Some adventures in teaching at the interface of physics and biology

William Bialek

Joseph Henry Laboratories of Physics
Lewis-Sigler Institute for Integrative Genomics
Princeton University

http://www.princeton.edu/~wbialek/wbialek.html

How many cultures?

> Literary intellectuals at one pole—at the other scientists, and as the most representative, physical scientists. Between the two a gulf of mutual incomprehension ... a curious distorted image of each other.

CP Snow, in *The Two Cultures and The Scientific Revolution*
(Cambridge University Press, 1959)

Could there be as large a gulf within science itself?

The "mathematical sciences"
(e.g, physics)

The "non-mathematical sciences"
(e.g, biology)

The flavors of freshman physics:
- Physics for physicists
- Physics for engineers
- Physics for biology and medicine
- Physics for Poets

There is a widespread (almost trite) sentiment that the gulf between the physical and biological sciences needs to bridged, and that now is the right time to do this.

"The biology of the 21st century will be a more quantitative science."
"The greatest challenges need to be met by interdisciplinary collaborations."
"The genome and the computer have revolutionized how we do biology."
"As we address system-level questions, we move beyond what we can do intuitively, and need more mathematical tools."
"Mathematics will be biology's new microscope."

There is much less agreement about what all of this actually means, even in principle.

"There's something happening here, what it is ain't exactly clear ... "
(Stephen Stills, Buffalo Springfield, 1967)

Two related but distinct goals

1. Educate biologists who find it natural to do quantitative experiments, sophisticated analyses of their data, and meaningful comparisons with theory.

    (because biology is so big, even incremental progress can have a big impact)
    (perhaps we shouldn't be shy to say "make biology more like physics")

2. Educate physicists who find it natural to bring the "physicist's style of thought" to study a broader class of systems, including biological systems.

    (this clearly can't be accomplished by learning less physics!)

(1.) is a service that physicists can do for the community.
It's important. It could even save your life someday. Only we can do it right.
Even if you don't care about this goal, you should care about the resources which will be devoted to achieving it, whether we step up to the challenge or not.

(2.) is something I really care about, as a physicist.

These goals are very different for advanced students.
If you wait too long, (1.) is nearly impossible to achieve.
(2.) can be addressed through advanced undergraduate and graduate courses, but it's much easier if we start earlier.

To make progress, we need to be brutally honest about the scale of the problem.

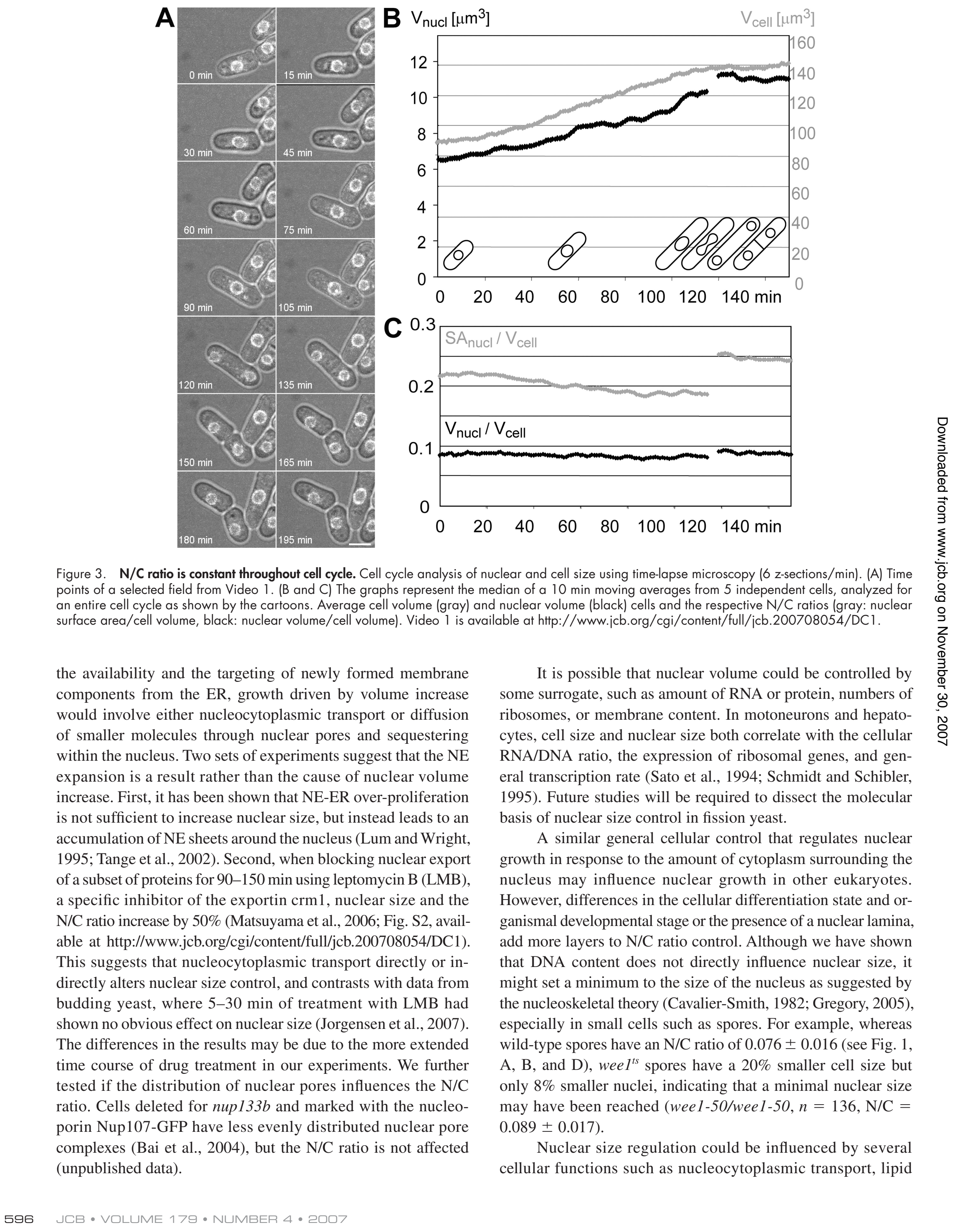


Figure 3. **N/C ratio is constant throughout cell cycle.** Cell cycle analysis of nuclear and cell size using time-lapse microscopy (6 z-sections/min). (A) Time points of a selected field from Video 1. (B and C) The graphs represent the median of a 10 min moving averages from 5 independent cells, analyzed for an entire cell cycle as shown by the cartoons. Average cell volume (gray) and nuclear volume (black) cells and the respective N/C ratios (gray: nuclear surface area/cell volume, black: nuclear volume/cell volume). Video 1 is available at http://www.jcb.org/cgi/content/full/jcb.200708054/DC1.

