Physics 446-546, Problem Set 6
(issued Apr 26, '05, due 5 May '05)

#1

5.4 *Coasting at low Reynolds*

The chapter asserted that tiny objects stop essentially at once when we stop pushing them. Let's see.

a. Consider a bacterium, idealized as a sphere of radius 1 μm, propelling itself at $1\,\mu\mathrm{m\,s^{-1}}$. At time zero the bacterium suddenly stops swimming and coasts to a stop, following Newton's Law of motion with the Stokes drag force. How far does it travel before it stops? Comment.

b. Our discussion of Brownian motion assumed that each random step was independent of the previous one; thus for example we neglected the possibility of a residual drift speed left over from the previous step. In the light of (a), would you say that this assumption is justified for a bacterium?

#2

5.9 **T2** *Friction as diffusion*

Section 5.2.1′ on page 166 claimed that viscous friction can be interpreted as the diffusive transport of momentum. The argument was that in the planar geometry, when the flux of momentum given by Equation 5.20 leaves the top plate it exerts a resisting drag force; when it arrives at the bottom plate it exerts an entraining force. So far the argument is quite correct.

Actually, however, viscous friction is more complicated than ordinary diffusion, because momentum is a vector quantity whereas concentration is a scalar. For example, Section 5.2.2 noted that the viscous force law (Equation 5.9 on page 149) needs to be modified for situations other than planar geometry. The required modification really matters if we want to get the correct answer for the spinning-rod problem (Figure 5.11b on page 160).

We consider a long cylinder of radius R with its axis along the $\hat{\mathbf{z}}$ direction and centered at $x = y = 0$. Some substance surrounds the cylinder. First suppose that this substance is *solid ice*. When we crank the cylinder, everything rotates as a rigid object with some angular frequency ω. The velocity field is then $\mathbf{v}(\mathbf{r}) = (-\omega y, +\omega x, 0)$. Certainly nothing is rubbing against anything, and there should be no dissipative friction—the frictional transport of momentum had better be zero. And yet if we examine the point $\mathbf{r}_0 = (r_0, 0, z)$ we find a nonzero gradient $\left.\frac{\mathrm{d}v_y}{\mathrm{d}x}\right|_{\mathbf{r}_0} = \omega$. Evidently our formula for the flux of momentum in planar geometry (Equation 5.20 on page 166) needs some modification for the non-planar case.

We want a modified form of Equation 5.20 that applies to cylindrically symmetric flows and vanishes when the flow is rigid rotation. Letting $r = \|\mathbf{r}\| = \sqrt{x^2 + y^2}$, we can write a cylindrically symmetric flow as

$$\mathbf{v}(\mathbf{r}) = (-yg(r), xg(r), 0)\,.$$

The case of rigid rotation corresponds to the choice $g(r) = \omega$. You are about to find $g(r)$ for a different case, the flow set up by a rotating cylinder. We can think of this flow field as a set of nested cylinders, each with a different angular velocity $g(r)$.

Near any point, say $\mathbf{r}_0$, let $\mathbf{u}(\mathbf{r}) = (-yg(r_0), xg(r_0)))$ be the rigidly rotating vector field that agrees with $\mathbf{v}(\mathbf{r})$ at $\mathbf{r}_0$. We then replace Equation 5.20 by

$$(j_{p_y})_x(\mathbf{r}_0) = -\eta \left(\left.\frac{\mathrm{d}v_y}{\mathrm{d}x}\right|_{\mathbf{r}_0} - \left.\frac{\mathrm{d}u_y}{\mathrm{d}x}\right|_{\mathbf{r}_0} \right). \qquad \text{cylindrical geometry} \qquad (5.21)$$

In this formula $\eta \equiv \nu\rho_{\mathrm{m}}$, the ordinary viscosity. Equation 5.21 is the proposed modification of the momentum-transport rule. It says that we compute $\left.\frac{\mathrm{d}v_y}{\mathrm{d}x}\right|_{\mathbf{r}_0}$ and *subtract off* the corresponding quantity with $\mathbf{u}$, in order to ensure that rigid rotation incurs no frictional resistance.

a. Each cylindrical shell of fluid exerts a torque on the next one, and feels a torque from the previous one. These torques must balance. Show that therefore the tangential force per area across the surface at fixed r is $\frac{\tau/L}{2\pi r^2}$, where τ is the external torque on the central cylinder and L is the cylinder's length.

b. Set your result from (a) equal to Equation 5.21 and solve for the function $g(r)$.

c. Find τ/L as a constant times ω. Hence find the constant C in Equation 5.18 on page 163.

