

2.2 Molecule counting and chemotaxis

Section last updated
April 5, 2009.

It is a remarkable fact that single celled organisms such as bacteria are endowed with sensory systems that allow them to move in response to a variety of signals from the environment, including the concentrations of various chemicals.

A classical observation (from the 19th century) is that some bacteria, swimming in water on a microscope slide, under a cover slip, will collect at the center of cover slip, while others will collect at the edges. Those with more refined tastes will form a tight band that traces the outlines of the square cover slip. Oxygen diffuses into the water through the edges of the cover slip, and by collecting along a square the bacteria have migrated to a place of constant (not maximal or minimal) oxygen concentration. It is plausible that this happens because they can sense the oxygen concentration and “know” the most comfortable value of this concentration, much as we might move to be the most comfortable distance from a fireplace in an otherwise unheated room.

That bacteria collect at nontrivial concentrations of different molecules really doesn't demonstrate that they sense the concentration. They might instead sense some internal consequences of the external variables, such as the accumulation of metabolic intermediates. In the 1960s Adler found mutants of *E coli* which cannot metabolize certain sugars or amino acids but will nevertheless migrate toward the sources of these molecules; also there are mutants that metabolize but can't migrate. This is convincing evidence that metabolism and sensing are separate systems, and thus begins the fruitful exploration of the sensory mechanisms of bacteria and the connection of these sensory mechanisms to motor output. This phenomenon is called *chemotaxis*.

I'll skip lots of the truly classical stuff and proceed with the modern biophysical approach, which begins ~ 1970 . To a large extent this modern approach rests on the work of Howard Berg and collaborators. The first key step taken by Berg and Brown was to observe the sensory motor behavior of individual bacteria. *E coli* are $\sim 1 \mu\text{m}$ in size, and can be seen relatively easily under the light microscope, but since the bacteria swim at ~ 20 body lengths per second they easily leave the field of view or the plane of focus; the solution is to build a tracking microscope [Berg & Brown 1972].

Observations in the tracking microscope, as in Fig 2.5, showed that the trajectories of individual bacteria consist of relatively straight segments interrupted but short intervals of erratic “motion in place.” These have come to be called runs and tumbles, respectively. Tumbles last ~ 0.1 seconds, but

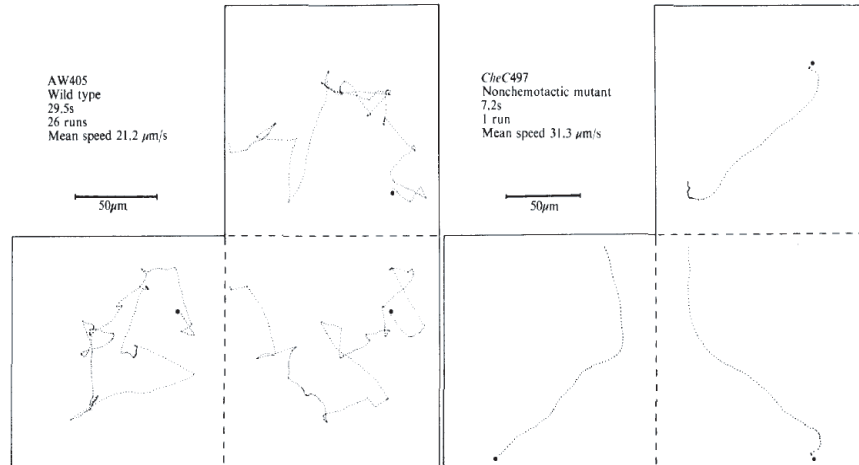


Figure 2.5: Paths of *E. coli* as seen in the original tracking microscope experiments [Berg & Brown 1972]. The three panels in each case are projections of the path onto the three orthogonal planes (imagine folding the paper into a cube along the dashed lines). At left, wild type bacteria, showing the characteristic runs and tumbles. At right, a nonchemotactic mutant that never manages to tumble.

the erratic motion during this brief time is sufficient to cause successive runs to be in almost random relative directions. Thus the bacterium runs in one direction, then tumbles and chooses a new direction at random, and so on. Runs themselves are distributed in length, as if the termination of a run is itself a random process.

Closer examination of the runs shows how it is possible for this seemingly random motion to generate progress up the gradient of attractive chemicals. When the bacterium runs up the gradient, the mean duration of the runs becomes longer, biasing the otherwise random walk. Interestingly, when bacteria swim down the gradient (or an attractant, or up the gradient of a repellent) there is relatively little change in the mean run length. Berg has described this as a form of optimism: If things are getting better, keep going, but if things are getting worse, don't worry.

Since runs get longer when bacteria swim along a positive gradient, it is natural to ask whether the cell is responding to the spatial gradient itself or

to the change in concentration with time along the path. As we will see, the spatial gradients to which the cell can respond are very small, and searching for a systematic difference (for example) between the front and back of the bacterium is unlikely to be effective just on physical grounds, independent of biological mechanisms. To search for a time domain mechanism one can expose the bacteria to concentrations which are spatially uniform but vary in time; if the sign of the change corresponds to swimming up a positive gradient, runs should be prolonged. The first such experiment used very large, sudden changes in concentration, and found that cells were trapped in extremely long runs [Macnab & Koshland 1972]. A more sophisticated experiment used enzymes to synthesize attractants from inert precursors, exposing the cells to gradual changes more typical of those encountered while swimming [Brown & Berg 1974]. Purely time domain stimuli were sufficient to generate modulations of run length that agree quantitatively with those observed for bacteria experiencing spatial gradients.

