

1. PHYSICS OF SENSORY TRANSDUCTION

PURPOSEFUL BEHAVIOR begins with sensory perception. We gather information about the sights, smells, tastes, and textures of our environments using different sensory systems. We respond to sensory information using our brains and motor systems.

At the sensory periphery, specialized detectors are tuned to different types of environmental stimuli (Figs. 2, 1). In our noses, for example, diverse *olfactory receptors* in the nasal epithelium are tuned to different volatile chemicals. In our retinas, specialized *photoreceptors* are tuned to light with different qualities. We have *trichromatic* vision because our retinas have three types of *cone cells* that are sensitive to long, medium, or short wavelengths of bright light. We have one type of rod photoreceptor that is sensitive to dim light. Because we only have one type of rod photoreceptor with one type of wavelength sensitivity, vision in dim light is *monochromatic*.

Information travels from sensory receptors to our brains along *afferent* neuronal processes. Our modern circuit-level understanding of sensory processing began with the anatomical studies of **Santiago Ramón y Cajal** (1852-1934) who visualized the detailed structure of many types of sensory neurons, as well as the wires and intercellular contacts (synapses) that carried information to the brain (Fig. 2). Cajal did this by perfecting Golgi's method for sparsely staining neural tissues with silver salts. The Golgi staining method made individual neurons visible at random. Each neuron could be categorized and reconstructed using light microscopy. Collecting neuronal structures across samples, Cajal worked out the basic structure and connectivity of many neural circuits, revealing many principles of information processing that remain valid today.

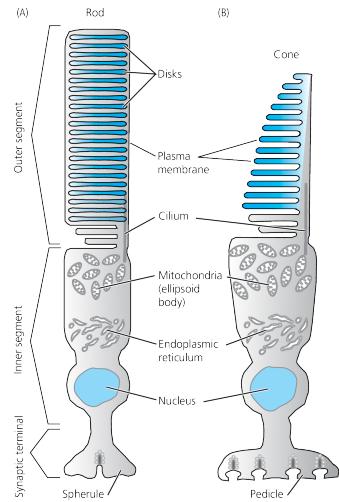


Figure 1: Vertebrate rods and cones. Principal structural features of vertebrate photoreceptors. (A) Rod. The outer segment is composed of disks detached from external plasma membrane. (B) Cone. The outer segment has membrane infoldings or lamellae instead of disks. The total area of sensory membrane is increased by these disk and lamellar structures, which are loaded with visual pigment. In a mammalian rod cell that is 0.025 cm long, roughly 2/3 of light is absorbed. (from Fain, Chapter 9).

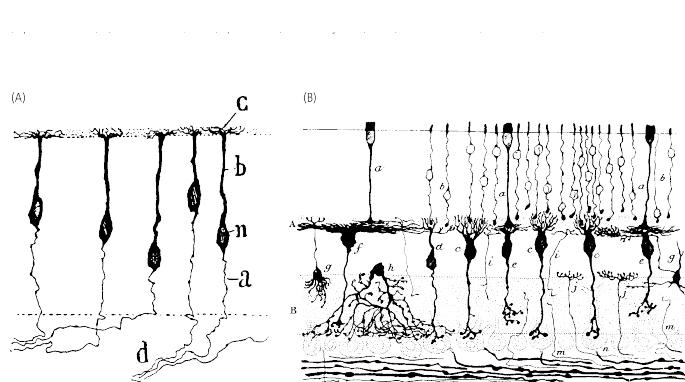


Figure 2: Santiago Ramón y Cajal has claim to be the father of modern neuroscience through his detailed anatomical studies of neural tissue. (A) Bipolar sensory neurons from mammalian olfactory mucosa. a, Axon; b, peripheral process; c, sensory dendrites; d, axon; n, nucleus. (B) Section of retina of an adult dog. A, Outer plexiform (synaptic) layer; B, inner plexiform (synaptic) layer; a, cone fiber; b, rod cell body and fiber; c, rod bipolar cell with vertical dendrites; d, cone bipolar cell with vertical dendrites; e, cone bipolar cell with flattened dendrites; f, giant bipolar cell with flattened dendrites; g, special cells stained very rarely (perhaps inter-plexiform cells); h, diffuse amacrine cell; i, ascendant nerve fibers (probably processes of cell not well stained); j, centrifugal fibers coming from central nervous system; m, nerve fiber (probably again of poorly stained cell); n, ganglion cell. (from Fain, Chapter 1)

THE FIVE SENSES recognized by Aristotle in his major treatise *de Anima* are hearing, sight, taste, smell, and touch (Fig. 3). These remain the five senses of conventional wisdom. But if the essence of sensory perception is gathering *information*, there are many more *bona fide* sensory systems throughout biology from microorganisms to man. Most organisms sense temperature, for example. Temperature sensing is particularly critical for small animals that do not regulate their own body temperatures – called poikilotherms or ectotherms or cold-blooded (if they have blood) (Fig. 4). Unlike endotherms, ectotherms use environmental temperatures to regulate their own body temperatures by moving from place to place. Pit vipers have evolved a way of using temperature sensing for infrared vision. Their pit organs are essentially *pinhole cameras* that form images of prey by sensing heat patterns caused by infrared radiation (Fig. 5). The molecular mechanisms of temperature sensing remain poorly understood – we know many of the molecules that sense temperature, but we don't know how they actually turn changes in temperature into changes in their own activity. The most sensitive thermoreceptors in biology can sense temperature changes as small as 0.001 °C/sec!

Most animals sense pain, called **nociception**. Larger animals sense the position and posture of their limbs and body parts to coordinate their own movements, called **proprioception**. Many animals sense the internal state of their own bodies, information that is sent to the brain both consciously or unconsciously, a process called **interoception**. Many animals have sensory modalities that we humans do not. Migratory birds are thought to sense earth's magnetic fields. Weakly electric fish sense the perturbations of self-generated electric fields as a sort of radar.

Some sensory modalities, like chemoreception, are so important and ancient that they evolved separately in many systems. For example, the receptors and signaling pathways that allow bacteria to 'smell' their environments evolved long before insects and other large animals evolved entirely different mechanisms for olfaction and gustation. Some sensory mechanisms are flexible such that they have diversified into fundamentally different functions (Fig. 6). Some G-protein coupled receptors (GPCRs) evolved into photoreceptors (rhodopsins). Other GPCRs evolved into chemoreceptors (olfactory and gustatory receptors).



Figure 3: Aristotle's Five Senses.

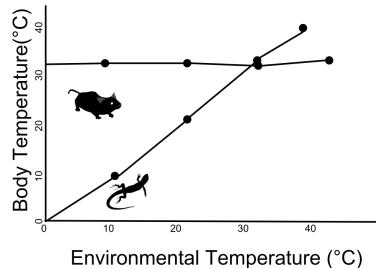


Figure 4: Ectotherms and endotherms. The mouse is endothermic and achieves **thermoregulation** through homeostasis. The lizard is ectothermic and its body temperature is dependent on the environment.



Figure 5: Thermal imaging cameras. The **pit organ** of vipers is an exquisitely sensitive thermal imaging camera, albeit with lower resolution than this one formed by a man-made camera.

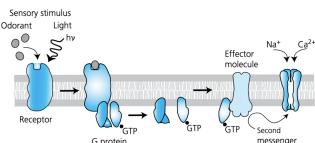


Figure 6: GPCRs. Both rhodopsin (light-sensing molecules) and olfactory receptors (chemical sensing molecules) evolved from G-protein coupled receptors that activate neurons through similar intracellular messaging systems.

STIMULUS ENERGY is the initiating event of all sensory systems. Photoreceptors transduce the energy of photons. Chemoreceptors transduce chemical binding energy. Mechanoreceptors transduce mechanical stretch or movement. Information is generated by some ‘energy absorbing’ primary event in sensory detection – a photoreceptor molecule absorbs a photon, a chemoreceptor binds a molecule, a mechanoreceptor is tugged open.

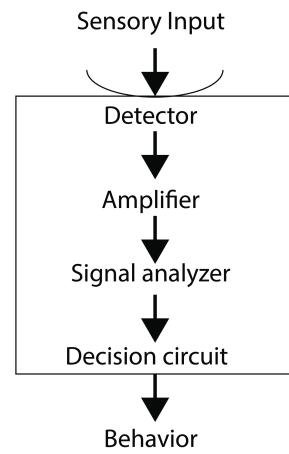
Sensory information eventually reaches mechanisms that shape organism behavior. For this to happen, the primary sensory signal – the energy from absorbing single photons or binding single molecules – must be filtered, amplified, and transmitted through intervening pathways. These different stages of sensory perception can be interpreted in engineering terms (Fig.). Every sensory system has some sort of ‘antenna’ that serve to detect primary signals. At the cellular level, the antenna for vision, for example, might be interpreted to be the photoreceptor cell. At the molecular level, the antenna might be molecular rhodopsin which directly absorbs each photon.

The primary sensory signal brings an amount of energy to the antenna that must be amplified. The energy of a single photon is calculated using Einstein’s relation $E = hc/\lambda$. The energy of one blue-green photon ($\lambda=500 \text{ nm}$) is $4 \times 10^{-19} \text{ J}$.

Primary signals are filtered or analyzed to create useful information to be sent to circuits that determine behavior. Filtering is needed to separate true signals from noise. Thermal energy constitutes a constant and unavoidable source of random noise that contaminates the sensory antenna. One of the most useful results The *Equipartition Theorem* from **Statistical Mechanics** states that all bodies in thermal equilibrium have $k_B T/2$ of average energy in every quadratic degree of freedom. For example, a gas molecule has an average kinetic energy of $k_B T/2$ along each axis of motion. Kinetic energy is quadratic in velocity components, three degrees of freedom in three-dimensional space:

$$\left\langle \frac{mv_x^2}{2} \right\rangle = \left\langle \frac{mv_y^2}{2} \right\rangle = \left\langle \frac{mv_z^2}{2} \right\rangle = \frac{k_B T}{2} \quad (1)$$

Thermal energy at room temperature ($\sim 25^\circ\text{C}$), roughly the temperature of most biology, is $4 \times 10^{-21} \text{ J}$. By comparison, the energy of one blue photon is $\sim 100 \times$ thermal energy. Although thermal energy fluctuates about $k_B T/2$, because a single photon is so much more energetic, its corresponding signal is much larger than this fluctuating noise. If the energy threshold for single photon detection is set to be much larger than $k_B T/2$ and below $\sim 100k_B T/2$, the energy of ‘true signals’ caused by single photons can be reliably filtered from ‘false signals’ caused by thermal energy fluctuations.



$$h = 6.6 \times 10^{-34} \text{ J s}$$

$$k_B = 1.4 \times 10^{-23} \text{ J K}^{-1}$$

THE BOLTZMANN DISTRIBUTION predicts that any detection threshold can be broached, albeit perhaps rarely, by thermal noise. For a body at thermal equilibrium, the probability of a thermal fluctuation with energy E_c is exponentially distributed:

$$P(E_c) \propto e^{-E_c/k_B T} \quad (2)$$

This means that the probability that a rhodopsin molecule reaches energies comparable to visible photons by thermal fluctuations, although vanishingly small, is also non-zero. Our eyes have hundreds of millions of photoreceptor cells (Fig. 1). Each photoreceptor cell is filled with billions of rhodopsin molecules. Although the likelihood that a single rhodopsin molecule reaches detection threshold because of temperature might be small, so many rhodopsin molecules mean that ‘dark noise’ events – the spontaneous activation of a rhodopsin molecule by temperature, not by photon absorption – can occur with significant frequency. The activation of rhodopsins by true photons have to be discriminated from the static noise of spontaneous rhodopsin activation by temperature. It is impossible to say whether a single given rhodopsin activation was triggered by a photon and not by thermal fluctuation. Multiple simultaneous rhodopsin activations are required to discriminate a true flash of light from a dark noise event.

WHAT IS THE SMALLEST NUMBER OF PHOTONS that a human observer can reliably detect? In the early history of the photon, Lorentz realized that a ‘just detectable’ flash of light delivered ~ 100 photons to the cornea. Most photons that reach the cornea do not reach the retina. The deeper question is: What is the threshold number of photons that is absorbed by photoreceptor cells to produce ‘seeing’? Hecht, Shlaer, and Pirenne (1942) did the classic experiment that established that single photons absorbed by single rod cells could be integrated into a perception of a flash of light. The threshold number of photons individually absorbed by a group of photoreceptor cells at the ‘threshold of seeing’ was 5-7 photons (Fig. 8).

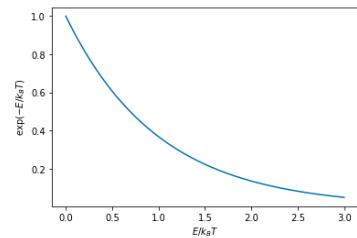


Figure 7: The Boltzmann Distribution

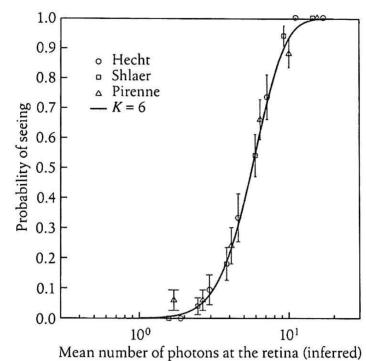


Figure 8: The threshold of seeing as the probability of seeing a flash plotted against the logarithm of the number of photons estimated to be absorbed by the retina at different flash strengths. Open symbols represent different measurements from three observers, the authors of the experiment.

SENSORY DISCRIMINATION is more complicated when different stimulus types must be separated and identified. Mammalian noses contain many different types of olfactory receptor neurons, each with its own molecular olfactory receptor (Fig. 9). When we whiff a scent, blends of different types and concentrations of molecules enter the nasal cavity. These molecules bind to different olfactory receptors with different chemical specificity. Odorant molecules can exhibit relatively small differences in their chemical structure and properties (Fig. 10). The *stimulus energies* that correspond to the binding of different molecules to a receptor might not be much different, perhaps varying from $\sim 2k_B T$ to $15 k_B T$. Different probabilities of odorant-receptor occupancy might be graded, not 0 or 1. Thus, one receptor can bind (and be activated) by many different kinds of molecule. One molecule can bind (and activate) many different kinds of receptors. Because typical smelly objects will emit many different odorant molecules, patterns of olfactory receptor activity can be expected to be complicated. A *combinatorial code* might be needed to discriminate odor molecules and identify smells (Fig. 11).

A revolution in structural biology, **cryo-electron microscopy**, has allowed visualization of the structural change caused by an odorant binding to an insect olfactory receptor. Unlike mammalian olfactory receptors which are GPCRs and *metabotropic* (meaning that they activate an intracellular signaling pathway that eventually changes the electrical activity of the cell) insect olfactory receptors are *ionotropic* (meaning that odorant binding directly opens an ion channel in the receptor itself).

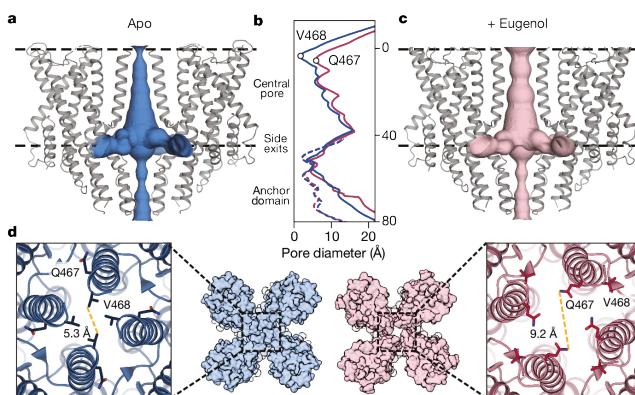


Figure 12: **Odorant-evoked opening of an ionotropic olfactory receptor.** a, c, The channel pores of unbound (a, blue) and eugenol-bound (c, pink). Black dashed lines, membrane boundaries. b, The diameter of the ion conduction pathway (solid lines) and along the anchor domain (dashed lines). d, Close-up view of the pore from the extracellular side. Marmol et al. (2021).

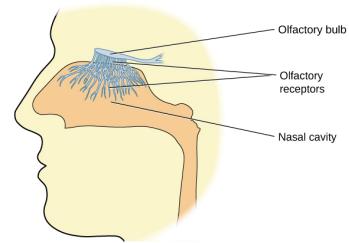


Figure 9: **Olfactory receptor neurons** innervate the olfactory epithelium in the human nasal cavity.

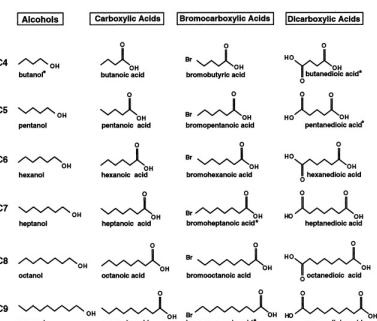


Figure 10: **Volatile odorant molecules** among the many thousands of chemicals that we can smell.

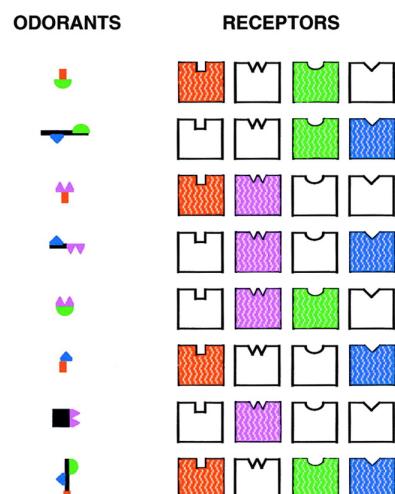


Figure 11: **Cartoon of the combinatorial code** where differently shaped molecules can bind to olfactory receptors with differently shaped binding pockets. Mammals can have 100s of different olfactory receptors that presumably detect many 1000s of different odorant molecules.

BACTERIAL CHEMOTAXIS is driven by smell. Bacteria are covered with chemoreceptors for molecules like amino acids that signify food. They use these chemoreceptors to assess surroundings and swim to favorable places. *E. coli* swims by rotating helical flagellar filaments attached to its $\sim 1 \mu\text{m}$ size body (Fig. 14). When all flagella rotate counterclockwise, as viewed from outside the cell, a bundle forms that pushes the cell forward at $\sim 25 \mu\text{m}/\text{s}$. These 'runs' typically last $\sim 1 \text{ s}$. Occasionally, one or more flagellar motors switch from CCW to CW rotation. This ends the run by disrupting the flagellar bundle. The cell stays in place and 'tumbles' until all flagella return to CCW rotation, and the cell starts a new run in a new direction.

Bacterial chemotaxis works by counting molecules. If the bacteria counts more attractant molecules over time during a run, it postpones the next tumble. Runs in favorable directions are thus longer than runs in unfavorable directions. Although each tumble randomly reorients each run, a *biased random walk* ensues that inexorably drives the bacteria where it wants to go.

Noise in counting molecules during chemotaxis has a thermal origin, but in another sense than thermal activation of unbound receptors. Consider a bacterium-sized volume, $L=1 \mu\text{m}$ on each side. If the cell instantaneously counted all molecules inside this volume, it would count $\sim 600,000$ molecules if the mean concentration was 1 mM , 600 molecules if the concentration was $1 \mu\text{M}$, and 60 molecules if the concentration was 10^{-7} M . But molecules constantly move in and out of the measurement volume, and so the number in an instantaneous count will fluctuate in a way governed by **Poisson statistics**. The standard deviation in the number of counted molecules will be the square root of the mean number of molecules. And so the relative error in estimating molecular concentration based on a single count within the $1 \mu\text{m}^3$ -sized measurement volume will be $600,000 \pm 800$ at 1 mM ($\sim 1.3\%$ error), 600 ± 24 at $1 \mu\text{M}$ ($\sim 4\%$ error), and 60 ± 8 molecules at 10^{-7} M ($\sim 13\%$ error). Berg and Purcell (1977) showed that bacteria do better than this by integrating measurements over time.

Understanding the biophysics of chemoreception is not idiosyncratic to bacterial chemotaxis. Virtually every process in intracellular and intercellular biological signaling involves the diffusive movement of molecules. Intracellular biochemical pathways involve the binding of ligands and enzymes (Fig. 15). Synaptic communication between nerve cells involves the diffusion of neurotransmitters from a presynaptic cell to the receptors on a postsynaptic cell. Understanding the biophysics of chemoreception is fundamental to understanding life.

Updated: September 12, 2023

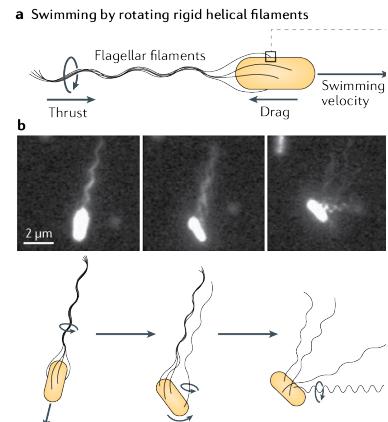


Figure 13: **Cartoon of the combinatorial code** where differently shaped molecules can bind to olfactory receptors with differently shaped binding pockets. Mammals can have 100s of different olfactory receptors that presumably detect many 1000s of different odorant molecules.

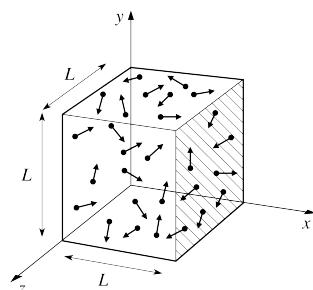


Figure 14: **Diffusion** is driven by the random and incessant motion of particles in and out of sampling volumes. At any point in time, the number of particles in a sampling volume fluctuates about a mean concentration owing to these 'Brownian movements'.

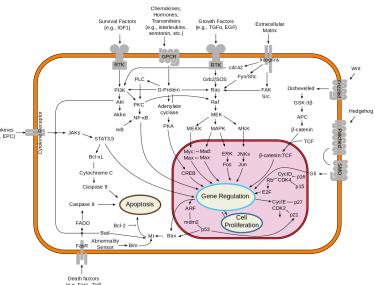


Figure 15: **Eukaryotic signal processing** is dominated by diffusion and the random binding and unbinding of molecules to receptors.

RECOMMENDED READING

- [Download paper](#)
- [Download paper](#)
- Chapter One. [Download paper](#)

ADDITIONAL READING

- [Download paper](#)
- [Download paper](#)

2. HOW BIOLOGISTS STUDY SENSORY SYSTEMS

2A. PSYCHOPHYSICS

THE QUANTITATIVE STUDY of sensory perception, *psychophysics*, began with 19th century experimental psychology. Conscious human responses to well-controlled stimuli began to be measured and analyzed. With a human subject, a scientist can directly obtain perceptual measurements by asking questions. Did the subject detect a stimulus – did she or didn't she see a dim flash of light? Can the subject tell the difference (discriminate) between two stimuli – which is heavier, block A or block B? For the answers to such questions to be meaningful, the scientist must adopt strict protocols, well-controlled stimuli and designing questions with clear answers. Even with rigorously designed experiments, interpreting the results of psychophysical experiments – often using *psychometric curves*, precise mathematical functions that interrelate sensory response and stimulus magnitude – is subtle (Fig. 16).

When studying sensory systems, it is useful to design experiments near *response thresholds*, using stimulus magnitudes where the sensory system switches between 'active' and 'inactive' states. Near response thresholds, sensory systems provide the most discriminatory information about stimuli. When a system is saturated (blinding lights) or far from threshold (stimuli too small to be detected), the sensory system reports little information. Near response thresholds, where a receptor might be 'on' or 'off' with equal probability, the amount of information that the receptor delivers to the organism is highest. This notion will be explored later with **Information Theory**.