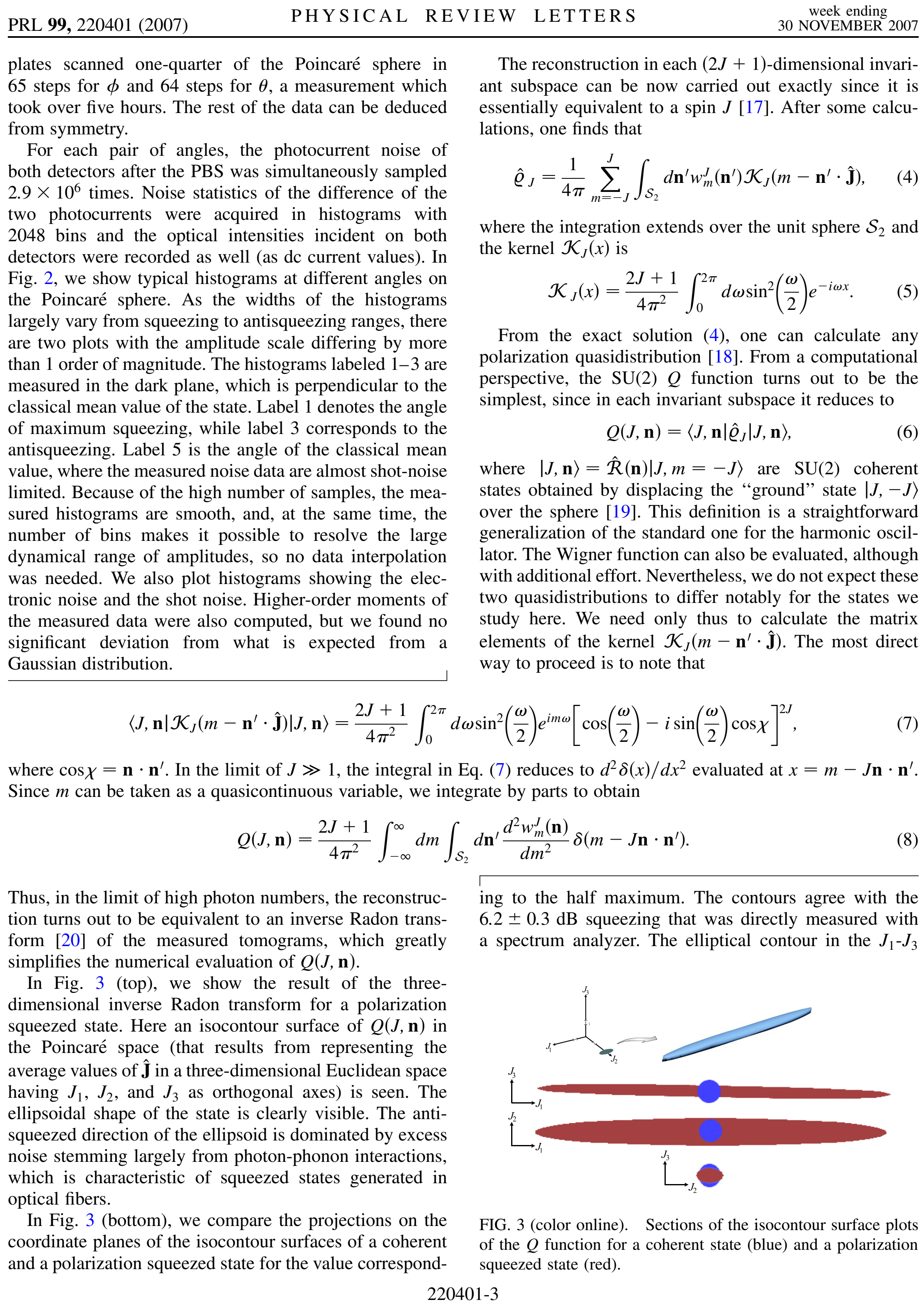
the availability and the targeting of newly formed membrane components from the ER, growth driven by volume increase would involve either nucleocytoplasmic transport or diffusion of smaller molecules through nuclear pores and sequestering within the nucleus. Two sets of experiments suggest that the NE expansion is a result rather than the cause of nuclear volume increase. First, it has been shown that NE-ER over-proliferation is not sufficient to increase nuclear size, but instead leads to an accumulation of NE sheets around the nucleus (Lum and Wright, 1995; Tange et al., 2002). Second, when blocking nuclear export of a subset of proteins for 90–150 min using leptomycin B (LMB), a specific inhibitor of the exportin crm1, nuclear size and the N/C ratio increase by 50% (Matsuyama et al., 2006; Fig. S2, available at http://www.jcb.org/cgi/content/full/jcb.200708054/DC1). This suggests that nucleocytoplasmic transport directly or indirectly alters nuclear size control, and contrasts with data from budding yeast, where 5–30 min of treatment with LMB had shown no obvious effect on nuclear size (Jorgensen et al., 2007). The differences in the results may be due to the more extended time course of drug treatment in our experiments. We further tested if the distribution of nuclear pores influences the N/C ratio. Cells deleted for *nup133b* and marked with the nucleoporin Nup107-GFP have less evenly distributed nuclear pore complexes (Bai et al., 2004), but the N/C ratio is not affected (unpublished data).

