Theory of continuum random walks and application to chemotaxis

Mark J. Schnitzer*
Lyman Laboratory of Physics, Harvard University, Cambridge, Massachusetts 02138
(Received 30 October 1992)

We formulate the general theory of random walks in continuum, essential for treating a collision rate which depends smoothly upon direction of motion. We also consider a smooth probability distribution of correlations between the directions of motion before and after collisions, as well as orientational Brownian motion between collisions. These features lead to an effective Smoluchowski equation. Such random walks involving an infinite number of distinct directions of motion cannot be treated on a lattice, which permits only a finite number of directions of motion, nor by Langevin methods, which make no reference to individual collisions. The effective Smoluchowski equation enables a description of the biased random walk of the bacterium *Escherichia coli* during chemotaxis, its search for food. The chemotactic responses of cells which perform temporal comparsions of the concentration of a chemical attractant are predicted to be strongly positive, whereas those of cells which measure averages of the ambient attractant concentration are predicted to be negative. The former prediction explains the observed behavior of wild-type (naturally occurring) cells; however, the latter behavior has yet to be observed, even in cells defective in adaption.

PACS number(s): 05.40.+j, 87.22.Nf, 87.10.+e

I. INTRODUCTION

How do we describe random walkers whose collision rate depends continuously on position and *direction of motion?* Although this question is within the realm of statistical physics, the investigation presented here was conducted with the aim of understanding the random walks of the bacterium *Escherichia coli* during chemotaxis, the search for chemical attractants. The particular cases we examine are chosen with *E. coli* in mind, but our derivations make no explicit reference to bacteria and apply to any particles that obey our assumptions. After we have established the theoretical foundations, we provide applications.

When a random walker has a different collision rate for each of the infinite possible directions of motion in two or more dimensions, the walk defies a lattice treatment, which involves only a finite number of directions of motion. For instance, if we try to model a continuum random walk with a direction-dependent collision rate on a two-dimensional square lattice, even if we correctly specify the collision rates for movements left, right, up, and down, knowledge of the collision rates for all other intermediate directions is lost. We also consider here two other features that must be modeled in continuum: correlations between the directions of motion before and after collisions, and rotational Brownian motion, which causes small gradual changes in direction of motion between collisions. An attempted lattice treatment of these two features would also suffer from the lack of intermediate directions. The former effect is a generalization of the "persistence" effect, studied by Goldstein [1], and Weis and collaborators [2]. Previous works have also treated correlations between successive directions of motion when the collision rate is constant [3,4].

The random walks considered here also defy a Langevin treatment, in which no explicit mention is made of distinct collisions. In the Langevin approach, the macroscopic effect of collisions is summarized through the inclusion of a random force; this technique will not suffice if microscopic features of individual collisions are to be included, as they are here. In contrast, the treatment here starts at smaller length scales and shorter time intervals; particles head in ballistic trajectories, and have specified collision rates. We have no need for a Langevin term in our equations, because we derive the effect of collisions, rather than summarize their effects through a random force. This alternative approach to diffusion involving telegraph equations originated with Goldstein [1] and Taylor [5], and was further studied by Kac [6] and Clayton [7]. The approach developed by these authors naturally generalizes to permit a description of walks involving direction-dependent collision rates, rotational Brownian motion, and correlations between successive directions of motion.

The continuum formulation has another important advantage over the lattice formulation. To derive diffusive motion on a lattice with characteristic length δ and time increment τ, one generally takes the unphysical limit $\delta^2/2\tau \to D$, the diffusion coefficient, as $\delta \to 0$ and $\tau \to 0$. However, in this limit the ballistic particle velocity becomes infinite, and it is no longer possible to explicitly relate a position-dependent ballistic velocity to the diffusion coefficient and the drift velocity. In continuum, this unphysical limit is avoided by looking, instead, at the long-time limit of the ballistic motion. An alternative approach circumvents this issue by avoiding all limits [8].

One random walker that exhibits position- and direction-dependent "collision" rate, correlations between directions of motion before and after "collisions," and rotational Brownian motion is the bacterium *E. coli,* a flagellated microorganism approximately 1 μm in diam-

eter by 2 μm long. During what are called "runs," cells of *E. coli* swim at roughly 20 μ/s in approximately straight paths, on average for about 1 s. Run durations exhibit a Poisson distribution, and are terminated by events called "tumbles." The Poisson rate constant is the "tumble rate," whose reciprocal is the mean run duration. In contrast to runs, during which the cell's flagella rotate counterclockwise synchronously, during tumbles the flagella rotate clockwise asynchronously. During tumbles, which last on average about 0.1 s, the bacterium's position is essentially fixed, but its orientation changes. When the cell finishes tumbling and resumes running, it swims off in a new direction chosen approximately at random. Such alternation between runs and tumbles produces a random walk in which tumbles are analogous to the collisions of molecular diffusion.

The random walk of *E. coli* is further characterized by several distinct features. To start, the new direction of motion after a tumble is more likely to be in the forward hemisphere: the distribution of angles between the new and the previous directions of motion is peaked at about 62°. Furthermore, the paths taken during runs are not completely straight. At 300 K, collisions with water molecules involve enough kinetic energy to gradually change the cell's orientation and direction of motion. This rotational Brownian motion is characterized by a mean-square angular deviation $\langle \theta^2 \rangle$, which grows linearly in time, $\langle \theta^2 \rangle = 2D_{\text{rot}}t$. For *E. coli*, $D_{\text{rot}} \approx 0.1$ rad^2/s, so the cell's direction of motion deviates by about 90° in about 10 s. These basic facts about *E. coli* have been taken from Refs. [9–11].

Perhaps the most fascinating aspect about the random walk of *E. coli* is that it is a biased one: on average, cells tend to drift toward chemical attractants. This behavior is called chemotaxis. Experiments by Adler showed that the cell detects attractants using chemoreceptors on the cell's body surface [12]. Using a tracking microscope, Berg and Brown discovered the exponential distribution of run durations [9]. Moreover, they found that when the cell is heading in a direction of increasing attractant concentration, the cell lowers the tumble rate from the usual ≈ 1 s^{-1}. When heading in a direction of decreasing attractant concentration, the cell maintains the tumble rate at 1 s^{-1} [9,13], or perhaps raises it very slightly [14]. This direction-dependent tumble rate enables the chemotactic drift toward favorable regions of high attractant concentration.

How does an *E. coli* cell determine whether its current direction of motion is favorable? In a classic work [15], Berg and Purcell showed that it was theoretically possible for the cell to perform temporal comparisons of ambient attractant concentration in a few seconds time, in spite of microscopic fluctuations in local attractant concentration. Subsequently, such temporal comparisons were found in experiments with wild-type (naturally occurring) cells of *E. coli* tethered by a single flagellum to a glass coverslip [16,17]. It was shown that *E. coli* modulates the rotational bias, the fraction of time a single flagellum spends rotating counterclockwise, as if it were a linear system. Wild-type cells filter the attractant concentration measured by its chemoreceptors in the recent past, $C(t)$, with a characteristic biphasic impulse response, $I(t)$ (Fig. 1, left axis), designed to perform temporal comparisons. Because the positive and negative lobes of the impulse response have equal area, the wild-type cell essentially averages the attractant concentration measured over the last ~ 1 s, and subtracts the average of the ~ 3 s prior to that. If the difference is positive, the bias is raised by an amount proportional to this difference. If the difference is negative, the bias is unaffected unless this negative difference is very large [18].

These findings with tethered cells are in accord with studies of swimming cells, which report little [14] to no [9,13] rise in the tumble rate when a wild-type cell swims down an attractant gradient. In run mode the cell's flagella all rotate counterclockwise, but it is not well understood how the biases of the cell's many flagella are related to the tumble rate [19]. However, it is known that when swimming in a favorable direction the cell lowers the tumble rate by an amount proportional to dP/dt, where P is the fraction of chemoreceptors bound to attractant molecules [13]. Because

$$\frac{dP}{dt} = \frac{K_D}{(K_D + C)^2} \frac{dC}{dt} , \tag{1.1}$$

where K_D is the dissociation constant of the chemoreceptor [13], when $C \ll K_D$

$$\frac{dP}{dt} \approx K_D^{-1} \frac{dC}{dt} , \tag{1.2}$$

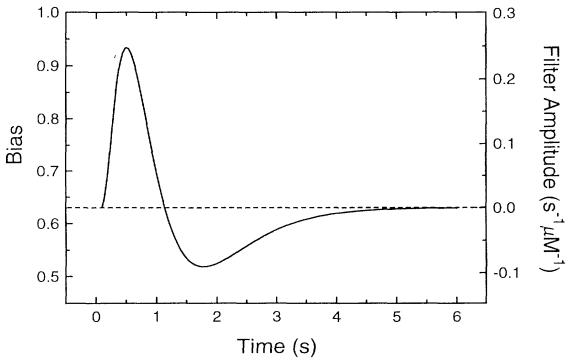

FIG. 1. Left axis: The impulse response of wild-type *E. coli*, based on a fit with six exponentials to the data of Refs. [15,16]. The bias, the fraction of time a single flagellum spends turning counterclockwise, is plotted as a function of time for cells suddenly exposed to a 1 mM increase in the attractant aspartate, delivered over 0.02 s starting at time zero. Right axis: The linear filter used by wild-type cells to detect changes in ambient attractant concentration and to modulate the tumble rate accordingly, (1.3). It is assumed that this filter has the same shape as the bias impulse response, shifted and scaled appropriately. Because the positive and negative lobes of the filter have equal area, the cell essentially averages the attractant concentration measured over the last ~ 1 s, and subtracts the average of the ~ 3 s prior to that, as a finite difference approximation to dC/dt. The scale has been chosen under the assumption that a cell swimming at 20 μm/s straight up an attractant gradient with $|\nabla C| = 0.1$ $\mu M/\mu$m lowers the tumble rate to $\alpha \approx 0.5$ s^{-1} from $\alpha_0 \approx 1$ s^{-1}.

and the tumble rate is lowered by an amount linearly proportional to dC/dt, if $dC/dt>0$. Because the cell cannot make an instantaneous measurement of dC/dt [15], an approximate measurement is presumably carried out by linear filtering with a filter $I(t)$ of the same characteristic biphasic shape shown in Fig. 1, shifted and rescaled appropriately (Fig. 1, right axis). This assumption is the simplest possible that is consistent with the data on both tethered and swimming cells and implies that the tumble rate $\alpha(t)$ is set as

$$\alpha(t)=\begin{cases}\alpha_0-\tilde{g}\int_0^\infty I(\tau)C(t-\tau)d\tau & \text{if } \int_0^\infty I(\tau)C(t-\tau)d\tau>0\\ \alpha_0 & \text{if } \int_0^\infty I(\tau)C(t-\tau)d\tau<0\ ,\end{cases} \tag{1.3}$$

when $C\ll K_D$. The gain $\tilde{g}$ sets the magnitude of the response to changes in attractant concentration.