Another reason to work near response thresholds is to study stimuli that are biologically relevant. All biological systems *evolved*. Natural selection drives sensory systems to provide maximal information about *natural* stimuli with minimal costs. Sensory receptors are tuned to detect and report stimuli that are useful in an organism's natural environment. In some cases, biological sensory systems approach the highest levels of performance allowed by physics. When this happens, biology can be viewed as 'optimized'. The smallest quantum of visual information is the photon, and our human rod photoreceptors are reliable detectors of single photons. We 'see' single photons because there must have been an evolutionary advantage to being able. The smallest 'ecologically relevant' visual stimulus is starlight, delivering photons at $10^{-2} \mu\text{m}^{-2} \text{ sec}^{-1}$. Seeing single photons matters in the night sky. Sensory systems are not always as good as they *can* be, but are only as good as they *need* to be for fitness and natural selection.

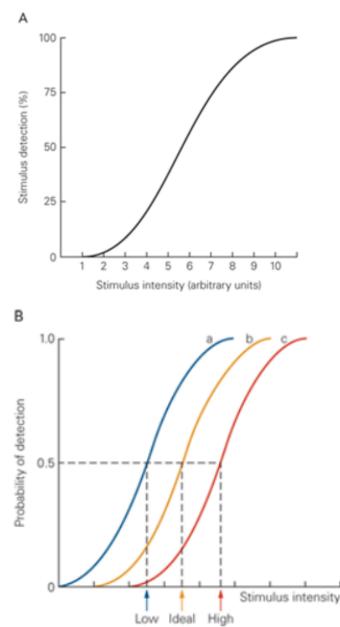


Figure 16: The psychometric curve. A. The psychometric function plots the percentage of stimuli detected by a human observer as a function of stimulus magnitude. Threshold is defined as the stimulus intensity detected on 50% of the trials. B. Detection and discrimination thresholds depend on the criteria used by individual subjects. Where an 'ideal' observer correctly detects the presence and absence of stimuli at the response threshold with equal probability (curve b), an observer who is told to respond to the slightest indication of a stimulus may report many false positives when no stimuli occur and has a lower response threshold (curve a). An observer who is told to respond only when very certain that a stimulus has occurred reports more hits than false positives and has a higher response thresholds (curve c). From *Principles of Neural Science*.

PROBABILITY AND STATISTICS is fundamental to understanding sensory responses near thresholds to account for the intrinsic variability in any experiment. **Stimuli** are intrinsically variable, weaker or stronger at random than a selected stimulus magnitude. For example, the quantum mechanical laws of physics does not allow an experimenter to make any device that reliably delivers single photons – one at a time in a stream of identical pellets – to the eye of a human observer. The best that can be done is to deliver packets of light at a controlled mean intensity, so that individual packets fluctuate at random about the mean. **Responses** can be variable. Sometimes a subject can think they detected a stimulus when none was delivered (*false positive*) or fail to see a stimulus that was delivered (*false negative*).

For the purpose of controlled, quantitative analysis of experimental results, studies are usually designed to measure *comparative judgments* of a stimulus property, such as stimulus magnitude or temporal frequency. To do this, most experimenters adopt *two-alternative forced-choice protocol* with two observation intervals and a pair of stimuli. Subjects are asked to report whether the second stimulus is stronger or weaker, higher or lower, larger or smaller, same or different than the first stimulus. In measurements of sensory thresholds the subject is asked whether the stimulus occurred in each interval. With this setup, there are only four types of responses – true positive, false positive, true negative, and false negative – which can be quantified and tabulated (Fig. 17).

THE VARIABILITY OF SENSATIONS evoked by a stimulus can be represented with a normal probability function with a mean (μ) and standard deviation (σ):

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$$

This allows using mathematics to predict the discriminability of pairs of normally distributed stimuli that differ along a physical dimension, such as the intensity of two flashes of light. The experimental subject adopts a *decision boundary* to classify stimuli. The frequencies of true and false positives (and true and false negatives) are demarcated by the probability functions corresponding to the stimuli and the decision boundary. Moving the decision boundary – by being a strict, lax, or an ideal observer who minimizes total error – will change these frequencies (Fig. 18).

		Response		Total stimuli
		Yes Red	No Blue	
Stimulus	Red	Hits (65)	Misses (35)	100
	Blue	False positives (20)	Correct rejections (80)	100
Total responses	85	115	200	

Figure 17: **Two-alternative forced choice tasks.** The stimulus-response matrix for a stimulus detection task (yes-no) or a categorical judgment task (red-blue). Although there are two possible stimuli and two possible responses, the data represent conditional probabilities in which the experimenter controls the stimuli and measures the subject's responses. The numbers provide examples of behavioral data obtained from a strict observer who responds "yes" less often than the actual frequency of occurrence of the stimulus.

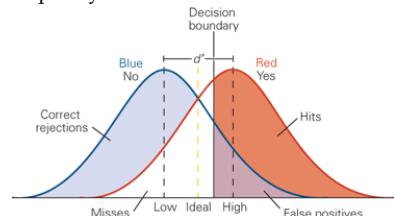


Figure 18: **Gaussian stimulus magnitudes.** Stimulus magnitudes can be represented by Gaussian curves with standard deviations that measure the fluctuation in sensations from trial to trial. The discriminability of a pair of stimuli is correlated with the distance between the two curves. When two stimuli are similar in magnitude, the two Gaussian curves overlap and no single criterion allows error-free responses. The frequency of true and false positives (and true and false negatives) is determined by the criteria used in the decision task. An ideal observer maximizes the number of correct responses and minimizes the total errors, setting the decision boundary at the intersection of the two curves. A strict observer minimizes the number of false positives but also reduces the total hits, setting the decision boundary to the right (solid line). A lax observer maximizes the number of hits but also increases the total false positives, setting the decision boundary to the left of the ideal subject.

BAYES' THEOREM is useful in interpreting psychophysical experiments in terms of *conditional probabilities*. Whether a stimulus is delivered in each trial has a probability that can be well-defined by experimental design. Whether a stimulus is detected has a probability that can be measured in the course of each experiment. Bayes' theorem is stated mathematically as the following:

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

where A and B are events and $P(B) \neq 0$.

- $P(A | B)$ is a *conditional probability*: the probability of event A occurring given that B is true.
 - $P(B | A)$ is also a conditional probability: the probability of event B occurring given that A is true.
 - $P(A)$ and $P(B)$ are the probabilities of observing A and B respectively without any given conditions; they are also known as the marginal probability or prior probability.
-

BAYES' THEOREM MAY BE DERIVED from the definition of conditional probability:

$$P(A | B) = \frac{P(A \cap B)}{P(B)}, \text{ if } P(B) \neq 0$$

where $P(A \cap B)$ is the probability of both A and B being true. Similarly,

$$P(B | A) = \frac{P(B \cap A)}{P(A)}, \text{ if } P(A) \neq 0$$

Solving for $P(A \cap B)$ and substituting into the above expression for $P(A | B)$ yields Bayes' theorem:

$$P(A | B) = \frac{P(B | A)P(A)}{P(B)} \text{ if } P(B) \neq 0$$

Surprising results can emanate from Bayes' theorem. A classic example is testing for disease in a population, a pertinent modern example. Suppose, a particular test for whether someone has COVID-19 is 90% sensitive, meaning that the test is correctly positive result for 90% of infected persons (and incorrectly negative for the other 10%). The test is also 80% specific, meaning that it is correctly negative for 80% of uninfected persons (and incorrectly positive for the other 20%). Assuming 5% of the population is infected, what is the probability that a random person who tests positive is really infected? This conditional probability is: $P(\text{Infected} | \text{Positive})$.

$$\begin{aligned} P(\text{Infected} | \text{Positive}) &= \frac{P(\text{Positive} | \text{Infected})P(\text{Infected})}{P(\text{Positive})} \\ &= \frac{P(\text{Positive} | \text{Infected})P(\text{Infected})}{P(\text{Positive} | \text{Infected})P(\text{Infected}) + P(\text{Positive} | \text{Uninfected})P(\text{Uninfected})} \\ &= \frac{0.90 \times 0.05}{0.90 \times 0.05 + 0.20 \times 0.95} = \frac{0.045}{0.045 + 0.19} \approx 19\% \end{aligned}$$

In other words, even if someone tests positive, the probability that they are infected is only 19% - this is because in this group, only 5% of people are infected, and most positives are false positives coming from the remaining 95%.

2B. ANATOMY

FORM AND FUNCTION are complementary in biology, just like any in area of man-made engineering – designing electrical circuits, machinery, automobiles, or buildings – where structure dictates performance. The difference is that ‘engineering’ in biology was done by evolution and natural selection. The study of ‘form’ in biology occurs at many levels from the anatomy of body parts to cells to molecules.

The functional study of sensory systems began with gross anatomy. Long before magnifying glasses and microscopes were invented, analysis was limited to manual dissection and unaided human observation. For example, Galen’s early studies of the eye marveled at its unique and specialized structures. But without a proper understanding of physics of optics, the ‘model’ that Galen built was inherently flawed. Fascinated by the lens, an object like none other in the animal body, Galen made it the central structure in the vision mechanism (Fig. 19). When physical optics was discovered in the Renaissance, the properties of refraction became understood. On the practical side, optics led to useful inventions like the magnifying glass, telescope, and microscope. On the conceptual side, physical optics also clarified the role of the lens in eye. The lens is only a device for focuses images onto the retina at the back of the eye. The retina, a thin sheet of brain tissue, is the central mechanism for detecting images and relaying information to the brain.

Galen’s model is a significant distortion of eye anatomy, and it is hard to imagine how it held sway for a millennium. Scientists needed to understand the physics before they could properly see the parts of the eye in their proper places. Seeing is believing. Believing is also seeing. Descartes’ model put the lens closer to its actual position and identified the muscles that change the shape of the lens during focal accommodation.

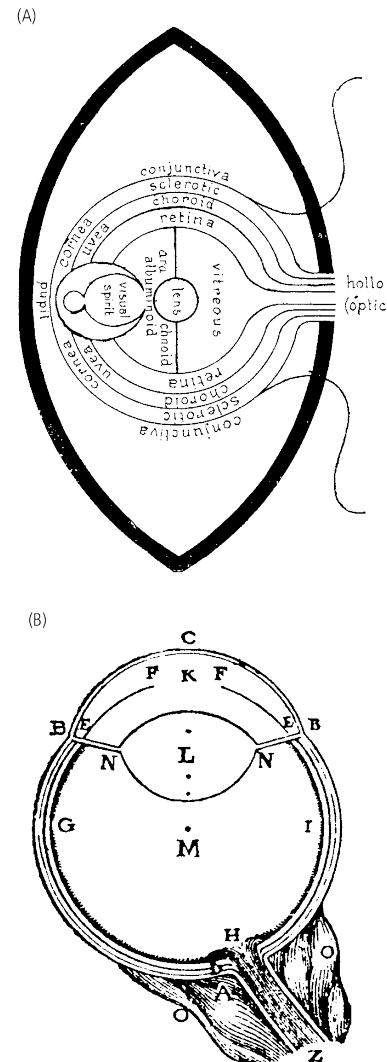


Figure 19: **Structure of the eye.** (A) Diagram of the eye from a ninth-century ad translation of Galen. (B) More anatomically correct diagram of cross-section of the eye made by Descartes. ABCB, Cornea and sclera; EF, iris (in actual fact closer to the lens than shown in Descartes’ diagram); K, aqueous humor; L, lens; EN, zonule fibers; M, vitreous humor; GHI, retina; H, optic nerve head; O, ocular muscles; and Z, optic nerve. From Fain, Chapter 1.

THE INVENTION OF THE OPTICAL MICROSCOPE led to anatomical investigations at the cellular level. Most of the sensory systems that we will study are in animals, vertebrate hearing, smell, and vision. The primary sensory receptors are specialized neurons for detecting sound, scent, and photons. The basic conceptual framework of neural circuit organization was largely established by Santiago Ramón y Cajal (1852-1934), the Spanish neuroscientist who was primarily a cellular-level anatomist (Fig. 20). Based on careful systematic analyses of sparsely labeled neurons in many brain tissues across animals and across developmental stages, Cajal identified basic principles about the organization of neural circuits in general and sensory circuits in particular. First, individual neurons interact with other neurons via contact or contiguity, synaptic contacts in today's language. Second, information flows with directionality through circuits, with dendrites and cell bodies on the input side and axons on the output side. These basic principles have been confirmed through anatomical investigations with higher spatial resolution using **electron microscopy**.

THE GOLGI STAINING METHOD was critical to Ramón y Cajal's success in resolving individual neurons. Neurons can be densely packed in brain tissue. Cell bodies are on the scale of micrometers. Synapses and nerve fiber thicknesses are on the scale of tens of nanometers. A theoretical limit to the resolution of any imaging system is the **Abbe diffraction limit**. Light with wavelength λ , traveling in a medium with refractive index n and converging to a spot with half-angle θ will have a minimum resolvable distance of

$$d = \frac{\lambda}{2n \sin \theta} = \frac{\lambda}{2NA}$$

The portion of the denominator $n \sin \theta$ is called the numerical aperture (NA), a property of the optical arrangement. The NA can reach about $1.4 - 1.6$ in the best optical microscopes, hence the Abbe limit is $d = \frac{\lambda}{2.8}$. Considering green light around 500 nm and an NA of 1 , the Abbe limit is roughly $d = 250$ nm. Cajal would not have been able to resolve labeled structures that were closer than 250 nm using his light microscope, but, by individually labeling neurons that were separated by much greater distances, he could perform his exquisite but laborious anatomical reconstructions, one cell at a time.

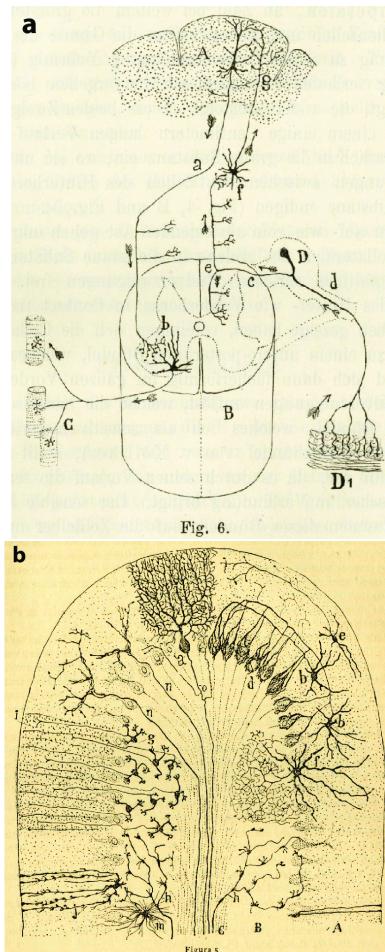


Figure 20: **Cellular level anatomical analyses.** (a) A nervous system-wide diagram of reflex and voluntary control of behavior by Cajal. Sensory information from the skin (D) is transmitted by dorsal root ganglion cells (d) to spinal cord (B) gray matter and to pyramidal neurons in the cerebral cortex (A), which in turn transmit impulses to motor neurons (b) in the spinal cord. For clarity, an interneuron between the spinal ending of (c) and an ipsilateral motoneuron are not shown. In this diagram, function (arrows) is predicted from structure. Neuron types in a gray matter region, the cerebellum. From Swanson and Lichtman (2016).

ELECTRON MICROSCOPES have greater spatial resolution because the electron wavelength, defined by quantum mechanics by the de Broglie equation:

$$\lambda = \frac{h}{p}$$

where h is Planck's constant and $p = mv$ is momentum, is so much smaller in scanning or transmission electron microscopes. In typical microscopes, electron velocities reach 20%-70% the speed of light. The electron wavelength reaches ~ 12 picometers in a 10 kV SEM and ~ 2 picometers in a 200 kV TEM.

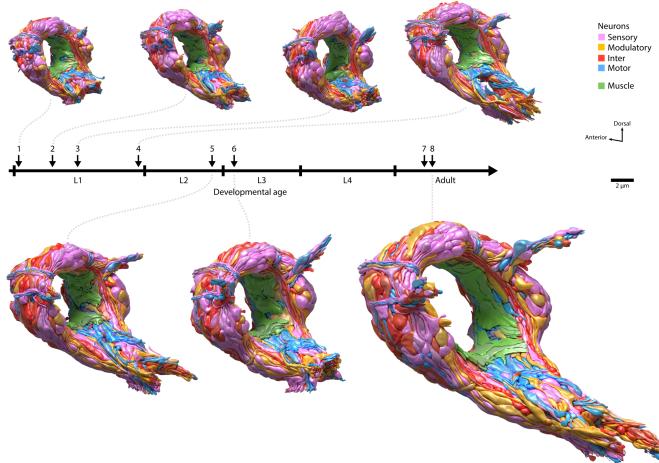


Figure 21: *C. elegans* *C. elegans* was the first animal to have its entire nervous system mapped by electron microscopy by John White and co-workers in the 1980s. More recently, eight *C. elegans* brains were reconstructed from birth to adulthood, mapping and comparing every chemical synapse and neuronal shape across whole-brain connectomes. From Witvliet et al. 2021

Electron microscopy imaging of brain sections has led to the characterization of synapses and neurons in diverse tissues. The main disadvantage of electron microscopy is that it requires ultrathin ($1\text{ }\mu\text{m}$) histological sections, which means that a single section is never adequate to generate full structural analysis of any cell. To overcome the problem of thin sections, scientists have used serial sectioning, in which each brain section is part of a sequence that transects a volume into hundreds or even tens of thousands of sections. Tracing objects from one section to the next reconstructs the geometry of the neurons and can also be used to identify the sites and cellular participants of each synapse. Connectomics has been used to reconstruct the entire nervous systems of small animals like nematodes and fruit flies (Fig. 21), and small parts of the nervous system of vertebrates (Fig. 22).

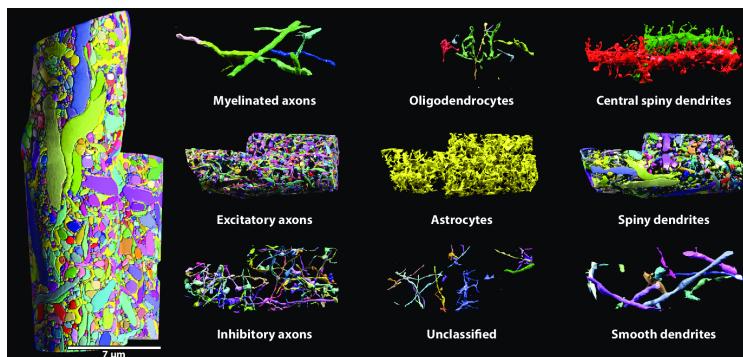


Figure 22: Reconstructions using modern, serial electron-microscopy connectomics approaches. On the far left is a fully reconstructed volume surrounding two apical dendrite segments of layer 5 pyramidal neurons in somatosensory cortex. Each colored object is a separate neuronal or glial cell process or extension. These processes are categorized in the images to the right. From Swanson and Lichtman, 2016.

ANATOMICAL ANALYSIS AT THE LEVEL OF INDIVIDUAL SENSORY RECEPTORS requires atomic resolution that greatly exceeds what is used for connectomics. X-ray crystallography has long been used to determine atomic-level structures of biomolecules. Crystallography only works for molecules that can be crystallized. Most sensory receptors are embedded in cell membranes, making them hard to crystallize in their native environments. Electron microscopy has atomic resolution, in principle, but sample degradation and poor contrast has long made it difficult to apply to molecules. Technical breakthroughs have led to cryo-electron microscopy (cryo-eM), which has allowed scientists to image many biomolecules for the first time, including membrane-bound, multimolecular sensory receptors. In cryo-eM, the biological sample is quickly submerged in liquid nitrogen. Rapid freezing forms vitreous ice, where water has an amorphous organization unlike crystalline ice, not causing problems associated with diffraction. Vitrification better preserves biological samples for the electron beam and the high vacuum of the electron microscope. Cryo-EM, combined with the computer-assisted approach of obtaining numerous images of the molecule of interest in different orientations, followed by algorithmic reconstruction, enables 3D structures of biomolecules at Angstrom-level resolution, such as that of Piezo, a major mechanosensory channel in animals (Fig. 23).

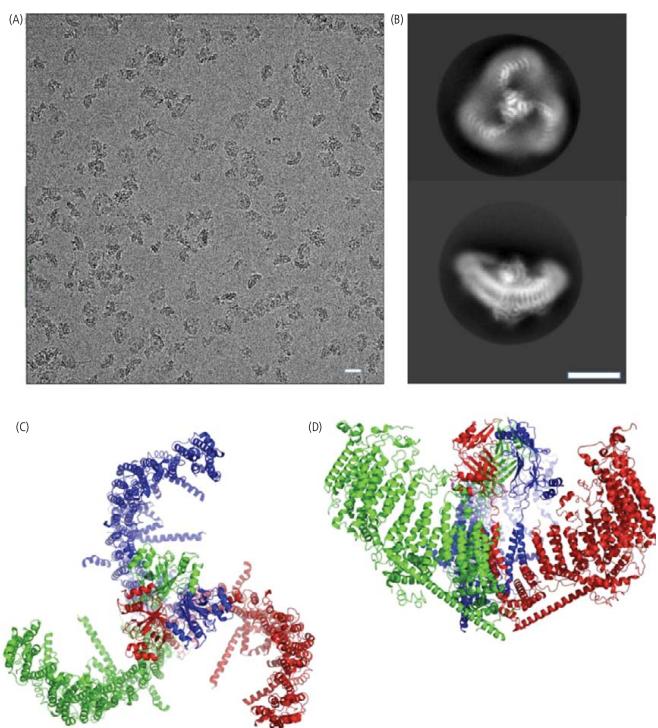


Figure 23: CryoEM of Piezo1 The gene for the Piezo1 protein was expressed in a cell line and purified. The protein was suspended on an electron-microscope grid and rapidly frozen by plunging the grid first into liquid ethane and then into liquid nitrogen. (A) Representative raw micrograph of the protein on the grid; scale bar is 20 nM. (B) Protein images like those in (A) were separated into groups according to their orientation, first manually to produce templates and then automatically under computer control. Images in each class were averaged, and representative averaged classes are shown viewed from the top (upper image) and side (lower image); scale bar is 100 nm. (C, D) Atomic model of the trimeric channel at an overall resolution of 3.7 Å shown as a ribbon diagram, viewed from the top (C) and side (D). The three subunits have been given different colors. From Fain, Chapter 1.

2C. PHYSIOLOGY

ANIMAL SUBJECTS can be asked whether they perceived a stimulus, whether directly (if a human) or by experimental design (quantitative analysis of behavioral responses in monkey, mouse, or smaller organisms). Knowing whether internal mechanisms – molecules, neurons, or brain circuits – perceived a stimulus requires *physiological analysis*. We will mostly discuss perception in animals, where sensory mechanisms are located in neurons. Neuronal activity can be measured in terms of electrical signals that travel along their ‘wires’ in a way that is mediated and amplified by ion channels in their cell membranes. Sensory cells have specialized ion channels that are directly (in the case of ionotropic receptors) or indirectly (in the case of metabotropic receptors) activated by the environmental stimulus itself.

E. D. Adrian recorded some of the first *extracellular recordings* from sensory neurons by placing the axons of touch receptor cells near wire electrodes. Skin pressure changed the frequency of *action potentials* in different ways depending on the stimulus. Thus, action potential firing patterns are a mechanism for encoding and communicating sensory information to the brain. The first photosensory responses were made by Hartline in the horseshoe crab, where action potential patterns also depended on the intensity and duration of light stimulation. Patterns of neuronal activity in sensory neurons encode the incoming stimulus.

THE BRAIN MUST SOLVE AN INVERSE PROBLEM. It must reconstruct a past incoming stimulus based on the present pattern of neuronal activity. Because neuron activity patterns can be intrinsically randomness, a physiologist must measure the *probability* of a specific neuronal response that might be conditioned on a specific stimulus: $P(\text{response} \mid \text{stimulus})$. In a sense, the animal must perform a Bayesian inference, estimating the probability that a specific stimulus occurred conditioned on a specific neuronal response: $P(\text{stimulus} \mid \text{response})$.