It is possible that nuclear volume could be controlled by some surrogate, such as amount of RNA or protein, numbers of ribosomes, or membrane content. In motoneurons and hepatocytes, cell size and nuclear size both correlate with the cellular RNA/DNA ratio, the expression of ribosomal genes, and general transcription rate (Sato et al., 1994; Schmidt and Schibler, 1995). Future studies will be required to dissect the molecular basis of nuclear size control in fission yeast.

A similar general cellular control that regulates nuclear growth in response to the amount of cytoplasm surrounding the nucleus may influence nuclear growth in other eukaryotes. However, differences in the cellular differentiation state and organismal developmental stage or the presence of a nuclear lamina, add more layers to N/C ratio control. Although we have shown that DNA content does not directly influence nuclear size, it might set a minimum to the size of the nucleus as suggested by the nucleoskeletal theory (Cavalier-Smith, 1982; Gregory, 2005), especially in small cells such as spores. For example, whereas wild-type spores have an N/C ratio of 0.076 ± 0.016 (see Fig. 1, A, B, and D), *wee1^{ts}* spores have a 20% smaller cell size but only 8% smaller nuclei, indicating that a minimal nuclear size may have been reached (*wee1-50/wee1-50*, $n = 136$, N/C = 0.089 ± 0.017).

Nuclear size regulation could be influenced by several cellular functions such as nucleocytoplasmic transport, lipid

Downloaded from www.jcb.org on November 30, 2007



plates scanned one-quarter of the Poincaré sphere in 65 steps for ϕ and 64 steps for θ, a measurement which took over five hours. The rest of the data can be deduced from symmetry.

For each pair of angles, the photocurrent noise of both detectors after the PBS was simultaneously sampled 2.9×10^6 times. Noise statistics of the difference of the two photocurrents were acquired in histograms with 2048 bins and the optical intensities incident on both detectors were recorded as well (as dc current values). In Fig. 2, we show typical histograms at different angles on the Poincaré sphere. As the widths of the histograms largely vary from squeezing to antisqueezing ranges, there are two plots with the amplitude scale differing by more than 1 order of magnitude. The histograms labeled 1–3 are measured in the dark plane, which is perpendicular to the classical mean value of the state. Label 1 denotes the angle of maximum squeezing, while label 3 corresponds to the antisqueezing. Label 5 is the angle of the classical mean value, where the measured noise data are almost shot-noise limited. Because of the high number of samples, the measured histograms are smooth, and, at the same time, the number of bins makes it possible to resolve the large dynamical range of amplitudes, so no data interpolation was needed. We also plot histograms showing the electronic noise and the shot noise. Higher-order moments of the measured data were also computed, but we found no significant deviation from what is expected from a Gaussian distribution.

The reconstruction in each $(2J + 1)$-dimensional invariant subspace can be now carried out exactly since it is essentially equivalent to a spin J [17]. After some calculations, one finds that

$$\hat{\varrho}_J = \frac{1}{4\pi} \sum_{m=-J}^{J} \int_{\mathcal{S}_2} d\mathbf{n}' w_m^J(\mathbf{n}') \mathcal{K}_J(m - \mathbf{n}' \cdot \hat{\mathbf{J}}), \tag{4}$$

where the integration extends over the unit sphere $\mathcal{S}_2$ and the kernel $\mathcal{K}_J(x)$ is

$$\mathcal{K}_J(x) = \frac{2J+1}{4\pi^2} \int_0^{2\pi} d\omega \sin^2\left(\frac{\omega}{2}\right) e^{-i\omega x}. \tag{5}$$

From the exact solution (4), one can calculate any polarization quasidistribution [18]. From a computational perspective, the SU(2) Q function turns out to be the simplest, since in each invariant subspace it reduces to

$$Q(J, \mathbf{n}) = \langle J, \mathbf{n} | \hat{\varrho}_J | J, \mathbf{n} \rangle, \tag{6}$$

where $|J, \mathbf{n}\rangle = \hat{\mathcal{R}}(\mathbf{n}) |J, m = -J\rangle$ are SU(2) coherent states obtained by displacing the "ground" state $|J, -J\rangle$ over the sphere [19]. This definition is a straightforward generalization of the standard one for the harmonic oscillator. The Wigner function can also be evaluated, although with additional effort. Nevertheless, we do not expect these two quasidistributions to differ notably for the states we study here. We need only thus to calculate the matrix elements of the kernel $\mathcal{K}_J(m - \mathbf{n}' \cdot \hat{\mathbf{J}})$. The most direct way to proceed is to note that

$$\langle J, \mathbf{n} | \mathcal{K}_J(m - \mathbf{n}' \cdot \hat{\mathbf{J}}) | J, \mathbf{n} \rangle = \frac{2J+1}{4\pi^2} \int_0^{2\pi} d\omega \sin^2\left(\frac{\omega}{2}\right) e^{im\omega} \left[\cos\left(\frac{\omega}{2}\right) - i \sin\left(\frac{\omega}{2}\right) \cos\chi \right]^{2J}, \tag{7}$$

where $\cos\chi = \mathbf{n} \cdot \mathbf{n}'$. In the limit of $J \gg 1$, the integral in Eq. (7) reduces to $d^2\delta(x)/dx^2$ evaluated at $x = m - J\mathbf{n} \cdot \mathbf{n}'$. Since m can be taken as a quasicontinuous variable, we integrate by parts to obtain

$$Q(J, \mathbf{n}) = \frac{2J+1}{4\pi^2} \int_{-\infty}^{\infty} dm \int_{\mathcal{S}_2} d\mathbf{n}' \frac{d^2 w_m^J(\mathbf{n})}{dm^2} \delta(m - J\mathbf{n} \cdot \mathbf{n}'). \tag{8}$$

Thus, in the limit of high photon numbers, the reconstruction turns out to be equivalent to an inverse Radon transform [20] of the measured tomograms, which greatly simplifies the numerical evaluation of $Q(J, \mathbf{n})$.