A mutant strain (*cheRcheB*) of *E. coli* lacking the normal methylation and demethylation enzymes required for performing temporal comparisons also acts like a linear system, but its impulse response $I(t)$ appears strictly positive (Fig. 2) [16,17]. Because the *cheRcheB* impulse response has only a positive lobe, the cell raises the bias and lowers the tumble rate, (1.3), in response to the average of the attractant concentration measured over the last ~ 1 s. If a wild-type cell is suddenly moved to a new environment of higher attractant concentration, the cell initially lowers the tumble rate, but after ~ 4 s it adapts to the new environment and returns the tumble rate to the baseline value of $\approx 1\ \mathrm{s}^{-1}$. In contrast, if a *cheRcheB* cell is moved to a new environment of higher attractant concentration, it does not adapt: the tumble rate is initially lowered and then maintained at the new lower rate. But in both wild-type and nonadaptive cells, the attractant concentration measured in the recent past is used to set the tumble rate, (1.3). Because the recent past of a swimming cell may be at least partly characterized by the cell's current direction of motion, the cell's tumble rate is a function of the direction of motion, the key ingredient of the random walks addressed in this article.

In Sec. II, we start off by considering a preliminary one-dimensional random walk, in which the collision rate and particle velocity depend on position but not direction of motion, leading to four cases, each with different generalizations of Fick's law. The results rule out a successful chemotaxis in which the microorganism speeds up in response to higher local attractant concentration. In Sec. III, we examine one-dimensional random walks in which the collision rate and particle velocity depend upon both position and direction of motion. This analysis yields two general classes of successful strategies for chemotaxis: when moving in a direction of increasing attractant concentration, the cell may either lower the tumble rate, speed up, or both.

In contrast with lattice models, in continuum models it is considerably harder to generalize to dimensions higher than one. Section IV illustrates the methods of the higher-dimensional derivations by presenting what appears to be a new derivation of Fick's law. In Sec. V, we generalize to the cases in which the collision rate and the particle velocity depend continuously upon spatial position and direction of motion. Sections VI and VII consider walks with correlations between the directions of motion before and after collisions, and with rotational Brownian motion, respectively. In Sec. VIII, we examine the effects of differences between members of the particle or cell population. Precisely known collision rates, velocities, rotational diffusion coefficients, correlations between successive runs, are all replaced by probability distributions for these quantities. In Sec. IX, we apply the effective Smoluchowski equation that results from a direction-dependent tumble rate to investigate possible successful strategies for chemotaxis using tumble rate modulation. We predict the chemotactic responses of the wild-type strain of *E. coli* to be strongly positive, and those of cells defective in adaption to be negative.

From tethering experiments, some evidence exists that *cheRcheB* cells may actually adapt partially over the course of a few minutes, rather than a few seconds [17]. This evidence and observations of weak chemotactic behavior [20–22] have led to conjectures of alternative adaption systems to the normal methylation system

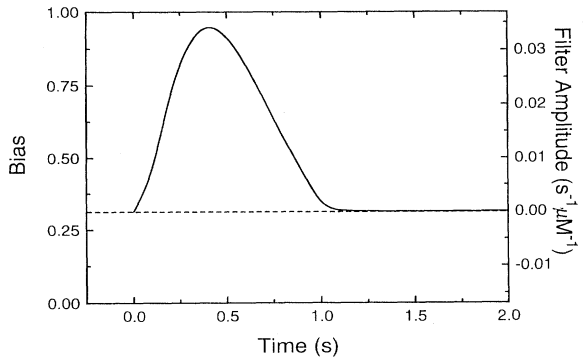

FIG. 2. Left axis: The impulse response of the *cheRcheB* strain of *E. coli*, modeled by eye from the data of Refs. [15,16]. The bias, the fraction of time a single flagellum spends turning counterclockwise, is plotted as a function of time for cells suddenly exposed to a 3–20 mM increase in the attractant aspartate, delivered over 0.01–0.39 s starting at time zero. Right axis: The linear filter used by *cheRcheB* cells to measure averages of ambient attractant concentration and to modulate the tumble rate accordingly, (1.3). It is assumed that this filter has the same shape as the bias impulse response, shifted and scaled appropriately. Because the filter has only a positive lobe, the cells cannot adapt. The scale has been chosen under the assumption that a cell exposed to an ambient attraction concentration of 100 mM lowers the tumble rate to $\alpha\approx 1\ \mathrm{s}^{-1}$ from $\alpha_0\approx 3\ \mathrm{s}^{-1}$.

[21,22]. In Sec. X, we examine whether it is actually necessary to invoke an alternative to methylation adaption to explain the observed weak chemotaxis of *cheRcheB* cells.

II. A SIMPLE EXAMPLE

If the diffusion coefficient D becomes spatially dependent, $D(\mathbf{x})$, how does Fick's law,

$$\mathbf{J}=-D\nabla\rho\ , \tag{2.1}$$

relating the particle flux $\mathbf{J}$ to the particle concentration ρ generalize? Is D simply replaced by $D(\mathbf{x})$, or does $-D\nabla\rho$ become $-\nabla(D\rho)$? As we shall now show, this question is badly posed, because in different cases there are different answers. Deciding upon the correct answer for a given situation requires partial knowledge of the microscopic mechanism for diffusion. For an alternative treatment and for discussion of particular confusions in the literature, see Ref. [8].

To illustrate that no single generalization of (2.1) exists, consider two identically sized closed boxes, each containing a very dilute gas. Each box also contains a dense mesh of iron wool, so dense that the free path length between adjacent strands of iron wool is the limiting factor on particle diffusion. If the two microscopic parameters, $\delta(\mathbf{x})$, the free path length, and $\tau(\mathbf{x})$, the free time interval, vary sufficiently slower over space, (i.e., $\delta\nabla\delta\ll\delta$ and $\delta\nabla\tau\ll\tau$), it is meaningful to speak of the macroscopic diffusion coefficient $D(\mathbf{x})$, equal to $\delta^2/2\tau$. Because two microscopic parameters fix one macroscopic parameter, many different pairs of $\delta(\mathbf{x})$ and $\tau(\mathbf{x})$ will yield the same $D(\mathbf{x})$. We will construct our two boxes of gas to exhibit the same $D(\mathbf{x})$, but for different reasons.

One box contains a uniform density of iron wool, so that δ is also uniform. This box also experiences an externally fixed temperature gradient, so that δ/τ, the mean particle speed, varies continuously throughout the box. The second box is held at uniform temperature, but the density of its iron wool is arranged to vary precisely in the fashion needed to duplicate the $D(\mathbf{x})$ within the first box. In the second box, by the ideal-gas law the equilibrium density of particles in the space not occupied by the iron must be uniform, because the box is held at uniform temperature. In the first box, the equilibrium particle concentration cannot be uniform, because the temperature is not uniform.

These two systems with the same $D(\mathbf{x})$ exhibit markedly different behavior when the diffusing particles approach equilibrium. How will the macroscopic diffusion equation, which makes no reference to δ and τ, only to $D(\mathbf{x})$, produce different equilibrium outcomes? When the diffusion coefficient varies over space, there is not just one macroscopic diffusion equation, but several; the choice of the correct equation requires partial knowledge about $\delta(\mathbf{x})$ and $\tau(\mathbf{x})$. We now exhibit these diffusion equations in the case of one dimension.

Although it is perhaps easiest to think intuitively in terms of $\delta(\mathbf{x})$ and $\tau(\mathbf{x})$, the best variables for a continuum treatment of diffusion are $v(\mathbf{x})$ and $\alpha(\mathbf{x})$, the particle speed and collision rate per particle, respectively. For the time being, we ignore distributions of particle speeds and imagine that the local particle speed is fixed separately at every point in space. Because the gas in our boxes is so dilute, collisions nearly always occur when particles bump into strands of iron wool, rather than each other. Thus, we also assume the collision rate is known precisely at each point in space. We treat distributions of speeds and collision rates in Sec. VIII and show that the speed and collision rate in the derivations below are simply replaced by their mean values.