The first question one might ask is whether this strange biased random walk *really* results in progress up a gradient. Maybe the way I have described things the answer is obvious, but in fact many vaguely similar strategies (including some proposed as biological mechanisms of chemotaxis) won't work. A nice discussion is by [Schnitzer et al 1990]. To address this let's look at a simple, one dimensional case. There are then two populations of bacteria, the + population that moves to the right and the - population that moves to the left, each at speed v . If this were the whole story then the probability of finding a + bacterium at position x would change with time according to

$$\frac{\partial P_+(x, t)}{\partial t} + v \frac{\partial P_+(x, t)}{\partial x} = 0, \quad (2.56)$$

and similarly for the - bacteria,

$$\frac{\partial P_-(x, t)}{\partial t} - v \frac{\partial P_-(x, t)}{\partial x} = 0. \quad (2.57)$$

In fact tumbles cause interconversion between + and - states. We'll assume that the rate of tumbling depends on the time derivative of the concentration along the bacterial trajectory as some function $r(\dot{c})$, where of course for the \pm bacteria we have $\dot{c} = \pm v dc/dx$. When a tumble occurs in three dimensions the bacterium effectively picks a new random direction. In 1D there are only two directions, so that (for example) a tumble removes bacteria from the + population and replaces them randomly in the - and + populations. This means that Eq's (2.56) and (2.57) are supplemented by new terms, to

become

$$\begin{aligned} \frac{\partial P_+(x,t)}{\partial t} + v \frac{\partial P_+(x,t)}{\partial x} &= -r \left(v \frac{dc}{dx} \right) P_+(x,t) \\ &+ \frac{1}{2} \left[r \left(v \frac{dc}{dx} \right) P_+(x,t) + r \left(-v \frac{dc}{dx} \right) P_-(x,t) \right] \end{aligned} \quad (2.58)$$

$$\begin{aligned} \frac{\partial P_-(x,t)}{\partial t} - v \frac{\partial P_-(x,t)}{\partial x} &= -r \left(-v \frac{dc}{dx} \right) P_-(x,t) \\ &+ \frac{1}{2} \left[r \left(v \frac{dc}{dx} \right) P_+(x,t) + r \left(-v \frac{dc}{dx} \right) P_-(x,t) \right] \end{aligned} \quad (2.59)$$

Note that in each case the first term on the right of the equation denotes population loss due to tumbles, and the second term represents the redistribution after the tumbles; in this approximation tumbles themselves are instantaneous, which isn't so bad (0.1 s vs the $\sim 1 - 10$ s for typical runs). More serious, perhaps, is that we have written the tumbles as being determined by the instantaneous values of the concentration gradient, when in fact there is important averaging over time (see below); on the other hand if the gradients are relatively smooth this shouldn't be too much of a problem.

To see if the bacteria really migrate toward high concentrations we look for the steady state of these equations (combining some of the terms to make things more compact),

$$v \frac{\partial P_+(x)}{\partial x} = -\frac{1}{2} \left[r \left(v \frac{dc}{dx} \right) P_+(x) - r \left(-v \frac{dc}{dx} \right) P_-(x) \right] \quad (2.60)$$

$$-v \frac{\partial P_-(x)}{\partial x} = +\frac{1}{2} \left[r \left(v \frac{dc}{dx} \right) P_+(x) - r \left(-v \frac{dc}{dx} \right) P_-(x) \right]. \quad (2.61)$$

If we just add the equations together we find that

$$v \left[\frac{\partial P_+(x)}{\partial x} - \frac{\partial P_-(x)}{\partial x} \right] = 0, \quad (2.62)$$

so the difference between P_+ and P_- must be constant, independent of x . If the region is open, we need to be able to normalize these probability distributions and so the constant has to be zero—in steady state there is a local balance between right movers and left movers. Thus we have $P_+(x) = P_-(x) = (1/2)P(x)$, where $P(x)$ is the total probability of finding

a bacterium at point x in steady state. This probability obeys the simple first order equation

$$v \frac{\partial P(x)}{\partial x} = -\frac{1}{2} \left[r \left(v \frac{dc}{dx} \right) - r \left(-v \frac{dc}{dx} \right) \right] P(x). \quad (2.63)$$

In particular, if we simplify and assume that the rate of tumbling is modulated linearly by the time derivative of the concentration,

$$r(\dot{c}) \approx r(0) + \frac{\partial r}{\partial \dot{c}} \dot{c} + \dots, \quad (2.64)$$

so that

$$r \left(\pm v \frac{dc}{dx} \right) \approx r(0) \pm \frac{\partial r}{\partial \dot{c}} \cdot v \frac{dc}{dx} + \dots, \quad (2.65)$$

then Eq (2.63) becomes

$$v \frac{\partial P(x)}{\partial x} = - \left[\frac{\partial r}{\partial \dot{c}} \cdot v \frac{dc}{dx} \right] P(x), \quad (2.66)$$

or more straightforwardly

$$\frac{\partial P(x)}{\partial x} = - \left[\frac{\partial r}{\partial \dot{c}} \cdot \frac{dc}{dx} \right] P(x). \quad (2.67)$$

It is easy to see that (since $\partial r / \partial \dot{c}$ is a constant), the solution is

$$P(x) = \frac{1}{Z} \exp \left[-\frac{\partial r}{\partial \dot{c}} c(x) \right]. \quad (2.68)$$

Thus, in these approximations, chemotaxis leads to a Boltzmann distribution of bacteria, in which the concentration acts as a potential. If the molecules are attractive then $\partial r / \partial \dot{c} < 0$ and hence maxima of concentration are minima of the potential, conversely for repellents. The stronger the modulation of the tumbling rate (as long as we stay in our linear approximation) the lower the effective temperature and the tighter the concentration of bacteria around the local maxima of concentration.

Problem 32: A little more about the effectiveness of chemotaxis. Be sure that you understand all of the details in the calculation that we have just done.

(a.) Within this one dimensional model, what happens if the tumbling rate is modulated not just by the time derivative, but also by the absolute concentration, so that the bacterium confuses “currently good” for “getting better”?

(b.) Can you generalize this discussion to three dimensions? Instead of having just two groups + and −, one now needs a continuous distribution $P(\boldsymbol{\Omega}, x, t)$, where $\boldsymbol{\Omega}$ denotes the direction of swimming. Derive an equation for the dynamics of $P(\boldsymbol{\Omega}, x, t)$ in the same approximations used above, and see if the Boltzmann-like solution obtains in this more realistic case.