In Fig. 3 (top), we show the result of the three-dimensional inverse Radon transform for a polarization squeezed state. Here an isocontour surface of $Q(J, \mathbf{n})$ in the Poincaré space (that results from representing the average values of $\hat{\mathbf{J}}$ in a three-dimensional Euclidean space having J_1, J_2, and J_3 as orthogonal axes) is seen. The ellipsoidal shape of the state is clearly visible. The antisqueezed direction of the ellipsoid is dominated by excess noise stemming largely from photon-phonon interactions, which is characteristic of squeezed states generated in optical fibers.

In Fig. 3 (bottom), we compare the projections on the coordinate planes of the isocontour surfaces of a coherent and a polarization squeezed state for the value correspond-

ing to the half maximum. The contours agree with the 6.2 ± 0.3 dB squeezing that was directly measured with a spectrum analyzer. The elliptical contour in the J_1-J_3

FIG. 3 (color online). Sections of the isocontour surface plots of the Q function for a coherent state (blue) and a polarization squeezed state (red).

(experimental papers with the same publication date in Journal of Cell Biology & Physical Review Letters)

At the risk of saying the obvious ...

The difference between physics and biology is not just that physics “makes more use of quantitative methods” (although it does).

In physics, we are searching for an understanding of Nature that we can summarize in mathematical terms.

Mathematics is not an optional accessory, nor is it merely one tool alongside many others.

> “La filosofia è scritta in questo grandissimo libro che continuamente ci sta aperto innanzi a gli occhi (io dico l'universo), ma non si può intendere se prima non s'impara a intender la lingua, e conoscer i caratteri, né quali è scritto. Egli è scritto in lingua matematica, e i caratteri sono triangoli, cerchi, ed altre figure geometriche, senza i quali mezi è impossibile a intenderne umanamente parola; senza questi è un aggirarsi vanamente per un'oscuro laberinto.”
>
> ~ “The book of Nature is written in the language of mathematics.”
>
> (Galileo Galilei, 1623)

So what should we do? I’m involved in two projects:
- A course for first year undergraduates, combining our two goals
- A course for Physics PhD students (http://www.princeton.edu/~wbialek/PHY562.html)

Boundary conditions on a new freshman course

(you could make different choices!)

First, do no harm.

For the first year, we want to create an alternative to the combination of freshman physics and chemistry. Hence, a double course (like our friends in the Humanities).

While we want lots of connections to biology, we don't want the responsibility of communicating the factual content of intro bio courses (save this for sophomores).

All relevant departments need to agree that we have delivered the equivalent of freshman physics and chemistry (+ a little CS) at some level.

Thus, students from our course will have access to the full range of majors.

As in our physics courses for majors, we build on previous mathematical experience, but will teach some of what we need as we go along.

We simplify our problem by taking students who have had a calculus course
at the level of AP Calculus BC.

You can't satisfy the boundary conditions without genuine collaboration among the departments.

We had the good fortune to have an interested group of faculty from all the relevant departments.
We worked from a "zero base budget."

Faculty (so far, not all at once!)

W Bialek (physics)
CG Callan (physics)
D Botstein (molecular biology)
EA Carter (chemistry/engineering)
B Chazelle (computer science)
JT Groves (chemistry)
M Hecht (chemistry)
L Hodges (teaching center)
L Kruglyak (evolutionary biology)
D Marlow (physics)
R Prud'homme (chem eng)
J Rabinowtiz (chemistry)
C Schutt (chemistry)
J Shaevitz (physics)
O Troyanskaya (computer science)
EF Wieschaus (molecular biology)
NS Wingreen (molecular biology)

Lewis-Sigler fellows

T Calhoun, M Dunham, E Pearlstein, WS Ryu & EM Schötz (experimentalists)
C Broedersz, M Desai, J England & M Kaschube (theorists)

Plus … <u>many</u> teaching assistants from all departments

Freshman physics topics

Newtonian mechanics
Electricity and magnetism (up to Maxwell)
Waves
Thermodynamics and a little statistical physics
(sometimes) “Modern physics”

Freshman chemistry topics

Thermodynamics and chemical equilibrium
Reaction kinetics
A tour of the periodic table
Chemical potential, electrochemistry, ...
Orbitals, bonds, ...

There are obvious commonalities, and some more subtle relationships through the common mathematical structures

Can we organize around these more general ideas?

What kinds of mathematical structures do we use in describing nature?

Functional relations

$V = IR$, $Q = CV$, $F = -kx$, $F = -\gamma v$, ...

Dynamical models (differential equations)

Elements of classical mechanics (more viscosity than usual!), chemical kinetics (including enzymes, approximations), ... stability and response in genetic switches, resonance in the cell membrane, ...

Probabilistic models

Boltzmann distribution, connections to thermodynamics (more complex examples, e.g. protein folding), but also genetics, ...

Fields

Electricity and magnetism, but also diffusion, ... pattern formation in development, ...