To start, we work in one dimension. During collisions, assumed to be instantaneous, particles have an even change of continuing in the same direction or reversing. Thus, particles switch directions at a rate of $\alpha/2$. Denote by $R(x,t)$ and $L(x,t)$ the density of right- and left-moving particles, respectively. We now require continuity in the flow of particles between collisions and conservation of particle number:

$$\frac{\partial R}{\partial t}=-\frac{\partial(vR)}{\partial x}-\frac{\alpha R}{2}+\frac{\alpha L}{2}\ , \tag{2.2a}$$

$$\frac{\partial L}{\partial t}=\frac{\partial(vL)}{\partial x}+\frac{\alpha R}{2}\frac{-\alpha L}{2}\ . \tag{2.2b}$$

Equations (2.2) comprise the standard starting point for the telegraph model of diffusion [1,6,7], trivially modified to account for a position-dependent speed. These equations may also be viewed as linearized Boltzmann equations, because deviation from continuous particle flow arises from collision terms linear in the particle densities (see, for instance, [23]). (Interestingly, when v is constant, (2.2a) and (2.2b) have recently been shown to be related to the Dirac equation in $1+1$ dimensions [24,25]). Equation (2.2a) gives the rate of change of the number of particles moving right at any point in space. The first term on the right-hand side arises from spatial inhomogeneities in the number of right-moving particles. The second and third terms represent the effect of collisions, which cause right movers to become left movers and vice versa. Define ρ and σ as $R+L$ and $R-L$, respectively. ρ is the local density of particles, and $v\sigma$ equals J, the particle flux. Adding and subtracting (2.2a) and (2.2b) yields

$$\frac{\partial\rho}{\partial t}=\frac{-\partial(v\sigma)}{\partial x}=\frac{-\partial J}{\partial x}\ , \tag{2.3a}$$

$$\frac{\partial\sigma}{\partial t}=\frac{-\partial(v\rho)}{\partial x}-\alpha\sigma\ . \tag{2.3b}$$

Differentiating Eqs. (2.3) and combining gives

$$\frac{\partial^2\rho}{\partial t^2}=\frac{\partial}{\partial x}\left[v\frac{\partial(v\rho)}{\partial x}\right]+\alpha\frac{\partial J}{\partial x}\ . \tag{2.4}$$

It is useful to distinguish three different regimes of behavior exhibited in systems governed by (2.4). At $t=0$, we imagine a configuration of particles distributed nonuniformly over a one-dimensional box of length L, and each particle is posed to move either left or right. For $t\ll 1/\alpha$, the particle distribution appears to have evolved through ballistic movements from the initial configuration, because not enough time has elapsed for

many collisions to occur. In this ballistic regime, the numbers of right movers may differ strongly from the number of left movers. For $1/\alpha \ll t \ll \alpha L^2/v^2$, many collisions have occurred, and the particle distribution appears to have evolved through diffusive movement from the initial configuration. However, particles at one end of the box have not yet had time to visit the other end, and so equilibrium is not yet attained. As equilibrium becomes closer, the number of right movers and the number of left movers at each point will approach the same value. For $t \gg \alpha L^2/v^2$, particles have had sufficient time to visit each side of the box many times, and at each point the number of right movers equals the number of left movers, yielding zero flux. Here we are only interested in the two regimes for which $t \gg 1/\alpha$, so we now neglect the second time derivative, as is standard when studying diffusive properties of telegraph equations [1]. If we were to Laplace transform the time variable in (2.4) in exchange for the imaginary angular frequency ω, this long-time approximation is *equivalent to neglecting terms quadratic and higher in* ω/α. Physically, by neglecting the second time derivative in (2.4), we only lose information about the ballistic, high-frequency behavior of particle movements, which occur over short time intervals, $\Delta t \ll 1/\alpha$. For more on the long-time behavior of telegraph equations, see Refs. [1,26]. Neglecting $\partial^2\rho/dt^2$, integrating (2.4), and ignoring constants of integration that produce only steady-state behavior, we find the particle flux:

$$J(x)=\frac{-v^2(x)}{\alpha(x)}\frac{\partial\rho}{\partial x}-\frac{v(x)\rho(x)}{\alpha(x)}\frac{\partial v}{\partial x} . \tag{2.5}$$

The term that is proportional to $\partial\rho/\partial x$ represents the flux contribution from purely random, diffusive movements, while the term that is proportional to ρ represents the flux contribution from drift at velocity

$$v_{\text{drift}}=\frac{-v(x)}{\alpha(x)}\frac{\partial v}{\partial x} . \tag{2.6}$$

Because of the minus sign in (2.6), particles tend to drift toward regions where the speed is low.

Equation (2.5) deserves several further comments.

(1) In the case that $v(x)$ and $\alpha(x)$ are constants, (2.5) reduces properly to (2.1), and the diffusion coefficient D equals v^2/α.

(2) In box two, in which $v(x)$ is constant but $\alpha(x)$ remains spatially dependent, the form of Fick's law is not modified, but D becomes $D(x)$ and equals $v^2/\alpha(x)$:

$$J(x)=-D(x)\frac{\partial\rho}{\partial x} . \tag{2.7}$$

The equilibrium solution of (2.7) is one of uniform particle density,

$$\rho=\rho_0 . \tag{2.8}$$

The spatial variations of the mean-free path do not affect the uniform distribution of particles at equilibrium, only the rate at which equilibrium is achieved.

(3) In box one, in which both $v(x)$ and $\alpha(x)$ have spatial dependence, but the mean-free path $v(x)/\alpha(x)$ is constant, (2.5) becomes a modified version of Fick's law

$$J(x)=-\frac{\partial(D\rho)}{\partial x} , \tag{2.9}$$

and the diffusion coefficient is $v^2(x)/\alpha(x)$. The equilibrium solution of (2.9) is

$$\rho(x)=\rho_0\left[\frac{D_0}{D(x)}\right]=\rho_0\left[\frac{\alpha(x)v_0^2}{\alpha_0 v^2(x)}\right]=\rho_0\left[\frac{v_0}{v(x)}\right] , \tag{2.10}$$

where the constant ρ_0 is the density at reference point x_0, at which the speed is v_0. The equilibrium density of a particle is inversely proportional to both the local particle speed and the resulting diffusion coefficient.

(4) In the case that $\alpha(x)$ is constant but $v(x)$ remains spatially dependent, (2.5) becomes yet a different generalization of Fick's law,

$$J(x)=-D(x)\frac{\partial\rho}{\partial x}-\frac{\rho}{2}\frac{\partial D}{\partial x} , \tag{2.11}$$

and the diffusion coefficient is $v^2(x)/\alpha$. The equilibrium solution to (2.11),

$$\rho(x)=\rho_0\left[\frac{D_0}{D(x)}\right]^{1/2}=\rho_0\left[\frac{v_0}{v(x)}\right] , \tag{2.12}$$

again dictates that equilibrium particle density is inversely proportional to the local particle speed. However, in this case the equilibrium particle density is inversely proportional to the *square root* of the diffusion coefficient.

(5) When both $v(x)$ and $\alpha(x)$ have spatial dependence and their ratio is not a constant, no proper description can be given in terms of $D(x)$ and a generalized Fick's law. $v(x)$ and $\alpha(x)$ are two microscopic parameters that vary independently, and except in the special cases (2)–(4) above, their effects cannot be summarized in terms of one macroscopic parameter $D(x)$. Nonetheless, in all cases (1)–(5), the equilibrium solution of (2.5) yields an equilibrium particle density that is inversely proportional to local particle speed:

$$\rho(x)=\rho_0\left[\frac{v_0}{v(x)}\right] . \tag{2.13}$$

This simple example illustrates that to pick the correct generalization of Fick's law, one must have partial knowledge about microscopic quantities. Equivalently, the question "What happens when the diffusion coefficient varies over space?" is nonsensical unless more information is provided. As for chemotaxis, (2.13) implies that cells that speed up in response to higher local attractant concentrations will aggregate in regions of *low* attractant concentration. This strategy represents a form of *reverse* chemotaxis. From (2.8), cells that lower the tumble rate in response to local attractant concentration aggregate uniformly at equilibrium.

However, as noted in [27,8], during the approach to equilibrium such cells are initially more likely to head toward regions of low tumble rate; this effect is called *pseudochemotaxis* and in a closed space gradually resides after cells have had sufficient time to visit all regions many times. Given a specified $D(x)$, the combination of (2.3a) and (2.7)

$$\frac{\partial \rho}{\partial t}=\frac{\partial}{\partial x}\left[D(x)\frac{\partial \rho}{\partial x}\right] \tag{2.14}$$

governs the particle density evolution over intervals $\Delta t \gg 1/\alpha$. For general $D(x)$, we may heuristically understand pseudochemotaxis through the following simple paradigm. A single particle moves in one dimension at constant speed v, and takes successive steps left and right with even probability. When the particle is to the right (left) of its origin, the duration of its steps is τ_R (τ_L). After T seconds, where $T \gg \tau_R$ and $T \gg \tau_L$, the particle's root-mean-square distance from the origin, $x_{\rm rms}$, is crudely approximated as

$$x_{\rm rms} \approx \begin{cases} (v^2 \tau_R T)^{1/2} & \text{if } x>0 \\ (v^2 \tau_L T)^{1/2} & \text{if } x<0 \, . \end{cases} \tag{2.15}$$

If $\tau_R > \tau_L$, the average particle position will gradually shift right. Equivalently, the effective collision rate α equals $1/\tau$, and so particles will head on average toward the region of lower collision rate. Note that this argument relies on the infinite extent of the one-dimensional space. In a closed space, particles bounce off walls, and the average position eventually returns to the origin; for illustrative graphs, see Refs. [8,27].

III. ONE-DIMENSIONAL THEORY

We now generalize (2.5) to the case when the collision rate and velocity depend on the particle's direction of motion; Rivero *et al.* have analyzed the case of constant velocity and direction-dependent collision rate [28]. The insights gained here will be useful in Secs. V through VIII for the treatment of higher-dimensional cases; here we can solve for the particle flux *exactly* in the regimes $t \gg \alpha$, whereas in dimensions two or more we perturb around the equilibrium solution.