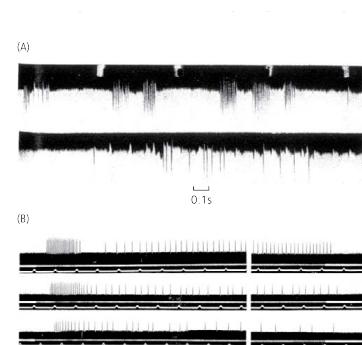
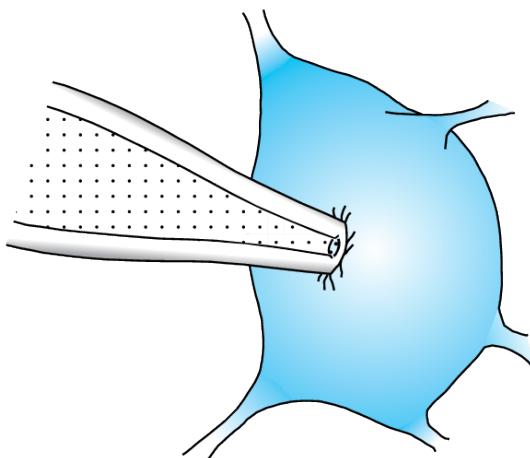


Figure 24: **Early electrical recordings of sensory responses.** (A) Action potentials recorded from single axons dissected from the cutaneous nerve of a frog. (B) Action potentials from the lateral eye of the horseshoe crab *Limulus*. Each trace gives the response to a different light intensity, which was systematically increased by an additional factor of ten from dimmest (bottom) to brightest (top). From Fain, Chapter One.

INTRACELLULAR RECORDINGS made it possible to record the detailed electrical responses of single neurons (Fig. 25). Electrodes are inserted into glass tubes that are melted and pulled to a fine point and filled with salt solution. Sharp electrode recordings can be made by penetrating individual cells. *Patch electrode recordings* can be made by polishing the tip of the glass electrode to be very smooth, such that when it is pressed against a cell membrane and a slight suction is applied, a very tight seal is formed, sometimes called a gigaseal with typical electrical resistances of $10\text{-}100\text{ G}\Omega$. High seal resistance ensures that most of the electrical current in each recording travels through the cell membrane that is being tested, not leaking through the seal.

(A)



(B)

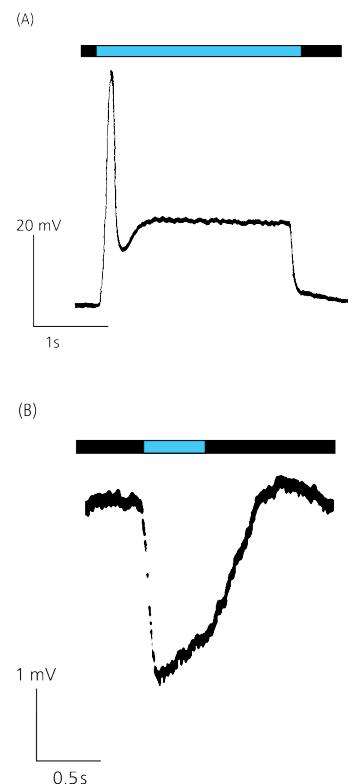
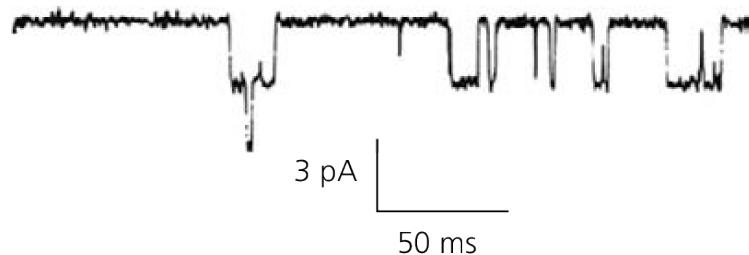


Figure 25: Intracellular recordings from sensory receptors. Bars above recordings show timing and duration of light flashes. (A) Depolarizing voltage response from photoreceptor of *Limulus* ventral eye. (B) Hyperpolarizing voltage response from photoreceptor (cone) of a fish. This is the first published recording of the response of a vertebrate photoreceptor. From Fain, Chapter One.

Figure 26: Patch-clamp recording from single channels. (A) The tip of a patch pipette is pushed against the cell body of a cell and slight suction is applied to form a seal. (B) Single-channel currents recorded from muscle acetylcholine receptors. The pipette contained $0.3\text{ }\mu\text{M}$ acetylcholine. Downward deflections indicate channel opening. At least two channels were present in this membrane patch. From Fain, Chapter One.

2D. MOLECULAR BIOLOGY

THE REVOLUTION IN MOLECULAR BIOLOGY deepened the analysis of sensory mechanisms (Fig. 27). Most sensory receptors are integral membrane proteins. Most known receptors were identified either by protein purification and sequencing, or by genetic analysis (finding a mutant that lacks a sensory modality, and working out the missing gene that was responsible for phenotype). Modern techniques to find sensory receptors include single-cell transcriptional and whole-genome sequencing and analysis. The genetic sequence of a putative receptor has characteristics that can betray its identity. Integral membrane proteins must have extensive sequences within the hydrophobic interior of the lipid bilayer. From amino acid sequences, which amino acids lie within the membrane and which face the cytoplasmic or extracellular solution can be inferred. Some amino acids (such as valine and isoleucine) are hydrophobic. Some amino acids (such as aspartate and lysine) are hydrophilic. Hydropathy analysis can be used to estimate the rough structure of putative receptors.

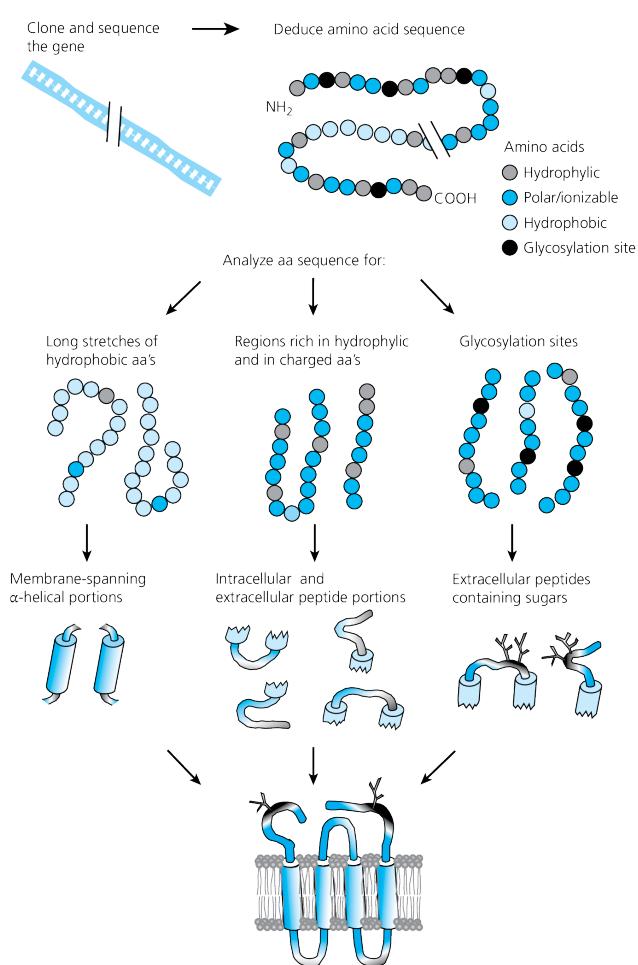


Figure 27: Analysis of hydropathy and the folding of membrane proteins. The amino acid sequence of a membrane protein can be used to make inferences about protein structure. From Fain Chapter One.

RECOMMENDED READING

- [Download](#)
- Chapter One. [Download paper](#)
- [Download](#)

ADDITIONAL READING

- This paper describes the modern use of high-throughput connectomics to measure the developmental dynamics of an entire animal brain from birth to adulthood.
 - [Download](#)
- The discovery of the Piezo mechanosensory channels involved a remarkable integration of modern techniques in molecular biology and electrophysiology to yield one of the most elusive sets of proteins in sensory perception.
 - [Download](#)

3. THE BOLTZMANN DISTRIBUTION

3A. RHODOPSIN

RHODOPSIN, also called visual purple, is the light-sensitive receptor protein in the retina (Fig. 29). Rhodopsin is also a membrane-bound photoswitchable G-protein-coupled receptor (GPCR) with both a protein component and covalently-bound cofactor, a photoreactive chromophore called *retinal*. Isomerization of 11-cis-retinal into all-trans-retinal by light triggers a conformational change that causes rhodopsin to activate another G protein called *transducin*. Transducin activation begins a signal transduction cascade that eventually modulates cyclic guanosine monophosphate (cGMP) levels. Changing cGMP levels change the electrical activity of the rod cell through cGMP-gated ion channels in the cell membrane of the photoreceptor cell. Changes in the electrical activity of the photoreceptor cell membrane lead to changes in their chemical synaptic outputs that are communicated to the rest of the retina as visual events.

THE ENERGY of one blue-green photon ($\lambda=500 \text{ nm}$) is $4 \times 10^{-19} \text{ J}$. Because rhodopsin is in thermal equilibrium with its environment ($\sim 25^\circ\text{C}$), the chromophore will have, on average, an amount of thermal energy around $4 \times 10^{-21} \text{ J}$. The energy of a blue photon is much higher than thermal energy. On average, the thermal energy fluctuations of the chromophore can be expected to be much lower than a threshold needed for its activation. But how frequently will the chromophore exceed the threshold of activation by thermal fluctuations alone? Answering this question requires knowing about probability distribution of energy levels of an object in thermal equilibrium, not just its mean thermal energy.

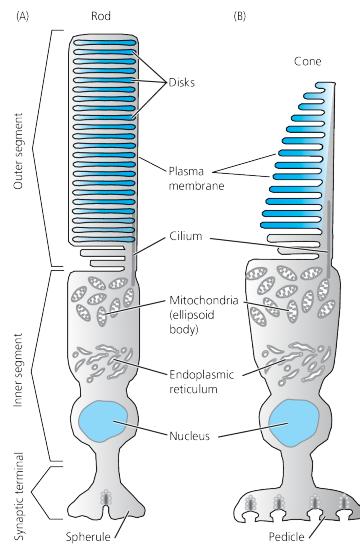


Figure 29: Rods and Cones.

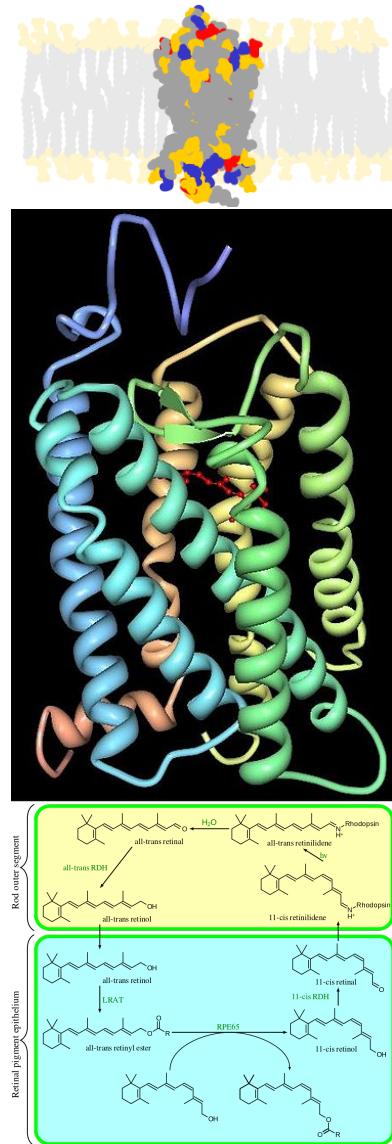


Figure 28: Rhodopsin. Three dimensional structure of bovine rhodopsin, a membrane bound protein. The chromophore, retinal, is embedded within the protein. Rhodopsin is embedded in detached discs in the outer segment of the rod photoreceptor cell and embedded in infolded lamellae in cones. Absorption of light energy, $h\nu$, causes a conformational change – 11-cis-retinal becomes all-trans-retinal – which thereby alters protein structure to activate G-protein-coupled signal transduction. Biochemistry in the rod cell restores the cis configuration.

3A. THE BOLTZMANN DISTRIBUTION

IN STATISTICAL MECHANICS, a Boltzmann distribution is a probability distribution that gives the probability that a system will be in a certain state as a function of state energy and temperature:

$$p_i \propto e^{-\varepsilon_i/(kT)}$$

where p_i is the probability of the system being in state i , ε_i is the energy of that state, and a constant kT of the Boltzmann constant and temperature T . The symbol \propto denotes proportionality.

System can have broad meaning; an macroscopic ensemble of components (a molecule) or a single atom. What matters is that the system has a *degree of freedom* that allows it to enter different measurable states, and that energy is freely exchanged into and out of the system. An exponential distribution means that states with lower energy have a higher probability of being occupied (Fig. 30).

THE RATIO OF PROBABILITIES of two states is the Boltzmann factor and depends on their energy difference:

$$\frac{p_i}{p_j} = e^{(\varepsilon_j - \varepsilon_i)/(kT)}$$

If a system has M possible states, and the sum of the probabilities of being in each state is normalized, $\sum_{j=1}^M p_i = 1$, one can compute the probability of being in each state:

$$p_i = \frac{1}{Q} e^{-\varepsilon_i/(kT)} = \frac{e^{-\varepsilon_i/(kT)}}{\sum_{j=1}^M e^{-\varepsilon_j/(kT)}}$$

THE EXPONENTIAL ATMOSPHERE is a classic example of a continuous Boltzmann distribution (Fig. 31). The altitude of gas molecules in Earth's atmosphere is one spatial degree of freedom. The altitude of each gas molecule is associated with a gravitational potential energy: $E = mgh$. Assuming a uniform atmospheric temperature T , the probability that a molecule is at height h is:

$$p(h) \propto e^{-mgh/kT}$$

For this probability density to be properly normalized:

$$p(h) = \frac{e^{-mgh/kT}}{\int_{h=0}^{\infty} e^{-mgh/kT} dh}$$

The expectation value for the altitude of a gas molecule depends on its mass: $\langle h \rangle = \frac{kT}{mg}$.

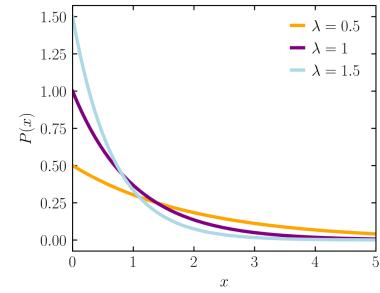


Figure 30: **Exponential Distributions.** Plot of the probability density function of the exponential distribution ($P(x) = \frac{e^{-\lambda x}}{\lambda}$) for rates $\lambda = 0.5, 1$ or 1.5 .

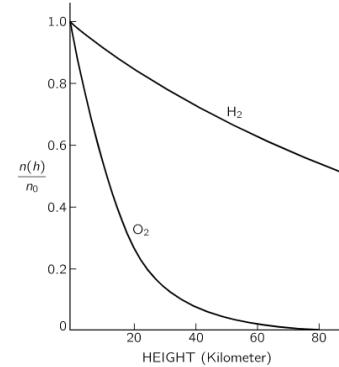


Figure 31: **Exponential atmosphere.** From Feynman's *Lectures in Physics*. The mass of one molecule of O_2 is 5×10^{-23} g. The mass of one molecule of H_2 is 3×10^{-24} g. Does Feynman's drawing make sense?

THE BOLTZMANN DISTRIBUTION governs the likelihood that a sensory receptor like rhodopsin has different states with different thermal energies. We describe the different conformational states of a molecule using a “reaction coordinate”, and assign an energy to each point along the reaction coordinate (Fig. 32).

Using Boltzmann factors, we can infer the relative likelihood of being in different states. Going from one stable state A to another state B might require transit through an activation state. The energy of this activation state creates a barrier that slows the reaction. At equilibrium, the relative likelihood of any two states is:

$$\frac{p_A}{p_B} = e^{-\frac{\Delta E}{kT}}$$

To estimate the reaction rate from A to B , we need the relative likelihood of being at the top of the activation barrier compared to state A . This estimates the fraction of particles in state A that are able to move over the barrier thanks to fluctuations in thermal energy:

$$k_{A \rightarrow B} \propto e^{-\frac{E_{act}}{k_B T}}$$

The Boltzmann distribution gives us insight into the fluctuating energies of sensory receptors before the arrival of stimulus energy, as well as the speed of the reactions that characterize sensory transduction.

3B. DERIVING THE BOLTZMANN DISTRIBUTION BY COUNTING

Say that you have N particles that are given a total amount of energy E . These particles are able to freely exchange energy among them, but are constrained to quantal energy levels. How many particles can you expect to find at each energy level? In other words, what is the probability of a particle having a certain amount of energy? To find the distribution of the particles over their possible energy states, we enumerate the states ($s = 1, 2, \dots$), associate each state with a discrete energy ($\varepsilon_1, \varepsilon_2, \dots$), and assign a number of particles to each state (n_1, n_2, \dots). The Boltzmann distribution should tell us how many particles, n_s , of the N total particles that we can expect to find in the s state with energy ε_s .

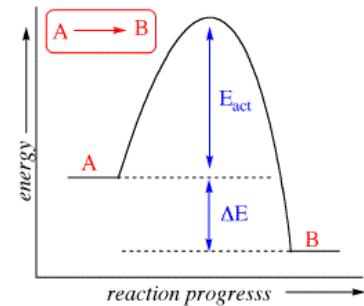


Figure 32: Reaction coordinate. From <http://butane.chem.uiuc.edu/pshapley/genchem2>

State	Energy	Number
1	ε_1	n_1
2	ε_2	n_2
3	ε_3	n_3
.	.	.
.	.	.
s	ε_s	n_s
.	.	.
.	.	.

Mass Conservation:
 $\sum_s n_s = N$.

Energy Conservation:
 $\sum_s n_s \varepsilon_s = E$.

The probability of any particle being in a given state s is n_s/N . Thus, the average energy of each particle is:

$$\langle \epsilon \rangle = \frac{\sum_s n_s \epsilon_s}{\sum_s n_s}$$

What is the probability of observing a specific distribution with say n_1 particles in state 1, n_2 in state 2, and so on? Our central assumption is that every possible distinct arrangement that satisfies conservation laws of mass and energy is equally possible. The probability of observing a particular distribution of n_s is thus proportional to the number of distinct arrangements that can be achieved with the N particles. The number of distinct particle arrangements that corresponds to the same distribution of particles among energy levels is a measure of the probability of that distribution. The number of ways a given distribution can be formed is a combinatorial problem:

$$W = \frac{N!}{n_1! n_2! n_3! \dots}$$

To find the distribution where W is largest, we need to maximize W with respect to n_s and with respect to the conservation laws. To do this, we need a convenient analytical expression for the factorial problem and we need to use Lagrange Multipliers.

We choose to maximize $\log W$ subject to the constraints of conservation of energy and mass. Thus,

$$d \left(\log W - \alpha \sum_s n_s - \beta \sum_s \epsilon_s n_s \right) = 0$$

The α and β are the Lagrange multipliers. Varying with respect to n_s and incorporating Stirling's Formula gives:

$$-\sum_s d n_s (\log n_s + \alpha + \beta \epsilon_s) = 0$$

Because this must hold true for every δn_s , every term in the sum must vanish. The values of n_s which do this are

$$\log n_s + \alpha + \beta \epsilon_s = 0$$

We can thus conclude that the occupancy of state s depends exponentially on its energy:

$$n_s \propto e^{-\beta \epsilon_s}$$

Actually, we have shown that the exponential dependence of state occupancy on energy is true only at the most probable W . We have not shown that every other arrangement can be ignored. The probability

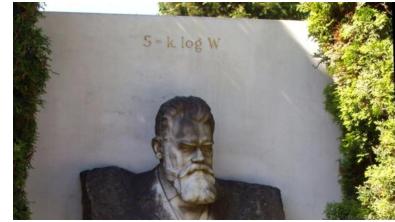
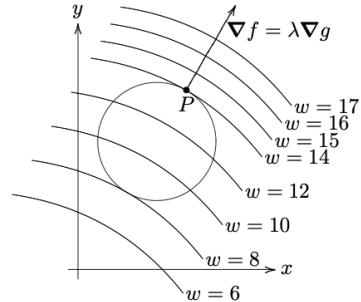


Figure 33: **Boltzmann's Tomb.** When you take Statistical Mechanics you will learn that $\log W = S/k_B$, the equation for entropy etched on his tomb.

Lagrange multipliers

Here, we sketch a geometric proof that builds intuition without worrying about rigor. For the function $w = f(x, y, z)$ constrained by $g(x, y, z) = c$, the maxima and minima are those points where ∇f is parallel to ∇g :

$$\nabla f - \lambda \nabla g = 0$$



For concreteness, we've drawn the constraint curve, $g(x, y) = c$, as a circle and some level curves for $w = f(x, y) = c$ with explicit (made up) values. Geometrically, we are looking for the point on the circle where w takes its maximum or minimum values.

Start at the level curve with $w = 17$, which has no points on the circle. Clearly, the maximum value of w on the constraint circle is less than 17. Move down the level curves until they first touch the circle when $w = 14$. Call the point where they first touch P. It is clear that P gives a local maximum for w on $g = c$, because if you move away from P in either direction on the circle you'll be on a level curve with a smaller value.

Since the circle is a level curve for g , we know ∇g is perpendicular to it. We also know ∇f is perpendicular to the level curve $w = 14$, since the curves themselves are tangent, these two gradients must be parallel. Q.E.D.

of alternative arrangements becomes negligible when N is very large, but this is hard to show and we will skip it.

We also have not shown that β , introduced here as a Lagrange multiplier, is related to temperature. For our purposes, we define $\beta = 1/k_B T$. When you take thermodynamics, you will learn that β behaves like $1/k_B T$ for various thermodynamic relationships and can have no other meaning. We are going to declare victory with the result that energy levels are exponentially distributed at thermal equilibrium.

3C. DERIVING THE BOLTZMANN DISTRIBUTION USING INFORMATION THEORY

CLAUDE SHANNON invented *information theory* starting with his own notion of the *entropy* of a probability distribution. Let X be a discrete random variable that can have different possible values, x . The probability density function, $p(x)$, is the likelihood of having different values.

SHANNON DEFINED HIS ENTROPY $H(x)$ of a discrete random variable X as:

$$H(X) = - \sum_x p(x) \log_2 p(x)$$

Say the random variable is the outcome of tossing a fair coin. X can either be heads ($p = 0.5$) or tails ($p = 0.5$). In this case, $H = 1$. Shannon's entropy is the amount of information in *bits* that you need to characterize the outcome of the toss.

What if the coin was biased and always came up heads? In this case, $H = 0$. You don't need any information to characterize the outcome of the coin toss (i.e., 0 bits), because the outcome is guaranteed. What if we vary p ? The entropy is minimum at $p = 0$ or $p = 1$ and maximum at $p = 0.5$ (Figure 34).

Roll a fair die. In this case, every outcome has $p = 1/6$ and $H = \log_2 6 \approx 2.5$. In Dungeons and Dragons, we have 8-sided die, and the number of bits needed to characterize one roll is 3.

If you know nothing about a probability distribution except the range of outcomes, the most unbiased *a priori* description of the probability distribution is the distribution that *maximizes* entropy. This does not mean that the maximum entropy distribution is the right one. But if it isn't, you lack salient facts about the distribution.