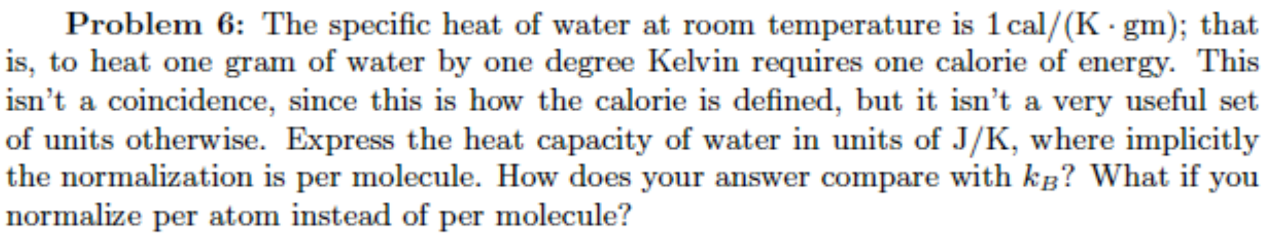
The quantum world

(example) Six weeks on probabilistic models

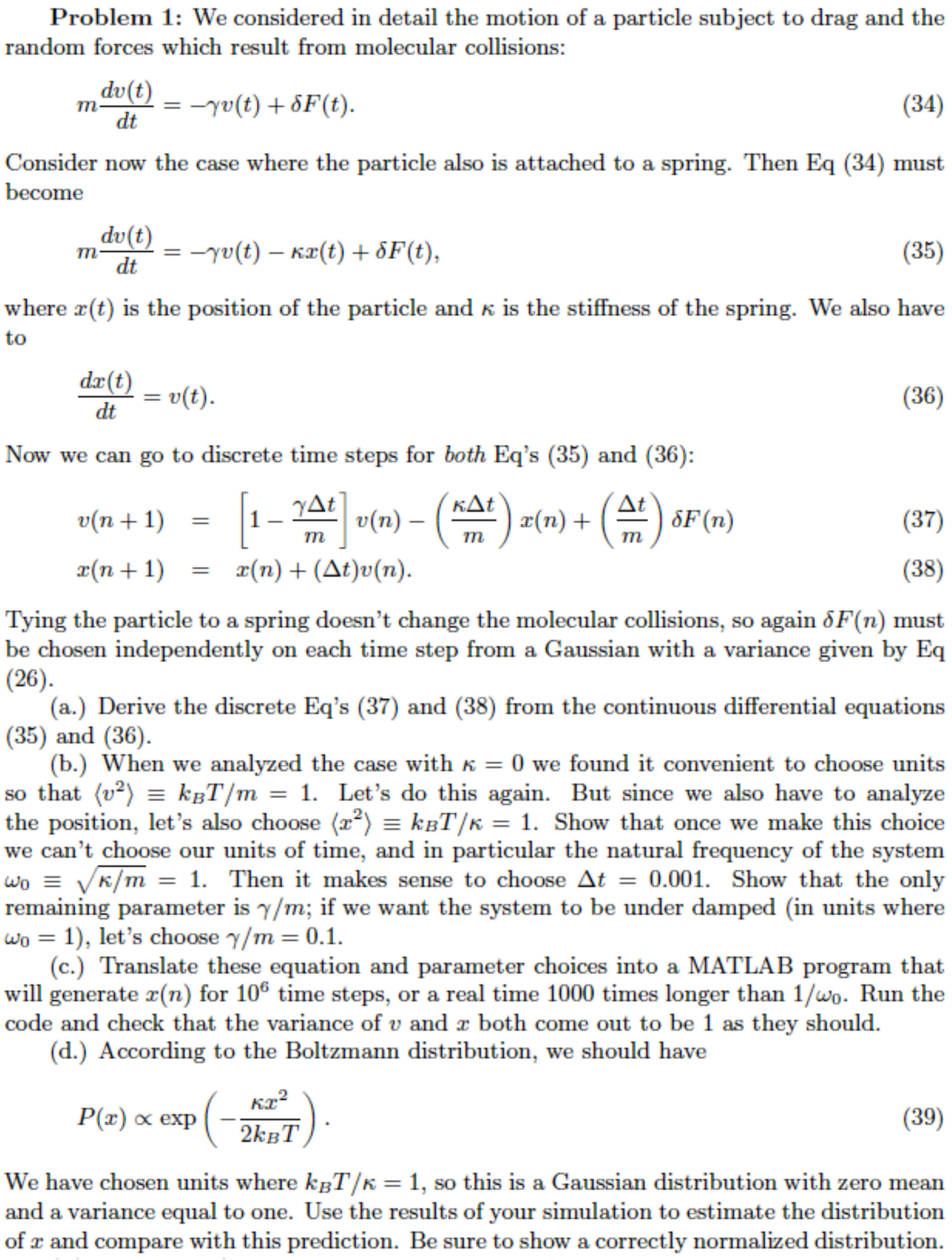
Genes, combinations and probability
    (with some inference)
Gas laws and the Boltzmann distribution
Brownian motion and the reality of molecules
Chemical equilibria and thermodynamics
Entropy, from Carnot to Shannon

We aimed high: mathematical sophistication at the level of our honors physics courses, but with more computing.

But also aim at order-of-magnitude reasoning.



Problem 6: The specific heat of water at room temperature is $1\,\text{cal}/(\text{K}\cdot\text{gm})$; that is, to heat one gram of water by one degree Kelvin requires one calorie of energy. This isn't a coincidence, since this is how the calorie is defined, but it isn't a very useful set of units otherwise. Express the heat capacity of water in units of J/K, where implicitly the normalization is per molecule. How does your answer compare with k_B? What if you normalize per atom instead of per molecule?



Problem 1: We considered in detail the motion of a particle subject to drag and the random forces which result from molecular collisions:

$$m\frac{dv(t)}{dt} = -\gamma v(t) + \delta F(t). \tag{34}$$

Consider now the case where the particle also is attached to a spring. Then Eq (34) must become

$$m\frac{dv(t)}{dt} = -\gamma v(t) - \kappa x(t) + \delta F(t), \tag{35}$$

where $x(t)$ is the position of the particle and κ is the stiffness of the spring. We also have to

$$\frac{dx(t)}{dt} = v(t). \tag{36}$$

Now we can go to discrete time steps for *both* Eq's (35) and (36):

$$\begin{aligned} v(n+1) &= \left[1 - \frac{\gamma\Delta t}{m}\right] v(n) - \left(\frac{\kappa\Delta t}{m}\right) x(n) + \left(\frac{\Delta t}{m}\right)\delta F(n) && (37)\\ x(n+1) &= x(n) + (\Delta t) v(n). && (38)\end{aligned}$$

Tying the particle to a spring doesn't change the molecular collisions, so again $\delta F(n)$ must be chosen independently on each time step from a Gaussian with a variance given by Eq (26).

(a.) Derive the discrete Eq's (37) and (38) from the continuous differential equations (35) and (36).