Before we proceed to the derivation, consider the following simple paradigm providing intuition. A single particle moves in one dimension and takes successive steps left and right with even probability. When moving left (right), the particle has velocity v_L (v_R) and the step lasts for duration τ_L (τ_R). After N steps, for N a very large number, the average position of the particle away from its origin, $\langle x \rangle$, and the time elapsed, T, are

$$T=\frac{N}{2}(\tau_R+\tau_L) \, , \tag{3.1a}$$

$$\langle x \rangle = \frac{N}{2}(v_R \tau_R - v_L \tau_L) \, . \tag{3.1b}$$

The left and right effective "collision" rates, α_L and α_R, are $1/\tau_L$ and $1/\tau_R$, respectively. Thus, the drift velocity is

$$v_{\rm drift}=\frac{\langle x \rangle}{T}=\frac{v_R \tau_R - v_L \tau_L}{\tau_R+\tau_L}=\frac{v_R \alpha_L - v_L \alpha_R}{\alpha_R+\alpha_L} \, . \tag{3.2}$$

Equation (3.2) indicates that particles tend to drift in the direction with the higher speed and the lower collision rate. These effects may compete.

Although the above simple paradigm does not allow for spatially dependent velocities and collision rates, our formal derivation shall allow for $\alpha_L(x)$, $\alpha_R(x)$, $v_L(x)$, and $v_R(x)$; (3.2) will turn out to be correct when we reduce to the case in which these four quantities are independent of space. Given particle velocities and collision rates that depend on both position and direction of motion, the requirements of continuous flow and conservation of particle number are modified from (2.2):

$$\frac{\partial R}{\partial t}=-\frac{\partial (v_R R)}{\partial x}-\frac{\alpha_R R}{2}+\frac{\alpha_L L}{2} \, , \tag{3.3a}$$

$$\frac{\partial L}{\partial t}=\frac{\partial (v_L L)}{\partial x}+\frac{\alpha_R R}{2}-\frac{\alpha_L L}{2} \, . \tag{3.3b}$$

The derivation proceeds similarly to that of Sec. II, and the flux is

$$J(x)=-\frac{2(v_R+v_L)}{(\alpha_L+\alpha_R)}\frac{\partial}{\partial x}\left[\frac{v_R v_L \rho}{v_R+v_L}\right]+\frac{(\alpha_L v_R-\alpha_R v_L)}{(\alpha_L+\alpha_R)}\rho \, . \tag{3.4}$$

When the velocities v_R and v_L are constants, the drift velocity from (3.4) reduces to that of (3.2), but when v_R and v_L do depend upon position, there is an extra contribution to the drift velocity from the derivative in the first term of (3.4). When the tumbling rates for left and right movement are equal, (3.4) reduces to (2.5). The equilibrium solution of (3.4) is

$$\rho(x)=\rho(x_0)\left[\frac{v_R(x_0)v_L(x_0)}{v_R(x)v_L(x)}\right]\left[\frac{v_R(x)+v_L(x)}{v_R(x_0)+v_L(x_0)}\right]\exp\left[\int_{x_0}^{x}\left[\frac{\alpha_L(x')v_R(x')-\alpha_R(x')v_L(x')}{2v_R(x')v_L(x')}\right]dx'\right] \, , \tag{3.5}$$

where x_0 is a reference point. Equations (3.4) and (3.5) tell us that two general classes of successful chemotactic strategies exist. To drift toward regions of high attractant concentration, a cell moving in a direction of increasing attractant concentration should either lower the tumble rate, speed up, or do both.

Equation (3.4) can also be used to estimate the importance of the finite duration of tumbles; we have assumed until now that collisions or tumbles were instantaneous. In reality, the average tumble duration of the wild-type cell (~0.1 s) is roughly 10% of the run duration (~1 s) when the cell swims in a direction of decreasing attractant concentration, and roughly 1% of the duration of the longest runs up steep gradients (~10 s) [9]. The percentage of time spent tumbling is direction dependent, so the cell has the lower effective swimming speed when moving in a direction of decreasing attractant concentration. If τ is the average tumble duration, the mean total duration of run and subsequent tumble is

$$\langle T_{\rm run}\rangle=\tau+1/\alpha\ , \tag{3.6}$$

while the mean distance travelled during a run is

$$\langle x_{\rm run}\rangle=v/\alpha\ . \tag{3.7}$$

The resulting effective swimming speed is then

$$v_{\rm eff}=\langle x_{run}\rangle/\langle T_{\rm run}\rangle=v(1+\tau\alpha)^{-1}\ . \tag{3.8}$$

From the values above, the effective swimming speed when the cell heads down the attractant gradient $\approx 0.9v$, is about 90% of the effective swimming speed when the cell heads up the steepest gradients $\approx 0.99v$. A simple calculation shows that the decrease in drift velocity from the overall lowering of the effective swimming speeds is always more important than the accompanying increase in drift velocity from the direction dependence of the effective swimming speed, and so the drift velocity is lowered by at most 1%. Thus, in wild-type cells the effects of finite tumble duration are small. However, some mutant strains have long tumble durations and spend more than 90% of their time tumbling [21]; if this behavior is a response to low absolute levels of attractant concentration, rather than changes in attractant concentration, then by (2.13) the long tumble duration would indeed lead to significant aggregation in regions of low effective swimming speed.

IV. FICK'S LAW IN CONTINUUM

Because there is a continuum of possible directions of motion in two or more dimensions, rather than just left and right, our higher-dimensional investigations will be considerably harder than those done in one dimension. To illustrate the techniques involved, we warm up by deriving Fick's law. The approach here is very different from the lattice or Langevin derivations with which the reader may be familiar.

The starting point for our higher-dimensional derivations will again be the requirements of continuous flow and conservation of particle number. Let $P(\mathbf{u},\mathbf{x},t)$ denote the density of particles at position $\mathbf{x}$, at time t, moving in the direction of the unit vector $\mathbf{u}$. If the particles move in a space of dimension n, $\mathbf{u}$ lies in the $(n-1)$-dimensional unit hypersphere centered at the origin. α is the collision rate and v is the particle speed, for now both assumed to be constant. At each point $\mathbf{x}$, we will be performing integrations over all possible directions of motion, denoted by $\oint d^{n-1}\mathbf{u}$. The surface integral over the unit hypersphere,

$$\Omega(n)=\oint d^{n-1}\mathbf{u}=\frac{2\pi^{n/2}}{\Gamma(n/2)}\ , \tag{4.1}$$

we denote simply as $\Omega(n)$. We write all vectors and tensors in boldface font. As before, define ρ, the particle density, and $\mathbf{J}$, the particle flux:

$$\rho(\mathbf{x},t)\equiv\oint P(\mathbf{u},\mathbf{x},t)d^{n-1}\mathbf{u}\ , \tag{4.2a}$$

$$v\boldsymbol{\sigma}(\mathbf{x},t)\equiv\mathbf{J}(\mathbf{x},t)\equiv v\oint P(\mathbf{u},\mathbf{x},t)\mathbf{u}\,d^{n-1}\mathbf{u}\ . \tag{4.2b}$$

The requirements of continuous flow and conservation of particle number lead to a separate equation for every $\mathbf{u}$:

$$\frac{\partial P}{\partial t}=-v\nabla P\cdot\mathbf{u}-\alpha P+\frac{\alpha}{\Omega(n)}\oint P(\mathbf{u}',\mathbf{x},t)d^{n-1}\mathbf{u}'$$

$$=-v\nabla P\cdot\mathbf{u}-\alpha P+\frac{\alpha\rho}{\Omega(n)}\ . \tag{4.3}$$

The first term on the right-hand side again represents the effects of spatial inhomogeneities in the particle density, and the second term represents the effect of collisions upon particles moving in direction $\mathbf{u}$, which subsequently cease to move in this direction. The third term represents the effects of particles that were moving in direction $\mathbf{u}'$, collided, and then moved off in direction $\mathbf{u}$; we integrate over all possible original directions $\mathbf{u}'$. For now, we assume that after a collision a particle has an equal chance of heading off in any direction, so in the third term we divide the total rate of collisions $\rho\alpha$ by the total area of possible new directions $\Omega(n)$.

The equilibrium solution of (4.3) is simply

$$\rho(\mathbf{x})=\rho_0\ , \tag{4.4a}$$

$$P(\mathbf{u},\mathbf{x})=\rho_0/\Omega(n)\ , \tag{4.4b}$$

in which ρ_0 is the total particle number divided by the volume of space. By equilibrium, we mean that not only are the time derivatives of all quantities zero, but also the particle flux $\mathbf{J}$ is zero. The equilibrium state is only one type of steady state, in which time derivatives of all quantities are zero, but the particle flux need not be zero. If we hold the particle gradient $\nabla\rho$ constant over time and space, there is also a steady-state solution of (4.3):

$$P(\mathbf{u},\mathbf{x})=\frac{\rho(\mathbf{x})}{\Omega(n)}-\frac{v\nabla\rho\cdot\mathbf{u}}{\alpha\Omega(n)}\ . \tag{4.5}$$

Using (4.2b) and the identities

$$0=\oint\mathbf{u}\,d^{n-1}\mathbf{u}\ , \tag{4.6a}$$

$$\Omega(n)/n=\oint\mathbf{u}\mathbf{u}\,d^{n-1}\mathbf{u}\ , \tag{4.6b}$$

the particle flux arising from (4.5) is

$$\mathbf{J}=\frac{-v^2\nabla\rho}{n\alpha}\ . \tag{4.7}$$

By restricting our attention to a steady-state solution with uniform particle gradient, we have found Fick's law, (4.7). What happens when $\nabla\rho$ is not uniform over time and space?

