ambiguous; you could try to solve the relevant differential equations, or to intuit the answer, or even do a small simulation [connect this problem to recent work by Escola & Paninski].

The need for energy dissipation and the apparently very large number of steps suggests a different physical picture. If there really are something like 25 steps, then if we plot the free energy of the rhodopsin molecule as a function of its atomic coordinates, there is a path from initial to final state that passes over 25 hills and valleys. Each valley must be a few $k_B T$ lower than the last, and the hills must be many $k_B T$ high to keep the rates in the right range. This means that the energy surface is quite rough [this needs a sketch]. Now when we take one solid and slide it over another, the energy surface is rough on the scale of atoms because in certain positions the atoms on each surface “fit” into the interatomic spaces on the other surface, and then as we move by an Ångström or so we encounter a very high barrier. If we step back and blur our vision a little bit, all of this detailed roughness just becomes friction between the two surfaces. Formally, if we think about Brownian motion on a rough energy landscape and we average over details on short length and time scales, what we will find is that the mobility or friction coefficient is renormalized and then the systems behaves on long time scales as if it were moving with this higher friction on a smooth surface.

So if the de-activation of rhodopsin is like motion on a rough energy surface, maybe we can think about the renormalized picture of motion on a smooth surface with high drag or low mobility. Suppose that the active and inactive states are separated by a distance ℓ along some direction in the space of molecular structures, and that motion in this direction occurs with an effective mobility μ . If there is an energy drop ΔE between the active and de-activated states, then the velocity of motion is $v \sim \mu \Delta E / \ell$ and the mean time to make the de-activation transition is

$$\langle t \rangle \sim \frac{\ell}{v} \sim \frac{\ell^2}{\mu \Delta E}. \quad (156)$$

On the other hand, diffusion over this time causes a spread in positions

$$\langle (\delta \ell)^2 \rangle \sim 2D \langle t \rangle = 2\mu k_B T \langle t \rangle, \quad (157)$$

where we make use of the Einstein relation $D = \mu k_B T$. Now (roughly speaking) since the molecule is moving in configuration space with typical velocity v , this spread in positions is equivalent to a variance in the time required to complete the transition to the de-activated state,

$$\langle (\delta t)^2 \rangle \sim \frac{\langle (\delta \ell)^2 \rangle}{v^2} \sim \frac{2\mu k_B T}{(\mu \Delta E / \ell)^2} \cdot \frac{\ell^2}{\mu \Delta E}. \quad (158)$$

If we express this as a fractional variance we find

$$\frac{\langle (\delta t)^2 \rangle}{\langle t \rangle^2} \sim \frac{2\mu k_B T}{(\mu \Delta E / \ell)^2} \cdot \frac{\ell^2}{\mu \Delta E} \cdot \left(\frac{\mu \Delta E}{\ell^2} \right)^2 \sim \frac{2k_B T}{\Delta E}. \quad (159)$$

Thus when we look at the variability of the lifetime in this model, the effective mobility μ and the magnitude ℓ of the structural change in the molecule drop out, and the reproducibility is just determined by the amount of energy that is dissipated in the de-activation transition. Indeed, comparing with the argument about multiple steps, our result here is the same as expected if the number of irreversible steps were $K \sim \Delta E / (2k_B T)$, consistent with the idea that each step must dissipate more than $k_B T$ in order to be effectively irreversible. To achieve a relative variance of 1/25 or 1/40 requires dropping $\sim 0.6 - 1$ eV (recall that $k_B T$ is 1/40 eV at room temperature), which is OK since the absorbed photon is roughly 2.5 eV.

Problem 29: Is there a theorem here? The above argument hints at something more general. Imagine that we have a molecule in some state, and we ask how long it takes to arrive at some other state. Assuming that the molecular dynamics is that of overdamped motion plus diffusion on some energy surface, can you show that the fractional variance in the time required for the motion is limited by the free energy difference between the two states?

How do we go about testing these different ideas? If saturation is important, one could try either by chemical manipulations or by genetic engineering to prolong the lifetime of rhodopsin and see if in fact the amplitude of the single photon response is buffered against these changes. If feedback is important, one could make a list of candidate feedback molecules and to manipulate the concentrations of these molecules. Finally, if there are multiple steps one could try to identify the molecular events associated with each step and perturb these events again either with chemical or genetic methods. All these are good ideas, and have been pursued by several groups.

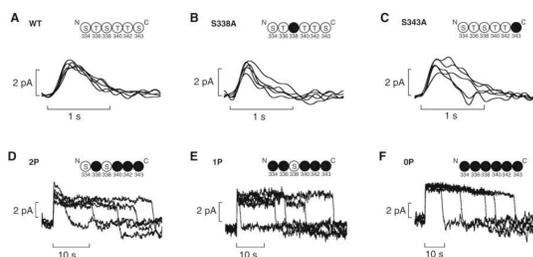


FIG. 25 Variability in the single photon response with genetically engineered rhodopsins. (A) Wild type responses from mouse rods. Schematic shows the six phosphorylation sites, which are serine or threonine residues. In the remaining panels, we see responses when the number of phosphorylation sites has been reduced by mutating alanine, leaving five sites (B & C), two sites (D), one site (E), or none (F). From Doan et al (2007).

An interesting hint about the possibility of multiple steps in the rhodopsin shutoff is the presence of multiple phosphorylation sites on the opsin proteins. In mice, there are six phosphorylation sites, and one can genetically engineer organisms in which some or all of these sites are removed. At a qualitative level it's quite striking that even knocking out one of these sites produces a noticeable increase in the variability of the single photon responses, along with a slight prolongation of the mean response (Figs 25B & C). When all but one or two sites are removed, the responses last a *very* long time, and start to look like on/off switches with a highly variable time in the 'on' state (Figs 25D & E). When there are no phosphorylation sites, rhodopsin can still turn off, presumably as a result of binding another molecule (arrestin). But now the time to shutoff is broadly distributed, as one might expect if there were a single step controlling the

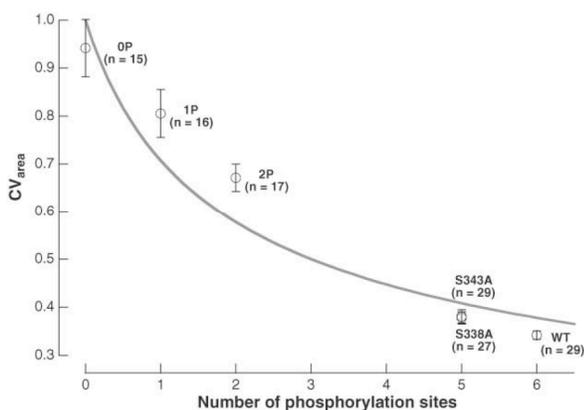


FIG. 26 Standard deviation in the integral of the single photon response, normalized by the mean. Results are shown as a function of the number of phosphorylation sites, from experiments as in Fig 25; error bars are standard errors of the mean. Solid line is $CV = 1/\sqrt{N_p + 1}$, where N_p is the number of phosphorylation sites. From Doan et al (2006).

Remarkably, if we examine the responses quantitatively, the variance of the single photon response seems to be inversely proportional the number of these sites, exactly as in the model where deactivation involved multiple steps, now identified with the multiple phosphorylations (Fig 26). This really is beautiful. One of the things that I think is interesting here is that, absent the discussion of precision and reproducibility, the multiple phosphorylation steps might just look like complexity for its own sake, the kind of thing that biologists point to when they want to tease physicists about our propensity to ignore details. In this case, however, the complexity seems to be the solution to a very specific physics problem.

Probably this section should end with some caveats. Do we really think the problem of reproducibility is solved?

A general review of the cGMP cascade in rods is given by Burns & Baylor (2001). Rieke & Baylor (1996) set out to understand the origins of the continuous noise in rods, but along the way provide a beautifully quantitative dissection of the enzymatic cascade; much of the discussion above follows theirs. For an explanation of how similarity to Rhodopsin (and other G-protein coupled receptors) drove the discovery of the olfactory receptors, see Buck (2004). For some general background on ion channels, you can try Aidley (see notes to Section 1.1), Johnston & Wu (1995), or Hille (2001). A starting point for learning about how different choices of channels shape the dynamics of responses in insect photoreceptors is the review by Weckström & Laughlin (1995). [There is much more to say here, and probably even some things left to do.]

Buck 2004: Unraveling the sense of smell. LB Buck, in *Les Prix Nobel: Nobel Prizes 2004*, T Frängsmyr, ed (Nobel Foundation, Stockholm, 2004).

Burns & Baylor 2001: Activation, deactivation and adaptation in vertebrate photoreceptor cells. ME Burns & DA Baylor, *Annu Rev Neurosci* **24**, 779–805 (2001).

Hille 2001: *Ion Channels of Excitable Membranes, 3rd Edition* B Hille (Sinauer, Sunderland MA, 2001).

Johnston & Wu 1995: *Foundations of Cellular Neurophysiology*. D Johnston & SM Wu (MIT Press Cambridge, 1995).

Rieke & Baylor 1996: Molecular origin of continuous dark noise in rod photoreceptors. F Rieke & DA Baylor, *Biophys J* **71**, 2553–2572 (1996).