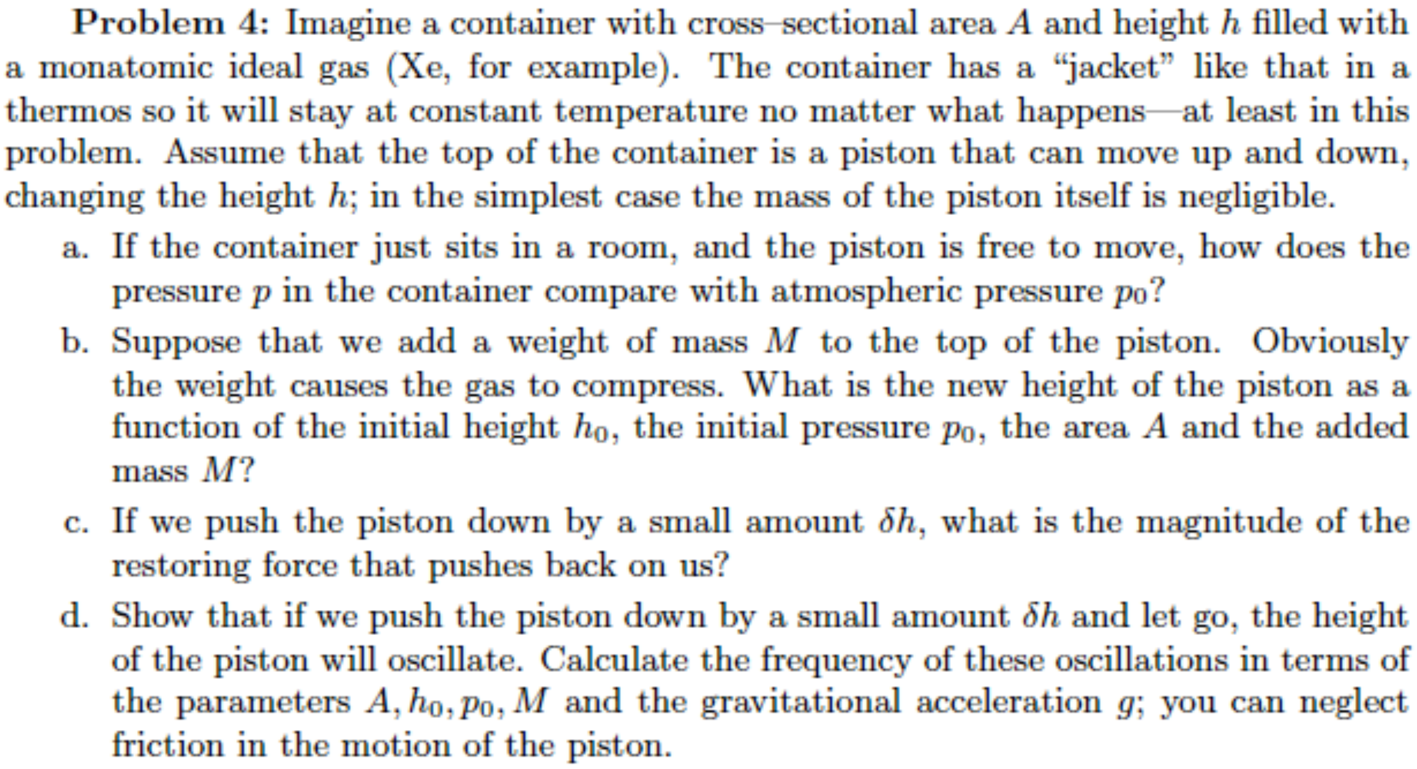
(b.) When we analyzed the case with $\kappa = 0$ we found it convenient to choose units so that $\langle v^2\rangle \equiv k_BT/m = 1$. Let's do this again. But since we also have to analyze the position, let's also choose $\langle x^2\rangle \equiv k_BT/\kappa = 1$. Show that once we make this choice we can't choose our units of time, and in particular the natural frequency of the system $\omega_0 \equiv \sqrt{\kappa/m} = 1$. Then it makes sense to choose $\Delta t = 0.001$. Show that the only remaining parameter is γ/m; if we want the system to be under damped (in units where $\omega_0 = 1$), let's choose $\gamma/m = 0.1$.

(c.) Translate these equation and parameter choices into a MATLAB program that will generate $x(n)$ for 10^6 time steps, or a real time 1000 times longer than $1/\omega_0$. Run the code and check that the variance of v and x both come out to be 1 as they should.

(d.) According to the Boltzmann distribution, we should have

$$P(x) \propto \exp\left(-\frac{\kappa x^2}{2k_BT}\right). \tag{39}$$

We have chosen units where $k_BT/\kappa = 1$, so this is a Gaussian distribution with zero mean and a variance equal to one. Use the results of your simulation to estimate the distribution of x and compare with this prediction. Be sure to show a correctly normalized distribution.



Problem 4: Imagine a container with cross–sectional area A and height h filled with a monatomic ideal gas (Xe, for example). The container has a "jacket" like that in a thermos so it will stay at constant temperature no matter what happens—at least in this problem. Assume that the top of the container is a piston that can move up and down, changing the height h; in the simplest case the mass of the piston itself is negligible.

a. If the container just sits in a room, and the piston is free to move, how does the pressure p in the container compare with atmospheric pressure p_0?

b. Suppose that we add a weight of mass M to the top of the piston. Obviously the weight causes the gas to compress. What is the new height of the piston as a function of the initial height h_0, the initial pressure p_0, the area A and the added mass M?