We now show that Fick's law is valid over intervals $\Delta t\gg 1/\alpha$, if the gradient does not change appreciably within the typical distance traveled between collisions, i.e., $v\nabla^2\rho\ll\alpha|\nabla\rho|$. When Fick's law is invoked as a phenomenological equation, the restriction is well known [29,30]. Integrating (4.3) over $d^{n-1}\mathbf{u}$ yields the general equation of continuity

$$\frac{\partial\rho}{\partial t}=-\nabla\cdot\mathbf{J}\ , \tag{4.8}$$

whereas multiplying (4.3) by $\mathbf{u}$ and then integrating gives

$$\frac{\partial\boldsymbol{\sigma}}{\partial t}=-v\nabla\cdot\oint(P\mathbf{u}\mathbf{u})d^{n-1}\mathbf{u}-\frac{\alpha\mathbf{J}}{v}\ . \tag{4.9}$$

We now perturb around the steady-state solution (4.5) by

writing $P(\mathbf{u},\mathbf{x},t)$ as

$$P(\mathbf{u},\mathbf{x},t)-\frac{\rho(\mathbf{x},t)}{\Omega(n)}+\frac{\nu\nabla\rho\cdot\mathbf{u}}{\alpha\Omega(n)}=\frac{f(\{\rho\},\mathbf{u},t)\mathbf{u}\cdot\mathbf{J}}{\nu}\ . \tag{4.10}$$

It makes sense to write $P(\mathbf{u},\mathbf{x},t)$ as in (4.10), because at steady state the function f equals zero and (4.10) becomes (4.5). Near steady state, if we define f as the small dimensionless functional of ρ that makes (4.10) true, we lose no generality. Strictly speaking, the most general f depends on the values of ρ all across space, and so is a functional of ρ. Note that (4.10) only holds for $t\gg 1/\alpha$. For $t<1/\alpha$, one cannot deduce how many particles are heading in each direction, $P(\mathbf{u},\mathbf{x},t)$, simply from the total density $\rho(\mathbf{x},t)$, because most particles are still heading in their initial directions of motion. Substituting (4.10) into (4.9), using

$$0=\oint \mathbf{uuu}\, d^{n-1}\mathbf{u}\ , \tag{4.11}$$

and combining with (4.8) yields

$$\frac{\partial^2\rho}{\partial t^2}=\frac{\nu^2\nabla^2\rho}{n}+\nu\nabla\cdot\left[\nabla\cdot\left[\mathbf{J}\cdot\oint f\mathbf{uuu}\, d^{n-1}\mathbf{u}\right]\right]+\alpha\nabla\cdot\mathbf{J}\ . \tag{4.12}$$

As in Sec. II, for the study of diffusive behavior we neglect the second time derivative when $t\gg 1/\alpha$, for it is $O(\omega^2/\alpha^2)$ if (4.12) is Laplace transformed. The second term on the right-hand side of (4.12) is also negligible if $\nabla\rho$ does not change much over the typical distance traveled between particle collisions, ν/α, by the following argument.

If we do neglect the second term on the right of (4.12) as well as the second time derivative, integrate, and discard the integration constant, which adds only an uninteresting uniform flux, we solve for $\mathbf{J}$ and find Fick's law, (4.7). To check when the corrections to Fick's law are important, we substitute (4.7) back into (4.12), and find that when

$$\nu|\nabla\mathbf{J}|\ll\alpha|\mathbf{J}| \tag{4.13a}$$

and

$$\nu\left|\nabla\cdot\left[\nabla\rho\cdot\oint f\mathbf{uuu}\, d^{n-1}\mathbf{u}\right]\right|\ll\alpha|\nabla\rho|\ , \tag{4.13b}$$

the corrections are small. (By the absolute value of a tensor, we mean here the square root of the sum of the squares of all its components.) When the rank-two tensor $\nabla\nabla\rho$ is zero and $t\gg 1/\alpha$, we are in the steady state described by (4.5) and so f is zero. Near the steady state, when $\nu|\nabla\nabla\rho|$ is everywhere small compared to $\alpha|\nabla\rho|$ and all higher derivatives of ρ are even less important, (4.13a) is satisfied and the leading correction to $f=0$ will be of order $O(\nu|\nabla\nabla\rho|/\alpha|\nabla\rho|)$. Thus, when $\nu|\nabla\nabla\rho|\ll\alpha|\nabla\rho|$, corrections to Fick's law are small. Physically, this means that when $\nabla\rho$ does not change much over the typical distance traveled between collisions, Fick's law is good.

The analysis of this section is easily extended to account for position-dependent speed $\nu(\mathbf{x})$, and collision rate $\alpha(\mathbf{x})$. As emphasized in Sec. II, one finds that the equilibrium density of particles is inversely proportional to $\nu(\mathbf{x})$, (2.12), and that the generalization of (4.7),

$$\mathbf{J}=\frac{-\nu(\mathbf{x})\nabla(\nu\rho)}{n\alpha(\mathbf{x})}\ , \tag{4.14}$$

is valid when

$$\nu|\nabla(\nabla(\nu\rho)/\alpha)|\ll|\nabla(\nu\rho)|\ . \tag{4.15}$$

Note that for (4.15) to be fulfilled, the condition

$$\nu|\nabla\alpha|\ll\alpha^2 \tag{4.16}$$

upon the collision rate must hold. Equation (4.16) says that the fractional change in the collision rate over the typical distance traveled between collisions must be small; a similar restriction is noted in [8].

V. DIRECTION-DEPENDENT COLLISION RATE

We now allow for a particle velocity that depends upon position, $\nu(\mathbf{x})$, and a collision rate that depends upon position and direction of motion, $\alpha(\mathbf{u},\mathbf{x})$. Both $\nu(\mathbf{x})$ and $\alpha(\mathbf{u},\mathbf{x})$ are scalar fields that are fixed in time. Keeping spatial dependencies implicit, the generalization of (4.3) is then

$$\frac{\partial P}{\partial t}=-\nabla(P\nu)\cdot\mathbf{u}-\alpha(\mathbf{u})P(\mathbf{u})+\frac{1}{\Omega(n)}\oint P\alpha(\mathbf{u}')d^{n-1}\mathbf{u}'\ . \tag{5.1}$$

This continuity equation is known [31], but the derivation below for the flux is new. For a review of this and other continuity equations, see [32]. Integrating (5.1) over $d^{n-1}\mathbf{u}$, we again obtain (4.8), but now

$$\mathbf{J}(\mathbf{x},t)=\nu(\mathbf{x})\oint P(\mathbf{u},\mathbf{x},t)\mathbf{u}\, d^{n-1}\mathbf{u}=\nu(\mathbf{x})\sigma(\mathbf{x},t)\ . \tag{5.2}$$

Multiplying (5.1) by $\mathbf{u}$ and integrating over $d^{n-1}\mathbf{u}$, we find

$$\frac{\partial\sigma}{\partial t}=-\nabla\cdot\left[\nu\oint(P\mathbf{uu})d^{n-1}\mathbf{u}\right]-\oint[P\alpha(\mathbf{u})\mathbf{u}]d^{n-1}\mathbf{u}\ . \tag{5.3}$$

The equilibrium condition for ρ is found by setting $\mathbf{J}$ and all time derivatives everywhere zero. Although $\rho(\mathbf{x})=\rho_0$ is no longer the equilibrium solution as in Sec. IV, it should be intuitively clear that at equilibrium,

$$P(\mathbf{u},\mathbf{x})=\rho(\mathbf{x})/\Omega(n)\ . \tag{5.4}$$

Like (4.4b), (5.4) implies that at any point in space, an equal number of particles are moving in each direction, resulting in zero flux. Unlike (4.4b), (5.4) implies that the actual density of particles is spatially dependent; however, unless the particle velocity has direction as well as spatial dependence [i.e., $\nu(\mathbf{u},\mathbf{x})$], then (5.4) will yield zero flux by (5.2).

At this point, it is useful to express $\alpha(\mathbf{x},\mathbf{u})$ as a power series in $\mathbf{u}$:

$$\alpha(\mathbf{x},\mathbf{u})=\alpha_0(\mathbf{x})+[\boldsymbol{\alpha}_1(\mathbf{x})\cdot\mathbf{u}]+[\boldsymbol{\alpha}_2(\mathbf{x})\cdot\mathbf{u}\cdot\mathbf{u}]+[\boldsymbol{\alpha}_3(\mathbf{x})\cdot\mathbf{u}\cdot\mathbf{u}\cdot\mathbf{u}]+\cdots\ . \tag{5.5}$$

The coefficients $\boldsymbol{\alpha}_k(\mathbf{x})$ are spatially dependent tensors of

rank k. To start, we consider the case in which

$$\alpha(\mathbf{x},\mathbf{u})=\alpha_0(\mathbf{x})+\boldsymbol{\alpha}_1(\mathbf{x})\cdot\mathbf{u} \ . \tag{5.6}$$

Equation (5.6) shall be useful in describing chemotactic *E. coli* (see Sec. VII). In this case, we now show that the equilibrium solution of (5.1) is (5.4), where $\rho(\mathbf{x})$ at equilibrium will be determined. The general case (5.5) is considered in Appendix A, where it is shown that if terms quadratic or higher exist in the series expansion (5.5), then (5.4) cannot describe equilibrium. Using (5.4), (5.6), and setting the time derivative equal to zero in (5.3) yields

$$\nabla(\nu\rho)=-\rho\boldsymbol{\alpha}_1 \ . \tag{5.7}$$

Substituting (5.4) and (5.7) into the right side of (5.1) indeed yields zero, showing that (5.4) and (5.7) comprise the equilibrium solution. Solving the differential equation (5.7) shall yield the generalization of (4.4a).