Reconsider the problem of N particles exchanging a total amount of energy E among them. Each particle is constrained to discrete energy levels $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_s$

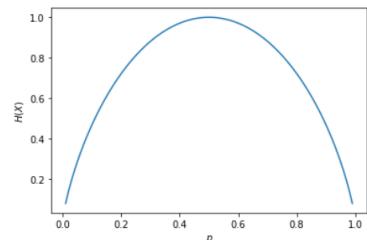


Figure 34: Entropy of a coin toss

What is the probability distribution that governs the energies of each particle? This problem is fully described. We are not missing any salient facts needed to calculate a distribution. The correct probability distribution must be the one that maximizes entropy. We have phrased another Lagrange multiplier problem:

$$\delta \left(H - \alpha \sum_s p_s - \beta \sum_s \epsilon_s p_s \right) = 0$$

We conclude that $p_s \propto e^{-\epsilon_s \beta}$. If the distribution were anything else, we would have required additional salient constraints on the physical problem. But because we set up the problem without any other constraints, only the exponential distribution of energies, the Boltzmann distribution, is possible. *Q.E.D.*

REFERENCES

- This is Shannon's original paper where he invented Information Theory.
 - [Download](#)
-

4. PHOTORECEPTOR CELLS

VERTEBRATES HAVE TWO TYPES OF PHOTORECEPTORS, rods and cones. Our human eyes contain ~ 130 million rod cells for scotopic (low light) vision and ~ 7 million cone cells for color vision. All rod cells have one peak wavelength sensitivity, whereas the three cone sub-types are tuned to long, medium, and short wavelengths (Fig. 36). Thus, rod vision is monochromatic whereas cone vision is trichromatic. Trichromacy means that any color that we can see can be built from three different wavelengths of visible light such as red/green/blue or cyan/magenta/yellow. The stimulus space of color vision has three dimensions.

The anatomy of rod cells is specialized for low-light detection. Each rod cell is ~ 2 microns in diameter, but ~ 100 microns long. Incoming photons travel along the long axis, the better the chance of photon capture by visual pigment absorption. Some animals have a reflecting layer of tissue behind the retina that acts as a retroreflector (Fig. 37). This *tapetum lucidum* reflects photons not captured by rods in the first pass through the retina, increasing the likelihood of absorption with a second pass. Thus, the likelihood of photon absorption increases by effectively doubling rod length without doubling the amount of visual pigment.

Starlight creates photon fluxes of $<10^{-2}$ photons $\mu\text{m}^{-2} \text{ sec}^{-1}$. Bright sunlight creates photon fluxes of $>10^8$ photons $\mu\text{m}^{-2} \text{ sec}^{-1}$. Cones mediate vision over the upper 7-8 log units of this range. Rods mediate vision at rates of photons per second. Rods are specifically engineered to detect and count single photons. Not only must rods be sensitive enough to absorb single photons, they must also reliably amplify each single photon signal to distinguish one photon from zero or two.

In photoreceptor cells, the visual pigments and signal transduction molecules that transduce photon detection into electrical activity are in the outer segments. In rods, rhodopsin molecules that contain the visual pigment is loaded into the membranes of tightly-packed disks (Fig. 38). Disk numbers vary, but are ~ 1000 in human rods. Each mammalian rod has about 10^8 rhodopsin molecules in its disks.

When a rod cell absorbs a photon, it creates and communicates a signal by changing synaptic output to downstream horizontal cells and bipolar cells in deeper retinal layers. Rods and cones lack the voltage-gated Na^+ needed to create fast all-or-none action potentials. Instead, photoreceptor cells have relatively slow, graded changes in membrane potential that modulate the rate of synaptic release. A biochemical signal transduction cascade connects changes in photon

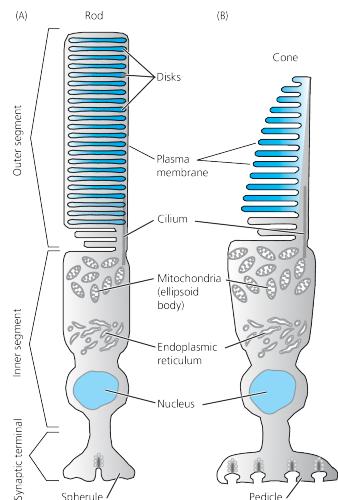


Figure 35: Rods and Cones.

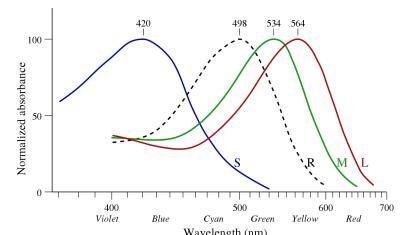


Figure 36: Color vision. Normalized human photoreceptor absorbances for different wavelengths of light.



Figure 37: Tapetum Lucidum. Choroid dissected from a calf's eye appearing iridescent blue.

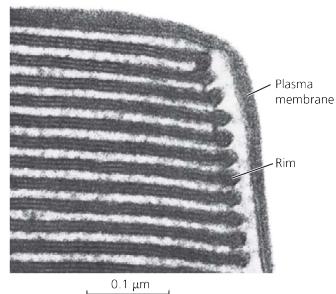


Figure 38: Rod disks. Low power electron micrograph of rod outer segment from Fain, Chapter 9.

absorption to changes in membrane potential.

ROD PHOTORECEPTOR CELLS are anatomically designed to efficiently capture photons, but not to capture *every* photon. The longer the rod, the more likely that an entering photon (at length $x = 0$) will be absorbed before exiting (at length $x = L$). What is the probability, P , that a rod captures each entering photon as a function of rod length? It does not matter where the photon is absorbed, as long as it is absorbed. Our answer will involve an integral over all distinct probabilities that the photon is absorbed in each interval, between x and $x + dx$, along rod length.

Consider a thin slice of rod with thickness Δx . A photon that enters this slice has a small finite probability of being absorbed. This probability is proportional to both rhodopsin concentration, C , and an absorption cross-section, σ :

$$p = \sigma C \Delta x$$

For a photon to be absorbed in the slice between x and $x + \Delta x$, it must have *not* been absorbed in any prior slice between 0 and x . What is the probability that the photon is not absorbed in any prior slice? Divide the distance from 0 to x into N slices. The probability that a photon entering any prior slice is absorbed by that slice is $\sigma C x / N$. So the probability that a photon is not absorbed by each prior slice is $q = 1 - \frac{\sigma C x}{N}$. These probabilities must be multiplied to achieve the total probability of not being absorbed by all N prior slices before reaching x :

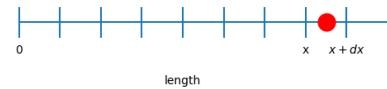
$$\left(1 - \frac{\sigma C x}{N}\right)^N$$

Hence, the probability that a photon that enters the rod at $x = 0$ is *first* absorbed between x and $x + dx$ is:

$$dP = \left(1 - \frac{\sigma C x}{N}\right)^N \sigma C dx$$

In the limit of large N , this differential probability is $dP = e^{-\sigma C x} \sigma C dx$. The total probability that an entering photon is absorbed by a rod within its length L is the sum of the probability that the photon is first absorbed somewhere between 0 and L :

$$P = \int_0^L dP = 1 - e^{-\sigma C L}$$



A useful definition of the exponential function in terms of a limit:

$$\lim_{N \rightarrow \infty} \left(1 - \frac{x}{N}\right)^N = e^{-x}$$

AS A ROD CELL INCREASES IN LENGTH, as L increases, the probability that each entering photon is absorbed and counted will exponentially converge to 1. Longer rod cells mean more rhodopsin molecules. Each rhodopsin molecule has a finite probability of spontaneous isomerization through thermal energy. These spontaneous isomerizations create “dark noise” events. There is a fundamental trade-off between the amount of signal (true absorbed photons) and the amount of noise (spontaneous isomerizations) as a function of rod length.

If the rate of spontaneous isomerization is r_{dark} , the number of dark-noise events in a given interval of time, τ , will be proportional to the total number of rhodopsin molecules in the cell, N_{rh} : $\langle n_{dark} \rangle = r_{dark}\tau N_{rh}$. If the number of spontaneous isomerizations were always exactly the average number of isomerizations, the rod cell could subtract this number from the total number of isomerized molecules to estimate the fraction corresponding to real photons. But spontaneous isomerizations occur randomly. A true photon can be seen only if its corresponding signal is larger than the fluctuating signals caused by spontaneous isomerization.

4B. BINOMIAL STATISTICS

Say that each rhodopsin molecule has a probability, p , of spontaneous isomerization in each trial. Say that we have N rhodopsin molecules. The mean number of ‘successful’ isomerizations in each trial is pN . What are the fluctuations around this mean? The answer comes from binomial statistics, although in the limit of small p and very large N . Start with binomial statistics. Instead of flipping rhodopsin molecules, consider flipping coins.

Let us start with a biased coin, which gives heads with probability p and tails with probability q . What is the probability that I will get k heads out of N flips? The probability that a single flip will give me one head is p . Since each flip is independent, I can multiply the probability of the desired results from each flip together: p^k for the heads, and q^{n-k} for the tails. I then have to account for all the possible permutations in order—HTHT is equally acceptable as a desired outcome as HHTT. Putting these three elements together gives me the desired probability:

$$P(k; N, p) = \binom{N}{k} p^k q^{N-k}, \quad (3)$$

$$\binom{N}{k} = \frac{N!}{k!(N-k)!}. \quad (4)$$

This is the **binomial distribution**. Can we prove that it is normalized? Remember that in algebra, we use the binomial distribution to take $(a + b)^N$. So:

$$\sum_{k=0}^N P(k; N, p) = \sum_{k=0}^N \binom{N}{k} p^k q^{N-k} \quad (5)$$

$$= (p + q)^N = 1^N = 1. \quad (6)$$

Let us derive the mean of the binomial distribution. This can equally be thought of as the expected value of k , the number of successes.

$$\langle k \rangle = \sum_{k=0}^N k P(k; N, p) \quad (7)$$

$$= \sum_{k=0}^N k \frac{N!}{k!(N-k)!} p^k q^{N-k} \quad (8)$$

$$= \sum_{k=1}^N \frac{N!}{(k-1)!(N-k)!} p^k q^{N-k} \quad (9)$$

$$= Np \sum_{k=1}^N \frac{(N-1)!}{(k-1)!(N-k)!} p^{k-1} q^{N-k}. \quad (10)$$

A change of variables to $m = N - 1$ and $s = k - 1$ makes it clear that the sum is just the binomial distribution, which we know is normalized to one. Therefore,

$$\langle k \rangle = Np \sum_{s=0}^m \frac{(m)!}{(s)!(m-s)!} p^m q^{m-s} \quad (11)$$

$$= Np. \quad (12)$$

This result matches our intuition about this system—the expected number of successes is equal to the probability of success multiplied by the number of trials. However, the mean is only one of half of characterizing a probability distribution—we also have to find the variance. **Variance** is a measure of the spread of a distribution about its mean, and is mathematically defined as follows:

$$\text{Var}(k) = \sigma_k^2 = \langle k^2 \rangle - \langle k \rangle^2. \quad (13)$$

The square root of variance, σ_k , is the **standard deviation**. To solve for the variance of the binomial distribution, we must compute another expected value, $\langle k^2 \rangle$.

$$\langle k^2 \rangle = \sum_{k=0}^N k^2 P(k; N, p) \quad (14)$$

$$= Np \sum_{k=1}^N k \frac{(N-1)!}{(k-1)!(N-k)!} p^{k-1} q^{N-k} \quad (15)$$

$$= Np \sum_{s=0}^m (s+1) \frac{(m)!}{(s)!(m-s)!} p^m q^{m-s} \quad (16)$$

$$= Np(sp+1) \quad (17)$$

$$= Np((N-1)p+1) \quad (18)$$

$$= N^2 p^2 + Np(1-p) \quad (19)$$

$$= N^2 p^2 + Npq. \quad (20)$$

$$\sigma_k^2 = \langle k^2 \rangle - \langle k \rangle^2 \quad (21)$$

$$= N^2 p^2 + Npq - (Np)^2 \quad (22)$$

$$= Npq. \quad (23)$$

4C. POISSON STATISTICS

Let us now try to take the limit of the binomial distribution as $N \rightarrow \infty$, but $\mu = Np$ is constant. Then:

$$P(k; N, p) = \frac{N!}{k!(N-k)!} p^k (1-p)^{N-k} \quad (24)$$

$$= \frac{N(N-1)(N-2)\dots(N-k+1)}{k!} \left(\frac{\mu}{N}\right)^k \left(1 - \frac{\mu}{N}\right)^{N-k} \quad (25)$$

$$= \frac{\mu^k}{k!} \frac{N(N-1)(N-2)\dots(N-k+1)}{N^k} \left(1 - \frac{\mu}{N}\right)^N \left(1 - \frac{\mu}{N}\right)^{-k}. \quad (26)$$

Taking the limit of each of the three N -dependent terms as $N \rightarrow \infty$, we get 1, $e^{-\mu}$, and 1, respectively. So:

$$P(k, \mu) = \frac{\mu^k}{k!} e^{-\mu}. \quad (27)$$

Since in the limit of small $p, q \rightarrow 1$, the variance $\sigma^2 = Np = \mu$. Poisson processes show up in all sorts of places, particularly in biology, where many phenomena are variations on the counting problem (binding processes, photon detection, random excitation, molecule synthesis etc.). Like the binomial distribution, the Poisson distribution is a discrete distribution. Note that the probability of having no events is:

$$P(0, \mu) = e^{-\mu}. \quad (28)$$

It follows that the probability of having at least one success is:

$$P(k \geq 1, \mu) = 1 - e^{-\mu}. \quad (29)$$

The Poisson distribution often shows up describing what is known as a Poisson process. A Poisson process is one where an event occurs with some constant probability per unit time λ , such that over some time t the mean number of events $\mu = \lambda t$. Thus, the probability that k events occur in time t is:

$$P(k, \lambda) = \frac{\mu^k}{k!} e^{-\lambda t}. \quad (30)$$

4D. OPTIMAL ROD LENGTH

We are now ready to calculate the optimal length of a rod cell in terms of signal to noise. If the mean number of spontaneous isomerizations n in a cell is $\langle n_{dark} \rangle = r_{dark} \tau N_{rh}$, the standard deviations of fluctuations in this number will be:

$$\delta = \sqrt{r_{dark} \tau N_{rh}}$$

A fraction of photons in each flash of light that enters a rod cell are absorbed by that rod cell. This constitutes the true signal:

$$N_{\text{flash}} \left(1 - e^{-\sigma CL} \right)$$

The total number of rhodopsin molecules in the rod cell is a function of rhodopsin concentration, C , and rod volume, AL . The ratio of signal to noise is thus:

$$SNR = \frac{N_{\text{flash}} (1 - e^{-\sigma CL})}{\sqrt{CALr_{\text{dark}}\tau}}$$

This function has a maximum at an intermediate value of L between 0 and ∞ . Its maximum is reached when $CL \sim 1.26/\sigma$. This means that the probability of an incident photon not being absorbed when signal to noise is maximum is:

$$1 - P = e^{-CL\sigma} \sim e^{-1.26} \sim 0.28$$

Thus, to maximize signal-to-noise ratio, nearly 30% of photons should pass through the rod without being absorbed.

REFERENCES

- Chapter Nine. [Download paper](#)
- [Download paper](#)

5. HOW MANY PHOTONS CREATE VISION

RODS ARE ENGINEERED TO DETECT PHOTONS, but how many photons are needed to create light that can be seen. The trouble is that there are so many rod cells in the human eye (100 million) and so many rhodopsin molecules per rod cells (100 million) that even if spontaneous isomerizations are rare per rhodopsin, they can occur with an appreciable rate in the retina as a whole. The eye has no way of telling the difference between an isomerization that occurs randomly and an isomerization that is triggered by a photon. The only way to see a flash of photons is if the corresponding signal is larger than the baseline noise caused by thermal isomerizations.

Scientists had been interested in the smallest see-able flash of light before photons were recognized. In 1881, Langley reported his *bolometer*, a device capable of measuring a difference in temperature as small as 0.00001°C . The bolometer detected the temperature-induced change in electrical resistance of a metal conductor. Increasing temperature increases metal resistance. Langley's bolometer compared the tiny differences in resistance of an illuminated and un-illuminated reference metal. With it, he could measure the thermal radiation from a cow from a quarter mile. He also estimated the minimum energy of a see-able flash: 3×10^{-9} ergs.

In 1905, Einstein explained the photoelectric effect in terms of the photon, the indivisible quantum of light energy. By then, the minimal see-able energy for vision had been reduced another 10-fold. Lorentz used the new equation for the energy of the photon, $E = h\nu$, to derive the minimum number of *photons* needed to see, about 100. This is the number of photons delivered to the cornea. Different experiments in different conditions made similar estimates. Hecht, Shlaer, and Pirenne (1934) noted the most reliable:

Hecht et al. knew that rhodopsin (then called visual purple) was the molecular photosensor. They asked a deeper question, how many rhodopsins must be activated to be seen? This number might be estimated using the measurable corneal reflectance, scattering of the vitreous humor, and rhodopsin absorption – an order of magnitude smaller than the number of photons at the cornea. They sought a direct measure.

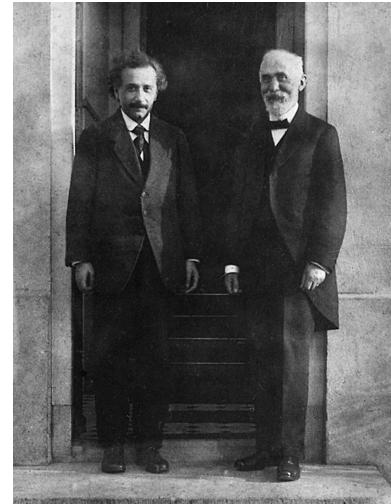


Figure 39: **Einstein and Lorentz.** In 1905, Einstein (left) published his paper on the photoelectric effect, the emission of electrons when electromagnetic radiation, such as light, hits a material. In classical electricity and magnetism, continuous light waves transfer energy to electrons, which would then be emitted when they accumulate enough energy. Einstein explained that the kinetic energy of emitted electrons was the difference in energy of single photons and a threshold voltage. There was no dependence on the number of photons, only the energy of single photons. The photoelectric effect could be fully explained to be a function of the frequency (energy) of single photons, quanta of light. His friend Lorentz (right) used the new equation for the energy of single photons to make the first estimate of the smallest see-able number of photons.

Wavelength	No. of quanta	Source
505	17-30	Chariton and Lea, 1929
507	34-68	von Kries and Eyster, 1907
530	40-90	Barnes and Czerny, 1932

HECHT, SHLAER, AND PIRENNE sought the smallest see-able light stimulus, and so maximized the chances that any human observer would see such a flash. To do this, they optimized experimental conditions. First, they measured scotopic, or rod-dominated, vision.

The first consideration is dark adaptation. In the dark, pupils will dilate to allow more incoming photons to reach the retina. This happens in a few seconds and increases sensitivity by about ten-fold. In bright light, all rhodopsin molecules in a rod cell will be "bleached" and non-absorptive because of the *all-trans* configuration of visual pigment. All rod cells will be in a state of hyperpolarization and low rates of synaptic release. The many additional orders of magnitude of light sensitivity in scotopic vision arise by biochemistry that restores all visual pigment to the *cis* configuration. This takes about 30 minutes.

Rods and cones are not evenly distributed in the retina (Fig. 40). Vision with highest-spatial resolution is at the fovea, the center-point of our visual field that only contains cones. Rod-dominated vision is in our peripheral vision. If you want to see something in dim light, you will do better by not looking straight at it.

Light energy is delivered in illumination that is spread over space and time, that is photons spread over square distances and over time. A larger test area that is illuminated weakly can be seen as easily as a smaller test area that is illuminated strongly, with similar numbers of photons in both cases. The reciprocal relationship between intensity and area is not perfect, and has a maximum sensitivity corresponding to a circular retinal area that spans about 500 rod cells.

It is easier to see small numbers of photons when they arrive in one visual 'moment' in which all the photons are counted at once. The visual 'moment' in both rod and cone photoreceptor cells is surprisingly long. Metabotropic receptors that require a biochemical signal transduction cascade to open and close ion channels (like vertebrate olfactory receptors and photoreceptors, both G-protein coupled receptors) are often slower than ionotropic receptors (like vertebrate Piezo mechanosensory channel or insect olfactory receptors) where the stimulus directly opens and closes ion channels. This is why the *flicker fusion threshold*, the frequency of visual stimuli when an intermittent stimulus appears to be steady, is so slow. Both rod and cone photoreceptors effectively integrate the responses to photons that arrive within 0.01-0.1 s. The sensitivity of human rod-mediated vision to flicker reaches a plateau at 15 Hz. Cone-mediated vision at very bright illumination reaches a plateau at 60 Hz. This is why television and computer monitors operate at 30 or 60 frames per second. Any

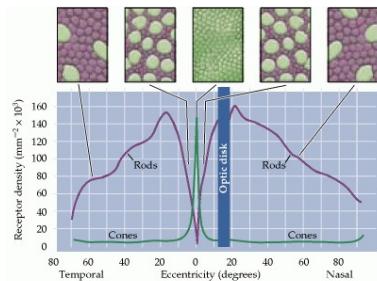


Figure 40: **Rod distribution.** Distribution of cones and rods in a typical human retina

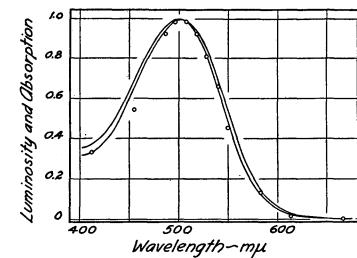


Figure 41: **Rod absorption.** Comparison of scotopic luminosity at the retina with visual purple absorption. The curves are the percentage absorption spectra of visual purple; the upper curve represents 20 per cent maximal absorption, and the lower one 5 per cent maximal absorption. All curves have been made equal to 1 at the maximum, 500 nm, for ease in comparison.

slower, and you would see motion jitter between frames. Any faster, and you would see no improvement in the smoothness of movement in video. The reciprocal relationship between intensity and time of exposure in scotopic vision holds perfectly for stimuli lasting <0.01 sec, and so Hecht, Shlaer, and Pirenne used 0.001 s flashes.

Finally, the color of the visual stimulus should be optimized for scotopic vision. The wavelength sensitivity of purified rhodopsin and scotopic vision are remarkably similar, both peaking near 510 nm (Fig. 41).

THE HSP EXPERIMENT involved a human observer who triggered the release of a flash of light that would fall on their retina in an area spanning ~ 500 rods at about 20° from the center of vision. Another person manipulated the filters and wedges that controlled the flash intensity. Because the observer triggered each flash, the observer also knew when to pay attention. This aspect of the experimental design, by maximizing the readiness and receptivity of the observer, might have increased the sensitivity of their measurements to minimal flashes. Honesty mattered. The observer had to be truthful about admitting to not seeing a flash of light that they themselves had delivered to their own retinas but failed to see.

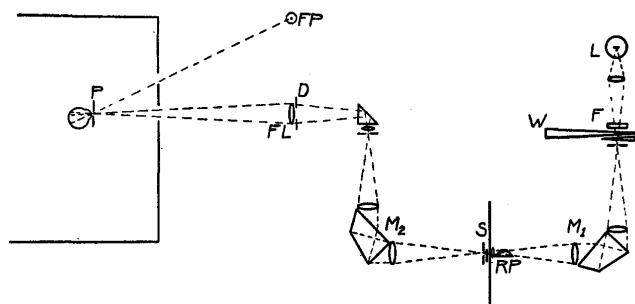


Figure 42: Apparatus for measuring minimum energies necessary for vision. The eye at the pupil P fixates the red point FP and observes the test field formed by the lens FL and the diaphragm D . The light for this field comes from the lamp L through the neutral filter F and wedge W , through the double monochromator M_1M_2 and is controlled by the shutter S .

A meaningful threshold for seeing is when an observer successfully perceives a majority of stimuli. The HSP experiment did not allow *false positives* – the observer cannot see a flash that was not triggered. But every failure to see a flash is a *false negative*, where the observer must admit to failing to see a self-administered light. Valid statistics require truthfulness of the observer, who must not pretend to see flashes, even when they knew that they occurred. HSP reduced the ‘threshold for seeing’ to one simple probability, when an observer performed at a rate of 60% true positive/40% false negative.