#3

(a) In class, we discussed the "occasionally dishonest casino" that used two kinds of dice: 99% were fair, but 1% were loaded so that a 6 came up 50% of the time. Thus $p(L) = 1/100$, and the conditional probabilities are $p(6|L) = 1/2$ and $p(6|\overline{L}) = 1/6$.

If we then pick a die at random, what are the joint probabilities $p(6, L)$ and and $p(6, \overline{L})$? What is the probability of rolling a 6 from the die we picked up?

If we rolled three 6's in a row, we saw that the posterior probability $p(L|6^3)$ that it was loaded was only 3/14. How many sixes in a row would we have to roll before concluding it was more likely to have been a loaded die? For what n does the probability $P(L|6^n)$, that the die is loaded given n consecutive sixes, begin to exceed 90%?

(b) We also discussed the case of Duchenne Muscular Dystrophy (DMD), regarded as a simple recessive sex-linked disease caused by a mutated X chromosome ($\overline{\mathrm{X}}$). An $\overline{\mathrm{X}}$Y male expresses the disease, whereas an $\overline{\mathrm{X}}$X female is a carrier but does not express the disease. If neither of a woman's parents expresses the disease, but her brother does, then the woman's mother must be a carrier, and the woman herself has an *a priori* 50/50 chance of being a carrier, $p(C) = 1/2$. Suppose she proceeds to give birth to n healthy sons (n h.s.). What now is her probability $p(C|n \text{ h.s.})$ of being a carrier?

(c) In clase, we touched upon the question of administering lie detector tests at a hypothetical national laboratory. We generously assumed that these tests have a 90% sensitivity, i.e., the probability of a spy failing the test is $p(f|S) = .9$, and equally generously assumed that the tests have a false positive rate of only 20%, i.e., the probability of a non-spy failing the test is $p(f|\overline{S}) = .2$. We assume that roughly one in a thousand laboratory employees is a spy, $p(S) = 10^{-3}$.

What is the probability $p(S|f)$ that someone who has failed the test is a spy?

Suppose someone takes the test 10 times and fails it 9 out of those ten times. What is the probability $p(S|9f + 1\overline{f})$ of being a spy?

#4 (from Bialek notes, physics/0205030, p. 32) [OPTIONAL, but easy]

Maximally informative experiments. Imagine that we are trying to gain information about the correct theory T describing some set of phenomena. At some point, our relative confidence in one particular theory is very high; that is, $P(T = T_*) > F \cdot P(T \neq T_*)$ for some large F. On the other hand, there are many possible theories, so our absolute confidence in the theory T_* might nonetheless be quite low, $P(T = T_*) << 1$. Suppose we follow the 'scientific method' and design an experiment that has a yes or no answer, and this answer is perfectly correlated with the correctness of theory T_*, but uncorrelated with the correctness of any other possible theory—our experiment is designed specifically to test or falsify the currently most likely theory. What can you say about how much information you expect to gain from such a measurement? Suppose instead that you are completely irrational and design an experiment that is irrelevant to testing T_* but has the potential to eliminate many (perhaps half) of the alternatives. Which experiment is expected to be more informative? Although this is a gross cartoon of the scientific process, it is not such a terrible model of a game like "twenty questions."