Weckström & Laughlin 1995: Visual ecology and voltage gated ion channels in insect photoreceptors. M Weckström & SB Laughlin, *Trends Neurosci* **18**, 17–21 (1995).

Rieke & Baylor (1998a) provide a review of photon counting rods with many interesting observations, including an early outline of the problem of reproducibility. An early effort to analyze the signals and noise in enzymatic cascades is by Detwiler et al (2000). The idea that restricted, saturable domains can arise dynamically and tame the fluctuations in the output of the cascade is described by the same authors (Ramanathan et al 2005). For invertebrate photoreceptors, it seems that reproducibility of the response to single photons can be traced to positive feedback mechanisms that generate a stereotyped pulse of concentration changes, localized to substructures analogous to the disks in vertebrate rods (Pumir et al 2008).

Detwiler et al 2000: Engineering aspects of enzymatic signal transduction: Photoreceptors in the retina. PB Detwiler, S Ramanathan, A Sengupta & BI Shraiman, *Biophys J* **79**, 2801–2817 (2000).

Pumir et al 2008: A Pumir, J Graves, R Ranganathan & BI Shraiman, Systems analysis of the single photon response in invertebrate photoreceptors. *Proc Nat'l Acad Sci (USA)* **105**, 10354–10359 (2008).

Ramanathan et al 2005: G–protein–coupled enzyme cascades have intrinsic properties that improve signal localization and fidelity. S Ramanathan, PB Detwiler, AM Sengupta & BI Shraiman, *Biophys J* **88**, 3063–3071 (2005).

Rieke & Baylor 1998a: Single–photon detection by rod cells of the retina. F Rieke & DA Baylor, *Revs Mod Phys* **70**, 1027–1036 (1998).

One of the early, systematic efforts to test different models of reproducibility was by Rieke & Baylor (1998b). Many of the same ideas were revisited in mammalian rods by Field & Rieke (2002b), setting the stage for the experiments on genetic engineering of the phosphorylation sites by Doan et al (2006). More recent work from the same group explores the competition between the kinase and the arrestin molecule, which binds to the phosphorylated rhodopsin to terminate its activity, showing this competition influences both the mean and the variability of the single photon response (Doan et al 2009).

Doan et al 2007: Multiple phosphorylation sites confer reproducibility of the rod's single–photon responses. T Doan, A Mendez, PB Detwiler, J Chen & F Rieke, *Science* **313**, 530–533 (2006).

Doan et al 2009: Arrestin competition influences the kinetics and variability of the single–photon responses of mammalian rod photoreceptors. T Doan, AW Azevedo, JB Hurley & F Rieke, *J Neurosci* **29**, 11867–11879 (2009).

Field & Rieke 2002: Mechanisms regulating variability of the single photons responses of mammalian rod photoreceptors. GD Field & F Rieke, *Neuron* **35**, 733–747 (2002b).

Rieke & Baylor 1998a: Oring of reproducibility in the responses of retinal rods to single photons. F Rieke & DA Baylor, *Biophys J* **75**, 1836–1857 (1998).

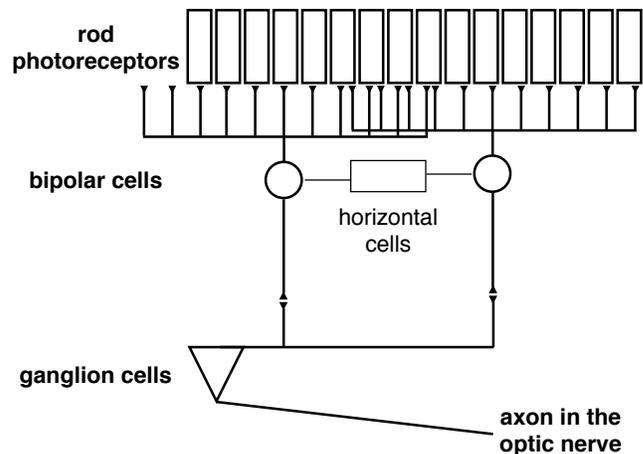


FIG. 27 A schematic of the circuitry in the retina. **Fill in caption.**

of the integration is achieved in the very first step of processing, as many rod cells converge and form synapses onto onto a single bipolar cell, as shown schematically in Fig 27 [maybe also need a real retina?]

If each cell generates an output n_i that counts the number of photons that have arrived, then it's trivial that the total photon count is $n_{\text{total}} = \sum_i n_i$. The problem is that the cells don't generate integers corresponding to the number of photons counted, they generate currents which have continuous variations. In particular, we have seen that the *mean* current in response to a single photon has a peak of $I_1 \sim 1$ pA, but this rests on continuous background noise with an amplitude $\delta I_{\text{rms}} \sim 0.1$ pA. In a single cell, this means that the response to one photon stands well above the background, but if we try to sum the signals from many cells, we have a problem, as illustrated in Fig 28.

To make the problem precise, let's use x_i to denote the peak current generated by cell i . We have

$$x_i = I_1 n_i + \eta_i, \quad (160)$$

where n_i is the number of photons that are counted in cell i , and η_i is the background current noise; from what we have seen in the data, each η_i is chosen independently from a Gaussian distribution with a standard deviation δI_{rms} . If we sum the signals generated by all the cells, we obtain

$$x_{\text{total}} \equiv \sum_{i=1}^{N_{\text{cells}}} x_i = I_1 \sum_{i=1}^{N_{\text{cells}}} n_i + \sum_{i=1}^{N_{\text{cells}}} \eta_i \quad (161)$$

$$= I_1 n_{\text{total}} + \eta_{\text{eff}}, \quad (162)$$

where the effective noise is the sum of N_{cells} independent samples of the η_i , and hence has a standard deviation

$$\eta_{\text{eff}}^{\text{rms}} \equiv \sqrt{\langle \eta_{\text{eff}}^2 \rangle} = \sqrt{N_{\text{cells}}} \delta I_{\text{rms}}. \quad (163)$$

D. The first synapse, and beyond

This is a good moment to remember a key feature of the Hecht, Shlaer and Pirenne experiment, as described in Section I.A. In that experiment, observers saw flashes of light that delivered just a handful of photons spread over an area that includes many hundreds of photoreceptor cells. One consequence is that a single receptor cell has a very low probability of counting more than one photon, and this is how we know that these cells must respond to single photons. But, it must also be possible for the retina to add up the responses of these many cells so that the observer can reach a decision. Importantly, there is no way to know in advance which cells will get hit by photons, so if we (sliding ourselves into the positions of the observer's brain ...) want to integrate the multiple photon counts we have to integrate over all the receptors in the area covered by the flash. This integration might be the simplest computation we can imagine for a nervous system, just adding up a set of elementary signals, all given the same weight. In many retinas, a large part

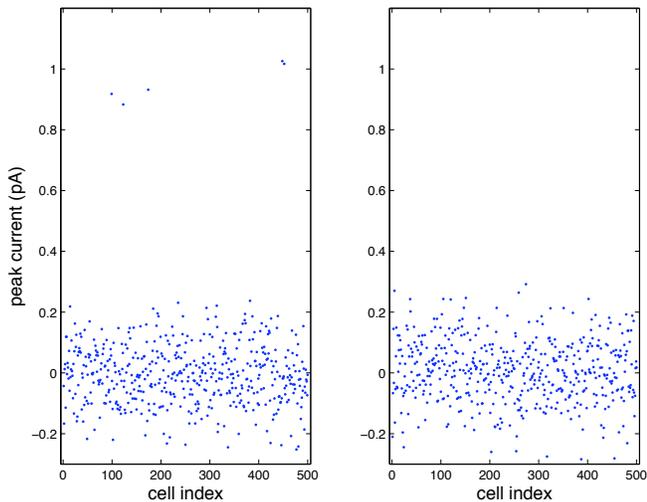


FIG. 28 Simulation of the peak currents generated by $N = 500$ rod cells in response to a dim flash of light. At left, five of the cells actually detect a photon, each resulting in a current $I_1 \sim 1$ pA, while at right we see the response to a blank. All cells have an additive background noise, chosen from a Gaussian distribution with zero mean and standard deviation $\delta I_{\text{rms}} \sim 0.1$ pA. Although the single photon responses stand clearly above the background noise, if we simply add up the signals generated by all the cells, then at left we find a total current $I_{\text{total}} = 1.85$ pA, while at right we find $I_{\text{total}} = 3.23$ pA—the summed background noise completely overwhelms the signal.

The problem is that with $\delta I_{\text{rms}} \sim 0.1$ pA and $N_{\text{cells}} = 500$, we have $\eta_{\text{eff}}^{\text{rms}} \sim 2.24$ pA, which means that there is a sizable chance of confusing three or even five photons with a blank; in some species, the number of cells over which the system integrates is even larger, and the problem becomes even more serious. Indeed, in primates like us, a single ganglion cell (one stage after the bipolar cells; cf Fig 27) receives input from ~ 4000 rods, while on a very dark night we can see when just one in a thousand rods is captures a photon [should have refs for all this]. Simply put, summing the signals from many cells buries the clear single photon response under the noise generated by those cells which did not see anything. This can't be the right way to do things!