Before we solve (5.7), note that it is only meaningful to talk about an equilibrium solution if the vector field $\boldsymbol{\alpha}_1(\mathbf{x})/\nu(\mathbf{x})$ can be written as the gradient of some scalar field, $\nabla\phi(\mathbf{x})$. If it cannot, then the equilibrium condition (5.7) is unattainable, for it would imply the contradiction

$$\frac{\nabla(\nu\rho)}{\nu\rho}=\nabla(\ln(\nu\rho))=-\frac{\boldsymbol{\alpha}_1}{\nu} \ . \tag{5.8}$$

Even if $\boldsymbol{\alpha}_1(\mathbf{x})/\nu(\mathbf{x})$ cannot be written as the gradient of some scalar field, this does not exclude the possibility of steady-state solutions with nonvanishing particle flux; however, we shall not consider such solutions here.

If equilibrium can be reached, then (5.7) says that two effects balance. The left side of (5.7) represents the effect of the gradient, increasing the net flow of particles towards regions of low particle density and low speed. The right side represents the effect of the direction-dependent collision rate, increasing the net flow of particles in the direction $-\boldsymbol{\alpha}_1(\mathbf{x})$. The solution to (5.7) is

$$\begin{aligned}\rho(\mathbf{x})&=\rho(\mathbf{x}_0)\frac{\nu(\mathbf{x}_0)}{\nu(\mathbf{x})}\exp\left[\int_{x_0}^{x}\left[\frac{-\boldsymbol{\alpha}_1(\mathbf{x}')}{\nu(\mathbf{x}')}\right]\cdot d\mathbf{x}'\right]\\&=\rho(\mathbf{x}_0)\frac{\nu(\mathbf{x}_0)}{\nu(\mathbf{x})}\exp\{-[\phi(\mathbf{x})-\phi(\mathbf{x}_0)]\} \ ,\end{aligned} \tag{5.9}$$

where $\mathbf{x}_0$ is a reference point and the line integral in (5.9) yields a single-valued $\rho(\mathbf{x})$ only because $\boldsymbol{\alpha}_1(\mathbf{x})/\nu(\mathbf{x})$ equals the gradient of a scalar field, $\nabla\phi(\mathbf{x})$, by assumption. As in Secs. II, III, and IV, the factor $\nu(\mathbf{x}_0)/\nu(\mathbf{x})$ increases particle aggregation in regions of low speed. In addition, the factor $e^{-\phi(\mathbf{x})}$ increases aggregation in regions of low ϕ. It is important to note that α_0 does not affect the equilibrium density, although it shall enter the expression for the flux.

Having found the equilibrium distribution, we now consider the approach to equilibrium. Combining (4.8), (5.2), and (5.3) yields

$$\begin{aligned}\frac{\partial^2\rho}{\partial t^2}&=\nabla\cdot\left[\nu\nabla\cdot\left[\nu\oint(P\mathbf{u}\mathbf{u})d^{n-1}\mathbf{u}\right]\right]\\&\quad+\nabla\cdot\left[\nu\oint(\alpha P\mathbf{u})d^{n-1}\mathbf{u}\right] \ .\end{aligned} \tag{5.10}$$

Substituting (5.6) into (5.10) and then using (5.2) gives

$$\begin{aligned}\frac{\partial^2\rho}{\partial t^2}&=\nabla\cdot\left[\nu\nabla\cdot\left[\nu\oint(P\mathbf{u}\mathbf{u})d^{n-1}\mathbf{u}\right]\right]+\nabla\cdot(\alpha_0\mathbf{J})\\&\quad+\nabla\cdot(\nu\boldsymbol{\alpha}_1\cdot\oint(P\mathbf{u}\mathbf{u})d^{n-1}\mathbf{u}) \ .\end{aligned} \tag{5.11}$$

At the analogous stage in the derivation of Sec. IV, we solved (4.12) by perturbing around the steady-state solution we had found, and concluded that Fick's law was valid when $\nu|\nabla\nabla\rho|\ll\alpha|\nabla\rho|$. In contrast, here we do not have such a steady-state solution. Instead, we perturb around the equilibrium solution by approximating $P(\mathbf{u},\mathbf{x},t)\approx\rho(\mathbf{x},t)/\Omega(n)$ in (5.11):

$$\frac{\partial^2\rho}{\partial t^2}=\nabla\cdot\left[\nu\nabla\left[\frac{\nu\rho}{n}\right]\right]+\nabla\cdot(\alpha_0\mathbf{J})+\nabla\cdot\left[\frac{\nu\boldsymbol{\alpha}_1\rho}{n}\right] \ . \tag{5.12}$$

Note that this approximation does *not* imply that we are already at equilibrium, because we allow for changes in $\rho(\mathbf{x},t)$ over time. Again, as in Sec. II, we neglect the second time derivative after $t\gg1/\alpha$. This long-time approximation yields the generalization of Fick's law, with both diffusive and drift contributions to the particle flux:

$$\mathbf{J}=\frac{-\nu\nabla(\nu\rho)}{n\alpha_0}-\frac{\nu\rho\boldsymbol{\alpha}_1}{n\alpha_0} \ . \tag{5.13}$$

If equilibrium can be attained, an identically zero $\mathbf{J}$ in (5.13) indeed yields (5.7). In an analysis similar to that of Sec. IV, (4.10)–(4.16), if we write

$$P(\mathbf{u},\mathbf{x},t)-\frac{\rho(\mathbf{x},t)}{\Omega(n)}=\frac{f(\{\rho\},\mathbf{u},t)\mathbf{u}\cdot\mathbf{J}}{\nu} \ , \tag{5.14}$$

it is found that (5.13) is valid close to equilibrium when $|\nabla(\nu\rho)+\rho\boldsymbol{\alpha}_1|\ll\alpha_0\rho$.

Combining (5.13) and the equation of continuity (4.8) yields an effective Smoluchowski equation:

$$\frac{\partial^2\rho}{\partial t^2}=\nabla\cdot\left[\frac{\nu\nabla(\nu\rho)}{n\alpha_0}+\frac{\nu\rho\boldsymbol{\alpha}_1}{n\alpha_0}\right] \ . \tag{5.15}$$

From (5.15), we see that if the particles have mass m, they each experience an effective force

$$F_{\text{eff}}=\frac{-m\nu(\nabla\nu+\boldsymbol{\alpha}_1)}{n} \ . \tag{5.16}$$

The effective force is not a true force, because it arises partly from direction-dependent collision rates, not accelerations. As evident in (5.16) and in the equilibrium distribution (5.9), the total effective force is a sum of two others, one in the direction of lowest collision rate, the other in minus the direction of the particle speed gradient.

VI. ANGULAR CORRELATIONS

In this section and the next, we consider two features unique to dimensions two and higher, which cannot be treated using Langevin or lattice methods. In contrast, the continuum formalism is naturally suited for these generalizations.

We had previously assumed that after a tumble or collision, the particle had an equal chance of heading off in any direction. Now, the outgoing trajectory is more like-

ly to be in the same hemisphere as the incoming trajectory. If the incoming trajectory is in direction $\mathbf{u}$, the conditional probability density that the outgoing trajectory is in direction $\mathbf{u}'$ is given by the angular correlation function $\gamma(\mathbf{u}',\mathbf{u},\mathbf{x})$. Angular correlations can be viewed as a generalization to dimensions two and higher of the "persistence" effect studied by Goldstein [1], and Weis and others [2]. For these authors, persistence is a one-dimensional effect characterized by the increased likelihood for a particle to head off in the same direction after a collision that it had before. Our results for angular correlations will reduce to known results for persistence in one dimension.

For *E. coli*, $\gamma(\mathbf{u}',\mathbf{u})$ depends only on the absolute difference $|\mathbf{u}'-\mathbf{u}|$, which is the sole case that we shall consider here. When $\gamma(\mathbf{u}',\mathbf{u})$ is normalized correctly, the integral over all possible outgoing directions should equal one, by conservation of particle number:

$$\oint \gamma(\mathbf{u}',\mathbf{u})d^{n-1}\mathbf{u}'=1 \ . \tag{6.1}$$

Here, we consider $\gamma(|\mathbf{u}'-\mathbf{u}|)$, and so also,

$$\oint \gamma(\mathbf{u}',\mathbf{u})d^{n-1}\mathbf{u}=1 \ . \tag{6.2}$$

The angular correlations are characterized by $\Theta(\mathbf{x})$, the mean cosine of the angle between successive directions of motion:

$$\oint \gamma(\mathbf{u}',\mathbf{u})\mathbf{u}'d^{n-1}\mathbf{u}'=\Theta\mathbf{u} \ . \tag{6.3}$$

To account for angular correlations, the third term in (5.1) is modified to reflect the nonuniform distribution of new directions of motion after collisions:

$$\frac{\partial P}{\partial t}=-\nabla(P\nu)\cdot\mathbf{u}-\alpha(\mathbf{u})P(\mathbf{u}) + \oint P(\mathbf{u}')\alpha(\mathbf{u}')\gamma(\mathbf{u},\mathbf{u}')d^{n-1}\mathbf{u}' \ , \tag{6.4}$$

where all spatial dependencies are kept implicit.