Perhaps some of this (and more) belongs in an appendix about the mechanics of bacterial swimming, but I'll leave it here for now.

All of this description so far is about the phenomenology of swimming. But how does it actually work? It is crucial that you appreciate the dramatic difference between swimming in bacteria and swimming of fish or humans. These differences have an expression in the biology, but fundamentally the issues are all about physics. The basic problem is that bacteria are too small to take advantage of inertia. When we swim, we can push off the wall of the pool and glide for some distance, even without moving our arms or legs; this gliding distance is on the order of one or two meters, roughly the length of our bodies. In contrast, if a bacterium stops running its motors, it will glide for a distance comparable not to its body length ($\sim 1 \mu\text{m}$) but to the diameter of an atom. To see this, think about a small particle moving through a fluid, subject only to drag forces (the motors are off). If the velocities are small, we know the drag will be proportional to the velocity, so Newton's equation is just

$$m \frac{dv}{dt} = -\gamma v. \quad (2.69)$$

For a spherical object of radius r , Stokes' law tells us that $\gamma = 6\pi\eta r$, where η is the viscosity of the fluid, and we also know that $m = 4\pi\rho r^3/3$, where ρ is the density. The result is that

$$v(t) = v(0) \exp(-t/\tau), \quad (2.70)$$

$$\tau = \frac{m}{\gamma} = \frac{2\rho r^2}{9\eta}. \quad (2.71)$$

If we assume that the density of bacteria is roughly that of water, then it is useful to recall that η/ρ has units of a diffusion constant, and for water $\eta/\rho = 0.01 \text{ cm}^2/\text{s}$. With $r \sim 1 \mu\text{m} = 10^{-4} \text{ cm}$, this gives $\tau \sim 5 \times 10^{-7} \text{ s}$. If the initial velocity is $v(0) \sim 10 \mu\text{m}/\text{s}$, the net displacement during this coasting is $\Delta x = v(0)\tau \sim 5 \times 10^{-12} \text{ m}$; recall that a hydrogen atom has a diameter of $\sim 1 \text{ \AA} = 10^{-10} \text{ m}$.

The conclusion from such simple estimates is that bacteria can't coast. More generally, mechanics on the scale of bacteria is such that inertia is

negligible, as if Aristotle (rather than Galileo and Newton) were right. This really about the nature of fluid flow on this scale.⁴ For an incompressible fluid (which is a good approximation here—surely the bacteria don't generate sound waves as they swim), the Navier–Stokes equations are

$$\rho \left[\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} \right] = -\nabla p + \eta \nabla^2 \mathbf{v}, \quad (2.72)$$

where \vec{v} is the local velocity of the fluid, p is the pressure, and as usual ρ is the density and η is the viscosity. The pressure isn't really an independent variable, but needs to be there so we can enforce the condition of incompressibility,

$$\nabla \cdot \mathbf{v} = 0. \quad (2.73)$$

These equations need to be supplemented by boundary conditions, in particular that the fluid moves with the same velocity as any object at the points where it touches that object. Thus the velocity should be zero at a stationary wall, and should be equal to the velocity of a swimmer at the swimmer's surface.

Problem 33: Understanding Navier–Stokes. This isn't a fluid mechanics course, but you should be sure you understand what Eq (2.72) is saying. In particular, this is nothing but Newton's $F = ma$. Explain.

Dimensional analysis is an enormously powerful tool in fluid mechanics. We are free to choose new units for length (ℓ) and time (t_0), and hence for velocity ($v_0 = \ell/t_0$), as well as for pressure p_0 , and this gives us

$$\rho \left[\frac{v_0}{t_0} \frac{\partial \tilde{\mathbf{v}}}{\partial \tilde{t}} + \frac{v_0^2}{\ell} \tilde{\mathbf{v}} \cdot \tilde{\nabla} \tilde{\mathbf{v}} \right] = -\frac{p_0}{\ell} \tilde{\nabla} \tilde{p} + \eta \frac{v_0}{\ell^2} \tilde{\nabla}^2 \tilde{\mathbf{v}}, \quad (2.74)$$

$$\frac{\rho \ell v_0}{\eta} \left[\frac{\partial \tilde{\mathbf{v}}}{\partial \tilde{t}} + \tilde{\mathbf{v}} \cdot \tilde{\nabla} \tilde{\mathbf{v}} \right] = -\frac{p_0 \ell}{\eta v_0} \tilde{\nabla} \tilde{p} + \tilde{\nabla}^2 \tilde{\mathbf{v}}, \quad (2.75)$$

⁴My experience is that most physics students don't know too much fluid mechanics, so although this is elementary I put it here. For a more thorough discussion, see, as usual, Landau and Lifshitz.

where $\tilde{t} = t/t_0$, and similarly for all the other variables. By choosing $p_0\ell/\eta v_0 = 1$ we can get rid of all the units, except we are left with a dimensionless combination

$$\text{Re} \equiv \frac{\rho\ell v_0}{\eta} \quad (2.76)$$

which is called the Reynolds' number. Notice that if we choose the unit of length to be the size of the objects that we are interested in, and v_0 to be the speed at which they are moving, then even the boundary conditions don't have any units, nor do they introduce any dimensionless factors that are far from unity. The conclusion is that all fluid mechanics problems with the same geometry (shapes) are the same if they have the same Reynolds' number. In this sense, being smaller (reducing ℓ) is the same as living at increased viscosity.⁵

To make a long story short, we live at high Reynolds' number, and bacteria live at low Reynolds' number. Turbulence is a high Reynolds' number phenomenon, as is the more mundane gliding through the pool after we push off the wall. At low Reynolds' number, life is very different.⁶ Inertia is absent, and so forces must balance at every instant of time. To say this more startlingly, time doesn't actually appear in the equations. This means that, as you swim, the distance that you move depends on the sequence of motions that you go through, but not on the dynamics with which you execute them.