HECHT, SHLAER, AND PIRENNE obtained remarkably consistent results for a threshold for seeing, the number of photons that hit the cornea, that are successfully seen 60% of the time. Seven different subjects saw flashes of light that delivered between 54 and 148 blue-green photons to the cornea. These numbers are comparable to the earlier ‘most reliable’ measurements.

Observer	Quanta
S.H.	126
	135
	107
	87
	79
	123
S.S.	148
	79
	54
	56
	62
	96
C.D.H.	99
	104
	65
	76
M.S.	58
	58
S.R.F.	81
	112
A.F.B.	120
M.H.P.	83
A.F.B.	79
	83
M.H.P.	138

A DEEPER QUESTION is how many photons are needed to activate rod cells to see? How many individual rods need to be activated for the retina to ‘see’ a flash? One approach is to estimate the fraction of photons that arrive at the cornea that reach the retina and are absorbed by rhodopsin. The cornea reflects ~4% of incident photons. The vitreous humor absorbs ~50% of transmitted photons. Even rods of optimal length will only absorb ~70% of photons that reach the retina. The retina absorbs a small fraction of the photons that arrive at the cornea. If the number of photons that arrive at the cornea is N , the number of absorbed photons is $a = \alpha N$, where α is unknown and takes into account all losses by reflection and absorption.

Say that flashes result, on average, in one absorbed photon at the retina: $a = 1$. Not every flash will result in one photon absorption. Some will result in zero absorptions. Some

will result in two or more absorptions. Poisson statistics describes the complete probability distribution of the number of absorbed photons in each trial, $P(k)$

$$P(k) = \frac{a^k}{k!} e^{-a}$$

, where a is the mean arrival number of photons across trials. Why Poisson statistics? No light source reliably delivers a fixed number of photons to the retina in every trial. Each photon that is released by a light source has a small probability of making its way to the retina and being absorbed. Large numbers of photons at the source and small probabilities that each photon might eventually be absorbed results in Poisson statistics. The probability of seeing curve can be plotted as a function of the mean stimulus size, but at each stimulus size, there is unavoidable trial-to-trial variability in the size of the stimulus that will affect the probability of seeing.

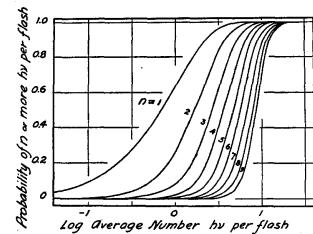


Figure 43: Probability of seeing, the theory. For any average number of quanta per flash, the y-axis gives the probabilities that the flash will deliver n or more quanta to the retina, with different values assumed for n .

If the threshold of seeing is one absorbed photon at the retina, then the ‘probability of seeing’ curve will be a smooth sigmoidal function. This sigmoidal curve can be calculated using the Poisson distribution and summing the probabilities of absorption of one, two, three photons and so on (Fig. 43). If the threshold of seeing is two absorbed photons, then Poisson distributions must be summed from the absorption of two photons and up. For a threshold of n photons, the probability of seeing curve is:

$$P_{see} = \sum_{k=n}^{\infty} \frac{a^k}{k!} e^{-a}$$

As the threshold increases, the sigmoidal curve shifts towards more photons being absorbed. As the curve shifts towards larger stimuli, it becomes steeper when plotted as probability versus the *logarithm* of stimulus size (Fig. 43).

There is merit in plotting P_{see} against the logarithm of stimulus size. Plot P_{see} as a function of the logarithm of the mean number of photons that arrive at the retina, $\log a$. Plot P_{see} as a function of the logarithm of the mean number of photons that arrive at the cornea, $\log N$. Because $a = \alpha N$, $\log a = \log \alpha + \log N$. The two P_{see} curves will have identical shapes except for a horizontal shift of $\log a$.

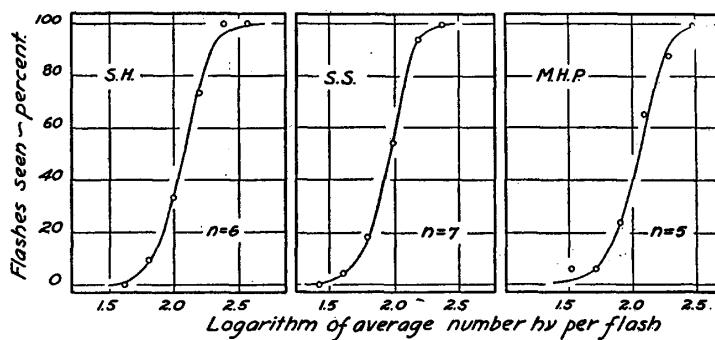


Figure 44: **Probability of seeing, the experiment.** Relation between the average energy content of a flash of light (in number of photons) and the frequency with which it is seen by three observers. Each point represents 50 flashes, except for S.H. where the number is 35. The curves are cumulative Poisson distributions with different thresholds.

Another merit in plotting against the logarithm of stimulus size (whether stimulus size is measured at the cornea or retina) is that the slope of the curve is a direct measure of the threshold:

$$\frac{dP_{see}}{d \log N} \approx \sqrt{n}$$

To derive this square root dependence, one needs to differentiate the probability of seeing, use Stirling’s formula, use the chain rule, and evaluate the slope near the inflection where $a \approx n$. As n increases, the steepness of the probability of seeing curve will increase. Whether the probability of seeing curve is plotted as the logarithm of photons at the cornea or the logarithm of photons at the retina does not matter to this steepness. Steepness is always a function of n , the threshold number of photons needed to see.

The steepness of the probability of seeing curve resembles the **signal-to-noise ratio** of sensory perception. The size of the signal – the numerator of the signal-to-noise ratio – is the number of absorbed photons. Because this signal obeys Poisson statistics, the standard deviation in stimulus size – the noise in the denominator of the signal-to-noise ratio – is its square root: $\sigma = \sqrt{n}$. For the three subjects for whom they had the most data – with initials S.S., S.H., and M.H.P. – HSP derived thresholds of 5,6, and 7 photons from the steepness of their respective probability of seeing curves.

WHY ARE MULTIPLE PHOTONS NEEDED TO SEE? HSP argued that probability of seeing curves depended on the statistical variability of any flash stimulus. But the probability of seeing curve can also depend on the statistical variability of internal noise. A weak signal corresponding to 5-8 photons must be distinguished from a background of spurious signals that will occur in the retina in total darkness. The retina cannot know whether a given rhodopsin activation was due to thermal activation or photon activation. The retina can only count rhodopsin activations. The threshold decision – whether a subject sees a flash of light in a given trial – must be based on the sum of weak and spurious signals. If the weak signal is ‘seen’, it is because this summed signal is larger than typical spurious signals in darkness.

The threshold for a perceptual decision is linked to the reliability of the response. If the threshold is lowered, then response reliability will be lowered. This is because more spurious signals can be interpreted as ‘seeing’, increasing the rate of false positives. If the threshold is raised, some weak signals will not be large enough to be distinguished from spurious signals alone, increasing the rate of false negatives. The probability of seeing and response reliability will be functions of signal size, noise, and the internal detection threshold used by the observer.

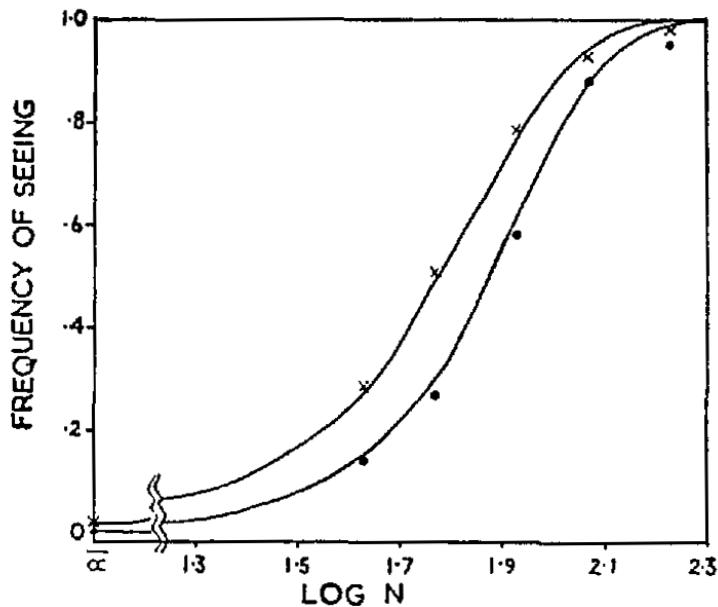


Figure 45: Poisson probability distributions. For any average number of quanta per flash, the y-axis gives the probabilities that the flash will deliver n or more quanta to the retina, with different values assumed for n .

HORACE BARLOW wanted to test the effect of response reliability on detecting weak signals. Human observers are conscious of the certainty of their own perceptions and the reliability of their own decisions. Barlow took advantage of this perceptual awareness to test the responses of one human observer, a fellow named Roy Rumble, at two different detection thresholds. Barlow delivered flashes of different intensities, including blanks where no flash was given, to Rumble. Rumble was encouraged to signal when a flash was “seen” and also when a flash was “possible”. Thus, the threshold number of absorbed photons “possible” would be lower than the threshold for “seen”. Because Barlow knew when blanks were delivered, he knew

when Rumble reported “false positives” and “false negatives”.

Indeed, the probability of seeing curve for “seen” events had a higher threshold (shifted to higher flash intensities) and higher steepness (required more photons) than probability of seeing for “possible” plus “seen” events. Response reliability differed for the two curves. The subject never said a blank was “seen”, but reported 3 of the 300 blanks as “possible”. Lowering the response threshold increased the rate of false positives to 1%. From these curves, Barlow extracted somewhat different parameters from his experiment than those of HSP.

- N , average number of photons at the cornea
- n , average number of photon absorptions, i.e., *bona fide* rod excitations
- x , average number of confusable events, i.e., spurious rod excitations
- $a = x + n$, total average of events (real photons plus noise)
- c , threshold number of events to “see”

Just as for HSP, the probability of seeing is a cumulative probability distribution involving these parameters.

$$P_{\text{see}} = \sum_{k=c}^{\infty} \frac{a^k}{k!} e^{-a}$$

The HSP experiment used the frequency of seeing curve as a function of the logarithm of flash intensity to estimate two values, 1) the fraction of incident photons at the cornea that activated rod cells, and 2) the threshold number of activated rod cells needed to see. These two values can be inferred from one frequency of seeing curve, because this sigmoidal curve has two quantifiable parameters, its horizontal shift and steepness. Barlow’s experiment had more parameters. Like the HSP experiment, a fraction of incident photons would activate rods (one unknown value). But two frequency of seeing curves in Barlow’s experiment meant two separate thresholds (two more unknown value). Lastly, every flash stimulus would be accompanied by a certain number of spurious rod activations (a fourth unknown, noise that has the same average amplitude in every trial). Barlow’s experimental measurement of two frequency of seeing curves contained enough information to calculate the four unknowns.

The steepness of the probability of seeing curve in Barlow’s formulation is somewhat different than the steepness in HSP’s formulation:

$$\frac{dP_{\text{see}}}{d \log N} \approx \frac{c - x}{\sqrt{c}}$$

As before, the steepness evokes the **signal-to-noise ratio**. In this case, the size of the signal in the numerator is the distance of the threshold c from confusable events x , and the size of the noise in the denominator is the expected fluctuations in that signal given by Poisson statistics, \sqrt{c} .

Parameters	Best fits
n/N	0.14
x	8.9
c for Possible or Seen	17
False positive rate for Possible or Seen	0.01
c for Seen	19
False positive rate for Seen	0.002

For “seen” events, meeting the threshold requires 10 events above the number of confusable events (spontaneous rhodopsin activation). In terms of signal-to-noise, events corresponding to “seen” *bona fide* flashes are ~ 2.3 standard deviations from the confusable events. Events corresponding to “possible” plus “seen” flashes are ~ 2 standard deviations from confusable events.

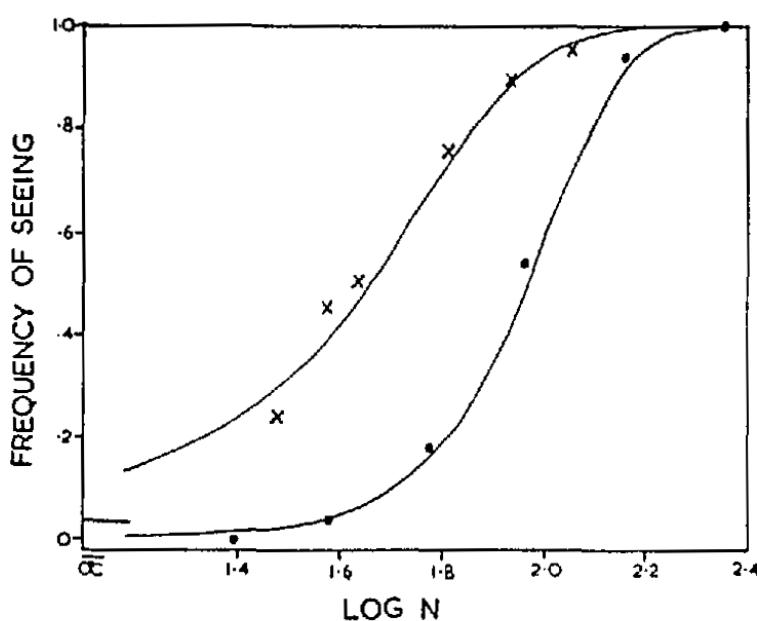


Figure 46: **Other experiments.** Data from Hecht et al. (dots) and van der Velden (crosses) fitted using Barlow's theoretical curves and parameters.

Is Barlow's formulation consistent with earlier measurements? From one probability of seeing curve in HSP (or one probability of seeing curve in another measurement by van der Velden), it is impossible to extract all the parameters that Barlow used. Earlier measurements did estimate the number of photons at the cornea, N , from which they inferred the fraction of photons that were lost and the threshold number of rhodopsin activations needed to see. If the number of spontaneous rhodopsin activations in earlier measurements was similar to that measured by Barlow, $x = 8.9$, these numbers could be inferred (Fig. 46). Fitting Barlow's parameters to HSP's experiments revealed roughly consistent numbers for the fraction of absorbed photons ($n/N = 0.13$) and the threshold of rhodopsin activations ($c = 21$). Fitting Barlow's parameters to another experiment by van der Velden revealed an unrealistically large fraction of absorbed photons ($n/N = 0.9$) and lower threshold of rhodopsin activations ($c = 15$).

REFERENCES

- [Download paper](#)
- [Download paper](#)

6. SINGLE PHOTONS AND SINGLE ROD CELLS

HOW IS PHOTON ABSORPTION TRANSDUCED INTO ROD ACTIVITY?

The predominant cation in the extracellular solution is sodium. The predominant cation inside the intracellular cytoplasm is potassium. In darkness, Na^+ channels in the outer segment are open, leading to inward ionic current. At the same time, K^+ channels in the inner segment are open, leading to outward ionic current. When a rhodopsin molecule is activated in the outer segment, a biochemical signal transduction cascade ensues. This cascade will close some Na^+ channels in the outer segment. By Ohms Law ($V = IR$), a net reduction in an inward current of positive ions means hyperpolarization of the membrane potential (which is typically reported as internal voltage minus external voltage).

The fortuitous anatomy of the rod cell allows light-evoked currents to be recorded by suction electrode recording. The outer segment is pulled into a tightly-fitting glass pipette. Any changes in the current loop between an electrode inside the pipette and an electrode in the bath can be recorded, and can indicate changes in light-evoked current through the outer segment.

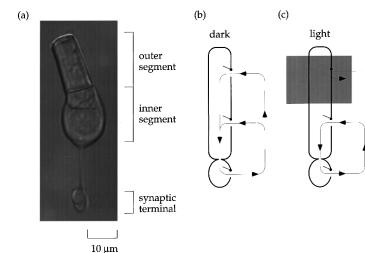


Figure 47: **Rod cells.** (a) Isolated rod photoreceptor from salamander. The cell membrane separates intracellular and extracellular salt solutions. The outer segment contains rhodopsin and transduces photons into rod activation. The inner segment keeps the cell alive. The synaptic terminal communicates signals to bipolar and horizontal cells. (b) In darkness Na^+ ions enter the outer segment through cGMP-gated channels. A current loop is completed by outward movement of K^+ ions through channels in the inner segment. Separate Na/K exchange pumps maintain the relatively lower intracellular concentration of Na^+ (which might be 15 mM inside a typical cell and 140 mM outside) and relatively higher intracellular concentration of K^+ (which might be 120 mM inside and 3 mM outside). (c) When a rod cell is exposed to light, some channels in the outer segment close. The cell hyperpolarizes, which reduces neurotransmitter release at the synaptic terminal.

REFERENCES

- Download paper

VISION WORKBOOK

ONE. COUNTING PRACTICE (FROM *Physical Biology of the Cell*)

Consider eight particles, four of which are black and four white. Four particles can fit left of a permeable membrane and four can fit to the right of the membrane. Imagine that due to random motion of the particles, every arrangement of the eight particles is equally likely. Some possible arrangements are BBBB|WWWW, BBW|BBWW, and BWBW|BWBW where membrane position is given by |.

1. How many different arrangements are there?
2. Calculate the probability of having all four black particles on the left of the membrane. What is the probability of having one white particle and three black particles on the left of the membrane? Finally, calculate the probability that two white and two black particles are left of the membrane. Compare these three probabilities. Which arrangement is most likely?
3. Imagine that, in one time instant, a random particle from the left side exchanges places with a random particle on the right side. Starting with three black particles and one white particle on the left of the membrane, compute the probability that after one time instant there are four black particles on the left. What is the probability that there are two black and two white particles on the left, after the same time instant? Which is the more likely scenario?

TWO. STIRLING'S APPROXIMATION (FROM *Physical Biology of the Cell*)

The Stirling approximation is useful in probability and combinatorics. Our task is to find a useful representation of $n!$.

1. Begin by showing that

$$n! = \int_0^\infty x^n e^{-x} dx$$

To demonstrate this, use repeated integration by parts. In particular, demonstrate the recurrence relation

$$\int_0^\infty x^n e^{-x} dx = n \int_0^\infty x^{n-1} e^{-x} dx$$

and then argue that repeated application of this relation leads to the desired result.

2. Make plots of the integrand $x^n e^{-x}$ for various values of n and observe the peak width and height of this integrand. We are interested in finding the value of x for which this function is a maximum. The idea is that we will then expand about that maximum. To carry out this step, consider $\ln(x^n e^{-x})$ and find its maximum – argue why it is acceptable to use the logarithm of the original function as a surrogate for the function itself, that is, show that the maxima of both the function and its logarithm are the same x . Also, argue why it might be a good idea to use the logarithm of the integrand rather than the integrand itself as the basis of our analysis. Call the value of x for which this function is maximized x_0 . Now expand the logarithm about x_0 . Examine

$$\ln \left[(x_0 + \delta)^n e^{-(x_0 + \delta)} \right] = n \ln(x_0 + \delta) - (x_0 + \delta)$$

and expand to second order in δ . Exponentiate your result and you should now have an approximation to the original integrand that is good in the neighborhood of x_0 . Plug this back into the integral (be careful with limits of integration) and, by showing that it is acceptable to send the lower limit of integration to $-\infty$, show that

$$n! \approx n^n e^{-n} \int_{-\infty}^{\infty} e^{-\delta^2/2n} d\delta$$

Evaluate the integral and show that in this approximation

$$n! = n^n e^{-n} (2\pi n)^{1/2}$$

Also, take the logarithm of this result and make an argument as to why most of the time we can get away with dropping the $(2\pi n)^{1/2}$ term.

THREE. THE BOLTZMANN DISTRIBUTION – COUNTING STATISTICS

1. Prove the identity

$$W = N! / (n_1! n_2! \dots n_s!)$$

where W is the number of partitions of N particles into s energy levels, with respectively n_1, n_2, \dots, n_s particles per level.

2. What is the probability of being in state s when W is maximized (use the derived Boltzmann distribution)?
3. If the energy levels correspond to kinetic energy of the particles such that

$$\epsilon_s = (mv_s^2)/2$$

calculate average speed and average kinetic energy of the particles.

FOUR. THE DENSITY OF RHODOPSIN

Using realistic numbers, estimate the mean separation between rhodopsin molecules in the disks of the human rod cell. How does this distance compare with the diffraction limit of visible light?

FIVE. THE HEIGHT OF THE ATMOSPHERE

Assuming that the gas molecules in Earth's atmosphere are in thermal equilibrium near 300K. This is a somewhat poor assumption, but correct to within an order of magnitude. Using realistic values, estimate the relative density of different gas molecules at different altitudes. How much less O₂ should you expect at the top of Mt. Everest with respect to sea level? Google the answer, and see if knowing the Boltzmann distribution led to a decent estimate.

SIX. INFORMATION THEORY AND AN UNFAIR DIE (FROM *Physical Biology of the Cell*)

Roll a fair 6-sided die. It will have equal probability of rolling 1, 2, 3, 4, 5, or 6. Thus, the probability distribution of possible outcomes is a flat: $P(n) = 1/6$ for any value of n . The mean value of a roll of a fair die is thus $\langle n \rangle = 3.5$.

Using information theory, calculate the ‘maximum entropy’ probability distribution, $P(n)$, for an unfair die. All you know about the unfair die is that the mean value of a roll is $\langle n \rangle = 2.5$.

SEVEN. MULTIPLE SOURCES OF RANDOMNESS

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. In this problem, you will show that if the photon arrivals constitute a Poisson process with rate r , and we detect a fraction f of these, selected at random, then the photon detections will also be a Poisson process, with rate rf .

1. Denote the cumulative number of photon arrivals and detections up to time t by $N(t)$ and $K(t)$, respectively. Explain why the conditional distribution of $K(t)$ given $N(t)$ is binomial.
2. Use the law of total probability in the form

$$\Pr[K(t) = k] = \sum_{n=0}^{\infty} \Pr[K(t) = k \mid N(t) = n+k] \Pr[N(t) = n+k]$$

to obtain the desired result. *Hint:* See what you can factor out of the sum.

EIGHT. THE VARIANCE OF A BINOMIAL DISTRIBUTION

Flip a biased coin n times. The probability of heads in one coin flip is p . The probability of k heads in n coin flips is given by the binomial distribution:

$$\Pr(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}$$

for $k = 0, 1, 2, \dots, n$, where

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Using the binomial distribution, show that the mean number of heads is $\langle k \rangle = np$. Show that the variance in the number of heads is $\langle k^2 \rangle - \langle k \rangle^2 = np(1-p)$. Sketch the shape of this probability distribution when n is large.

NINE. THE VARIANCE OF A POISSON DISTRIBUTION

For a Poisson process where the expectation value of successes is $\langle k \rangle = \mu$, the probability distribution of any number of successes is:

$$P(k, \mu) = \frac{\mu^k}{k!} e^{-\mu}. \quad (31)$$

Show that the variance is equal to the mean:

$$\langle k^2 \rangle - \langle k \rangle^2 = \langle k \rangle$$

TEN. EXPONENTIAL FUNTIONS

Show that the exponential function can be written in terms of the following limit:

$$\lim_{N \rightarrow \infty} \left(1 - \frac{x}{N}\right)^N = e^{-x}$$

ELEVEN. POOLING DATA

In the paper by Hecht et al., psychometric curves for each of the three subjects are analyzed independently. What would have happened if the authors had instead pooled their data? Generate a single plot of detection frequency against log stimulus intensity after pooling the data in Table 5 for all 3 subjects. If the authors had followed this procedure, would they have been able to draw the correct conclusions? Discuss.