Before we start trying to do something formal, let's establish some intuition. Since the single photon signals are clearly detectable in individual rod cells, we could solve our problem by making a 'decision' for each cell—is there a photon present or not?—and then adding up the tokens that represent the outcome of our decision. Roughly speaking, this means passing each rod's signal through some fairly strong nonlinearity, perhaps so strong that it has as an output only a 1 or a 0, and then pooling these nonlinearly transformed signals. In contrast, a fairly standard schematic of what neurons are doing throughout the brain is adding up their inputs and then passing this sum through a nonlinearity (Fig 29).

So perhaps the problems of noise in photon counting are leading us to predict that this very first step of neural computation in the retina has to be different from this standard schematic. Let's try to do an honest calculation that makes this precise. [Is “nonlinearity” clear enough here?]

Formally, the problem faced by the system is as follows. We start with the set of currents generated by all the rod cells, $\{x_i\}$. We can't really be interested in the currents themselves. Ideally we want to know about what is happening in the outside world, but a first step would be to estimate the total number of photons that arrived, n_{total} . What is the best estimate we can make? To answer this, we need to say what we mean by “best.”

One simple idea, which is widely used, is that we want to make estimates which are as close as possible to the right answer, where closeness is measured by the mean square error. That is, we want to map the data $\{x_i\}$ into an estimate of n_{total} through some function $n_{\text{est}}(\{x_i\})$ such that

$$\mathcal{E} \equiv \left\langle [n_{\text{total}} - n_{\text{est}}(\{x_i\})]^2 \right\rangle \quad (164)$$

is as small as possible. To find the optimal choice of the function $n_{\text{est}}(\{x_i\})$ seems like a hard problem—maybe we have to choose some parameterization of this function, and then vary the parameters? In fact, we can solve this problem once and for all, which is part of the reason that this definition of ‘best’ is popular.

When we compute our average error, we are averaging over the joint distribution of the data $\{x_i\}$ and the actual

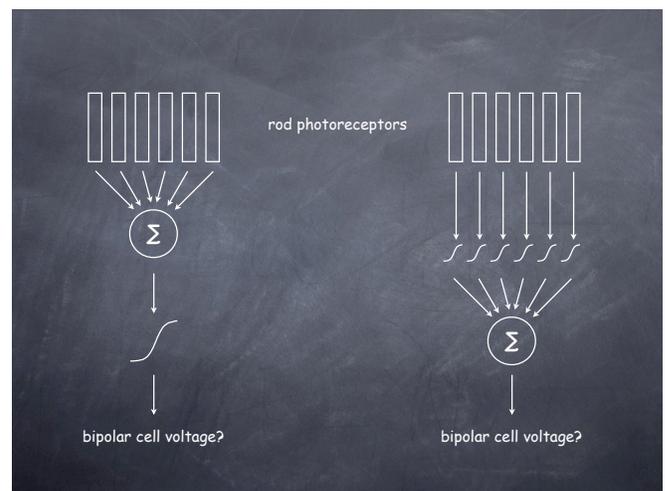


FIG. 29 Schematic of summation and nonlinearity in the initial processing of rod cell signals. At left, a conventional model in which many rods feed into one bipolar cell; the bipolar cell sums its inputs and passes the results through a saturating nonlinearity. At right, an alternative model, suggested by the problems of noise, in which nonlinearities precede summation.

photon count n_{total} . That is,

$$\begin{aligned} \mathcal{E} &\equiv \left\langle [n - n_{\text{est}}(\{x_i\})]^2 \right\rangle \\ &= \int \left[\prod_{i=1}^{N_{\text{cells}}} dx_i \right] \sum_{n_{\text{total}}} P(n, \{x_i\}) [n - n_{\text{est}}(\{x_i\})]^2 \end{aligned} \quad (165)$$

where, to simplify the notation, we drop the subscript total. Now to minimize the error we take the variation with respect to the function $n_{\text{est}}(\{x_i\})$ and set the result equal to zero. We have

$$\frac{\delta \mathcal{E}}{\delta n_{\text{est}}(\{x_i\})} = - \sum_n P(n, \{x_i\}) 2[n - n_{\text{est}}(\{x_i\})], \quad (166)$$

so setting this to zero gives (going through the steps carefully):

$$\sum_n P(n, \{x_i\}) n_{\text{est}}(\{x_i\}) = \sum_n P(n, \{x_i\}) n \quad (167)$$

$$n_{\text{est}}(\{x_i\}) \sum_n P(n, \{x_i\}) = \sum_n P(n, \{x_i\}) n \quad (168)$$

$$n_{\text{est}}(\{x_i\}) P(\{x_i\}) = \sum_n P(n, \{x_i\}) n \quad (169)$$

$$n_{\text{est}}(\{x_i\}) = \sum_n \frac{P(n, \{x_i\})}{P(\{x_i\})} n, \quad (170)$$

and, finally,

$$n_{\text{est}}(\{x_i\}) = \sum_n P(n|\{x_i\}) n. \quad (171)$$

Thus the optimal estimator is the mean value in the conditional distribution, $P(n|\{x_i\})$. Since we didn't use any special properties of the distributions, this must be true in general, as long as 'best' means to minimize mean square error. We'll use this result many times, and come back to the question of whether the choice of mean square error is a significant restriction.

Notice that the relevant conditional distribution is the distribution of photon counts given the rod cell currents. From a mechanistic point of view, we understand the opposite problem, that is, given the photon counts, we know how the currents are being generated. More precisely, we know that, given the number of photons in each cell, the currents will be drawn out of a probability distribution, since this is (implicitly) what we are saying when we write Eq (160). To make this explicit, we have

$$P(\{x_i\}|\{n_i\}) \propto \exp \left[-\frac{1}{2} \sum_{i=1}^{N_{\text{cells}}} \left(\frac{x_i - I_1 n_i}{\delta I_{\text{rms}}} \right)^2 \right]. \quad (172)$$

Again, this is a model that tells us how the photons generate currents. But the problem of the organism is to use the currents to draw inferences about the photons.

We expect that since the signals are noisy, this inference will be probabilistic, so really we would like to know $P(\{n_i\}|\{x_i\})$.

Problem 30: Just checking. Be sure that you understand the connection between Eq (172) and Eq (160). In particular, what assumptions are crucial in making the connection?

The problem of going from $P(\{x_i\}|\{n_i\})$ to $P(\{n_i\}|\{x_i\})$ is typical of the problems faced by organisms: given knowledge of how our sensory data is generated, how do we reach conclusions about what really is going on in the outside world? In a sense this is the same problem that we face in doing physics experiments. One could argue that what we have posed here is a very easy version of the real problem. In fact, we probably don't really care about the photon arrivals, but about the underlying light intensity, or more deeply about the identity and movements of the objects from which the light is being reflected. Still, this is a good start.

The key to solving these inference problems, both for organisms and for experimental physicists, is Bayes' rule. Imagine that we have two events A and B ; to be concrete, we could think of A as some data we observe, and B as a variable in the world that we really want to know. There is some probability $P(A, B)$ that both of these are true simultaneously, i.e. that we observe A and the world is in state B . In the usual view, the states of the world cause the data to be generated in our instruments, so we can say that the state of the world is chosen out of some distribution $P(B)$, and then given this state the data are generated, with some noise, and hence drawn out of the conditional distribution $P(A|B)$. By the usual rules of probability, we have

$$P(A, B) = P(A|B)P(B). \quad (173)$$

We could also imagine that we have just seen the data A , drawn out of some distribution $P(A)$, and then there must be some distribution of things happening in the world that are consistent with our observation. Formally,

$$P(A, B) = P(B|A)P(A). \quad (174)$$

But these are just two different ways of decomposing the joint distribution $P(A, B)$, and so they must be equal:

$$P(A, B) = P(B|A)P(A) = P(A|B)P(B) \quad (175)$$

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}. \quad (176)$$

This last equation is called Bayes' rule, and tells us what we need to know. It is useful to rewrite this, taking seriously the case where A refers to measurable data and B refers to the state of the world:

$$P(\text{world}|\text{data}) = \frac{P(\text{data}|\text{world})P(\text{world})}{P(\text{data})}. \quad (177)$$

Equation (177) is telling us that the probability of the world being in a certain state is proportional to the probability that this state could generate the data we have seen, but this is multiplied by the overall probability that the world can be in this state. This term often is referred to as the 'prior' probability, since it summarizes our knowledge prior to the observation of the data. Put another way, our inference about the world should be both consistent with the data we have observed in this one experiment *and* with any prior knowledge we might have from previous data.²³

Applied to our current problem, Bayes' rule tells us how to construct the probability distribution of photon counts given the rod currents:

$$P(\{n_i\}|\{x_i\}) = \frac{P(\{x_i\}|\{n_i\})P(\{n_i\})}{P(\{x_i\})}. \quad (178)$$