To use Purcell's evocative example, at high Reynolds' number a scallop can propel itself by snapping shut, expelling a jet of water, and then opening slowly. The jet will propel the scallop forward, and the drag of reopening can be made small by moving slowly. At low Reynolds' number this doesn't work, and the forward displacement generated by snapping shut will be exactly compensated by the drag on reopening. To have net movement from a cycle, the *sequence* of shapes that the swimmer goes through in the cycle must break time reversal invariance, not just the trajectory.

⁵It is worth reflecting on the level of universality that we have here. We start with the molecular description of fluids, and eventually figure out that all we need to know are the density and viscosity. Now we see that even these quantities are tied up with our choice of units. If we want to know what happens in natural units (i.e., scaling to the size and speed of the objects we are looking at), then all that matters is a single dimensionless combination, Re.

⁶This fact surely was known to many people, for many years. But Berg's experiments on *E coli* provided a stimulus to think about this, and it resulted in a beautiful exposition by EM Purcell, Life at low Reynolds' number, *Am J Phys* **45**, 3–11 (1977). This has been hugely influential.

So, how do bacteria do evade the ‘scallop theorem’? If you watch them swimming, you can see that they have long filaments sticking out, and these seem to be waving. I emphasize that “see” it tough here. These filaments are very small, much thinner than the wavelength of light. To see them, the easiest thing is to use dark field microscopy, in which the sample is illuminated from the side and what you see is the light scattered by $\sim 90^\circ$. These apparently waving appendages are called flagella, and remind us of what we see on other small swimming cells, such as sperm. The difference is that the flagella in these other cases are huge by comparison with the bacterial flagella. If you slice through the tail of a sperm and take an electron micrograph, you find an enormously complex structure, and if you try to analyze the system biochemically you find it is made from several different proteins. Importantly, some of these proteins act as enzymes and eat ATP, which we know is a source of energy, for example in our muscles. In contrast, the bacterial flagellum is small, with a relatively simple structure, and the biochemistry suggests that it is little more than a very long polymer made from one kind of protein; this protein is not an enzyme. How can this simple structure, with no ATPase activity, generate motions?

In experiments that aimed at better ways to see the flagella, one can attach “flags” to them using viruses that would stick to the flagella via antibodies. Once in a while, a virus with antibodies on both ends would stick to two flagella from different bacteria. When this happened, you could see the bacterial cells rotating, which one can imagine was a huge surprise. Eventually people figured out how to break off the flagella and stick the bacteria to a glass slide by the remaining stump, and then the bacterium rotates. Rotation can look like a wave if the flagellum is shaped like a corkscrew, and it is. Rotating a corkscrew obviously violates time reversal invariance. If you have several corkscrews and you rotate them with the correct handedness, they can fit together into a bundle. If you rotate the other way, the corkscrews clash, and any bundle will be blown apart by this clashing. So, with many flagella projecting from their surface, we can imagine that by switching the direction of rotation, the bacterium switches between a bundle that can smoothly propel the cell forward, and many independently moving flagella that would cause the cell to tumble in place—runs and tumbles correspond to counterclockwise and clockwise flagellar rotation.⁷ If you find mutants that never tumble, and stick them down by

⁷Which way this association goes of course depends on our convention for defining the handedness of rotation; it doesn’t matter (and I have trouble remembering it!) as long as you are consistent.

their stumps, then they all rotate one way; similarly, mutants that tumble too often rotate the other way.

There is much more to say about the rotary engine itself, sitting at the base of the flagella. It is powered not by ATP but by a difference in chemical potential for hydrogen ions between the inside and the outside of the cell. This is an energy source that all cells use, albeit in different ways, because it allows chemical events at very different spatial locations to be coupled. Thus, for example, photosynthetic organisms use the energy of the absorbed photons to move electrons across a membrane, and then compensate the charges by moving protons; the resulting difference in chemical potential can be used by other membrane-spanning enzymes to make ATP, without being anywhere near the molecules that absorb the photon.⁸ In fact, these enzymes that synthesize ATP also rotate as they let protons move down the gradient in their chemical potential, and these same enzymes are responsible for ATP synthesis in all cells. So, proton driven rotary motors are at the heart of energy conversion in all organisms.

There is also more to say about mechanics at low Reynolds' number. Swimming involves changing shape, and this provides the boundary conditions on the Navier–Stokes equations. A cycle of changing boundary conditions should lead to a net displacement. There is some subtlety here, since the space of shapes is not so easy to parameterize. If we think, for example, about a closed surface, shape is three dimensional position as a function of the two coordinates on the surface (e.g., latitude and longitude), but there is an arbitrariness in how we choose these coordinates; of course any physical quantity, such as the amount by which the swimmer moves forward, must be invariant to this choice. Looking more closely, the freedom to choose coordinates means that the natural formulation of the problem includes a gauge symmetry [Shapere & Wilczek 1987]. Reluctantly, let's leave all this and go back to the problem of chemotaxis itself.

Problem 34: Switching in tethered bacteria. As noted above, one way of studying bacterial motility and chemotaxis is to “tether” a bacterium by the stump of one flagellum, observing the rotation of the whole cell rather than the rotation of the flagellum. The file `omega.txt` contains a very long time series of the angular velocity from such an experiment done by Will Ryu here at Princeton. The samples are taken sixty times per second, and the units of velocity are not quite arbitrary but not really important either; you should be able to load this into MATLAB (`load omega.txt`).

⁸You can imagine how confusing this was before people figured it out! It looked like a mysterious action at a distance.