TWELVE. OPTIMAL LENGTH OF ROD CELLS FROM BIALEK's *Biophysics*

In lecture we discussed how optimizing the signal-to-noise ratio (SNR) in a rod cell gives the counterintuitive result that 30% of photons should pass through undetected. The SNR can be expressed as

$$\text{SNR} = \frac{n_{\text{flash}}(1 - e^{-\sigma\rho L})}{\sqrt{\rho L A r_{\text{dark}} \tau}},$$

where σ is the absorption cross section of rhodopsin, ρ is the concentration of rhodopsin, L and A are respectively the length and cross-sectional area of the rod cell, n_{flash} is the number of incident photons, r_{dark} is the rate of spontaneous thermal isomerization, and τ is the effective integration time.

1. Explain why $1 - e^{-\sigma\rho L}$ is just the probability for a photon to be absorbed.
2. Show that the SNR is maximized when $\rho L \approx 1.26/\sigma$.
3. Cats, dogs, and many other animals have a reflective layer of tissue in their eye known as the *tapetum lucidum*, which gives them superior night vision to humans. Assuming that the *tapetum lucidum* acts as a mirror reflecting the incoming photons with a ratio $R (0 \leq R \leq 1)$ at the end of the rod cell, repeat the SNR optimization. What fraction of photons now pass through undetected and How does the *tapetum lucidum* improve the SNR, as a function of R ? Discuss.

THIRTEEN. CONVOLUTIONS

Let X, Y be two independent, continuous random variables. In this problem, you will show that the distribution of their sum $Z = X + Y$ is given by the convolution of their respective distributions:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z-x)dx =: (f_X \circledast f_Y)(z)$$

1. Find $F_Z(z) := \Pr(Z \leq z)$, the cumulative distribution of Z , by integrating the joint distribution $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ over the region of the (x,y) -plane where $x + y \leq z$. You should obtain an expression for $F_Z(z)$ in the form of a single integral over the real line.
2. Compute $f_Z(z) = F'_Z(z)$. Hint: It is fine to bring the derivative inside the integral as a partial derivative with respect to z , since the integration variable should be x (or y) rather than z .

FOURTEEN. LIFETIME OF ACTIVATED RHODOPSIN

Photoreceptors exhibit remarkably stereotyped electrical responses to single photons because the distribution for the lifetime of activated rhodopsin is less dispersed than an exponential. It has been argued that the sub-exponential variance is due to the multistep biochemical mechanism by which rhodopsin is deactivated; in this problem you will explore that idea.

1. Let $S_n = T_1 + T_2 + \dots + T_n$, where the T_i are i.i.d. exponential random variables with common rate parameter r . Define the Erlang distribution $E(r, n)$ by the density function

$$f(t) = \frac{r^n t^{n-1} e^{-rt}}{(n-1)!}. \quad (32)$$

Use the result of the problem on convolution (above) to show that $S_n \sim E(r, n)$ implies $S_n + T_{n+1} = S_{n+1} \sim E(r, n+1)$. Then, since $E(r, 1)$ is just the exponential distribution, it follows by induction that a sum of i.i.d. exponential random variables is Erlang. Hint: When setting the limits of the convolution integral, consider the range of the integration variable on which the integrand is nonzero.

2. Suppose that the phosphorylation state of a rhodopsin molecule is initially n . Every dephosphorylation event reduces this state by 1 until it reaches 0, at which point the molecule is deactivated. Suppose, further, that dephosphorylation is a Poisson process with rate $r = nr_0$. With $n = 7$ and $r_0 = 0.1 \text{ s}^{-1}$, generate a histogram of lifetimes for an ensemble of $N = 1000$ rhodopsin molecules. Plot the prediction of Eq. (32) on the same set of axes for comparison.
3. Now simulate $n = 1, 2, \dots, 10$ for $r_0 = 0.1 \text{ s}^{-1}$. Plot the coefficient of variation (i.e., the ratio of the standard deviation to the mean) of the lifetime of activated rhodopsin as a function of n . How could you have gotten this result without simulating (or integrating!) the Erlang distribution?

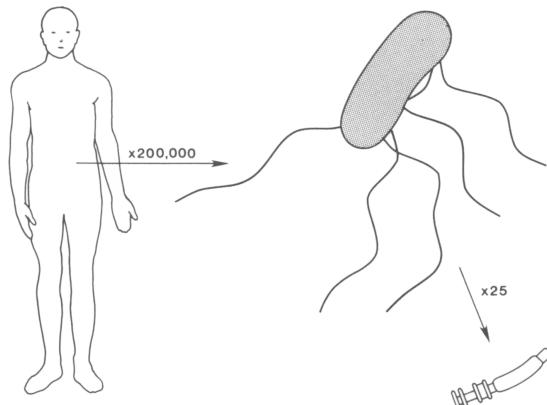
FIFTEEN. MODEL OF DARK NOISE

Suppose that we observe some number n of visual ‘events’ drawn from a Poisson distribution whose mean $\langle n \rangle$ can be written as the sum of a contribution from dark noise $\langle n_{\text{dark}} \rangle$ a contribution from actual light flashes $\langle n_{\text{flash}} \rangle = N$ if there is a flash and 0 otherwise. As an observer, you declare a flash to have occurred if and only if the number of events that you observe exceeds some threshold θ . Your sensitivity to light is maximized when the threshold θ is small, but so is your rate of false positives (and vice versa when θ is large). The conventional way to describe this trade-off is to plot the true positive rate (i.e., the probability that a given flash was indeed called a flash) against the false positive rate (i.e., the probability of ‘seeing’

a flash that never happened) while varying the value of θ . Visualize this so-called receiver operating characteristic (ROC) curve for the case that $\langle n_{dark} \rangle = 10$ and $N = 10$. Then adjust N while holding $\langle n_{dark} \rangle$ fixed and explore 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 Figure 7 in their paper.

7. OLFACTION AND DIFFUSION

BACTERIA SENSE THEIR ENVIRONMENTS USING THEIR SENSE OF SMELL. *E. coli* does not eat molecules in order to judge how many molecules are in their neighborhood. It uses chemoreceptors that transiently bind molecules in the environment. These metabotropic receptors communicate signals through a biochemical signal transduction cascade to control the direction of flagellar rotation. An organism as small as *E. coli* that is trying to ‘smell’ its environment encounters challenges. To understand these challenges, we have to think about physics on small length scales. How small is *E. coli*? Each bacterial cell is about 1/20,000 the size of the host that it inhabits. The bacterial flagellar motor that drives its swimming motility is another 50-fold smaller.



In the absence of chemical cues (an isotropic environment without gradients of attractants or repellants), *E. coli* wanders around. It swims smoothly by rotating several helical flagella in the counter-clockwise direction (when viewed from the outside looking in). Runs are random in duration and exponentially distributed. Run length obeys a Poisson interval distribution. During a run, the bacterium has a constant probability per unit time of ending the run and starting a ‘tumble’. During a tumble, one or more flagellar rotate in the CW direction. When this happens, the flagellar bundle flies apart and the cell moves erratically. The tumble continues until all flagella resume CCW rotation and coalesce to form a new bundle. Runs typically last ~ 1 sec and tumbles last ~ 0.1 sec.

Figure 48: **A sense of scale.** Comparisons of man, *E. coli*, and the flagellar rotary motor that drives swimming

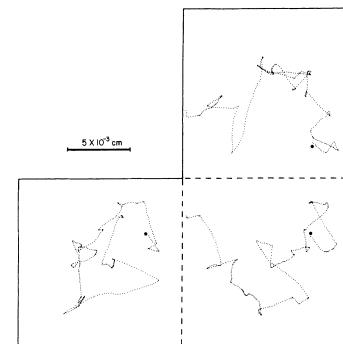


Figure 49: **Biased random walks.** A digital plot (12.6 points/sec) of the displacement of a wild-type cell (*E. coli* strain AW405) executing a random walk in a homogeneous, isotropic medium. These are planar projections of a three-dimensional track: If the left and upper panels are folded out of the page along the dotted lines, the projections appear in proper orientation on three adjacent faces of the cube. Tracking began at the large dot and continued for about 30 sec. The cell swam at the speed of 2×10^{-3} cm/sec. There were 26 runs and tumbles.

TO PERFORM CHEMOTAXIS, bacteria must be able to count molecules in their environment, and modulate their movements accordingly.

When *E. coli* swims in a spatial gradient of a chemical attractant, runs that happen to carry it up the gradient are extended, whereas those that happen to carry it down the gradient are not. Thus, the cell drifts in a favorable direction by executing a biased random walk.

So the bacteria has to figure out whether the ambient concentration of attractant molecules is increasing or decreasing in about ~ 1 sec. One problem is that the molecules that the bacteria is trying to count are distributed at random. Even if the cell was a ‘perfect monitor’ of the molecules within its volume, the number of molecules that it will count in a given period of time will be random. If the volume of the perfect monitor is V , the average number of particles that it will detect is $\langle N \rangle = c_0 V$, where c_0 is molecular concentration. But each measurement will fluctuate as molecules move in and out of the volume. These fluctuations will be governed by Poisson statistics. In the environment surrounding the cell, there are a huge number of molecules, and each has only a tiny chance of being within the measurement volume at any time. Hence, Poisson statistics governs fluctuations in the number of molecules inside the perfect monitor. The variance in the number of counted molecules will be equal to the mean number of counted molecules:

$$\delta N^2 = \langle N \rangle$$

Fractional error varies with concentration:

$$\frac{\delta c^2}{c_0^2} = \frac{\delta N^2}{\langle N \rangle^2} = \frac{1}{\langle N \rangle} = \frac{1}{c_0 V}$$

If a μm -sized bacteria counted all the molecules inside it, it would count 600,000 molecules if the concentration of molecules were 1 mM. It would count 600 molecules if the concentration were 1 μM and 60 molecules if the concentration were 10^{-7} M. For these concentrations, measurement errors would be ± 800 , 25, and 8 molecules. Relative errors are 10% at the low end of the scale, but we know bacteria perform chemotaxis at 10^{-7} M. If bacteria made M independent measurements within each run, fractional error would decrease. In time T , $M \approx T/\tau_D$ independent measurements, ($\tau_D \approx a^2/D$ is turnover time). This reduces uncertainty:

$$\frac{\delta c^2}{c_0^2} = \frac{1}{M \langle N \rangle} = \frac{1}{(T/\tau_D)c_0 V} \approx \frac{1}{D a c_0 T}$$

To understand bacteria, we need to understand diffusion.

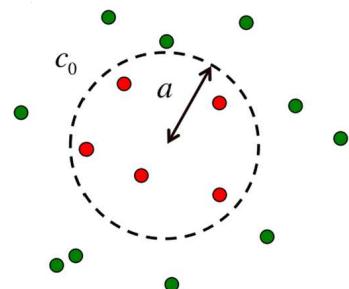


Figure 50: **The perfect monitor.** The perfect monitor is permeable to ligand molecules and estimates the concentration c_0 by counting the molecules in its volume during time T .

Diffusion coefficient for a small molecule $D \approx 10^{-5} \text{ cm}^{-2} \text{ s}^{-1}$

Diffusion: Microscopic Theory

Diffusion is the random migration of molecules or small particles arising from motion due to thermal energy. A particle at absolute temperature T has, on the average, a kinetic energy associated with the movement along each axis of $kT/2$.

A particle of mass m and velocity v_x on the x axis has a kinetic energy mv_x^2 . This quantity fluctuates, but on the average $\langle mv_x^2/2 \rangle$. We can thus estimate the instantaneous velocity of a small particle, for example, a molecule of the protein lysozyme. Lysozyme has a molecular weight 1.4×10^4 g. This is the mass of one mole, or 6×10^{23} molecules; the mass of one molecule is $m = 2.3 \times 10^{-20}$ g. The value of kT at 300°K is 4×10^{-14} g cm $^2/\text{sec}^2$. Therefore, $\langle v_x^2 \rangle^{1/2} = 1.3 \times 10^3$ cm/sec. This is *fast*. If the molecule faced no obstruction, it would move quickly across a room. Since the protein is immersed in water, it does not go far before it bumps into other molecules that slow it down and change its direction. The molecule is forced to undergo a random walk. The overall movement is called diffusion.

One-dimensional random walk



In order to characterize diffusive spreading, it is convenient to reduce the problem to its essentials, and consider the motion of particles along one axis only such as shown in Figure 1. The particles start at time $t = 0$ at position $x = 0$ and execute a random walk according to a set of rules.

- Each particle steps to the right or to the left once every τ seconds, moving at velocity $\pm v_x$ a distance $\delta = \pm v_x \tau$. For simplicity, we treat τ and δ as constants. In practice, they will depend on the size of the particle, the liquid, and the absolute temperature T .
- The probability of going to the right at each step is $1/2$ and the probability of going to the left at each step is $1/2$. The particles forget what they did on the previous leg of the journey. Successive steps are statistically independent. The walk is not biased.
- Each particle moves independently of all the other particles. The particles do not interact with one another. This will be true if the suspension of particles is reasonably dilute.

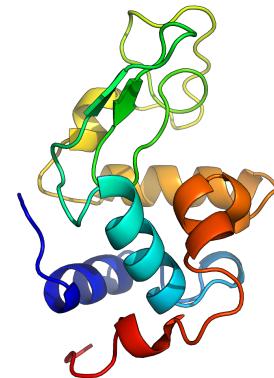


Figure 51: **Lysozyme.** An antimicrobial enzyme produced by animals that forms part of the innate immune system.

Figure 52: **1-D random walk.** Particles executing a one-dimensional random walk start at the origin, o , and move in steps of length δ , occupying positions $o, \pm\delta, \pm 2\delta, \pm 3\delta, \dots$

THE RULES OF THE RANDOM WALK have two striking consequences. The first is that the particles go nowhere on the average. The second is that their root-mean-square displacement is proportional not to the time, but to the square-root of the time. It is possible to establish these propositions by using an iterative procedure. Consider an ensemble of N particles. Let $x_i(n)$ be the position of the i th particle after the n th step. According to the first rule, the position of a particle after the n th step differs from its position after the $(n - 1)$ th step by $\pm \delta$:

$$x_i(n) = x_i(n - 1) \pm \delta \quad (33)$$

The + sign will apply to roughly half of the particles. The – sign will apply to the rest. The mean displacement of the particles after the n th step can be found by summing over the particle index i and dividing by N :

$$\langle x(n) \rangle = \frac{1}{N} \sum_{i=1}^N x_i(n) \quad (34)$$

On expressing $x_i(n)$ in terms of $x_i(n - 1)$, we find:

$$\langle x(n) \rangle = \frac{1}{N} \sum_{i=1}^N [x_i(n - 1) \pm \delta] \quad (35)$$

$$= \frac{1}{N} \sum_{i=1}^N x_i(n - 1) \quad (36)$$

$$= \langle x(n - 1) \rangle \quad (37)$$

The second term in the brackets ($\pm \delta$) averages to zero, because its sign is positive for roughly half of the particles and negative for the other half. Because $\langle x(n) \rangle = \langle x(n - 1) \rangle$, the mean position does not change from step to step. If all the particles start at the origin, the mean position remains at the origin. The spreading of particles is symmetric about the origin.

How much do the particles spread? A convenient measure of spreading is the root-mean-square displacement $\langle x^2(n) \rangle^{1/2}$. Here, we average the square of the displacement rather than the displacement itself. To find $\langle x^2(n) \rangle$, we write $x_i(n)$ in terms of $x_i(n - 1)$ and take the square:

$$x_i^2(n) = x_i^2(n - 1) \pm 2\delta x_i(n - 1) + \delta^2 \quad (38)$$

Then we compute the mean,

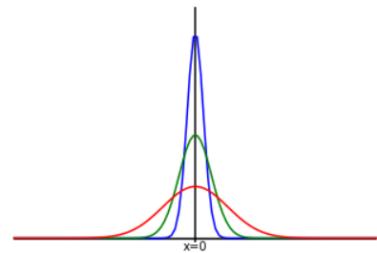


Figure 53: **Gaussian Distributions.**
The probability of finding parti-
cles at different points x at times
 $t = 1, 4$, and 16 sec. The standard devi-
ations increase with the square root of
time.

$$\langle x^2(n) \rangle = \frac{1}{N} \sum_{i=1}^N x_i^2(n) \quad (39)$$

which is

$$\langle x^2(n) \rangle = \frac{1}{N} \sum_{i=1}^N [x_i^2(n-1) \pm 2\delta x_i(n-1) + \delta^2] \quad (40)$$

$$= \langle x^2(n-1) \rangle + \delta^2 \quad (41)$$

As before, the second term in the brackets averages to zero. Since $x_i(0) = 0$ for all particles i , $\langle x^2(0) \rangle = 0$. Thus, $\langle x^2(1) \rangle = \delta^2$, $\langle x^2(2) \rangle = 2\delta^2$, $\langle x^2(3) \rangle = 3\delta^2$, ..., and $\langle x^2(n) \rangle = n\delta^2$. We conclude that the mean-square-displacement increases with the step number n , the root-mean-square displacement with the square-root of n . The particles execute n steps in a time $t = n\tau$; n is proportional to t . It follows that the mean-square displacement is proportional to t , the root-mean-square displacement is proportional to the square-root of t .

Note that $n = t/\tau$, so that:

$$\langle x^2(t) \rangle = (t/\tau)\delta^2 = (\delta^2/\tau)t \quad (42)$$

where we write $x(t)$ rather than $x(n)$ to denote the fact that x is being considered as a function of t . For convenience, we define a diffusion coefficient, $D = \delta^2/2\tau$, in units cm^2/sec . This gives us

$$\langle x^2 \rangle = 2Dt \quad (43)$$

and

$$\langle x^2 \rangle^{1/2} = (2Dt)^{1/2} \quad (44)$$

The diffusion coefficient, D , characterizes the migration of particles of a given kind in a given medium at a given temperature.

For a small molecule in water at room temperature, $D = 10^{-5}\text{cm}^2/\text{sec}$. A particle with a diffusion coefficient of this order will diffuse a distance $x = 10^{-4}\text{cm}$ in a time $t \approx x^2/2D = 5 \times 10^{-4}\text{sec}$. It diffuses a distance $x = 1\text{cm}$ in a time $t \approx 5 \times 10^4\text{sec}$ or 14 hours.

Two- and three-dimensional random walk

Let the motions in the x , y , and z directions be statistically independent. If $\langle x^2 \rangle = 2Dt$, then $\langle y^2 \rangle = 2Dt$ and $\langle z^2 \rangle = 2Dt$. In two

dimensions, the square of the distance from the origin to the point (x, y) is $r^2 = x^2 + y^2$; therefore

$$\langle r^2 \rangle = 4Dt \quad (45)$$

In three dimensions, $r^2 = x^2 + y^2 + z^2$, and

$$\langle r^2 \rangle = 6Dt \quad (46)$$

George Polya discovered the **Recurrence Theorem** about one-, two-, and three-dimensional random walks. In one and two-dimensions, a random walker on a lattice is guaranteed to eventually revisit every point in space when given infinite time, no matter how far that point might be. In one and two-dimensions, a random walker that leaves the origin is guaranteed to eventually revisit the origin.

The case of three-dimensions is qualitatively different. In three-dimensions or more, a random walker *can* escape. A particle that leaves the origin, only has a finite probability of returning to that origin (and thus also a finite probability of escaping to infinity and never returning.) Even given infinite time, every particle has a finite probability of reaching each point in space.

The binomial distribution

We have learned that particles undergoing free diffusion have a zero mean displacement and a root-mean-square displacement that is proportional to the square root of time. What can we say about the shape of the distribution of particles? To find out, we have to work out the probabilities that the particles step different distances to the right or to the left. It is convenient to generalize the one-dimensional random walk and suppose that a particle steps to the right with a probability p and to the left with a probability q . The probability that such a particle steps exactly k times to the right in n trials is given by the binomial distribution

$$P(k; n, p) = \frac{n!}{k!(n-k)!} p^k q^{n-k} \quad (47)$$

The displacement of the particles in n trials, $x(n)$, is equal to the number of steps to the right minus the number of steps to the left times the step length, δ :

$$x(n) = [k - (n - k)] \delta = (2k - n)\delta \quad (48)$$

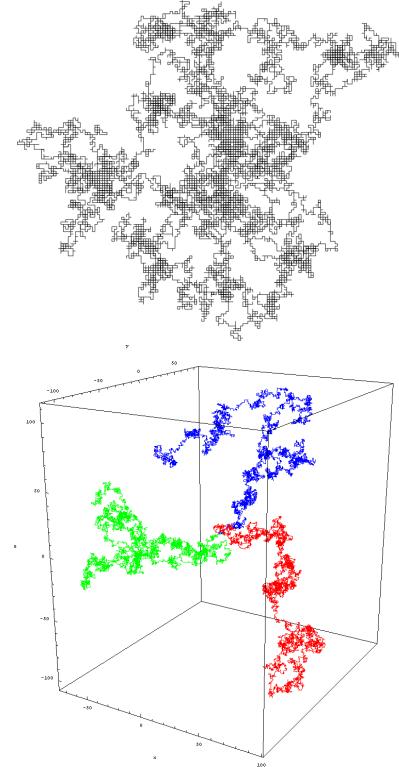


Figure 54: **Random walk simulations.** Top. One random walker in two dimensions. Bottom. Three random walkers in three dimensions.

Since we know the distribution of k , we know the distribution of x .
The two distributions have the same shape.

The mean displacement of the particle is:

$$\langle x(n) \rangle = (2 \langle k \rangle - n) \delta \quad (49)$$

where

$$\langle k \rangle = np. \quad (50)$$

The mean-square displacement is

$$\langle x^2(n) \rangle = \langle [2k - n] \delta]^2 \rangle \quad (51)$$

$$= (4 \langle k^2 \rangle - 4 \langle k \rangle n + n^2) \delta^2 \quad (52)$$

where

$$\langle k^2 \rangle = (np)^2 + npq \quad (53)$$

For the case $p = q = 1/2$, $\langle x(n) \rangle = 0$ and $\langle x^2(n) \rangle = n \delta^2$ as expected.

The Gaussian Distribution

When n and np are both very large, the binomial distribution, $P(k; n, p)$ is equivalent to:

$$P(k)dk = \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-(k-\mu)^2/2\sigma^2} dk \quad (54)$$

where $P(k)dk$ is the probability of finding a value of k between k and $k + dk$, $\mu = \langle k \rangle = np$, and $\sigma^2 = npq$. This is the Gaussian or normal distribution. By substituting $x = (2k - n)\delta$, $dx = 2\delta dk$, $p = q = 1/2$, $t = n/\tau$, and $D = \delta^2/2\tau$,

$$P(x)dx = \frac{1}{(4\pi Dt)^{1/2}} e^{-x^2/4Dt} dx \quad (55)$$

where $P(x)dx$ is the probability of finding a particle between x and $x + dx$. The variance of this distribution is $\sigma_x^2 = 2Dt$. Its standard deviation is $\sigma_x = (2Dt)^{1/2}$.

Fick's Laws

Suppose we know the number of particles at each point along the x axis at time t . How many particles will move across unit area in unit time from the point x to the point $x + \delta$? What is the net flux in the x direction, J_x ? At time $t + \tau$, after the next step, half the particles at x will have stepped across the dashed line from left to right and half the particles at $x + \delta$ will have stepped across the dashed line from right to left.