Assuming (5.6), we now show that the equilibrium solution of (6.4) is again given by (5.4) and (5.7), but with an effective $\boldsymbol{\alpha}_{1,\text{eff}}$ replacing $\boldsymbol{\alpha}_1$. Integrating (6.4) over $d^{n-1}\mathbf{u}$ and using (6.1) yields the equation of continuity (4.8). Multiplying (6.4) by $\mathbf{u}$, integrating over $d^{n-1}\mathbf{u}$, and using (6.3) yields

$$\frac{\partial\boldsymbol{\sigma}}{\partial t}=-\nabla\cdot\left[\nu\oint[P(\mathbf{u})\mathbf{u}\mathbf{u}]d^{n-1}\mathbf{u}\right] -(1-\Theta)\oint[P(\mathbf{u})\alpha(\mathbf{u})\mathbf{u}]d^{n-1}\mathbf{u} \ . \tag{6.5}$$

Except for the factor $(1-\Theta)$, (6.5) is identical to (5.3), and it is straightforward to show that the equilibrium solution of (6.4) is given by (5.4) and

$$\nabla(\nu\rho)=-\rho\boldsymbol{\alpha}_1(1-\Theta) \ . \tag{6.6}$$

Equilibrium may only be attained when the condition

$$\frac{\boldsymbol{\alpha}_1(x)[1-\Theta(x)]}{\nu(x)}=\nabla\phi(x) \tag{6.7}$$

is fulfilled, for some function $\phi(x)$. If this condition is fulfilled, then $\boldsymbol{\alpha}_1(1-\Theta)$ acts like an effective $\boldsymbol{\alpha}_{1,\text{eff}}$ and replaces $\boldsymbol{\alpha}_1$ in the equation for the equilibrium distribution, (5.9). Perturbing around the equilibrium solution, the particle flux is found to be

$$\mathbf{J}=\frac{-\nu\nabla(\nu\rho)}{n\alpha_0(1-\Theta)}-\frac{\nu\rho\boldsymbol{\alpha}_1}{n\alpha_0} \ , \tag{6.8}$$

when $|\nabla(\nu\rho)+\rho(1-\Theta)\boldsymbol{\alpha}_1| \ll \alpha_0\rho$.

In Secs. IV and V, we implicitly considered the case when no correlation between successive trajectories exists, i.e., $\gamma(\mathbf{u}',\mathbf{u})$ is $1/\Omega(n)$ for all $\mathbf{u}'$ and $\mathbf{u}$, and so Θ is everywhere zero. If correlation between successive trajectories is total, in other words, no changes of direction are allowed, then $\gamma(\mathbf{u}',\mathbf{u})$ will equal $\delta^{n-1}(\mathbf{u}'-\mathbf{u})$, Θ will equal 1, and $\boldsymbol{\alpha}_{1,\text{eff}}$ will be zero. Consequently, in this case the equilibrium particle density will be uniform over space, (6.6). However, when $\gamma(\mathbf{u}',\mathbf{u})$ equals $\delta^{n-1}(\mathbf{u}'-\mathbf{u})$, particles never change their direction of motion, so if the initial particle distribution is not already at equilibrium, equilibrium will not be attained. In all cases, the factor $1-\Theta$ reduces the relative contribution of the drift term in (6.8) to the particle flux.

In one dimension, we replace the integral in (6.3) with a sum, and (6.8) reduces to the flux equation with persistence found by Goldstein [1]. Furthermore, (6.8) agrees with a formula found by arguments originating with Flory [3]. Lovely and Dahlquist extended Flory's treatment of polymers to allow for distributions of constituent monomer lengths and angles between successive monomers [4]. When the monomer lengths are exponentially distributed, they apply the result to chemotaxis and find the diffusion coefficient

$$D=\frac{\nu^2}{n\alpha(1-\Theta)} \ . \tag{6.9}$$

Equation (6.8) generalizes (6.9) to the case when the particle speed and collision rate vary spatially, and the collision rate depends upon direction of motion; the result is that angular correlations do not affect the drift velocity, only the diffusion coefficient. It is not hard to build an intuitive picture of why this should be so.

If we consider a random walk at constant speed, without angular correlations, and with $\boldsymbol{\alpha}_1=0$, adding angular correlations cannot introduce a drift velocity. However, if we imagine an initial configuration of particles all moving in the same direction, adding angular correlations certainly increases the relaxation time to equilibrium, because it takes longer for the system to attain a random distribution of particles moving in every direction. Hence, adding angular correlations must lower the effective collision rate of the diffusive term, as (6.8) and (6.9) state.

Furthermore, consider a random walk without angular correlations but with a drift term due to nonzero $\boldsymbol{\alpha}_1$. Now add angular correlations with a mean cosine Θ of roughly 0.5, corresponding to a 60° angle between successive directions. For each direction of motion, the effective correlation time is essentially doubled, because now it will take roughly two collisions instead of one to wipe out any correlation between previous directions of motion. However, this change is true for all directions; all effective tumble rates are divided by two. But the drift

velocity depends only on the ratio α_1/α_0 and is therefore unchanged, as (6.8) says.

VII. ROTATIONAL BROWNIAN MOTION

In addition to the large and sudden changes in particles' directions of motion caused by collisions, there may also exist small gradual random changes from rotational Brownian motion; such changes are characterized by the rotational diffusion coefficient $D_{\rm rot}$. Between collisions, a particle will, on average, change its direction of motion by an angle $\sqrt{2D_{\rm rot}\Delta t}$ after a short interval Δt. More precisely, if we consider a species of particle that does not experience collisions, and if $\Psi(\mathbf{u})$ is the total number of particles moving in direction $\mathbf{u}$,

$$\Psi(\mathbf{u},t)=\int P(\mathbf{u},\mathbf{x},t)d^n\mathbf{x} \ , \tag{7.1}$$

then rotational Brownian motion implies

$$\frac{\partial\Psi}{\partial t}=D_{\rm rot}\nabla_{\mathbf{u}}^2\Psi \ , \tag{7.2}$$

where $\nabla_{\mathbf{u}}^2$ is the Laplacian in direction space. Equation (7.2) is analogous to the standard Fick's law describing diffusion in real space found by combining (4.7) and (4.8)

$$\frac{\partial\rho}{\partial t}=\frac{\nu^2\nabla^2\rho}{n\alpha} \ . \tag{7.3}$$

For particles that do experience collisions, we modify (5.1) to account for rotational Brownian motion in a fashion similar to (7.2):

$$\frac{\partial P}{\partial t}=-\nabla(\nu P)\cdot\mathbf{u}-\alpha P+\frac{1}{\Omega(n)}\oint P\alpha\, d^{n-1}\mathbf{u}'+D_{\rm rot}\nabla_u^2 P \ . \tag{7.4}$$

Assuming (5.6), from (7.4), it is straightforward to show that the equilibrium state is still described by (5.4), (5.7), and (5.9). Manipulations similar to those of Secs. IV and V yield

$$\frac{-\nu\nabla(\nu\rho)}{n}=\alpha_0\mathbf{J}+\frac{\nu\rho\boldsymbol{\alpha}_1}{n}-D_{\rm rot}\nu\left[\oint P\nabla_{\mathbf{u}}^2\mathbf{u}\, d^{n-1}\mathbf{u}\right] \ . \tag{7.5}$$

As shown in Appendix B, the Laplacian in direction space, ∇_u^2, acting on $\mathbf{u}$, merely gives $(1-n)\mathbf{u}$. Using this fact, (7.5) yields

$$\mathbf{J}=-\left[\frac{\nu\nabla(\nu\rho)+\nu\rho\boldsymbol{\alpha}_1}{n[\alpha_0+(n-1)D_{\rm rot}]}\right] \ . \tag{7.6}$$

Equations (7.5) and (7.6) are only valid in the long-time approximation and when $|\nabla(\nu\rho)+\rho\boldsymbol{\alpha}_1| \ll[\alpha_0+n-1)D_{\rm rot}]\rho$.

In some sense, rotational Brownian motion has exactly the opposite effects of angular correlations. Whereas angular correlations decrease the effective tumbling rate in the diffusive term, rotational Brownian motion increases it. When particle speed is constant, angular correlations affect the equilibrium particle distribution, but not the drift velocity. In contrast, rotational Brownian motion does not affect the equilibrium configuration, but does diminish the drift velocity.

It makes good physical sense that rotational Brownian motion diminishes the drift velocity (when particle speed is constant). Suppose $D_{\rm rot}$ is of the same order of magnitude as a typical $\alpha(\mathbf{u})$. Directions that would otherwise be characterized by a long correlation time, i.e., directions with a low collision rate, have the correlation time cut short by the upper cap of roughly $D_{\rm rot}^{-1}$. Directions with very high collision rates and correlation times short compared to $D_{\rm rot}^{-1}$ are hardly affected. Thus, the drift velocity should decrease, as (7.6) says it does.

Why do angular correlations affect the equilibrium configuration but rotational Brownian motion does not? Rotational Brownian motion takes place constantly, but correlations appear only at collisions and hence have more importance when a particle is moving in a direction characterized by a high collision rate.