To make progress (and see how to use these ideas), let's start with the simple case of just one rod cell, so we can drop the indices:

$$P(n|x) = \frac{P(x|n)P(n)}{P(x)}. \quad (179)$$

To keep things really simple, let's just think about the case where the lights are very dim, so either there are zero photons or there is one photon, so that

$$P(1|x) = \frac{P(x|1)P(1)}{P(x)}, \quad (180)$$

and similarly for $P(0|x)$. In the denominator we have $P(x)$, which is the probability that we will see the current x , without any conditions on what is going on in the world. We get this by summing over all the possibilities,

$$P(x) = \sum_n P(x|n)P(n) \quad (181)$$

$$= P(x|1)P(1) + P(x|0)P(0), \quad (182)$$

where in the last step we use the approximation that the lights are very dim. Putting the terms together, we have

$$P(1|x) = \frac{P(x|1)P(1)}{P(x|1)P(1) + P(x|0)P(0)}. \quad (183)$$

Now we can substitute for $P(x|n)$ from Eq (172),

$$P(x|n) = \frac{1}{\sqrt{2\pi}(\delta I_{\text{rms}})^2} \exp\left[-\frac{(x - I_1 n)^2}{2(\delta I_{\text{rms}})^2}\right]. \quad (184)$$

Going through the steps, we have

$$P(1|x) = \frac{P(x|1)P(1)}{P(x|1)P(1) + P(x|0)P(0)} = \frac{1}{1 + P(x|0)P(0)/P(x|1)P(1)} \quad (185)$$

$$= \frac{1}{1 + [P(0)/P(1)] \exp\left[-\frac{(x)^2}{2(\delta I_{\text{rms}})^2} + \frac{(x - I_1)^2}{2(\delta I_{\text{rms}})^2}\right]} = \frac{1}{1 + \exp(\theta - \beta x)}, \quad (186)$$

where

$$\theta = \ln \left[\frac{P(0)}{P(1)} \right] + \frac{I_1^2}{2(\delta I_{\text{rms}})^2} \quad (187)$$

$$\beta = \frac{I_1}{(\delta I_{\text{rms}})^2}. \quad (188)$$

The result in Eq (186) has a familiar form—it is as if the two possibilities (0 and 1 photon) are two states of a physical system, and their probabilities are determined by a Boltzmann distribution; the energy difference between the two states shifts in proportion to the data x , and the temperature is related to the noise level in the system. In the present example, this analogy doesn't add much, essentially because the original problem is so simple, but we'll see richer cases later on in the course.

Equation (186) tells us that, if we observe a very small current x , the probability that there really was a photon present is small, $\sim e^{-\theta}$. As the observed current becomes larger, the probability that a photon was present goes up, and, gradually, as x becomes large, we become certain [$P(1|x) \rightarrow 1$]. To build the best estimator of n from this one cell, our general result tells us that we should compute the conditional mean:

$$n_{\text{est}}(x) = \sum_n P(n|x)n \quad (189)$$

$$= P(0|x) \cdot (0) + P(1|x) \cdot (1) \quad (190)$$

$$= P(1|x). \quad (191)$$

Thus, the Boltzmann-like result [Eq (186)] for the probability of a photon being counted is, in fact, our best

estimator of the photon count in this limit where photons are very rare. Further, in this limit one can show that the optimal estimator for the total photon count, which after all is the sum of the individual n_i , is just the sum of the individual estimators.

Problem 31: Summing after the nonlinearity. Show that the optimal estimator for the total number of photons is the sum of estimators for the photon counts in individual rods, provided that the lights are very dim and hence photons are rare. The phrasing here is deliberately vague—you should explore the formulation of the problem, and see exactly what approximations are needed to make things come out right.

The end result of our calculations is that the optimal estimator of photon counts really is in the form shown at the right in Fig 29: nonlinearities serve to separate signal from noise in each rod cell, and these ‘cleaned’ signals are summed. How does this prediction compare with experiment? Careful measurements in the mouse retina show that the bipolar cells respond nonlinearly even to very dim flashes of light, in the range where the rods see single photons and respond linearly, with two photons producing twice the response to one photon. The form of the nonlinearity is what we expect from the theory, a roughly sigmoidal function that suppresses noise and passes signals only above an amplitude threshold. Importantly, this nonlinearity is observed in one class of bipolar cells but not others, and this is the class that, on other grounds, one would expect is most relevant for processing of rod outputs at low light levels.

Looking more quantitatively at the experiments [show some of the data, perhaps replotted in different forms ... go back and look at the original papers and clean up this paragraph!], we can see discrete, single photon events in the bipolar cells. Although the details vary across organisms, in this retina, one bipolar cell collects input from ~ 20 rod cells, but the variance of the background noise is larger than in the lower vertebrates that we first saw in Fig 4. As a result, if we sum the rod inputs and pass them through the observed nonlinearity—as in the model at left in Fig 29—we would not be able to resolve the single photon events. Field and Rieke considered a family of models in which the nonlinearity has the observed shape but the midpoint (analogous to the threshold θ above) is allowed to vary, and computed the signal to noise ratio at the bipolar cell output for the detection of flashes corresponding to a mean count of $\sim 10^{-4}$ photons/rod cell, which is, approximately, the point at which we can barely see something on a moonless night. Changing the threshold by a factor of two changes the signal to noise ratio by factors of several hundred. The measured value

of the threshold is within 8% of the predicted optimal setting, certainly close enough to make us think that we are on the right track.

The discussion thus far has emphasized separating signals from noise by their amplitudes.²⁴ We also can see, by looking closely at the traces of current vs time, that signal and noise have different frequency content. This suggests that we could also improve the signal to noise ratio by filtering. It’s useful to think about a more general problem, in which we observe a time dependent signal $y(t)$ that is driven by some underlying variable $x(t)$; let’s assume that the response of y to x is linear, but noisy, so that

$$y(t) = \int_{-\infty}^{\infty} d\tau g(\tau)x(t - \tau) + \eta(t), \quad (192)$$

where $g(\tau)$ describes the response function and $\eta(t)$ is the noise. What we would like to do is to use our observations on $y(t)$ to estimate $x(t)$.

Problem 32: Harmonic oscillator revisited. Just to be sure you understand what is going in Eq (192), think again about the Brownian motion of a damped harmonic oscillator, as in Problem [***], but now with an external force $F(t)$,

$$m \frac{d^2x(t)}{dt^2} + \gamma \frac{dx(t)}{dt} + \kappa x(t) = F(t) + \delta F(t). \quad (193)$$

Show that

$$x(t) = \int_{-\infty}^{\infty} d\tau g(\tau)F(t - \tau) + \eta(t). \quad (194)$$

Derive an explicit expression for the Fourier transform of $g(\tau)$, and find $g(\tau)$ itself in the limit of either small or large damping γ .

Since the y is linearly related to x , we might guess that we can make estimates using some sort of linear operation. As we have seen already in the case of the rod currents, this might not be right, but let’s try anyway—we’ll need somewhat more powerful mathematical tools to sort out, in general, when linear vs nonlinear computations are the most useful. We don’t have any reason to prefer one moment of time over another, so we should do something that is both linear and invariant under time translations, which means that our estimate must be of the form

$$x_{\text{est}}(t) = \int_{-\infty}^{\infty} dt' f(t - t')y(t'), \quad (195)$$

²⁴ Need to be a little careful here, since the analysis from Fred’ lab actually involves applying the nonlinearity to voltages that have already been filtered. Presumably this will be clearer when I am pointing to the real data .. come back and fix this!

where $f(t)$ is the ‘filter’ that we hope will separate signal and noise. Following the spirit of the discussion above, we’ll ask that our estimate be as close as possible to the right answer in the sense of mean-square error. Thus, our task is to find the filter $f(t)$ that minimizes

$$\mathcal{E} = \left\langle \left[x(t) - \int_{-\infty}^{\infty} dt' f(t-t')y(t') \right]^2 \right\rangle. \quad (196)$$

In taking the expectation value of the mean-square error, we average over possible realizations of the noise and the variations in the input signal $x(t)$. In practice this averaging can also be thought of as including an average over time.²⁵ Thus we can also write

$$\mathcal{E} = \left\langle \int_{-\infty}^{\infty} dt \left[x(t) - \int_{-\infty}^{\infty} dt' f(t-t')y(t') \right]^2 \right\rangle. \quad (197)$$

This is useful because we can then pass to the Fourier domain. We recall that for any function $z(t)$,

$$\int_{-\infty}^{\infty} dt z^2(t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} |\tilde{z}(\omega)|^2, \quad (198)$$

and that the Fourier transform of a convolution is the product of transforms,

$$\int_{-\infty}^{\infty} dt e^{i\omega t} \int_{-\infty}^{\infty} dt' f(t-t')y(t') = \tilde{f}(\omega)\tilde{y}(\omega). \quad (199)$$

Putting things together, we can rewrite the mean-square error as

$$\mathcal{E} = \left\langle \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \left| \tilde{x}(\omega) - \tilde{f}(\omega)\tilde{y}(\omega) \right|^2 \right\rangle. \quad (200)$$

Now each frequency component of our filter $\tilde{f}(\omega)$ appears independently of all the others, so minimizing \mathcal{E} is straightforward. The result is that

$$\tilde{f}(\omega) = \frac{\langle \tilde{y}^*(\omega)\tilde{x}(\omega) \rangle}{\langle |\tilde{y}(\omega)|^2 \rangle}. \quad (201)$$

Problem 33: Details of the optimal filter. Fill in the steps leading to Eq (201). Be careful about the fact that $f(t)$ is real, and so the transform $\tilde{f}(\omega)$ is not arbitrary. Hint: think about positive and negative frequency components.