- a. You should see that the velocity switches between positive and negative values, but these values are fairly constant. This is consistent with swimming by switching between runs and tumbles, with little or no modulation of the swimming speed. What is the distribution of times spent with during each segment of positive or negative (clockwise or counterclockwise) velocity?
- b. It usually is said that switching is a Poisson process, so that (as you remember from the discussion of photon counting) the distribution of intervals between switches should be exponential. Are your results in [a] consistent with this prediction?
- c. Look carefully at the velocity vs. time in the data set. Are the data statistically stationary (time–translation invariant)? If you focus on segments of the data that are more clearly stationary, does that change your conclusions in [b]?
- d. Sometimes the angular velocity makes a “partial switch,” a brief excursion away from the typical positive or negative value but not quite a full switch to the opposite direction of rotation. Qualitatively, what is happening in these cases? What would be the simplest model to describe the velocity vs. time during such an event? Can you give a quantitative analysis of the data, fitting to your model? This is a bit open ended.

We are interested in the question of how sensitively the bacterium can respond to small concentration gradients. We suspect that, since individual molecular motions are random, there must be a limit, analogous to the shot noise in counting photons. In a classic paper, Berg and Purcell provided a clear intuitive picture of the noise in ‘measuring’ chemical concentrations [Berg & Purcell 1977]. Their argument was that if we have a sensor with linear dimensions a , then effectively the sensor samples a volume a^3 . In this volume we expect to count an average of $N \sim ca^3$ molecules when the concentration is c . Each such measurement, however, is associated with a noise $\delta N_1 \sim \sqrt{N}$. Since the count of molecules is proportional to our estimate of the concentration, the fractional error will be the same, so from one observation we obtain a precision

$$\left. \frac{\delta c}{c} \right|_1 = \frac{\delta N}{N} = \frac{1}{\sqrt{N}} = \frac{1}{\sqrt{ca^3}}. \quad (2.77)$$

But we know that we can make more accurate measurements by averaging over time, although this is a bit tricky—we won’t get a better estimate of the concentration around us by counting the same molecules over and over again. Thus if we are willing to average over a time τ_{avg} , we can only really make K independent measurements, where $K \sim \tau_{\text{avg}}/\tau_c$, and the correlation time τ_c is the time we have to wait in order to get an independent sample of molecules.

How do we get independent samples? If we look in a small volume, the molecules that we are looking at exchange with the surroundings through diffusion. Thus the time required to get an independent collection of molecules is the time required for molecules to diffuse in and out of the volume, $\tau_c \sim a^2/D$. Putting everything together we have

$$\frac{\delta c}{c} = \frac{1}{\sqrt{K}} \cdot \frac{\delta c}{c} \Big|_1 \quad (2.78)$$

$$= \sqrt{\frac{\tau_c}{\tau_{\text{avg}}}} \cdot \frac{1}{\sqrt{ca^3}} \quad (2.79)$$

$$= \sqrt{\frac{a^2}{D\tau_{\text{avg}}}} \cdot \frac{1}{\sqrt{ca^3}} \quad (2.80)$$

$$= \frac{1}{\sqrt{Dac\tau_{\text{avg}}}}. \quad (2.81)$$

This is a lovely result. It says that the limit to the accuracy of measurements depends on the absolute concentration (more molecules \rightarrow more accuracy), on the size the detector (bigger detectors \rightarrow more accuracy), on the time over which we are willing to average (more time \rightarrow more accuracy), and finally on the diffusion constant of the molecules we are sensing, because faster diffusion lets us see more independent samples in the same amount of time. All these parameters combine simply, essentially in the only way allowed by dimensional analysis.

One way of understanding the Berg and Purcell result on the limits to precision is to think about the rate at which molecule find their target. If we have molecules at concentration c moving with diffusion constant D , we expect that the rate (number of molecules per second) that arrive at a target of size a should be proportional both to c and to D , and then by dimensional analysis we need one factor of length, so the rate is $\sim Dac$ molecules per second. This result is used most often to talk about the “diffusion limited rate constant” for a chemical reaction; if we have



then the second order rate constant k_+ can never be bigger than $\sim Da$, where D is the diffusion constant of the molecules and a is their size, or more precisely the size of the region where they have to hit in order to react. But if the rate of molecular arrivals is $\sim Dac$, in a time τ_{avg} we will count $\sim Dac\tau_{\text{avg}}$ molecules, and if these molecules are arriving at random then there will be the usual square root fluctuations, which leads us to Eq (2.81).

In this view, the Berg–Purcell limit is nothing but shot noise in molecular arrivals, and thus is completely analogous to shot noise in photon arrivals. Photons propagate and molecules diffuse, but under most conditions they both arrive at random, hence there is shot noise in counting.

Problem 35: Diffusion limited rates, more carefully. One can try a more careful calculation of the rate at which molecules find their target by diffusion. Imagine a sphere of radius a such that all molecules which hit the surface are immediately absorbed. Outside the sphere, the concentration profile must obey the diffusion equation, and the absorption means that on the spherical surface the concentration will be zero. Far from the sphere, the concentration should be equal to c . Thus we have

$$\frac{\partial c(\mathbf{x}, t)}{\partial t} = D\nabla^2 c(\mathbf{x}, t); \quad (2.83)$$

$$c(|\mathbf{x}| = a, t) = 0, \quad (2.84)$$

$$c(\mathbf{x} \rightarrow \infty, t) = c. \quad (2.85)$$

The number of molecules arriving per second at the surface of the sphere is given by an integral of the diffusive flux over the surface

$$\text{rate} = \int d^2s \hat{\mathbf{n}} \cdot [-D\nabla c(\mathbf{x}, t)] \Big|_{|\mathbf{x}|=a}, \quad (2.86)$$

where d^2s is an element of the surface area on the sphere, and $\hat{\mathbf{n}}$ is the unit vector normal to the sphere.

(a.) Solve Eq (2.83), with the boundary conditions in Eqs (2.84) & (2.85), in steady state. Note that as a first step you should go to spherical coordinates; recall that in three dimensions the Laplacian can be written as

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2} \left[\frac{1}{\sin^2 \phi} \frac{\partial^2}{\partial \theta^2} + \frac{1}{\sin \phi} \frac{\partial}{\partial \phi} \left(\sin \phi \frac{\partial}{\partial \phi} \right) \right], \quad (2.87)$$

where as usual r is the radius and θ and ϕ are the polar and azimuthal angles, respectively.