The net number crossing to the right will be:

$$-\frac{1}{2} [N(x + \delta) - N(x)] \quad (56)$$

To obtain the net flux, we divide by the area normal to the x axis and by the time interval, τ ,

$$J_x = -\frac{1}{2} [N(x + \delta) - N(x)] / A\tau \quad (57)$$

Multiplying by δ^2/δ^2 and rearranging, we obtain:

$$J_x = -\frac{\delta^2}{2\tau} \frac{1}{\delta} \left[\frac{N(x + \delta)}{A\delta} - \frac{N(x)}{A\delta} \right] \quad (58)$$

The quantity $\delta^2/2\tau$ is the diffusion coefficient, D . $N(x + \delta)/A\delta$ is the number of particles per unit volume at the point $x + \delta$, i.e., the concentration $C(x + \delta)$. $N(x)/A\delta$ is the concentration $C(x)$. Therefore,

$$J_x = -D \frac{1}{\delta} [C(x + \delta) - C(x)] \quad (59)$$

But δ is very small. In the limit $\delta \rightarrow 0$,

$$J_x = -D \frac{\partial C}{\partial x} \quad (60)$$

This is Fick's first equation. It states that the net flux (at x and t) is proportional to the slope of the concentration function (at x and t). The constant of proportionality is D . If the particles are uniformly distributed, the slope is 0, i.e., $\partial C / \partial x = 0$ and $J_x = 0$. If the slope is constant, i.e., if $\partial C / \partial x$ is constant, J_x is constant. This occurs when C is a linear function of x .



Figure 55: At time t there are $N(x)$ particles at position x , $N(x + \delta)$ particles at $x + \delta$. At time $t + \tau$, half of each set will have stepped to the right and half to the left.

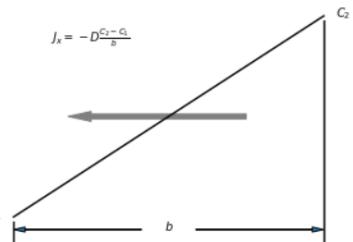


Figure 56: The flux due to a linear concentration gradient $(C_2 - C_1)/b$. There is net movement of particles from right to left solely because there are more particles at the right than at the left.

FICK'S SECOND EQUATION follows from the first, provided that the total number of particles is conserved, i.e., that the particles are neither created nor destroyed. Consider the box shown in Fig. 57. In a period of time τ , $J_x(x)A\tau$ particles will enter from the left and $J_x(x + \delta)A\tau$ particles will leave from the right. The volume of the box is $A\delta$. If particles are neither created nor destroyed, the number of particles per unit volume in the box must increase at the rate

$$\frac{1}{\tau} [C(t + \tau) - C(t)] = -\frac{1}{\tau} [J_x(x + \delta) - J_x(x)] \frac{A\tau}{A\delta} \quad (61)$$

$$= -\frac{1}{\delta} [J_x(x + \delta) - J_x(x)] \quad (62)$$

In the limit $\tau \rightarrow 0$ and $\delta \rightarrow 0$, this means that

$$\frac{\partial C}{\partial t} = -\frac{\partial J_x}{\partial x} \quad (63)$$

or, when we combine Fick's First Law:

$$\frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial x^2} \quad (64)$$

Fick's second equation states that the time rate of change in concentration (at x and t) is proportional to the curvature of the concentration function (at x and t); the constant of proportionality is D . The diffusion equation tells us how a nonuniform distribution of particles will redistribute itself over time. If we know the initial distribution and other boundary conditions, we can figure out all later distributions.

In three dimensions, we have $J_x = -D\partial C/\partial x$, $J_y = -D\partial C/\partial y$, and $J_z = -D\partial C/\partial z$. These are components of a flux vector:

$$\mathbf{J} = -D \nabla C \quad (65)$$

The concentration changes with time as

$$\frac{\partial C}{\partial t} = D \nabla^2 C \quad (66)$$

where ∇^2 is the three dimensional Laplacian.

If the problem is spherically symmetric, the flux is radial,

$$J_r = -D\partial C/\partial r \quad (67)$$

and

$$\frac{\partial C}{\partial t} = D \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial C}{\partial r} \right) \quad (68)$$

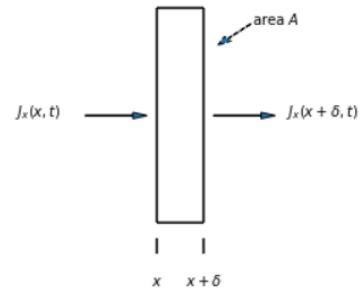


Figure 57: Fluxes through the faces of a thin box extending from position x to position $x + \delta$. The area of each face is A . The faces are normal to the x axis.

Diffusion to a spherical adsorber

Consider a spherical adsorber of radius a in an infinite medium. Every particle reaching the surface of the sphere is gobbled up, so the concentration at $r = a$ is 0. The concentration at $r = \infty$ is C_0 . With these boundary conditions, the diffusion equation has the solution:

$$C(r) = C_0 \left(1 - \frac{a}{r}\right) \quad (69)$$

The flux is

$$J_r = -DC_0 \frac{a}{r^2} \quad (70)$$

The net migration of molecules is radially inward, as shown by the dashed arrows in Fig. 58. The particles are adsorbed by the sphere at a rate equal to the area, $4\pi a^2$ times the inward flux $-J_r(a)$:

$$I = 4\pi DaC_0 \quad (71)$$

We will refer to this adsorption rate, I , as a diffusion current. Note that this current is proportional not to the area of the sphere, but its radius. As the radius increases, the area increases as a^2 but the concentration gradient decreases as $1/a$.

Probability of capture

Suppose a particle is released near a spherical adsorber of radius a at a point $r = b > a$? What is the probability that the particle will be adsorbed at $r = a$ rather than wander away for good?

Consider a spherical shell source of radius b between a spherical adsorber of radius a and a spherical shell adsorber of radius c as shown in **Figure 7**. The concentration rises from 0 at $r = a$ to a maximum value C_m at $r = b$ and then falls again to 0 at $r = c$. With these boundary conditions, the diffusion equation has the solution:

$$C(r) = \begin{cases} \frac{C_m}{1-a/b} \left(1 - \frac{a}{r}\right) & \text{if } a \leq r \leq b, \\ \frac{C_m}{c/b-1} \left(\frac{c}{r} - 1\right) & \text{if } b \leq r \leq c \end{cases} \quad (72)$$

The radial flux is

$$J_r(r) = \begin{cases} \frac{-DC_m}{1-a/b} \frac{a}{r^2} & \text{if } a \leq r \leq b, \\ \frac{DC_m}{c/b-1} \frac{c}{r^2} & \text{if } b \leq r \leq c \end{cases} \quad (73)$$

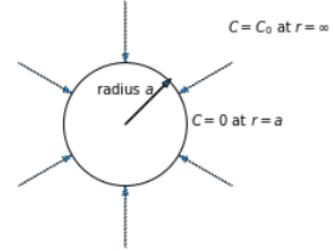


Figure 58: A spherical adsorber of radius a in an infinite medium containing particles at an initial concentration C_0 . The dashed arrows are lines of flux.

Thus, the diffusion current from the spherical shell source to the inner adsorber is

$$I_{in} = 4\pi DC_m \frac{a}{1 - a/b} \quad (74)$$

and the diffusion current from the spherical shell source to the outer adsorber is

$$I_{out} = 4\pi DC_m \frac{c}{c/b - 1} \quad (75)$$

The ratio

$$\frac{I_{in}}{I_{in} + I_{out}} = \frac{a(c - b)}{b(c - a)} \quad (76)$$

is the probability that a particle released at $r = b$ will be adsorbed at $r = a$. In the limit $c \rightarrow \infty$, this probability is just a/b . This is the probability of capture for the sphere of radius a immersed in an infinite medium. As b increases, this probability decreases as $1/b$.

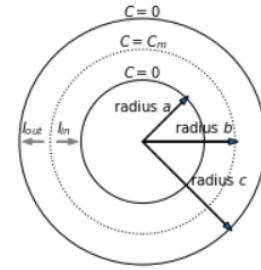


Figure 59: A spherical shell source, radius b , between a spherical adsorber of radius a and a spherical shell adsorber of radius c . Particles released at $r = b$ move inward and are adsorbed at $r = a$ at rate I_{in} or move outward and are adsorbed at $r = c$ at rate I_{out} . Their steady-state concentration rises from 0 at $r = a$ to C_m at $r = b$ and then falls again to 0 at $r = c$.

REFERENCES

- Download paper
- Download paper
- Download paper
-

8. THE DIFFUSION COEFFICIENT

THE MICROSCOPIC THEORY OF DIFFUSION gives us an intuitive sense of the dynamics of random walks. A key concept is that the ‘root mean square’ displacement of a particle grows as the square root of time (as opposed to the ballistic movements of particles in free space where displacement grows linearly with time in proportion to speed).

$$\langle x^2 \rangle^{1/2} = (2Dt)^{1/2} \quad (77)$$

The diffusion coefficient is rigorously defined with respect to random walks on a lattice, $D = \delta^2/2\tau$. But the ingredients of the definition of the diffusion coefficient in a lattice model, involving a step size δ and step interval τ , are difficult to relate to the physics of real particles moving in real liquids under the influence of thermal forces. Einstein discovered how to calculate the diffusion coefficient for a real particle in one of his three major papers in 1905, known as his *annus mirabilis* (the other two papers were about the photoelectric effect and special relativity).

Before 1905, it was not known how to relate the Brownian movements and fluid dynamics that were observed under the microscope with statistical mechanics that had been developed by Boltzmann with the mathematics of lattice random walks. Einstein discovered the connections between these previously separated areas of mathematics and physics. In doing this, he first derived what is now known as the Einstein-Smoluchowski relation:

$$D = \frac{k_B T}{f_r} \quad (78)$$

Here, $k_B T$ is from statistical mechanics, the product of Boltzmann’s constant and temperature. The denominator is the *frictional drag coefficient*. If an object is under the influence of a steady force in a liquid, it will eventually reach terminal velocity. The frictional drag coefficient is the proportionality between terminal velocity and applied force.

To CALCULATE THE FRICTIONAL DRAG COEFFICIENT, we need to invoke fluid mechanics. Fluid mechanics is governed by a particularly complicated set of dynamical equations called the Navier-Stokes Equations. Even for an incompressible fluid for which both fluid density and fluid viscosity are constants throughout the volume, the Navier-Stokes Equations are daunting:

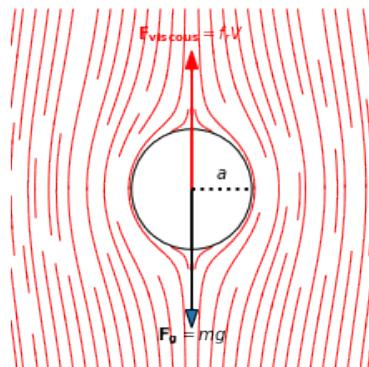


Figure 60: **Viscous drag.** If an object of mass m falls under the influence of gravity in a fluid with low density and high viscosity, it will experience both a gravitational force (mg) that is balanced by viscous resistance, and achieve a terminal velocity V . The proportionality between applied force and terminal velocity is the frictional drag coefficient, f_r that depends on the size and shape of the object and fluid viscosity. Calculating the frictional drag coefficient requires solving Navier-Stokes equations for fluid movements outside the cell. At low Reynolds numbers, when viscous forces are much larger than inertial forces, the equations of fluid dynamics are easier to solve. The frictional drag coefficient of a sphere of radius a at low Reynolds number is $f_r = 6\pi\eta a$, where η is fluid viscosity. For water, $\eta = 0.01 \text{ g cm}^{-1} \text{ s}^{-1}$

$$\rho_0 \left(\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} \right) = \eta \nabla^2 \mathbf{u} - \nabla p \quad (79)$$

The Navier-Stokes equations are less daunting if one breaks them down into components. The inertial terms, forces due to mass and acceleration discovered by Newton, are on the left. The terms on the right are forces due to viscous shear and pressure gradients. Solving the Navier-Stokes equations means calculating the scalar field of pressure at all points in space and time $p(x, y, z, t)$ and the vector field of fluid velocity at all points in space and time $\mathbf{u}(x, y, z, t)$. They can be solved, in a few cases, with substantial effort in solving differential equations or numerical calculation. To do this, one often solves a boundary value problem, setting the movements of the surface of an object, and thereby solving for the movements of the fluid throughout the surrounding medium. Once you know all about the movements of the fluid, you can calculate the forces on the object that are consistent with its surface movements.

The magnitude of forces due to viscous shear and inertia can be estimated based on the scales of size (say of order of magnitude a) and velocity (order of magnitude v), as well as the viscosity (η) and density (ρ) of the surrounding fluid. The ratio between inertial forces and viscous forces in a given situation is a dimensionless number called the *Reynolds number*:

$$Re = \frac{\rho v a}{\eta} \quad (80)$$

For bacteria, Reynolds numbers are very small. Inertial forces are negligible in comparison to forces due to viscous shear. When inertial terms are excluded from the Navier-Stokes equations, they become simpler:

$$\eta \nabla^2 \mathbf{u} - \nabla p = 0 \quad (81)$$

Solving these equations can still be involved, but their simple form allows intuitive inferences about the physics of fluid dynamics at low Reynolds number. The first thing to remember is that these are ‘force balance’ equations. The units of the Stokes equations (what the Navier-Stokes equations are called at low Reynolds numbers) are force (with dimensions of mass \times length \times time $^{-2}$). So, forces are proportional to both velocities and viscosities. How much force is needed to move a sphere of radius a at velocity v ? Given that the units of viscosity are [$\text{g cm}^{-1} \text{ s}^{-1}$] and the units of velocity are [cm s^{-1}], the only way to get the dimensions of force in such a way that

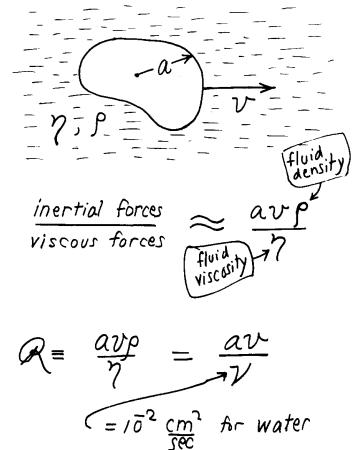


Figure 61: **The Reynolds number.**
An object moves through a fluid with velocity v . It has dimension a . In Stoke's law, the object is a sphere, but it can be anything. η and ρ are the viscosity and density of the fluid. The ratio of the inertial forces to the viscous forces, as Osborne Reynolds pointed out, is $\rho v a / \eta$. For water, $\eta \approx 0.01 \text{ g cm}^{-1} \text{ s}^{-1}$ and $\rho \approx 1 \text{ g cm}^{-3}$.

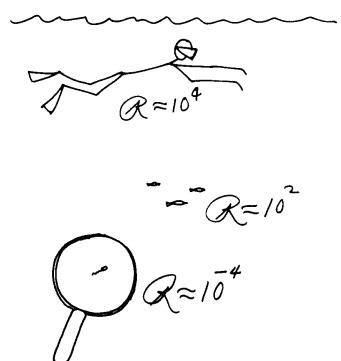


Figure 62: **Viscous drag.** We swim at high Reynolds numbers. We accelerate water behind us, which accelerates us forward. Bacteria cannot accelerate the fluid around them. They swim at low Reynolds number. They swim by taking advantage of viscous shear.

force is proportional to velocity and viscosity is for the frictional drag coefficient of the sphere to be proportional to its radius:

$$f_r \sim \eta a \quad (82)$$

If you solve the differential equations, you will find that the proportionality constant is 6π . The force needed to tow a sphere at radius v is thus:

$$F = 6\pi\eta av \quad (83)$$

Earlier, we considered lysozyme, a small protein of typical size. We can now calculate its diffusion coefficient by estimating its frictional drag coefficient. The order of magnitude for the size of a small biological molecule is ~ 1 nm (Fig. 63). At room temperature, the diffusion coefficient of a small biological molecule can be estimated using the Einstein relation: $D \sim 10 \text{ cm}^2 \text{ s}^{-1}$.

HOW LONG DOES IT TAKE for a molecule to diffuse the length of a bacterial cell? Mean square displacements grow linearly with time. We are asking for the time required for diffusion to occur over a length $x = 0.0001 \text{ cm}$, $t \sim \frac{D}{x^2}$, which is $t \sim 1 \text{ millisecond}$.

The efficiency of diffusion on the micrometer-scale of the bacterial cell can benefit the cell. One benefit is that rapid diffusion allows the cell to make about 1000 independent measurements of the ambient concentration of attractants in the one second of a typical run. Thus, for the micrometer-sized bacteria, the signal-to-noise of measuring ambient chemical concentration by integrating measurements over one second is ~ 30 -fold better than measuring chemical concentration based on an instantaneous snapshot of local molecules. Another benefit is that it allows signaling within the cell to be fast and efficient. Intracellular signaling molecules – the computational machinery that interacts with chemoreceptors that count ambient molecules – can move from end-to-end of the bacterial cell in a millisecond. The signal transduction from chemoreceptors to flagellar motors can be *fast*.

But the efficiency of diffusion also directly affects how the bacteria might make measurements of ambient molecules. For example, it makes it difficult for bacteria to make spatial comparisons. Can the bacterial cell compare the binding of molecules on one end to the binding of molecules on its other end? It only takes a millisecond for a molecule that the bacteria might have counted on its ‘head’ to diffuse to its ‘tail’ (and vice-versa). It would be difficult to make spa-

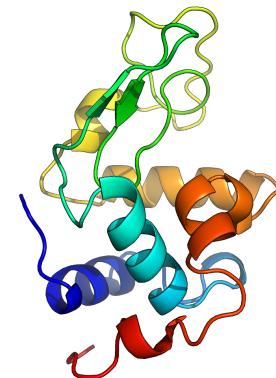


Figure 63: **Lysozyme.** An antimicrobial enzyme produced by animals that forms part of the innate immune system.

tial comparisons on the size scale of micrometers because diffusion rapidly erases spatial gradients on this scale within a millisecond.

For the bacterial cell to count ‘new’ molecules by the end of each run, it has to out-swim the molecules it counted at the beginning of the run (Fig. 64). Because the displacement of the swimming bacteria grows linearly with time, it eventually outruns the diffusive displacement of ambient molecules that grows as the square root of time. Bacteria outrun the molecules that it is trying to count in ~ 1 second, which partly explains why Berg discovered that runs last ~ 1 second on average (Fig. 65).

THE DIFFUSION COEFFICIENT was first derived by Einstein. But before we present Einstein’s derivation, we’ll go through a later derivation by Langevin in 1908 that is conceptually interesting in its own way. Langevin approached the problem from the point of view of Newton’s equations of motion. First, consider the 1-dimensional motion of a particle immersed in fluid, with nothing acting upon it except inertial and viscous forces. If there are no other forces, inertial and viscous forces will sum to zero.

$$F = m \frac{d^2x}{dt^2} + f_r \frac{dx}{dt} = 0 \quad (84)$$

Substituting $v = dx/dt$ creates a simple first-order differential equation for velocity:

$$\frac{dv}{dt} + \frac{f_r}{m} v = 0 \quad (85)$$

If the initial velocity is v_0 , this differential equation is easily solved:

$$v(t) = v_0 \exp\left(\frac{-f_r t}{m}\right) \quad (86)$$

Velocity will exponentially decay with a time constant that depends on particle mass and the frictional drag coefficient. For molecules, the time constant of exponential decay is on the order of picoseconds, far shorter than any reasonable observation time. Viscous damping rapidly quenches all inertial movements. And yet, particles move incessantly as they exhibit Brownian movement. To explain Brownian movement, we need to add another force to our force-balance equation:

$$m \frac{d^2x}{dt^2} + f_r \frac{dx}{dt} = X(t) \quad (87)$$

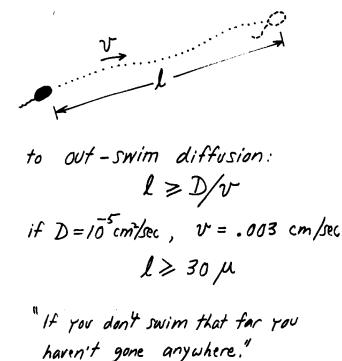


Figure 64: **Greener pastures.** For a bacterial cell to tell whether it has found a greener pasture, it has to move to that pasture. The cell has to ‘out-swim’ diffusion. This happens when it travels a distance $l \sim D/v$.

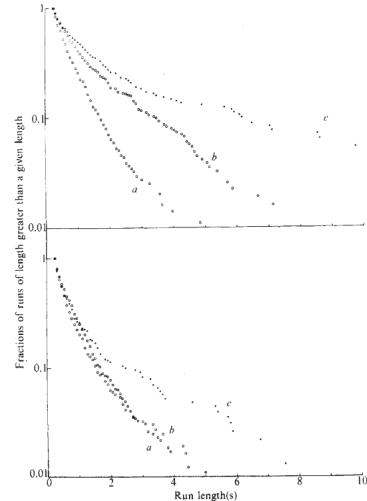


Figure 65: Runs are longer in gradients, but runs up gradients are even longer than runs down gradients. This difference is dramatic when run-length distributions are plotted for different attractants, serine and aspartate. A log-linear plot reveals that run length distributions are exponential in time, as one would expect from a Poisson interval distribution. Mean run lengths are ~ 1 sec.

where $X(t)$ is a random force due to thermal movements. Multiplying the new force balance equation by x and rearranging gives:

$$\frac{m}{2} \frac{d^2x^2}{dt^2} - mv^2 = -\frac{f_r}{2} \frac{dx^2}{dt} + xX(t) \quad (88)$$

But the equipartition theorem tells us the average kinetic energy of the particle.

$$\langle mv^2 \rangle = k_B T \quad (89)$$

Moreover, because the random thermal force is indifferently positive or negative, and because the diffusing particle is just as likely to drift towards positive and negative values of x , $\langle xX(t) \rangle = 0$. Taking the ensemble average of the force balance equation (an ensemble average is an idealization consisting of a large number of virtual copies of the system, each representing a possible state that the system might be in), we get:

$$\frac{m}{2} \frac{d^2 \langle x^2 \rangle}{dt^2} - k_B T = -\frac{f_r}{2} \frac{d \langle x^2 \rangle}{dt} \quad (90)$$

To make this differential equation look less daunting, substitute $\frac{d \langle x^2 \rangle}{dt} = z$:

$$\frac{m}{2} \frac{dz}{dt} + \frac{f_r}{2} z = k_B T \quad (91)$$

The general solution of this “first-order inhomogeneous differential equation” is:

$$z = \frac{2k_B T}{f_r} + C e^{-\frac{f_r}{m} t} \quad (92)$$

Note that the second term exponentially decays with time. For small molecules, the time constant of this exponential decay, as we have already argued, is tiny. This term can be neglected, giving:

$$\frac{d \langle x^2 \rangle}{dt} = \frac{2k_B T}{f_r} \quad (93)$$

Hence, $\langle x^2 \rangle = 2 \frac{k_B T}{f_r} t$. The mean square displacement grows linearly with time. The proportionality constant can be identified as the diffusion coefficient, $\langle x^2 \rangle = 2Dt$, which means that:

$$D = \frac{k_B T}{f_r} \quad (94)$$

HOW DID EINSTEIN DERIVE THE EINSTEIN RELATION? He started with Fick's Law, which states that any concentration gradient of diffusing particles will create fluxes which try to diminish those gradients. He then considered the exponential atmosphere that we considered earlier. In a gravitational field at non-zero temperatures, particles are unevenly distributed in the vertical dimension, z . Particle density (and hence particle concentration) is exponentially distributed in the vertical density because the potential energy of each gas particle with mass m is in simple proportion to altitude, mgz .

$$\rho(z) = \rho_0 e^{-\frac{mgz}{k_B T}} \quad (95)$$

Fick's First Law tells us that there must be an upward diffusive flux because of Brownian movements:

$$J_z^{diff} = D \times \frac{mg}{k_B T} \times \rho_0 e^{-\frac{mgz}{k_B T}} \quad (96)$$

But all gas particles are also constantly falling because of Earth's gravity. The sedimentation velocity of each particle is the ratio between gravitational force (mg) and the frictional drag coefficient (f_r). The downward flux due to sedimentation is the product of particle density and sedimentation velocity:

$$J_z^{sed} = -\frac{mg}{f_r} \times \rho_0 e^{-\frac{mgz}{k_B T}} \quad (97)$$

At steady-state, the downward flux due to sedimentation must balance the upward flux due to diffusion, $J_z^{diff} + J_z^{sed} = 0$, which gives:

$$D \times \frac{mg}{k_B T} \times \rho_0 e^{-\frac{mgz}{k_B T}} - \frac{mg}{f_r} \times \rho_0 e^{-\frac{mgz}{k_B T}} = 0 \quad (98)$$

After canceling terms, we are left with:

$$\frac{D}{k_B T} - \frac{1}{f_r} = 0 \quad (99)$$

which gives us the *original* derivation of the Einstein relation:

$$D = \frac{k_B T}{f_r} \quad (100)$$

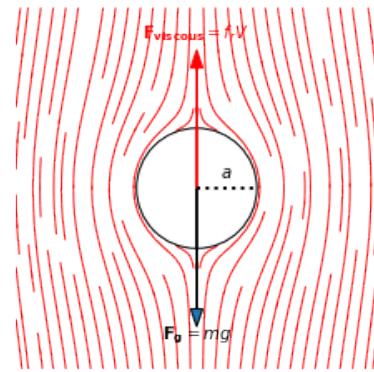


Figure 66: Sedimentation velocity.