To finish our calculation, we go back to Eq (192), which in the frequency domain can be written as

$$\tilde{y}(\omega) = \tilde{g}(\omega)\tilde{x}(\omega) + \tilde{\eta}(\omega). \quad (202)$$

Thus

$$\langle \tilde{y}^*(\omega)\tilde{x}(\omega) \rangle = \tilde{g}^*(\omega)\langle |\tilde{x}(\omega)|^2 \rangle \quad (203)$$

$$\langle |\tilde{y}(\omega)|^2 \rangle = |\tilde{g}(\omega)|^2\langle |\tilde{x}(\omega)|^2 \rangle + \langle |\tilde{\eta}(\omega)|^2 \rangle. \quad (204)$$

If all of these variables have zero mean (which we can have be true just by choosing the origin correctly), then quantities such as $\langle |\tilde{x}(\omega)|^2 \rangle$ are the variances of Fourier components, which we know (see Appendix B) are proportional to power spectra. Finally, then, we can substitute into our expression for the optimal filter to find

$$\tilde{f}(\omega) = \frac{\tilde{g}^*(\omega)S_x(\omega)}{|\tilde{g}(\omega)|^2S_x(\omega) + S_\eta(\omega)}, \quad (205)$$

where, as before, S_x and S_η are the power spectra of x and η , respectively.

In the case that noise is small, we can let $S_\eta \rightarrow 0$ and we find

$$\tilde{f}(\omega) \rightarrow \frac{1}{\tilde{g}(\omega)}. \quad (206)$$

This means that, when noise can be neglected, the best way to estimate the underlying signal is just to invert the response function of our sensor, which makes sense. Notice that since \tilde{g} generally serves to smooth the time dependence of $y(t)$ relative to that of $x(t)$, the filter $\tilde{f}(\omega) \sim 1/\tilde{g}(\omega)$ undoes this smoothing. This is important because it reminds us that smoothing in and of itself does not set a limit to time resolution—it is only the combination of smoothing with noise that obscures rapid variations in the signal.

Guided by the limit of high signal to noise ratio, we can rewrite the optimal filter as

$$\tilde{f}(\omega) = \frac{1}{\tilde{g}(\omega)} \cdot \frac{|\tilde{g}(\omega)|^2S_x(\omega)}{|\tilde{g}(\omega)|^2S_x(\omega) + S_\eta(\omega)} \quad (207)$$

$$= \frac{1}{\tilde{g}(\omega)} \cdot \frac{SNR(\omega)}{1 + SNR(\omega)}, \quad (208)$$

where we identify the signal to noise ratio at each frequency, $SNR(\omega) = |\tilde{g}(\omega)|^2S_x(\omega)/S_\eta(\omega)$. Clearly, as the signal to noise ratio declines, so does the optimal filter—in the limit, if $SNR(\omega) = 0$, everything we find at frequency ω must be noise, and so it should be zeroed out if we want to minimize its corrupting effects on our estimates.

In the case of the retina, x is the light intensity, and y are the currents generated by the rod cells. When it’s very dark outside, the signal to noise ratio is low, so that

$$\tilde{f}(\omega) \rightarrow \frac{\tilde{g}^*(\omega)}{S_\eta(\omega)} \cdot S_x(\omega). \quad (209)$$

²⁵ More formally, if all the relevant random variations are ergodic, then averaging over the distributions and averaging over time will be the same.

The filter in this case has two pieces, one of which depends only on the properties of the rod cell,

$$\tilde{f}_1(\omega) = \frac{\tilde{g}^*(\omega)}{S_\eta(\omega)}, \quad (210)$$

and another piece that depends on the power spectrum of the time dependent light intensity, $S_x(\omega)$. With a bit more formalism we can show that this first filter, $\tilde{f}_1(\omega)$, has a universal meaning, so that if instead of estimating the light intensity itself, we try to estimate something else—e.g., the velocity of motion of an object across the visual field—then the first step in the estimation process is still to apply this filter. So, it is a natural hypothesis that this filter will be implemented near the first stages of visual processing, in the transfer of signals from the rods to the bipolar cells.

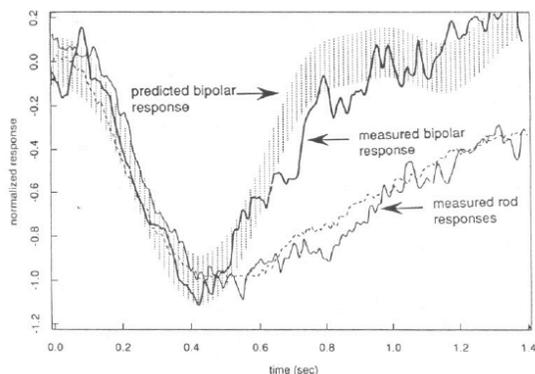


FIG. 30 Voltage responses of rod and bipolar cells in the salamander retina, compared with theory, from Rieke et al (1991). The theory is that the transmission from rod currents to bipolar cell voltage implements the optimal filter as in Eq (210). Measured responses are averages over many presentations of a flash at $t = 0$ that results in an average of five photons being counted. The predicted filter is computed from measured signal and noise properties of the rod cell, with no adjustable parameters.

Problem 34: Filtering the real rod currents. The raw data that were used to generate Fig 4 are available at <http://www.princeton.edu/~wbialek/PHY562/data.html>, in the file `rodcurrents.mat`. The data consist of 395 samples of the rod current in response to dim flashes of light. The data are sampled in 10 ms bins, and the flash is delivered in the 100th bin. If these ideas about filtering are sensible, we should be able to do a better job of discriminating between zero, one and two photons by using the right filter. Notice that filtering of a response that is locked to a particular moment in time is equivalent to taking a weighted linear combination of the currents at different times relative to the flash. Thus you can think of the current in response to one flash as a vector, and filtering amounts to taking the dot product of this vector with some template. As a first step, you should reproduce the results of Fig 4, which are based just on averaging points in the neighborhood of the peak. Under some conditions, the best

template would just be the average single photon response. How well does this work? What conditions would make this work best? Can you do better? These data are from experiments by FM Rieke and collaborators at the University of Washington, and thanks to Fred for making them available.

The idea that the rod/bipolar synapse implements an optimal filter is interesting not least because this leads us to a prediction for the dynamics of this synapse, Eq (210), which is written entirely in terms of the signal and noise characteristics of the rod cell itself. All of these properties are measurable, so there are no free parameters in this prediction.²⁶ To get some feeling for how these predictions work, remember that the noise in the rod cell has two components—the spontaneous isomerizations of rhodopsin, which have the same frequency content as the real signal, and the continuous background noise, which extends to higher frequency. If we have only the spontaneous isomerizations, then $S_\eta \sim |\tilde{g}|^2$, and we are again in the situation where the best estimate is obtained by ‘unsmoothing’ the response, essentially recovering sharp pulses at the precise moments when photons are absorbed. This unsmoothing, or high-pass filtering, is cut off by the presence of the continuous background noise, and the different effects combine to make \tilde{f}_1 a band-pass filter. By the time the theory was worked out, it was already known that something like band-pass filtering was happening at this synapse; among other things this speeds up the otherwise rather slow response of the rod. In Fig 30 we see a more detailed comparison of theory and experiment.

Problem 35: Optimal filters, more rigorously. Several things were left out of the optimal filter analysis above; let’s try to put them back here.

(a.) Assume that there is a signal $s(t)$, and we observe, in the simplest case, a noisy version of this signal, $y(t) = s(t) + \eta(t)$. Let the power spectrum of $s(t)$ be given by $S(\omega)$, and the power spectrum of the noise $\eta(t)$ be given by $N(\omega)$. Further, assume that both signal and noise have Gaussian statistics. Show that the distribution of signals given our observations is

$$P[s(t)|y(t)] = \frac{1}{Z} \exp \left[-\frac{1}{2} \int \frac{d\omega}{2\pi} \frac{|\tilde{s}(\omega) - \tilde{y}(\omega)|^2}{N(\omega)} - \frac{1}{2} \int \frac{d\omega}{2\pi} \frac{|\tilde{s}(\omega)|^2}{S(\omega)} \right]. \quad (211)$$

²⁶ We should be a bit careful here. The filter, as written, is not causal. Thus, to make a real prediction, we need to shift the filter so that it doesn’t have any support at negative times. To make a well defined prediction, we adopt the minimal delay that makes this work. One could perhaps do better, studying the optimal filtering problem with explicitly causal filters, and considering the tradeoff between errors and acceptable delays.

(b.) Show that the most likely function $\tilde{s}(\omega)$ given the data on y is also the best estimate in the least squares sense, and is given by

$$\tilde{s}_{\text{est}}^{(\text{nc})}(\omega) = \frac{S(\omega)}{S(\omega) + N(\omega)} \tilde{y}(\omega); \quad (212)$$

the superscript (nc) reminds us that this estimate does not respect causality. Show that this is consistent with Eq (205). Notice that you didn't assume the optimal estimator was linear, so you have shown that it is (!). Which of the assumptions here are essential in obtaining this result?