(b.) Use your steady state solution to evaluate the rate at which molecules arrive at the sphere, using Eq (2.86). Also, explain why simple dimensional analysis of these equations yields $\text{rate} \sim Dac$.

(c.) What happens if you try to give a dimensional analysis argument for the rate in one or two dimensions? If there are problems, can you explain how these problems either go away or are made more precise by trying to solve the diffusion equation with appropriate boundary conditions? As a hint, the two dimensional case is a bit delicate; focus first on one dimension.

In their original work, Berg and Purcell applied this simple formula to the case of bacterial chemotaxis by identifying the linear dimensions of the

probably should go back and review the numbers here

sensor with the size of the bacterium itself. In effect this is equivalent to assuming that *E coli* can count all the attractant or repellent molecules that would have appeared within its volume, which certainly seems like a generous assumption—presumably it can only count molecules at the surface, where the receptors are located, and we know that the receptors are nowhere near being so numerous as to cover the whole surface. Nonetheless, as Berg and Purcell discuss in detail, this estimate of the limiting noise level is very close to the smallest concentration differences that *E coli* can detect reliably.

When we talk about the the smallest change that the bacterium can detect, we should be careful. In a spatial gradient, the relevant difference is the difference measured across the length of the cell, or roughly one micron. But if the bacterium can make comparisons over time, then it can compare concentration differences across the length of a run, which is ten or twenty microns. What Berg and Purcell found is that, for the gradients that bacteria can just barely detect, this difference, roughly an order of magnitude, is a big difference. With reasonable integration times—roughly the duration of a run—temporal comparisons generate differences larger than the noise, while spatial comparisons do not. Thus, in order to achieve the observed reliability of detection, bacteria must make temporal comparisons, and this is a matter of physics not biology. Further, to see whether they are going up or down the gradient they must be measuring the time derivative of concentration along their trajectory, and they need to integrate their measurement of the derivative for a few seconds in order to beat down the noise.

Why don't the bacteria integrate for longer, and reduce the noise further? If you look closely at the trajectories of the bacteria, you can see that the longer runs curve a bit. In fact, the bacteria are sufficiently small that their own rotational Brownian motion disorients them on a time scale of ten or fifteen seconds. So, if you integrate for longer than this, you are no longer integrating something related to the gradient in a particular direction, or even your current direction of motion. This suggests that there is a physical limit setting the longest useful integration time.

Berg and Purcell also argued that there is a minimum useful integration time. Recall that molecules moving via diffusion traverse a distance $x_{\text{diff}} \sim \sqrt{Dt}$ in a time t ; in contrast, swimming at velocity v moves the bacterium by a distance $x_{\text{swim}} \sim vt$. For short times, diffusion, with its square root dependence on time, goes farther than ballistic swimming motion. This means that on short time scales, the molecules that the bacterium sees along its path are the same molecules, and hence it really isn't combining

statistically independent measurements.⁹ So, there is a minimum useful integration time (assuming you want to improve the signal-to-noise ratio by integrating) of $\tau \sim D/v^2$, and this works out to be about one second.

So, the strategy of *E coli* for measuring gradients is incredibly constrained by physics. To reach the observed performance, it has to count nearly every molecule that arrives at its surface. Even with this near ideal behavior, it can work only by making comparisons across time, not space, and estimates of time derivatives have to be averaged for a few seconds, not more and not less. This set of predictions about chemotactic strategy is almost parameter free, even if not precisely quantitative.

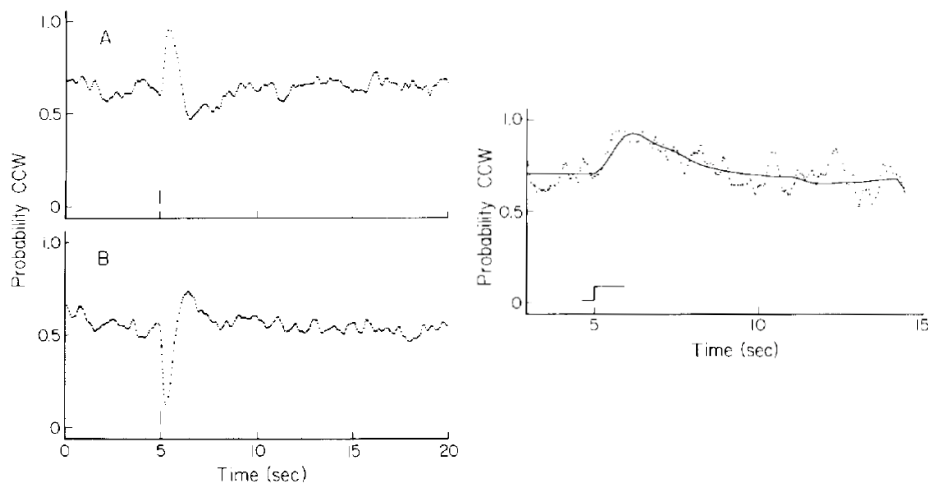


Figure 2.6: Impulse responses in bacterial chemotaxis [Block et al 1982]. At left, changes in the probability of counterclockwise rotation of the motor, corresponding to running, as a function of time in response to a pulse of attractant (top) or repellent (bottom). We see that the form of the response is equivalent to integrating the time derivative of the input over a window of several seconds. At right, the response to a step of attractant again has the form expected if we integrate the derivative over a short window. The real data are compared with a prediction based on integrating the response to impulses shown at left, and the agreement is good, as if the system were linear.

⁹This is a very clear intuition, but I am not sure I know an honest calculation that gives this result. I should go back and try to clean this up, although it's slightly puzzling that it has remained obscure since 1977. Suggestions welcome.