(It is interesting to ask whether people play such question games following strategies that might seem irrational but nonetheless serve to maximize information gain [Ginzburg I., & Sejnowski, T. J. (1996). Dynamics of Rule Induction by Making Queries: Transition Between Strategies, in *18th Annual Conference of the Cognitive Science Society,* pp. 121–125 (Lawrence Erlbaum, Mahwah NJ). See also http://www.cnl.salk.edu/CNL/annual–reps/annual–rep95.html]. Related but distinct criteria for optimal experimental design have been developed in the statistical literature [Fedorov, V. V. (1972). *Theory of Optimal Experimental Design,* translated and edited by Studden, W. J., & Klimko, E. M. (Academic Press, New York)].)

#5 (from Bialek notes, physics/0205030, p. 33)

Positivity of information.

$$I(D;W) = S(W) - \sum_d p(d) S(W|d) \tag{43}$$

$$= \sum_w \sum_d p(w,d) \log_2 \left[\frac{p(w,d)}{p(w)p(d)} \right], \tag{44}$$

where $S(W) = -\sum_w p(w) \log_2 p(w)$ and $S(W|d) = -\sum_w p(w|d) \log_2 p(w|d)$.

a) Prove the above formula and show that the mutual information $I(D;W)$ is positive.

(Hint: prove and use the inequality: $\ln z < z - 1$.)

b) The usual information theoretic notation for these formulae is as follows. For two random variables x, y with probability distributions, $p(x)$, $x \in X$ and $p(y)$, $y \in Y$, and joint probability distribution $p(x,y)$, we write the entropies $H(X) = -\sum_x p(x) \log_2 p(x)$, $H(Y) = -\sum_y p(y) \log_2 p(y)$, and the joint entropy $H(X,Y) = -\sum_{x,y} p(x,y) \log_2 p(x,y)$.

The conditional entropy is then defined as $H(X|Y) \equiv H(X,Y) - H(Y)$, and the average mutual information as $I(X;Y) \equiv H(X) - H(X|Y)$.

Show that this formula for I agrees with the above for $X = D$, $Y = W$. Show that $I(X;Y) = H(Y) - H(Y|X)$ and also show the fully symmetric form $I(X;Y) = H(X) + H(Y) - H(X,Y)$.

Write down the straightforward generalization of the last formula for the case of n variables $I(X_1; X_2; \ldots; X_n)$, giving a natural measure of their dependence (or equivalently the reduction in potential code length considered as a vector $\vec{X}$ rather than n independent variables).

#6 (from Bialek notes, physics/0205030, p. 35)

> The problem of finding the maximum entropy given some constraint again is familiar from statistical mechanics: the Boltzmann distribution is the distribution that has the largest possible entropy given the mean energy. More generally, let us imagine that we have knowledge not of the whole probability distribution $P(d)$ but only of some expectation values,
>
> $$\langle f_{\rm i} \rangle = \sum_d P(d) f_{\rm i}(d),$$
>
> where we allow that there may be several expectation values known $({\rm i} = 1, 2, ..., K)$. Actually there is one more expectation value that we always know, and this is that the average value of one is one; the distribution is normalized:
>
> $$\langle f_0 \rangle = \sum_d P(d) = 1.$$
>
> Given the set of numbers $\{\langle f_0 \rangle, \langle f_1 \rangle, \cdots, \langle f_K \rangle\}$ as constraints on the probability distribution $P(d)$, we would like to know the largest possible value for the entropy, and we would like to find explicitly the distribution that provides this maximum.
>
> The problem of maximizing a quantity subject to constraints is formulated using Lagrange multipliers. The result is that
>
> $$P(d) = \frac{1}{Z} \exp\left[-\sum_{{\rm i}=1}^{K} \lambda_{\rm i} f_{\rm i}(d) \right], \tag{53}$$
>
> where $Z = \exp(1 + \lambda_0)$ is a normalization constant.
>
> **Details.** Derive Eq. (53). In particular, show that Eq. (53) provides a probability distribution which genuinely *maximizes* the entropy, rather than being just an extremum.

(Recall "Your Turn 6G" on p.224 of the course text, and 6.1′ on p.232, for a similar derivation of the Boltzmann distribution.)