(c.) The non-causal estimator is Eq (212) is constructed by assuming that we have access to the entire function $y(t)$, with $-\infty < t < \infty$, as we try to estimate, for example $s(t = 0)$. If we want our estimator to be something that we can build, then we must impose causality: the estimate of $s(t)$ can be based only on the history $y_- \equiv y(t' < t)$. Another way of saying this is that we don't really know $y_+ \equiv y(t' > t)$, so we should average over this part of the trajectory. But the average should be computed in the distribution $P[y_+|y_-]$. To construct this, start by showing that

$$P[y_+, y_-] \equiv P[y(t)] = \frac{1}{Z_0} \exp \left[-\frac{1}{2} \int \frac{d\omega}{2\pi} \frac{|\tilde{y}(\omega)|^2}{S(\omega) + N(\omega)} \right]. \quad (213)$$

$$P[y_+, y_-] = \frac{1}{Z_0} \exp \left[-\frac{1}{2} \int_{-\infty}^0 dt \left| \int \frac{d\omega}{2\pi} e^{-i\omega t} \tilde{y}_-(\omega) \tilde{\psi}(\omega) \right|^2 - \frac{1}{2} \int_0^{\infty} dt \left| \int \frac{d\omega}{2\pi} e^{-i\omega t} (\tilde{y}_-(\omega) + \tilde{y}_+(\omega)) \tilde{\psi}(\omega) \right|^2 \right], \quad (217)$$

and that

$$P[y_+|y_-] \propto \exp \left[-\frac{1}{2} \int_0^{\infty} dt \left| \int \frac{d\omega}{2\pi} e^{-i\omega t} (\tilde{y}_-(\omega) + \tilde{y}_+(\omega)) \tilde{\psi}(\omega) \right|^2 \right]. \quad (218)$$

Explain why averaging over the distribution $P[y_+|y_-]$ is equivalent to imposing the "equation of motion"

$$\int \frac{d\omega}{2\pi} e^{-i\omega t} (\tilde{y}_-(\omega) + \tilde{y}_+(\omega)) \tilde{\psi}(\omega) = 0 \quad (219)$$

at times $t > 0$.

(f.) Write the non-causal estimate Eq (212) in the time domain as

$$s_{\text{est}}^{(\text{nc})}(t) = \int \frac{d\omega}{2\pi} e^{-i\omega t} \tilde{\psi}^*(\omega) \tilde{\psi}(\omega) \tilde{y}(\omega). \quad (220)$$

But the combination $\tilde{\psi}(\omega) \tilde{y}(\omega)$ is the Fourier transform of $z(t)$, which is the convolution of $\psi(t)$ with $y(t)$. Show that Eq (219) implies that the average of $z(t)$ is the distribution $P[y_+|y_-]$ vanishes for $t > 0$, and hence the averaging over y_+ is equivalent to replacing

$$\tilde{\psi}(\omega) \tilde{y}(\omega) \rightarrow \int_{-\infty}^0 d\tau e^{i\omega\tau} \int \frac{d\omega'}{2\pi} \tilde{\psi}(\omega') \tilde{y}(\omega') e^{-i\omega'\tau} \quad (221)$$

in Eq (212). Put all the pieces together to show that there is a causal estimate of $s(t)$ which can be written as

$$s_{\text{est}}(t) = \int \frac{d\omega}{2\pi} e^{-i\omega t} \tilde{k}(\omega) \tilde{y}(\omega), \quad (222)$$

where

$$\tilde{k}(\omega) = \tilde{\psi}(\omega) \int_0^{\infty} d\tau e^{i\omega\tau} \int \frac{d\omega'}{2\pi} e^{-i\omega'\tau} S(\omega') \tilde{\psi}^*(\omega'). \quad (223)$$

Verify that this filter is causal.

(d.) Recall that when we discuss causality, it is useful to think about the frequency ω as a complex variable. Explain why we can write

$$\frac{1}{S(\omega) + N(\omega)} = |\tilde{\psi}(\omega)|^2, \quad (214)$$

where $\tilde{\psi}(\omega)$ has no poles in the upper half of the complex ω plane. Verify that, with this decomposition,

$$\psi(t) = \int \frac{d\omega}{2\pi} e^{-i\omega t} \tilde{\psi}(\omega) \quad (215)$$

is causal, that is $\psi(t < 0) = 0$. Consider the case where the signal has a correlation time τ_c , so that $S(\omega) = 2\sigma^2\tau_c/[1 + (\omega\tau_c)^2]$, and the noise is white $N(\omega) = N_0$; construct $\tilde{\psi}(\omega)$ explicitly in this case.

(e.) Putting Eq (213) together with Eq (214), we can write

$$P[y_+, y_-] = \frac{1}{Z_0} \exp \left[-\frac{1}{2} \int \frac{d\omega}{2\pi} \left| \tilde{y}(\omega) \tilde{\psi}(\omega) \right|^2 \right]. \quad (216)$$

Show that

filtering. In the other, the same separation is achieved by a static nonlinearity, applied in practice to a linearly filtered signal. Presumably there is some more general nonlinear dynamic transformation that really does the best job. We expect that the proper mix depends on the detailed spectral structure of the signals and noise, and on the relative amplitudes of the signal and noise, which might be why the different effects are clearest in retinas from very different species. Indeed, there is yet another approach which emphasizes that the dynamic range of neural outputs is limited, and that this constrains how much information the second order neuron can provide about visual inputs; filters and nonlinearities can be chosen to optimize this information transmission across a wide range of background light intensities, rather than focusing only on the detectability of the dimmest lights. This approach has received the most attention in invertebrate retinas, such as the fly that we met near the end of Section I.A, and we will return to these ideas in Chapter 4. It would be nice to see this all put together correctly, and this is an open problem, surely with room for some surprises.

So far we have followed the single photon signal from the single rhodopsin molecule to the biochemical network that amplifies this molecular event into a macroscopic current, and then traced the processing of this electrical signal as it crosses the first synapse. To claim that we have said anything about *vision*, we have to at least follow the signal out of the retina and on its way to the brain. [By now we should have said more about retinal anatomy—optic nerve, made up of the axons from 'retinal ganglion cells,' and the stereotyped action potentials

It is worth noting that we have given two very different analyses. In one, signals and noise are separated by linear

that propagate along these axons. Should also discuss techniques for picking up the signals, up to current work with electrode arrays. Show a modern figure, e.g. from Berry's lab.]

The classic experiments on single photon responses in retinal ganglion cells were done well before it was possible to measure the responses of single rods. The spikes from single ganglion cells are relatively easy to record, and one can try to do something like the Hecht, Shlaer and Pirenne experiment, but instead of “seeing” (as in Fig 2), you just ask if you can detect the spikes. There were a number of hints in the data that a single absorbed photon generated more than one spike, so some care is required. As shown in Fig 31, there are neurons that seem to count by threes—if you wait for three spikes, the probability of seeing is what you expect for setting a threshold of $K = 1$ photon, if you wait for six spikes it is as if $K = 2$, and so on. This simple linear relation between photons and spikes also makes it easy to estimate the rate of spontaneous photon-like events in the dark. Note that if photons arrive as a Poisson process, and each photon generates multiple spikes, then the spikes are *not* a Poisson process; this idea of Poisson events driving a second point process to generate non-Poisson variability has received renewed attention in the context of gene expression, where the a single messenger RNA molecule (perhaps generated from a Poisson process) can be translated to yield multiple protein molecules.

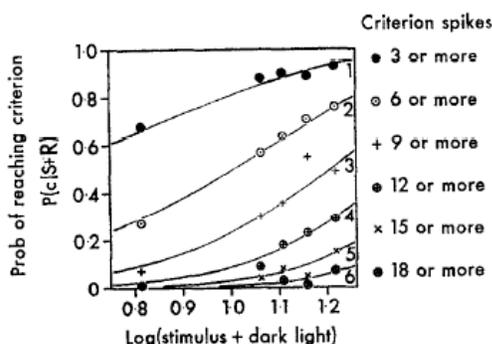


FIG. 31 A frequency of seeing experiment with spikes, from Barlow et al (1971). Recording from a single retinal ganglion cell, you can say you “saw” a flash when you detect 3, 6, 9, ... or more spikes within a small window of time (here, 200 ms). The probability of reaching this criterion is plotted vs the log of the flash intensity, as in the original Hecht, Shlaer and Pirenne experiments (Fig 2), but here the intensity is adjusted to include a background rate of photon-like events (“dark light”). Curves are from Eq (2), with the indicated values of the threshold K . Notice that three spikes corresponds to one photon.

Problem 36: Poisson-driven bursts. A characteristic feature of events drawn out of a Poisson process is that if we count the number of events, the variance of this number is equal to the mean. Suppose that each photon triggers exactly b spikes. What is the ratio of variance to mean (sometimes called the Fano factor) for spike counts in response to light flashes of fixed intensity? Suppose the the burst of spikes itself is a Poisson process, with mean b . Now what happens to the variance/mean ratio?