REFERENCES

- Download paper
- Download paper
- Download paper

9. COUNTING MOLECULES

THE PERFECT COUNTER AND THE PERFECT ADSORBER are good starting points to better understand the problem of counting molecules in olfaction. The perfect counter is somehow able to immediately count all the molecules in a small volume near the cell. The perfect adsorber is able to capture all the molecules in its vicinity by diffusion. The rate at which the adsorber captures molecules can be calculated using Fick's Laws.

$$I = 4\pi DaC_0 \quad (101)$$

where D is the diffusion coefficient, a is the sphere radius, and C_0 is the steady concentration at infinity. If the cell were able to count the rate of molecule adsorption, it could estimate the far-field molecule concentration.

But actual cells are neither perfect counters nor adsorbers. *E. coli* counts molecules by smell. Odorant molecules bind transiently to receptors on the surface, and a signal is propagated through the cell membrane to a sensory transduction network within the cell. To be a perfect adsorber, the cell would have to be completely covered by molecular receptors, and these receptors would have to permanently capture each odorant molecule. This is not what happens. Most cells, whether *E. coli* trying to do chemotaxis or any cell in an organism that is trying to count extracellular signaling molecules, do so with a distribution of surface-bound receptors. How many receptors does a cell need to accurately assess its environment? How should these receptors be distributed? To answer these questions, we need a new model, the “patchy” cell.

Each molecule that a patchy cell tries to count undergoes a random walk. When the molecule is near the cell, it will randomly bump along the surface of the cell and sometimes bump into a receptor. Not every molecule that is near the cell will find the cell. It is possible, in three dimensions, for a molecule undergoing a random walk to *never* bump into the cell, and instead wander away to infinity without ever being counted. This is due to George Polya's **Recurrence Theorem** about one-, two-, and three-dimensional random walks. For molecules and adsorbers in one or two dimensions, every molecule will eventually find the adsorber.

Thinking about the random movements of a molecule near the cell surface can give us intuition about how effectively a small number of receptors can detect that molecule. Say each receptor has a radius s . A molecule that starts its random walk a distance s from the cell

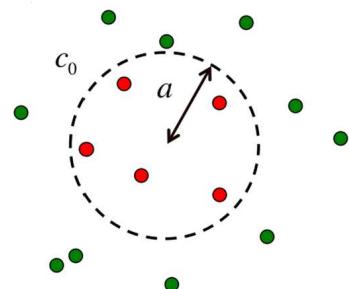


Figure 67: **The perfect monitor.** The perfect monitor is permeable to ligand molecules and estimates the concentration c_0 by counting the molecules in its volume during time T .

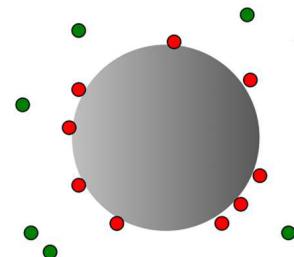


Figure 68: **The perfect adsorber.** The perfect adsorber estimates the ligand concentration from the number of molecules incident on its surface during time T .

will have a high probability of making its way to the surface. This probability is the “probability of capture” that we calculated earlier using concentric spheres:

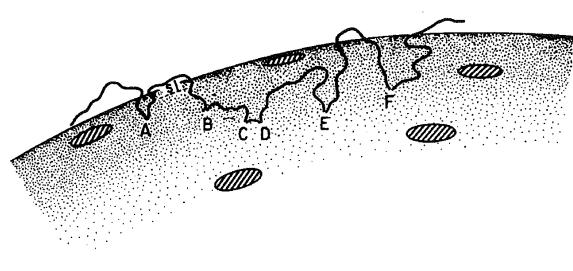


Figure 69: Random walk near a patchy sphere Path of a diffusing molecule that touches the surface of a cell at a sequence of points A, B, ... F. The cell has radius a . The receptor patches, shown shaded, are of radius s . A and B constitute independent tries at hitting a patch, but C and D do not. Note between A and B the excursion of distances perpendicular to the surface of the sphere.

$$P_s = \frac{a}{a+s} \quad (102)$$

After a molecule that starts at radius $a + s$ gets to the cell at radius a , it will eventually wander back to its starting radius. By the time it returns to its starting radius, it would have roughly explored an s -sized region of the cell surface. If the molecule does not encounter a receptor in its first exploration, it would have the chance to explore a new neighboring region of space on the cell surface that may or may not have a receptor. Thus, each time that the molecule diffuses to the cell and back from a radius $a + s$ will represent a new exploration of a new s -sized portion of cell surface. At some point, the molecule starting at $a + s$ will not actually get to the surface, and instead of exploring for receptors will wander away to infinity, never to be seen again. The probability of failure to explore in a given excursion is $1 - P_s$. How many times will a molecule explore the cell surface before diffusing away to infinity? How many s -sized regions on the cell surface will be encountered? The more regions on the cell surface that are encountered, the more likelihood that the molecule will be recognized by cellular receptors. How fast does this likelihood increase with the number of receptors?

The probability that a molecule at $r = a + s$ executes exactly n excursions to the surface separated by reappearances at $r = a + s$ before escaping to infinity is $P_s^n(1 - P_s)$. The average number of excursions is thus:

$$\langle n \rangle = \sum_{n=0}^{\infty} n P_s^n (1 - P_s) = \frac{P_s}{1 - P_s} = \frac{a}{s} \quad (103)$$

For a molecule exploring nanometer-sized patches on a micrometer sized cell starting a nanometer away from the surface, the average

number of excursions is ~ 1000 . The molecule has 1000 independent chances to find a receptor and be counted.

What is the probability that such molecule is eventually counted? To estimate this probability, start by assuming that the receptors are fully adsorbing. If a molecule is adsorbed, it can never escape to infinity. What is the total chance of escaping? In each excursion, the molecule must (a) travel to the surface with probability P_s and (b) not get absorbed with probability $\beta = 1 - (Ns^2/4a^2)$. The probability of not getting absorbed, β , is the fraction of the surface that lacks receptors. N receptors, each with area πs^2 , cover a total surface area of $N\pi s^2$ on a sphere with surface area $4\pi a^2$. A molecule can escape after 0, 1, 2, or more excursions. It does not matter how many excursions, n , are made, only that the molecule eventually escapes. The total escape probability is thus:

$$P_{esc} = \sum_{n=0}^{\infty} \beta^n P_s^n (1 - P_s) = \frac{1 - P_s}{1 - \beta P_s} \quad (104)$$

$$= \frac{4a}{4a + Ns} \quad (105)$$

We can now estimate the fraction of all molecules that arrive at the surface that are counted by receptors before diffusing away to infinity, $1 - P_{esc}$. Let J_{max} be the maximum rate of counting molecules by a sphere that is completely covered by receptors. The rate at which molecules are counted by the patchy sphere is a fraction of this maximum rate:

$$\frac{J}{J_{max}} = \frac{Ns}{4a + Ns}$$

IS THERE ANOTHER WAY TO CALCULATE THE COUNTING EFFICIENCY OF A PATCHY SPHERE? Although modeling the trajectory of a diffusing particle near the surface has intuitive appeal, we made various simplifying assumptions to complete this “back of the envelope” calculation. Is there another way to calculate diffusive fluxes to differently shaped adsorbers – spheres, disc-like adsorbers, patchy spheres – in a more rigorous manner?

Steady-state flux to adsorbing objects, where one fixes the far-field concentration of the diffusing particles to c_∞ , is calculated by solving the time-independent diffusion equation for $c(x, y, z)$ (i.e., $\partial c / \partial t = 0$):

$$\nabla^2 c = 0$$

The flux of particles is a vector field based on this concentration field:

$$\mathbf{J} = -D \nabla c$$

The total diffusive current entering a closed surface is calculated by integrating the flux around the surface:

$$I = \int_S \mathbf{J} \cdot d\mathbf{s}$$

So far, we have solved this problem in the one special case of a spherical adsorber, exploiting the spherical symmetry to simplify the calculation. Developing de novo solutions for other geometries would be arduous. But others have done the work for us.

Consider the equations for electrical potential and electric field in charge free space, using cgs units where these equations have a particularly simple form. Laplace's Equation gives the electrical potential in charge-free space:

$$\nabla^2 \phi = 0$$

The gradient of the electrical potential gives a simple form for the electric field in charge-free space:

$$\mathbf{E} = -\nabla \phi$$

Gauss's Law tells us that the total electric charge on any closed surface is a surface integral of the electric field:

$$Q = \frac{1}{4\pi} \int_S \mathbf{E} \cdot d\mathbf{s}$$

Now, consider the electrical field and electrical potential surrounding a simple charged conductor, setting the potential at the conductor equal to zero and the potential at infinity equal to ϕ_∞ . The linearity of the equations for electrostatics means that the far-field electrical potential will be proportional to the charge on the conductor:

$$Q = C\phi_\infty$$

where C is the capacitance of the object. In c.g.s. units, the capacitance of a sphere, for example, is just the radius of the sphere, $C = a$.

Note the side-by-side similarity between the equations for diffusive flux and electrostatics. I is analogous to Q . \mathbf{J} is analogous to \mathbf{E} . ϕ_∞ is analogous to c_∞ . The two sets of equations differ by factors of

4π and D , but are otherwise the same. Thus, solutions for electrostatics can be used as solutions for diffusive flux. One only needs to correct for the missing multiplicative factors.

The flux of diffusing particles to any adsorber will be proportional to the far-field concentration, $I \sim c_\infty$. If we know the capacitance of the same-shaped object in electrostatics, we can use this capacitance to solve for the diffusive flux to the object. Correcting for the missing multiplicative factors, the flux to this object is:

$$I = 4\pi C D c_\infty$$

where C is its electrical capacitance in cgs units.

Thus, the flux to a complete adsorbing sphere (where the capacitance is equal to its radius in cgs units, $C = a$) is what we found earlier:

$$I = 4\pi a D c_\infty$$

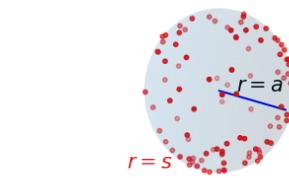


Figure 70: The patchy sphere.

The capacitance of a thin, conducting disk of radius b is $2\pi/b$. This gives the flux to both sides of the disk as $I = 8b D c_\infty$. If we model one adsorbing receptor as one one-sided disk on an insulating sphere, we can estimate the flux to a receptor:

$$I = 2\pi a D c_\infty$$

If more disks are added randomly to the surface, the total flux will start to increase linearly with the number of disks, N . But the flux can never exceed the flux to a totally adsorbing sphere. As the number of receptors increases, it must asymptotically reach the flux to a complete adsorbing sphere. What is this functional relationship?

Purcell and Berg carefully calculated the capacitance of an insulating sphere of radius a covered with a random distribution of conducting disks of radius s , where the disks were interconnected by vanishingly thin conducting wires. The capacitance of this object was:

$$C = \frac{Ns}{Ns + \pi a}$$

Thus, the flux to a patchy sphere can be calculated using its capacitance (and substituting term for maximum flux to a completely adsorbing sphere, $I_{max} = 4\pi D c_\infty a$):

$$\begin{aligned} I &= 4\pi D c_\infty \frac{Ns}{Ns + \pi a} \\ &= I_{max} \frac{Ns}{Ns + \pi a} \end{aligned}$$

How quickly does this function reach the asymptote? Again, assume that the patch radius is 1 nm and the cell radius is 1 μm . The flux reaches its half-maximal value when $N = a\pi/s$ or ~ 3100 . The receptor patches occupy about 0.1% of the cell surface. The average separation of these randomly distributed patches is about 60 times the patch radius.

The ability of cells to count molecules does not require large numbers of receptors all over their surfaces. Because diffusion gives each molecule many chances to visit the surface, a few dispersed receptors are as likely to count each molecule as well as a large number of condensed receptors. If the receptor patches were condensed into one patch that covers the same surface area on the cell as distributed receptors that reach half-maximal efficiency, counting efficiency would be reduced, from $I_{max}/2$ to $I_{max}/\sqrt{3100}$. This is good for many cells. Cells often have to count many different molecules using myriad receptors. Because a cell only has to commit a small fraction of its surface area to counting any one type of molecule with one type of receptor, it has plenty of room to devote to additional types of receptors. This is a deep insight into how cells sense their environments. Diffusion helps cells diversify the range of molecules that they can sense, by increasing the speed and efficiency of molecular detection on micrometer-length scales.

It turns out, however, that *E. coli* does not work this way. When electron microscopy was finally used to detect the distribution of receptors on the cell surface, it turned out that receptors were tightly packed into single patches! The physics is not wrong. These cells with single receptor patches accepted a lower counting efficiency of ambient molecules. It was later discovered that receptor patches provide other benefits to the cell that presumably offset the loss of counting efficiency.¹ Tightly-packed receptors can communicate with one another, amplifying chemoreceptor signals in a way that Berg and Purcell did not imagine.

There are more things in heaven and Earth, Horatio,
Than are dreamt of in your philosophy.

— *Hamlet*, Shakespeare

REFERENCES

- Download paper

OLFACTION WORKBOOK

ONE: LIFE AT LOW REYNOLDS NUMBER

1. An average *E. coli* has a diameter of $1 \mu\text{m}$ and a speed of $20 \mu\text{m}/\text{s}$. What is the Reynolds Number for a swimming *E. coli*?
2. Estimate the Reynolds Number of a swimming human. How does it compare to that of a bacterium?

TWO: DIFFUSION

Consider a spherical cell of radius a immersed in an infinite medium. The medium contains a low concentration of molecules of species B with diffusion constant D . The local concentration of B is $c(\mathbf{x}, t)$, and obeys the diffusion equation:

$$\frac{\partial c}{\partial t} = D \nabla^2 c. \quad (106)$$

1. Suppose the cell is a perfect sink for B molecules. Solving the steady state diffusion equation in spherical coordinates, show that the current of molecules to the cell is:

$$J = 4\pi a D c_\infty, \quad (107)$$

where c_∞ is the concentration of B infinitely far from the cell.

2. Let a molecule of B be released at a point in the medium a distance r from the center of the cell. The probability that the molecule will be eventually be captured by the cell is $P_c = a/r$. Derive this result by solving the time-independent diffusion equation using concentric spheres.
3. Note that this probability diminishes as $1/r$, and not as $1/r^2$. Is this surprising? Give reasons for why one might naively expect the probability to scale with $1/r^2$. Then, explain physically why this expectation is incorrect.

THREE: A 1-D RANDOM WALK

1. Simulate a one-dimensional random walk in MATLAB, or another programming language of your choice. Plot mean-squared distance as a function of number of steps.
2. Compare your result to the theoretical prediction. How does this comparison change as the number of steps increases?
3. Compare Brownian motion to ballistic motion of a particle with the same speed. How quickly does the distance traveled by Brownian motion drop off compared to the distance traveled by ballistic motion?

FOUR: SINGLE RECEPTOR BINDING

When discussing a single receptor, Berg and Purcell (1977) absorb the probability of binding α into an effective definition of s . Why is this valid? Under the assumptions made in Berg and Purcell, derive an expression for $s_{\text{effective}}$ as a function of s and α .

FOUR: SINGLE RECEPTOR BINDING

Consider a sphere of radius a in water. Due to random collisions, the sphere will rotationally diffuse. The diffusion law for rotational motion is analogous to the diffusion law for translational motion:

$$\langle \theta^2 \rangle = 2D_r t. \quad (108)$$

1. What are the units of the rotational diffusion coefficient D_r ? Use the Einstein relation to write down a formula for D_r , given that the rotational frictional drag coefficient for a sphere is $f_r = 8\pi\eta R^3$.
2. Estimate the time it takes for an *E. coli* to diffuse through an angle of 1 radian. Assuming an *E. coli* has a diameter of 1 μm and a speed of 20 $\mu\text{m}/\text{s}$, what is the distance traveled by the cell during this time?
3. What are the consequences of rotational diffusion on the navigational problem faced by *E.coli*?

FIVE: ROTATIONAL DIFFUSION

Consider a sphere of radius a in water. Due to random collisions, the sphere will rotationally diffuse. The diffusion law for rotational motion is analogous to the diffusion law for translational motion:

$$\langle \theta^2 \rangle = 2D_r t. \quad (109)$$

1. What are the units of the rotational diffusion coefficient D_r ? Use the Einstein relation to write down a formula for D_r , given that the rotational frictional drag coefficient for a sphere is $f_r = 8\pi\eta R^3$.
2. Estimate the time it takes for an *E. coli* to diffuse through an angle of 1 radian. Assuming an *E. coli* has a diameter of 1 μm and a speed of 20 $\mu\text{m}/\text{s}$, what is the distance traveled by the cell during this time?
3. What are the consequences of rotational diffusion on the navigational problem faced by *E.coli*?
- 4.

SIX: CURRENT OF A PATCHY CELL

In lecture, we found that a spherical cell of radius a with N disk-shaped receptors each of radius s , the current is:

$$J = J_{\max} \frac{Ns}{Ns + \pi a}. \quad (110)$$

Take $s = 1 \text{ nm}$ and $a = 1 \mu\text{m}$.

1. At what value of N is J half maximal? Find the mean distance on the surface of the cell between receptors at this value of J , and determine the fraction of the cell surface occupied by receptors.
2. What happens to the flux if you pack all the receptors together into one patch with an area equal to the sum of all the individual receptor areas?
3. What are the pros and cons to having receptors concentrated on one area of the cell surface?

SEVEN: SELF AVOIDING RANDOM WALK

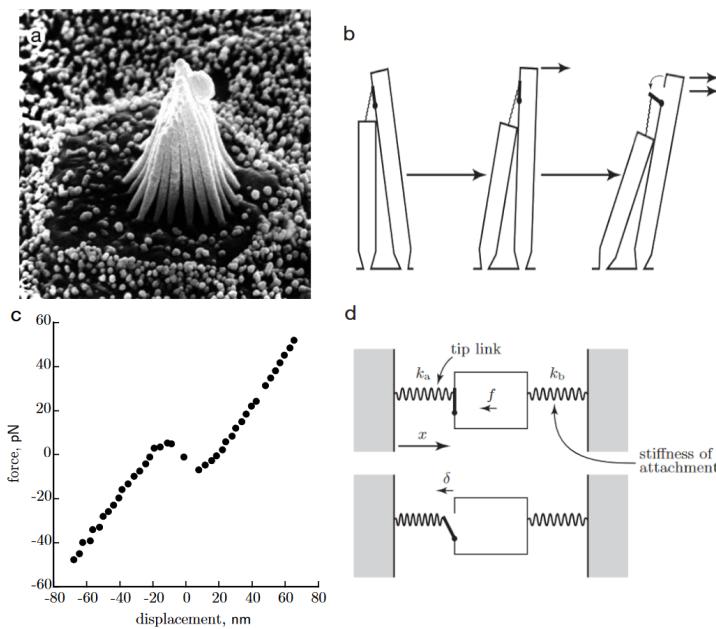
1. Simulate a random walking organism on a 2D grid which avoids traces of itself by leaving a pheromone. It always takes a grid point it never visited unless all surrounding points are visited, in which case it takes a random move as usual. Visualize 4 runs.

2. Run many instances of the above for 300 time steps. and get the ensemble mean of the squared displacement depending on time. Does it follow the linear scaling of diffusion? If so, could you get the empirical diffusion constant?
3. Now, simulate a brood of such organisms. On each simulation, you should dispatch N_{org} organisms from the same point, while the pheromones are all shared: each organism avoid traces of all other organisms. It is fine to set up an arbitrary order of who takes the step first. Visualize simulations with 2,5,10,50 and 100 organisms.
4. Simulate with brood sizes of 1,3,5,10,15,20,25,30,35,40,80,100,250. Try to get enough animals to get statistics. We recommend getting roughly 300 total animals for each brood size (i.e. run 50 times for the 5 animal case, once for the 250 case). Introduce a global batch of 10, i.e. do the whole thing 10 times, to get errorbars. Plot the ensemble mean of the **final** squared displacement depending on the brood size, with errorbars from the 10 batch sizes. Plot the ensemble mean squared displacement as a function of time for all brood sizes, with error bars.
5. Do you see a maximum squared displacement at some value between 1 and 250 ? Or do you see a monotonic curve? Qualitatively explain the result.

HEARING WORKBOOK

ONE: GATING COMPLIANCE

We can think of the bundle of stereocilia projecting from an auditory hair cell as a collection of N elastic units in parallel. Each element consists of two springs representing the tip link (stiffness k_a , equilibrium position x_a) and the actin filaments at the attachment point (stiffness k_b , equilibrium position x_b), respectively. The first spring attaches via a “trap door” which, when open, extends a distance δ from the body of the stereocilium. We may regard the trap door as an equilibrium system with two states (open and closed) separated by an energy difference ΔE_0 . See the figure below.



- Derive the formula $f_{\text{closed}} = k_a(x - x_a) + k_b(x - x_b)$ for the net force on the stereocilium in the closed state. Rewrite this in the more compact form $f_{\text{closed}} = k(x - x_1)$, and find the effective parameters k, x_1 in terms of the original quantities. Then obtain the analogous formula for the net force $f(x, y)$ when the trap door is open a distance $y < \delta$.
- The total force f_{tot} is the sum of N copies of the formula you just found. In NP_{open} of these terms the trap door is open, while in the remainder it is closed. To obtain P_{open} from the Boltzmann factor, you will need to compute $\Delta E = \Delta E_0 + \int [0][\delta]f(x, y)2y$, where $f(x, y)$ is your answer from (a).
- Assemble the pieces of your answer to get the force $f_{\text{tot}}(x)$ in terms of the parameters N, k, x_1, δ , and $\Delta E_1 = \Delta E_0 + \frac{1}{2}k_a\delta^2$. Plot your solution with $N = 65$, $k = 0.017 \text{ nm}^{-1}$, $x_1 = 23 \text{ nm}$, and $\Delta E_1 = 15 \text{ nm}$ for various values of $\delta > 0$. What choice of δ yields a curve resembling the observed nonlinearity? Is this a reasonable number?

TWO: MEASURING THE HUMAN EAR SENSITIVITY DEPENDING ON FREQUENCY

In this problem, we will going to measure the sensitivity of our ears depending on frequency. Use the notebook, soundmaker.ipynb to see how to write numerical values to a wav file.

1. Generate .wav files containing 1 s of sine waves of the following frequencies: [200 Hz, 400 Hz, 800 Hz, 1600 Hz, 3200 Hz, 6400 Hz, 12800 Hz], with different amplitudes enough to span the range of "cannot be detected" to "pretty loud".
2. For each frequency, try to find the minimal amplitude you can hear. It is helpful to ask a friend to trigger the signal so that you are minimally affected by psychological effects. Repeat the experiment two times, at least 4 hours apart.(Nobody will check this but try to do this.)
3. Make a plot of frequency to minimal amplitude. Comment on your experiment.