What do real bacteria do? We have already seen that they make temporal comparisons. Does the detailed form of these comparisons agree with the Berg–Purcell predictions? Although one could probably do better with modern experimental techniques, the best test was done in the early 1980s [Block et al 1982, 1983]. In these experiments, bacteria were tethered to a glass slide and exposed to changing concentrations of attractants or repellents; a long series of such observations is then combined to measure the probability that the flagellar motor is rotating counterclockwise (corresponding to running) as function of time relative to the changing concentration. A summary of these experiments is shown in Fig 2.6. We see that the probability of running is modulated by the time derivative of the concentration, averaged over a window of a few seconds, exactly as predicted by the Berg–Purcell argument.

Being sensitive to a derivative means that the response to a step comes back almost exactly to the baseline before the step, as seen at right in Fig 2.6, so that the constant signal is ignored at times long after it was turned on. This gradual ‘forgetting’ of a constant signal is common in biological systems, and such phenomena are called ‘adaptation.’ All of our sensory systems exhibit adaptation, the most familiar being the experience of stepping into a dark movie theater or out into the bright sunlight; at first we are acutely aware of the large difference in overall light intensity, but after a while everything looks normal and we are insensitive to the absolute photon flux. The case of bacteria is interesting because it seems that the adaptation is nearly perfect. We return to this point in the next Chapter.

Experiments of the sort pictured in Fig 2.6 also make it possible to estimate the absolute sensitivity of the system in perhaps more compelling units. We now know how many receptors there are on the cell’s surface, and so we can convert changes in concentration into changes in the number of occupied receptors. Indeed, one extra occupied receptor leads to a significant change in the probability of running vs tumbling. So, as expected, the bacterium is responding to individual molecular events.

At this point we should step back and see where we are. On the one hand, this is a great success: much of bacterial behavior is understandable, semi-quantitatively, as a response to the physical constraints posed by life at low Reynolds’ number and the noise in molecular counting; one can go further and say that bacterial behavior is near optimal in relation to this noise. On the other hand, many questions are left hanging:

1. Can we do an honest calculation that leads to the Berg–Purcell limit on the precision of concentration sensing?

should put the numbers here,
maybe reproduce a figure

2. Similarly, can we turn the ideas about maximum and minimum useful integration times into a theory of optimal filtering that would predict, quantitatively, the form of the impulse responses in Fig 2.6?
3. What is the mechanism of this extreme sensitivity?

The first question is incredibly important. What Berg and Purcell wrote down makes absolutely no reference to the messy details of what actually happens to molecules as they are counted. This could be wonderful, because it would mean that can say something about the limits to precision in *all* biochemical signaling systems, regardless of details. Alternatively, the absence of details might be a disaster, a clue that we have simply missed the point. This is a big subject, however, so we'll leave it for the next lecture.

From a practical point of view, I think the second question is easy: we should be able to do this, but I don't think anyone has really managed to get it right. There have been some serious attempts [Strong et al 1998, Andrews et al 2006], but I think the issue still is open.¹⁰ One might also wonder whether it even makes sense to formulate this problem for individual bacteria, as opposed to looking at competition or cooperation in a population; this is related to the question of what, precisely, one thinks is being optimized by the behavior. It seems likely that any theory of optimal strategies will predict that this optimum is context dependent, suggesting that adaptation is richer than just subtracting off a constant, which is a theme that we return to in relation to noise in neurons and perception; here we should note that quantitative characterization of chemotactic behavior has not been pursued under a very wide range of stimulus conditions, so we may be missing the data we need to test such theories when they emerge.

check for other references

The third and last question is one where much progress has been made, although again some issues are open. As with the rod cell, there is a cascade of biochemical event that leads from input (here, binding to receptors on the cell surface) to output (direction of motor rotation). Since input and output are spatially separated, it is not surprising to find that there is an internal signaling molecule that diffuses through the cell. In rods, this is a small molecule (cGMP), but for bacterial chemotaxis it is a protein called CheY. More precisely, this protein can be phosphorylated, and in its phosphorylated form CheY~P it binds to the motor and favors clockwise rotation. Working backward from the output, we would like to know how the rotational bias of the motor depends on the concentration of CheY~P.

¹⁰Full disclosure: I am one of the et al in the earlier reference. Maybe this adds weight to my claim that we don't understand the problem, maybe not.

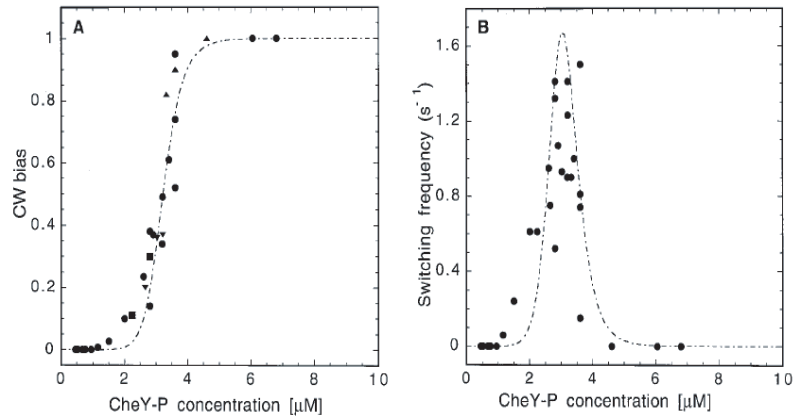


Figure 2.7: Response of the flagellar motor to changes in concentration of CheY~P [Cluzel et al 2000]. CW denotes the counterclockwise rotation of the motor, and the switching frequency is the mean rate of transitions between the CW and CCW states.

To measure the bias vs CheY~P, one has to do many tricks [Cluzel et al 2000]. It's relatively easy to measure the bias of the motor, either in experiments where the cell is tethered or where it is laying on a slide and one motor stump is sticking up with a bead attached. To know the concentration of a protein in a single cell, we need to make the protein visible, and so this is done by genetic engineering, replacing the normal CheY with a fusion between this protein and the green fluorescent protein. To adjust the concentration, one puts the expression of the CheY-GFP fusion under the control of a promoter that turns on in response to external signals. Finally, we need to know the concentration of the phosphorylated form of the protein, and this is very difficult. But once phosphate groups are attached to a protein, they stay there until removed by another enzyme (the phosphatase). So, if we genetically engineer the bacterium to remove the phosphatase, we will surely screw up the overall chemotactic response, but we can then be sure that all the CheY will be in its phosphorylated state. The result of all this is shown in Fig 2.7.

not sure where this first comes up ... should be sure to have a clear discussion of gfp somewhere!