Before tracing the connections between individual spikes and photons, it was possible to do a different experiment, just counting spikes in response to flashes of different intensities, and asking what is the smallest value of the difference ΔI such that intensities I and $I + \Delta I$ can be distinguished reliably. The answer, of course, depends on the background intensity I [show figure from Barlow (1965)?]. For sufficiently small I , the just noticeable different ΔI is constant. For large I , one finds $\Delta I \propto I$, so the just noticeable fractional change in intensity is constant; this is common to many perceptual modalities, and is called Weber’s law. At intermediate intensities one can see $\Delta I \propto \sqrt{I}$. This last result, predicted by Rose and de Vries (cf Section 1.1), is what you expect if detecting a change in intensity just requires discriminating against the Poisson fluctuations in the arrival of photons. At high intensities, we are counting many photons, and probably the system just can’t keep up; then fluctuations in the gain of the response dominate, and this can result in Weber’s law. At the lowest intensities, the photons delivered by the flash are few in comparison with the thermal isomerizations of Rhodopsin, and this constant noise source sets the threshold for discrimination. Happily, the rate of spontaneous isomerizations estimated from these sorts of experiments agrees with other estimates, including the (much later) direct measurements on rod cells discussed previously. This work on discrimination with neurons also is important because it represents one of the first efforts to connect the perceptual abilities of whole organisms with the response of individual neurons.

If retinal ganglion cells generate three spikes for every photon, lights wouldn’t need to be very bright before the cells should be generating thousands of spikes per second, and this is impossible—the spikes themselves are roughly one millisecond in duration, and all neurons have a ‘refractory period’ that defines a minimum time (like a hard core repulsion) between successive action potentials. The answer is something we have seen already in the voltage responses of fly photoreceptors (Fig 13): as the background light intensity increases, the retina adapts and turns down the gain, in this case generating fewer spikes per photon. Of course this takes some time, so if we

suddenly expose the retina to a bright light there is very rapid spiking, which then adapts away to a much slower rate. [Need a figure about light/dark adaptation.] If we imagine that our perceptions are driven fairly directly by the spikes, then our impression of the brightness of the light should similarly fade away. This is true not just for light (as you experience whenever you walk outside on a bright sunny day); almost all constant sensory inputs get adapted away—think about the fact that you don't feel the pressure generated by your shoes a few minutes after you tie them. But there are more subtle issues as well, involving the possibility that the coding strategy used by the retina adapts to the whole distribution of inputs rather than just the mean; this is observed, and many subsequent experiments are aimed at understanding the molecular and cellular mechanisms of these effects. The possibility that adaptation serves to optimize the efficiency of coding continuous signals into discrete spikes is something we will return to in Chapter 4.

The problem of photon counting—or any simple detection task—also hides a deeper question: how does the brain “know” what it needs to do in any given task? Even in our simple example of setting a threshold to maximize the probability of a correct answer, the optimal observer must at least implicitly acquire knowledge of the relevant probability distributions. Along these lines, there is more to the ‘toad cooling’ experiment than a test of photon counting and dark noise. The retina has adaptive mechanisms that allow the response to speed up at higher levels of background light, in effect integrating for shorter times when we can be sure that the signal to noise ratio will be high. The flip side of this mechanism is that the retinal response slows down dramatically in the dark [connect back to photoreceptor responses; a figure here would be good, including τ vs I relevant to Aho et al]. In moderate darkness (dusk or bright moonlight) the slowing of the retinal response is reflected directly in a slowing of the animal's behavior. It is as if the toad experiences an illusion because images of its target are delayed, and it strikes at the delayed image. It is worth emphasizing that we see a closely related illusion.

Problem 37: Knowing where to look. Give a problem to illustrate the role of uncertainty in reducing performance.

Imagine watching a pendulum swinging while wearing glasses that have a neutral density filter over one eye, so the mean light intensity in the two eyes is different. The dimmer light results in a slower retina, so the signals from the two eyes are not synchronous, and recall that differences in the images between our right and left eyes

are cues to the depth of an object. As we try to interpret these signals in terms of motion, we find that even if the pendulum is swinging in a plane parallel to the line between our eyes, what we see is motion in 3D. The magnitude of the apparent depth of oscillation is related to the neutral density and hence to the slowing of signals in the ‘darker’ retina. This is called the Pulfrich effect.

If the pattern of delay vs light intensity continued down to the light levels in the darkest night, it would be a disaster, since the delay would mean that the toad inevitably strikes behind the target! In fact, the toad does not strike at all in the first few trials of the experiment in dim light, and then strikes well within the target. It is hard to escape the conclusion that the animal is learning about the typical velocity of the target and then using this knowledge to extrapolate and thereby correct for retinal delays.²⁷ Thus, performance in the limit where we count photons involves not only efficient processing of these small signals but also learning as much as possible about the world so that these small signals become interpretable.

If you'd like a general overview of the retina, a good source is Dowling (1987). For the experiments on nonlinear summation at the rod–bipolar synapse, along with a discussion of the theoretical issues of noise and reliability, see Field & Rieke (2002a). The analysis of optimal filtering is presented in Bialek & Owen (1990) and Rieke et al (1991). For a discussion how our experience of a dark night translates into photons per rod per second, see Walraven et al (1990).

- Bialek & Owen 1990:** Temporal filtering in retinal bipolar cells: Elements of an optimal computation? W Bialek & WG Owen, *Biophys J* **58**, 1227–1233 (1990).
- Dowling 1987:** *The Retina: An Approachable Part of the Brain* JE Dowling (Harvard University Press, Cambridge, 1987).
- Field & Rieke 2002a:** Nonlinear signal transfer from mouse rods to bipolar cells and implications for visual sensitivity. GD Field & F Rieke, *Neuron* **34**, 773–785 (2002).
- Rieke et al 1991:** Optimal filtering in the salamander retina. F Rieke, WG Owen & W Bialek, in *Advances in Neural Information Processing 3*, R Lippman, J Moody & D Touretzky, eds, pp 377–383 (Morgan Kaufmann, San Mateo CA, 1991).
- Walraven et al 1990:** The control of visual sensitivity. J Walraven, C Enroth–Cugell, DC Hood, DIA MacLeod & JL Schnapf, in *Visual Perception: The Neurophysiological Foundations*, L Spillmann & SJ Werner, eds, pp 53–101 (Academic Press, San Diego, 1990).

The classic presentations of filtering, estimation and prediction are by Kolmogorov (1939, 1941) and Wiener (1949). The long problem about optimal filtering is based on Potters & Bialek (1994).

Kolmogoroff 1939: Sur l'interpolation et extrapolations des suites stationnaires. A Kolmogoroff, *C R Acad Sci Paris* **208**, 2043–2045 (1939).

²⁷ As far as I know there are no further experiments that probe this learning more directly, e.g. by having the target move at variable velocities.

Kolmogorov 1941: Interpolation and extrapolation of stationary random sequences (in Russian). AN Kolmogorov, *Izv Akad Nauk USSR Ser Mat* **5**, 3–14 (1941). English translation in *Selected Works of AN Kolmogorov, Vol II*, AN Shiryagev, ed, pp 272–280 (Kluwer Academic, Dordrecht, 1992).

Potters & Bialek 1994: Statistical mechanics and visual signal processing. M Potters & W Bialek, *J Phys I France* **4**, 1755–1775 (1994); arXiv:cond-mat/9401072 (1994).

Wiener 1949: *Extrapolation, Interpolation and Smoothing of Time Series* N Wiener (Wiley, New York, 1949).

The idea of maximizing information transmission across the first visual synapse is something we will discuss at greater length in Chapter 4. Still, you might like to look ahead, so here are some references to how these ideas developed in the context of fly vision.

Hateren 1992: Real and optimal neural images in early vision. JH van Hateren, *Nature* **360**, 68–70 (1992).

Laughlin 1981: A simple coding procedure enhances a neuron’s information capacity. SB Laughlin, *Z Naturforsch* **36c**, 910–912 (1981).

Srinivasan et al 1982: Predictive coding: A fresh view of inhibition in the retina. MV Srinivasan, SB Laughlin & A Dubs, *Proc R Soc Lond Ser B* **216**, 427–459 (1982).

The classic paper about single photon responses in retinal ganglion cells is Barlow et al (1971); it has quite a lot of detail, and still makes great reading. [Mastronade 1983?; might also need pointers to more modern recordings] The idea that single molecular events can drive bursts, generating non-Poisson statistics, reappears thirty years later in the context of gene expression; see for example Ozbudak et al (2002). The early papers on intensity discrimination using spikes from single neurons are Barlow (1965) and Barlow & Levick (1969); see also the even earlier work from FitzHugh (1957, 1958).

Barlow 1965: Optic nerve impulses and Weber’s law. HB Barlow, *Cold Spring Harb Symp Quant Biol* **30**, 539–546 (1965).

Barlow & Levick 1969: Three factors limiting the reliable detection of light by retinal ganglion cells of the cat. HB Barlow & WR Levick, *J Physiol (Lond)* **200**, 1–24 (1969).

Barlow et al 1971: Responses to single quanta of light in retinal ganglion cells of the cat. HB Barlow, WR Levick & M Yoon, *Vision Res Suppl* **3**, 87–101 (1971).

FitzHugh 1957: The statistical detection of threshold signals in the retina. R FitzHugh, *J Gen Physiol* **40**, 925–948 (1957).