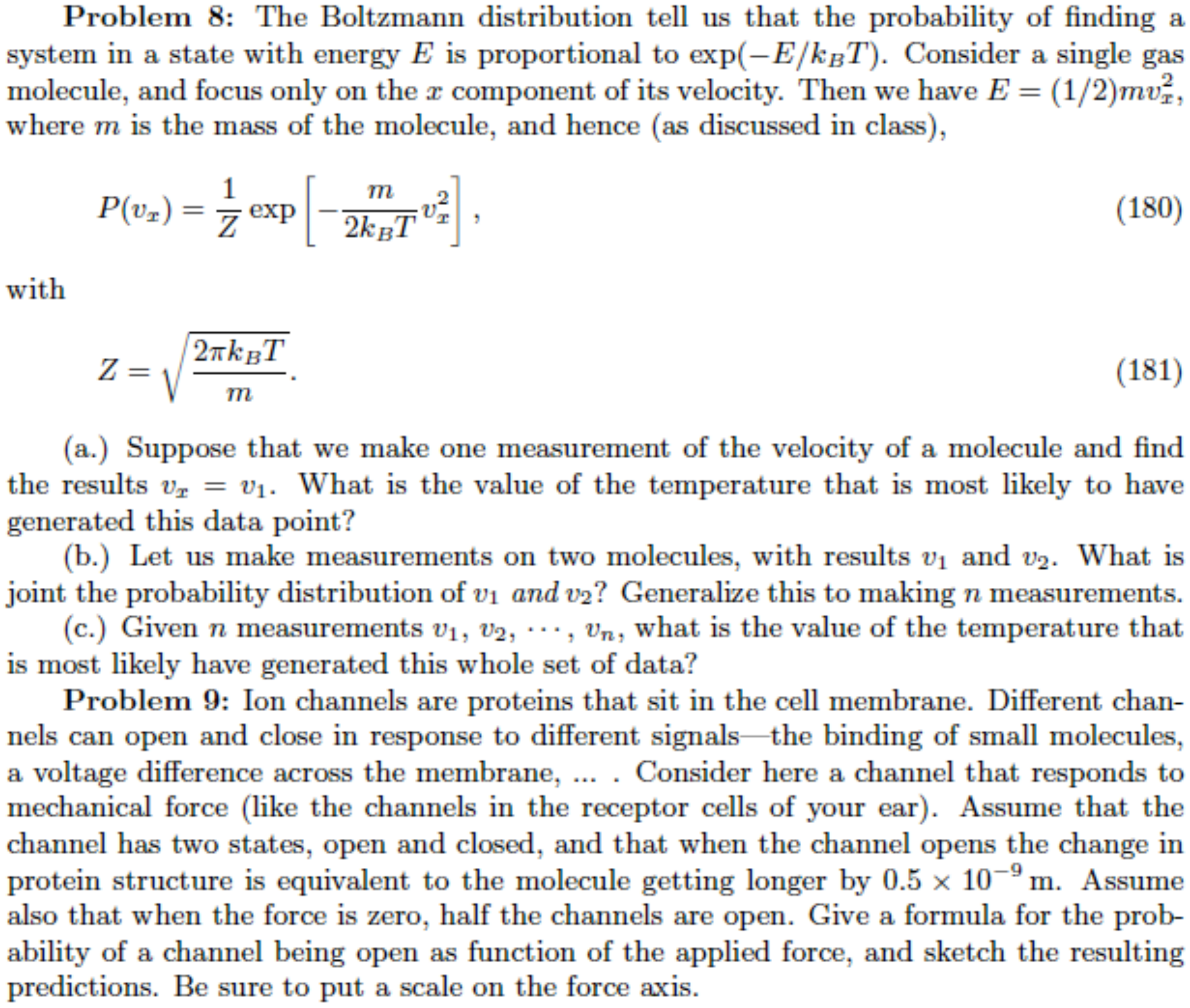
c. If we push the piston down by a small amount δh, what is the magnitude of the restoring force that pushes back on us?

d. Show that if we push the piston down by a small amount δh and let go, the height of the piston will oscillate. Calculate the frequency of these oscillations in terms of the parameters A, h_0, p_0, M and the gravitational acceleration g; you can neglect friction in the motion of the piston.

Connect across modules of the course.

Integrate statistical physics with probabilistic inference.

Encourage thinking about biological systems with simple models.



Problem 8: The Boltzmann distribution tell us that the probability of finding a system in a state with energy E is proportional to $\exp(-E/k_BT)$. Consider a single gas molecule, and focus only on the x component of its velocity. Then we have $E = (1/2)mv_x^2$, where m is the mass of the molecule, and hence (as discussed in class),

$$P(v_x) = \frac{1}{Z}\exp\left[-\frac{m}{2k_BT}v_x^2\right], \tag{180}$$

with

$$Z = \sqrt{\frac{2\pi k_BT}{m}}. \tag{181}$$

(a.) Suppose that we make one measurement of the velocity of a molecule and find the results $v_x = v_1$. What is the value of the temperature that is most likely to have generated this data point?

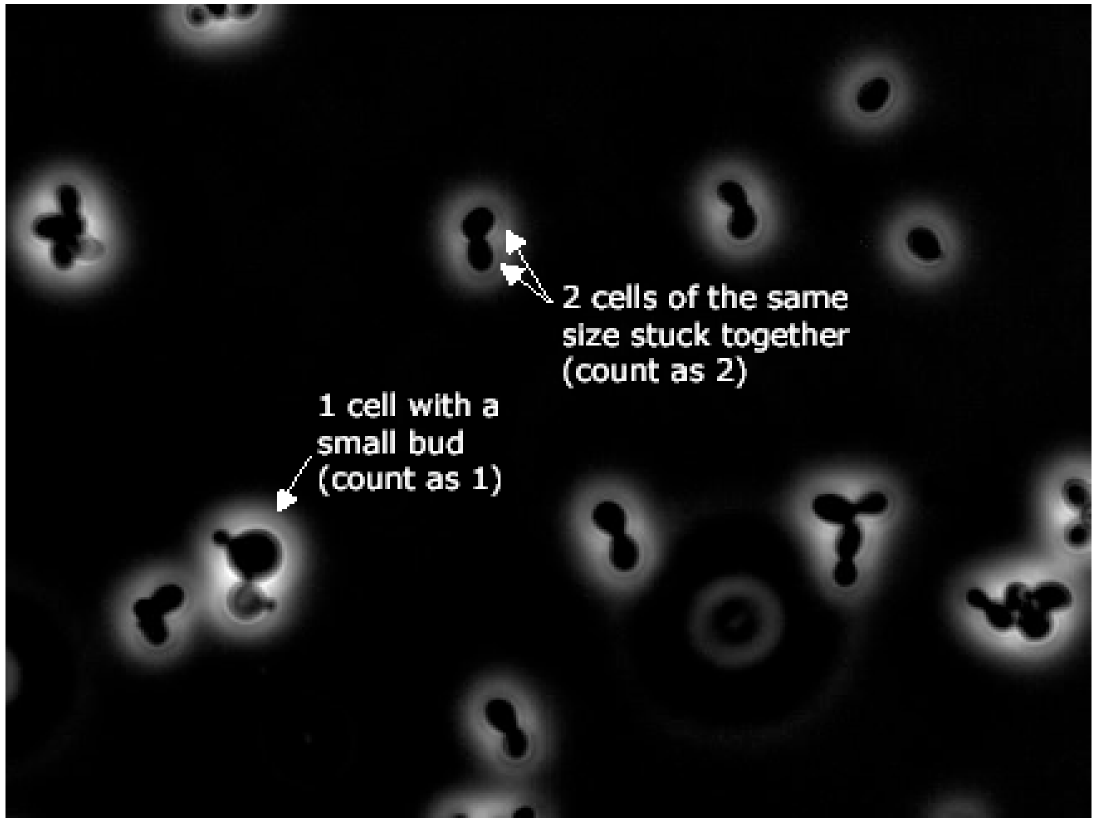
(b.) Let us make measurements on two molecules, with results v_1 and v_2. What is joint the probability distribution of v_1 *and* v_2? Generalize this to making n measurements.