Problem 36: Absolute concentration measurements. In this problem you should try to understand how Cluzel et al were able to put the CheY~P concentration on

an absolute scale in Fig 2.7. Bacteria can be engineered to make a fluorescent version of many naturally occurring proteins. While the fluorescence signal that we then see under a microscope is proportional to the number of molecules under illumination, it can be difficult to measure the proportionality constant in an independent experiment. One can circumvent this problem by watching small numbers of molecules diffusing randomly in and out of an illuminated volume inside an individual cell and using the *variance* in the fluorescence intensity, along with its mean value, to make an absolute measurement of the concentration of the molecules.

(a.) Explain (qualitatively) how this measurement might work. What do you gain by using both the variance and the mean of this signal? How can the fluctuating fluorescence signal be analyzed further to give an estimate of the protein diffusion constant?

(b.) Now let's convert the above intuition into a quantitative framework for analysis of the data. Consider the concentration $c(\vec{x}, t)$ of fluorescent molecules at different points in space and time. It fluctuates and the deviation δc of the concentration from its average value \bar{c} is uncorrelated between different points in space (but the same instant of time). Show that the analytic statement

$$\langle \delta c(\vec{x}, t) \delta c(\vec{x}', t) \rangle = \bar{c} \delta(\vec{x} - \vec{x}') \quad (2.88)$$

of this fact is equivalent to the 'intuitive' remark that the variance of the number of molecules in a volume is equal to the mean number.

(c.) If the system starts with some fluctuation in the concentration $c(\vec{x}, 0) = \bar{c} + \delta c(\vec{x}, 0)$, this profile will relax according to the diffusion equation. Since the diffusion equation is linear, this means that the profile of fluctuations at time t , $\delta c(\vec{x}, t)$, can be written as a linear operator acting on the initial condition $\delta c(\vec{x}, 0)$. Show that this linear relationship can be written as

$$\delta c(\vec{x}, t) = \int d^3 y \left(\frac{1}{\sqrt{4\pi Dt}} \right)^3 \exp(-|\vec{x} - \vec{y}|^2 / 4Dt) \delta c(\vec{y}, 0) \quad (2.89)$$

where D is the diffusion constant.

(d.) When we bring light to a focus under the microscope, we effectively weight the points around the focus with a Gaussian function, so that the light intensity collected from the fluorescent molecules will be proportional to

$$s(t) = \int d^3 x c(\vec{x}, t) \exp(-|\vec{x}|^2 / \ell^2) \quad (2.90)$$

where ℓ is the size of the focal region (roughly the size of the wavelength of light). Using the results above, show that the temporal correlation function of this signal is given by

$$\langle \delta s(t) \delta s(0) \rangle \propto (|t| + \tau)^{-3/2}, \quad (2.91)$$

and relate the correlation time τ to the diffusion constant D and the size of the focal region ℓ . As a hint, note that in doing the multidimensional Gaussian convolution integrals that show up in the last step of this computation, it is a good idea to do them Cartesian coordinate by Cartesian coordinate. This gives a precise method for extracting the diffusion constant from the fluctuating fluorescence signal.

What we see most clearly from Fig 2.7 is that the motor is remarkably sensitive to small changes in concentration of the internal signaling molecule. One can fit a function of the form

$$P_{ccw} = \frac{c^n}{c^n + K^n}, \quad (2.92)$$

with $K \sim 3 \mu\text{M}$ and $n \sim 10$, but the data are almost within errors of being a step function. So part of the answer to how the the bacterium is so sensitive to small changes in the external concentration of attractants or repellents is that the motor is very sensitive to small changes in the concentration of CheY~P. This is not implausible, since the structure of the motor (which is complicated) suggests locations for as many as 34 sites where CheY~P could bind around a ring of radius $R \sim 45 \text{ nm}$ [Thomas et al 1999].

Having such strong sensitivity to the CheY~P concentration means that, in roughly the one second it takes for the motor to switch once, one can be sure whether the concentration was $\delta c/c \sim 1/n \sim 10\%$ above or below the critical value $c = K$. But from Berg and Purcell we might expect that there is a limit on this precision set by random arrival of the CheY~P molecules at the motor, and this should be $\delta c/c \sim 1/\sqrt{DRc\tau_{\text{avg}}}$, treating the whole motor ring as one big receptor. With diffusion constants for proteins, including CheY, in the range of $D \sim \mu\text{m}^2/\text{s}$, this suggests that the limit with one second of integration is not much smaller than 10% (see more details in the next lecture). So, cooperative action of many signaling molecules generates a steep slope, but the system still has to suppress other sources of noise since even this last step in the cascade of events is operating close to the fundamental limits set by noise considerations.

need to discuss measurements on diffusion constants for proteins .. here or earlier?

The observations on the sensitivity of the motor tell us that the bacterium can generate a significant response even from a small fractional change in the concentration of CheY~P. Still, we need to understand the biochemical processes that lead from essentially single molecular events to these quasi-macroscopic changes in molecule number.¹¹ Figure 2.7 shows that this extreme sensitivity must coexist with a very tight regulation, since if the concentration of CheY~P drifts far away from K the cell loses all sensitivity to changes. This combination of sensitivity to small changes without accumulation of large variations poses significant problems, which we will take up in the next Chapter.

¹¹At $c \sim 3 \mu\text{M}$, a cell with volume $\sim 1 \mu\text{m}^3$ has ~ 2000 molecules of CheY~P, so even a ten percent change in concentration involves hundreds of molecules.