FitzHugh 1958: A statistical analyzer for optic nerve messages. R FitzHugh, *J Gen Physiol* **41**, 675–692 (1958).

Ozbudak et al 2002: Regulation of noise in the expression of a single gene. E Ozbudak, M Thattai, I Kurtser, AD Grossman & A van Oudenaarden, *Nature Gen* **31**, 69–73 (2002).

The observation that neurons gradually diminish their response to constant stimuli goes back to Adrian’s first experiments recording the spikes from single cells; he immediately saw the connection to the fading of our perceptions when inputs are constant, and this sort of direct mapping from neural responses to human experience is now the common language we use in thinking about the brain and mind. An early paper about adaptation to the distribution of inputs is Smirnakis et al (1997). Since then a number of papers have explored more complex versions of this adaptation, as well as trying to tease apart the underlying mechanisms; some examples are Rieke (2001), Kim & Rieke (2001, 2003), and Baccus & Meister (2002).

Adrian 1928: *The Basis of Sensation* ED Adrian (Christopher’s, London, 1928).

Baccus & Meister 2002: Fast and slow adaptation in retinal circuitry. SA Baccus & M Meister, *Neuron* **36**, 909–919 (2002).

Kim & Rieke 2001: Temporal contrast adaptation in the input and output signals of salamander retinal ganglion cells. KJ Kim & F Rieke, *J Neurosci* **21**, 287–299 (2001).

Kim & Rieke 2003: Slow Na⁺ inactivation and variance adaptation in salamander retinal ganglion cells. *J Neurosci* **23**, 1506–1515 (2003).

Rieke 2001: Temporal contrast adaptation in salamander bipolar cells. F Rieke, *J Neurosci* **21**, 9445–9454 (2001).

Smirnakis et al 1997: Adaptation of retinal processing to image contrast and spatial scale. S Smirnakis, MJ Berry II, DK Warland, W Bialek & M Meister, *Nature* **386**, 69–73 (1997).

There is a decent demonstration of the Pulfrich effect available on the web (Newbold 1999). The experiments on reaction times in toads and the connection to retinal delays are from the work of Aho et al (1993).

Aho et al 1993: Visual performance of the toad (*Bufo bufo*) at low light levels: Retinal ganglion cell responses and prey-catching accuracy. A-C Aho, K Donner, S Helenius, LO Larsen & T Reuter, *J Comp Physiol A* **172**, 671–682 (1993).

Newbold 1999: The Pulfrich illusion. M Newbold, <http://dogfeathers.com/java/pulfrich.html> (1999).

E. Perspectives

What have we learned from all of this? I think the first thing to notice is that we have at least one example of a real biological system that is susceptible to the sorts of reproducible, quantitative experiments that we are used to in the rest of physics. This is not obvious, and runs counter to some fairly widespread prejudices. Although things can get complicated,²⁸ it does seem that, with care, we can speak precisely about the properties of cells in the retina, not just on average over many cells but cell by cell, in enough detail that even the noise in the cellular response itself is reproducible from cell to cell, organism to organism. It’s important that all of this is not guaranteed—removing cells from their natural milieu can be traumatic, and every trauma is different. If you dig into the original papers, you’ll see glimpses of the many things which experimentalists need to get right in order to achieve the level of precision that we have emphasized in our discussion—the things one needs to do in order to turn the exploration of living systems into a physics experiment.

The second point is that the performance of these biological systems—something which results from mechanisms of incredible complexity—really is determined by the physics of the “problem” that the system has been selected to solve. If you plan on going out in the dark of night, there is an obvious benefit to being able to detect dimmer sources of light, and to making reliable discriminations among subtly different intensities and, ultimately, different spatiotemporal patterns. You can’t do

²⁸ We have not explored, for example, the fact that there are many kinds of ganglion cells.

better than to count every photon, and the reliability of photon counting by the system as a whole can't be better than the limits set by noise in the detector elements. The fact that real visual systems reach these limits is extraordinary.

The last point concerns the nature of the explanations that we are looking for. We have discussed the currents generated in response to single photons, the filter characteristics and nonlinearities of synapses, and the spiking outputs of ganglion cells, and in each case we can ask why these properties of the system are as we observe them to be. Importantly, we can ask analogous questions about a wide variety of systems, from individual molecules to the regulation of gene expression in single cells to the dynamics of neural networks in our brains. What are we doing when we look for an "explanation" of the data?

When we ask "why" in relation to a biological system, we can imagine (at least) two very different kinds of answers.²⁹ First, we could plunge into the microscopic mechanisms. As we have seen in looking at the dynamics of biochemical amplification in the rod cell, what we observe as functional behavior of the system as a whole depends on a large number of parameters: the rates of various chemical reactions, the concentrations of various proteins, the density of ion channels in the membrane, the binding energies of cGMP to the channel, and so on. To emphasize the obvious, these are not fundamental constants. In a very real sense, almost all of these

parameters are under the organism's control.

Our genome encodes hundreds of different ion channels, and the parameters of the rod cell would change if it chose to read out the instructions for making one channel rather than another. Further, the cell can make more or less of these proteins, again adjusting the parameters of the system essentially by changing the concentrations of relevant molecules. A similar line of argument applies to other components of the system (and many other systems), since many key molecules are members of families with slightly different properties, and cells choose which member of the family will be expressed. More subtly, many of these molecules can be modified, e.g. by covalent attachment of phosphate groups as with the shut-off of rhodopsin, and these modifications provide another pathway for adjusting parameters. Thus, saying that (for example) the response properties of the rod cell are determined by the parameters of a biochemical network is very different from saying that the absorption spectrum of hydrogen is determined by the charge and mass of the electron—we would have to go into some alternative universe to change the properties of the electron, but most of the parameters of the biochemical network are under the control of the cell, and these parameters can and do change in response to other signals.

An explanation of functional behavior in microscopic terms, then, may be correct but somehow unsatisfying. Further, there may be more microscopic parameters than phenomenological parameters, and this may be critical in allowing the system to achieve nearly identical functional behaviors via several different mechanisms. But all of this casts doubt on the idea that we are 'explaining' the functional behavior in molecular terms.

A second, very different kind of explanation is suggested by our discussion of the first synapse in vision, between the rod and bipolar cells. In that discussion (Section I.D), we promoted the evidence of near optimal performance at the problem of photon counting into a principle from which the functional properties of the system could be derived. In this view, the system is the way it is because evolution has selected the best solution to a problem that is essential in the life of the organism. This principle doesn't tell us how the optimum is reached, but it can predict the observable behavior of the system. Evidently there are many objections to this approach, but of course it is familiar, since many different ideas in theoretical physics can be formulated as variational principles, from least action in classical mechanics to the minimization of free energy in equilibrium thermodynamics, among others.

Organizing our thinking about biological systems around optimization principles tends to evoke philosophical discussions, in the pejorative sense that scientists use this term. I would like to avoid discussions of this flavor. If we are going to suggest that "biological systems maximize X" is a principle, then rather than having ev-

²⁹ My colleague Rob de Ruyter van Steveninck has an excellent way of talking about closely related issues. He once began a lecture by contrasting two different questions: Why is the sky blue? Why are trees green?. The answer to the first question is a standard part of a good, high level course on electromagnetism: when light scatters from a small particle—and molecules in the atmosphere are much smaller than the wavelength of light—the scattering is stronger at shorter wavelengths; this is called Rayleigh scattering. Thus, red light (long wavelengths) moves along a more nearly straight path than does blue light (short wavelength). The light that we see, which has been scattered, therefore has been enriched in the blue part of the spectrum, and this effect is stronger if we look further away from the sun. So, the sky is blue because of the way in which light scatters from molecules. We can answer the question about the color of trees in much the same way that we answered the question about the color of the sky: leaves contain a molecule called chlorophyll, which is quite a large molecule compared with the oxygen and nitrogen in the air, and this molecule actually absorbs visible light; the absorption is strongest for red and blue light, so what is reflected back to us is the (intermediate wavelength) green light. Unlike the color of the sky, the color of trees could have a different explanation. Imagine trees of other colors—blue, red, perhaps even striped. Microscopically, this could happen because their leaves contain molecules other than chlorophyll, or even molecules related to chlorophyll but with slightly different structures. But trees of different colors will compete for resources, and some will grow faster than others. The forces of natural selection plausibly will cause one color of tree to win out over the others. In this sense, we can say that trees are green because green trees are more successful, or more fit in their environment.

everyone express their opinion about whether this is a good idea, we should discipline ourselves and insist on criteria that allow such claims to be meaningful and predictive. First, we have to understand why X can't be arbitrarily large—we need to have a theory which defines the physical limits to performance. Second, we should actually be able to measure X, and compare its value with this theoretical maximum. Finally, maximizing X should generate some definite predictions about the structure and dynamics of the system, predictions that can be tested in independent, quantitative experiments. In what follows, we'll look at three different broad ideas about what X might be, and hopefully we'll be able to maintain the discipline that I have outlined here. Perhaps the most important lesson from the example of photon counting is that we can carry through this program and maintain contact with real data. The challenge is to choose principles (candidate Xs) that are more generally applicable than the very specific idea that the retina maximizes the reliability of seeing on a dark night.