(c.) Given n measurements $v_1, v_2, \cdots, v_n$, what is the value of the temperature that is most likely have generated this whole set of data?

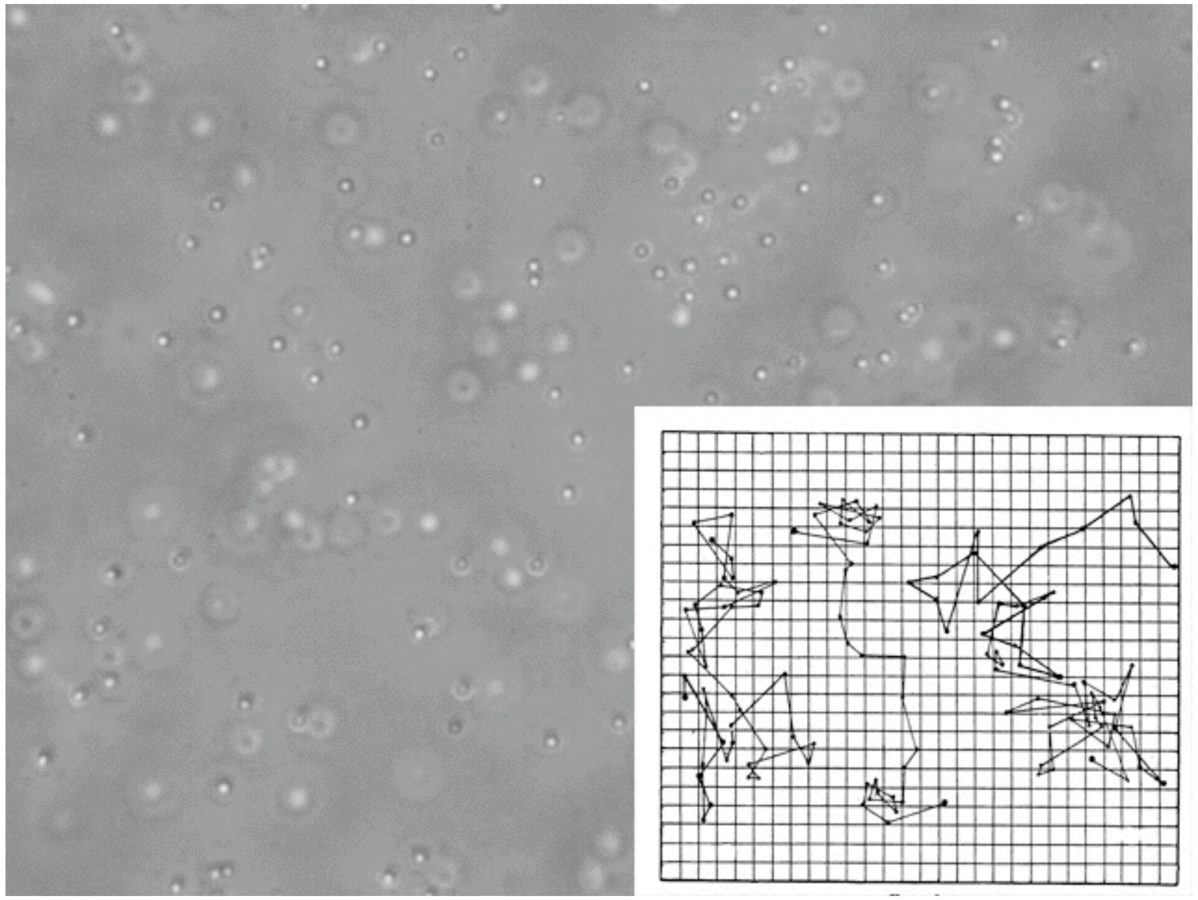
Problem 9: Ion channels are proteins that sit in the cell membrane. Different channels can open and close in response to different signals—the binding of small molecules, a voltage difference across the membrane, ... Consider here a channel that responds to mechanical force (like the channels in the receptor cells of your ear). Assume that the channel has two states, open and closed, and that when the channel opens the change in protein structure is equivalent to the molecule getting longer by 0.5×10^{-9} m. Assume also that when the force is zero, half the channels are open. Give a formula for the probability of a channel being open as function of the applied force, and sketch the resulting predictions. Be sure to put a scale on the force axis.

Meanwhile, in the lab ...

Delbrück-Luria experiment with yeast



we try to keep labs, lectures, problem sets, ... all tied together



Direct measurements of Brownian motion

How well is it working?

Fall 2011: 50+ students, reasonably stable year-to-year.
Semester-to-semester attrition is low (~10%).
Consistently > 1/3 women.

1/3 go to physical sciences
1/3 go to biological sciences
1/6 go to comp sci + engineering
1/6 scatter across campus

22% of physics majors come through this course
(small numbers, but more or less uniformly distributed performance)

huge impact on biology majors

The pioneers (now, mostly, PhD students)

and this is <u>after</u> the final exam