VIII. POPULATION DISTRIBUTIONS

Until this point, we have assumed that the particle speed, collision rate, mean angle between runs, and rotational diffusion coefficient are all uniquely determined at each position in space for a given direction of motion. We now relax this assumption and allow different particles of the population to have different characteristics; all quantities that characterize the particles' biased random walk will now be chosen from a probability distribution $\psi(\mathbf{x},\nu,\alpha(\mathbf{u}),\gamma(\mathbf{u}',\mathbf{u}),D_{\rm rot})$. We assume this probability distribution is fixed in time and also local, i.e., the probability that a given particle at position $\mathbf{x}$ has a speed ν, mean cosine of the angle between successive runs Θ, tumble rate $\alpha(\mathbf{u})$, and rotational diffusion coefficient $D_{\rm rot}$, is independent of the probability that when at location $\mathbf{x}'$, the particle will take on some other values of these functions. However, at a fixed location $\mathbf{x}$, different particle characteristics are not independent; for example, $\langle\nu(\mathbf{x})\rangle\langle\alpha(\mathbf{x},\mathbf{u})\rangle\neq\langle\nu(\mathbf{x})\alpha(\mathbf{x},\mathbf{u})\rangle$, where $\langle\ \rangle$ means averaged over the distribution $\psi(\mathbf{x},\nu,\alpha(\mathbf{u}),\gamma(\mathbf{u}',\mathbf{u}),D_{\rm rot})$. To avoid clutter, when computing averages we shall keep integrations that do not affect the averaged quantity, as well as spatial dependencies, implicit. In other words, $\langle\nu(\mathbf{x})\rangle$ shall be written as

$$\int\nu\psi(\nu)d\nu \ ,$$

rather than as

$$\int\int\int\int\nu\psi(\mathbf{x},\nu,\alpha(\mathbf{u}),\gamma,D_{\rm rot})d\nu\, dD_{\rm rot}\mathcal{D}\alpha\,\mathcal{D}\gamma \ .$$

Note that the distribution $\psi(\mathbf{x},\nu,\alpha(\mathbf{u}),\gamma(\mathbf{u}',\mathbf{u}),D_{\rm rot})$ is a functional of $\gamma(\mathbf{u}',\mathbf{u})$ and $\alpha(\mathbf{u})$, so we use the script $\mathcal{D}$ to indicate functional integration. As before, we only consider $\gamma(\mathbf{u}',\mathbf{u})$ that obey (6.1)–(6.3) and $\alpha(\mathbf{u})$ of the form (5.6). Thus,

$$\begin{aligned}\langle\alpha(\mathbf{u})\rangle&=\int\alpha(\mathbf{u})\psi(\alpha(\mathbf{u}))\mathcal{D}\alpha\\&=\int\alpha(\mathbf{u})\psi(\alpha_0,\boldsymbol{\alpha}_1)d\alpha_0 d^n\boldsymbol{\alpha}_1\\&=\int(\alpha_0+\boldsymbol{\alpha}_1\cdot\mathbf{u})\psi(\alpha_0,\boldsymbol{\alpha}_1)d\alpha_0 d^n\boldsymbol{\alpha}_1\\&=\langle\alpha_0\rangle+\langle\boldsymbol{\alpha}_1\rangle\cdot\mathbf{u}\end{aligned} \tag{8.1}$$

and

$$\oint \langle \gamma(\mathbf{u}',\mathbf{u})\rangle d^{n-1}\mathbf{u}' = \oint \left[\int \gamma(\mathbf{u}',\mathbf{u})\psi(\gamma(\mathbf{u}',\mathbf{u}))\mathcal{D}\gamma \right] d^{n-1}\mathbf{u}' = \int \left[\oint \gamma(\mathbf{u}',\mathbf{u})\psi(\gamma(\mathbf{u}',\mathbf{u}))d^{n-1}\mathbf{u}' \right] \mathcal{D}\gamma = 1 \ . \tag{8.2}$$

Similarly,

$$\oint \langle \gamma(\mathbf{u}',\mathbf{u})\rangle d^{n-1}\mathbf{u} = 1 \ , \tag{8.3}$$

$$\oint \langle \gamma(\mathbf{u}',\mathbf{u})\rangle \mathbf{u}' d^{n-1}\mathbf{u}' = \langle \Theta \rangle \mathbf{u} \ . \tag{8.4}$$

In the generalization of (6.4) and (7.4), each term is now averaged over the distribution ψ. For instance, $\nabla(\nu P)\cdot\mathbf{u}$ is replaced with $\nabla(P\int \nu\psi(\nu))\cdot\mathbf{u}$. These replacements yield

$$\frac{\partial P}{\partial t} = -\nabla(P\langle \nu \rangle)\cdot\mathbf{u} - P(\mathbf{u})\langle \alpha(\mathbf{u})\rangle + \oint P(\mathbf{u}')\langle \alpha(\mathbf{u}')\gamma(\mathbf{u},\mathbf{u}')\rangle d^{n-1}\mathbf{u}' + \langle D_{\text{rot}}\rangle \nabla_{\mathbf{u}}^2 P \ . \tag{8.5}$$

Using (8.1)–(8.4), it is straightforward by now to proceed from (8.5) and to find that the equilibrium distribution is characterized by

$$\nabla(\langle \nu \rangle \rho) = -\rho(\langle \boldsymbol{\alpha}_1 \rangle - \langle \boldsymbol{\alpha}_1 \Theta \rangle) \ , \tag{8.6}$$

and so $(\langle \boldsymbol{\alpha}_1 \rangle - \langle \boldsymbol{\alpha}_1 \Theta \rangle)$ replaces $\boldsymbol{\alpha}_1$ in (5.9). Of course, the condition

$$\frac{(\langle \boldsymbol{\alpha}_1 \rangle - \langle \boldsymbol{\alpha}_1 \Theta \rangle)}{\langle \nu \rangle} = \nabla\phi \tag{8.7}$$

for some $\phi(\mathbf{x})$ must be fulfilled for equilibrium to be possible. Using

$$\mathbf{J} = \int v \left[\oint P\mathbf{u}\, d^{n-1}\mathbf{u} \right] \psi(\nu) d\nu = \langle \nu \rangle \oint P\mathbf{u}\, d^{n-1}\mathbf{u} \ , \tag{8.8}$$

and perturbing around the equilibrium solution, the flux is found to be

$$\mathbf{J} = -\left[\frac{\langle \nu \rangle \nabla(\langle \nu \rangle \rho) + \langle \nu \rangle \rho(\langle \boldsymbol{\alpha}_1 \rangle - \langle \boldsymbol{\alpha}_1 \Theta \rangle)}{n(\langle \boldsymbol{\alpha}_0 \rangle - \langle \boldsymbol{\alpha}_0 \Theta \rangle + (n-1)\langle D_{\text{rot}} \rangle)} \right] . \tag{8.9}$$

Equation (8.9) is valid in the long-time approximation close to equilibrium when

$$|\nabla(\langle \nu \rangle \rho) + \rho(\langle \boldsymbol{\alpha}_1 \rangle - \langle \boldsymbol{\alpha}_1 \Theta \rangle)| \ll (\langle \boldsymbol{\alpha}_0 \rangle - \langle \boldsymbol{\alpha}_0 \Theta \rangle + (n-1)\langle D_{\text{rot}} \rangle) \ .$$

Equation (8.9) can be applied to three different types of effective population differences:

(1) *Variation within a species or strain.* Each individual cell of a particular strain of *E. coli* will obviously be slightly different from all others of the same strain. To find the cell flux from (8.9), we use the strain average speed, tumble rates, etc. In treating the variations between individuals, the assumption of local probability distributions is quite restrictive. Although certain deviations of an individual cell's behavior from the mean may be uncorrelated from point to point, other deviations may not be. For instance, certain cells may always swim faster than the strain average speed; such deviation from the mean is certainly not local, because there are strong correlations between different locations. In contrast, a cell exhibiting local deviations from the mean would swim faster than the strain average speed sometimes, but slower other times. The formulation presented here is not sufficient to treat nonlocal probability distributions. Cells that always swim faster than the strain average speed are effectively a different species with their own strain average speed. For such problems, a multispecies formulation is necessary, rather than the single-species formulation presented here.

(2) *Fluctuations in attractant concentration.* At a microscopic level, the concentration of the chemical attractant constantly fluctuates, even though it may appear unchanged macroscopically. If the cell's method for detecting changes in the macroscopic attractant concentration is to be insensitive to these fluctuations, the cell's chemoreception system must be able to perform averages. Fluctuations in attractant concentration are assumed to exhibit a Gaussian distribution with zero mean. Under these conditions, Berg and Purcell showed that the root-mean-square percentage error in the detection of attractant concentration is inversely proportional to the square root of both the duration of the measurement and the average attractant concentration itself [15]. As a result, cells modulate the tumble rate based on averages of attractant concentration measurements taken over a second or more [16,17]. However, from (8.9), if (8.1) remains valid, then *the rms percentage error does not affect the net chemotactic flux of cells.* This conclusion has been verified by computer simulation [33]. Of course, a poor chemoreception system might drastically affect an individual cell, preventing it from reaching a region of high food concentration before the food is injested by competing cells. But over a full lifetime (if the cell makes it that long), even a cell with a poor chemoreception system should experience as many accidental windfalls as disappointments. However, this conclusion is false if (8.1) no longer holds; such scenarios shall be examined in depth elsewhere [33].

(3) *Trajectory averages.* As described in Sec. I, *E. coli* uses attractant concentrations measured in the recent past to modulate the tumble rate, and so the tumble rate is direction dependent. However, the tumble rate is not only a function of direction and position. Two cells at the same spot heading in the same direction have not traveled identical trajectories. The cells' prior meanderings from rotational Brownian motion will not have been identical, nor will the times of the cells' last tumbles. Therefore, the attractant concentrations measured during the last $\sim$4 s will not have been identical, so neither will the cells' tumble rates be identical, even if the physiological and measurement differences described above were somehow eliminated. However, if we consider a cell at a given spot heading in a certain direction, and average over all possible trajectories the cell could have taken to arrive in such a state, weighted by the probability of each trajectory's occurrence, then the trajectory-averaged

tumble rate is a strict function of position and direction. It is such trajectory-averaged quantities that should enter (8.9).

IX. STRATEGIES FOR CHEMOTAXIS

Using (1.3) and (8.9), we shall compare the chemotactic strategies of cells that respond to temporal changes in attractant concentration and those that respond to the averages of absolute concentration of attractant. When moving in a favorable direction, the wild-type cell lowers the tumble rate by an amount proportional to the difference of the average attractant concentrations measured during the last ~ 1 s and the ~ 3 s prior to that. If the attractant concentration gradient does not change much over the course of a single run, i.e., $v\nabla^2 C \ll \alpha|\nabla C|$, this finite difference is an approximation to the time derivative $\partial C/\partial t$, and so we approximate (1.3) as

$$\alpha(t)=\alpha_0-g\frac{dC}{dt}=\alpha_0-gv\nabla C\cdot\mathbf{u} , \tag{9.1}$$

for a cell swimming at constant speed v in direction $\mathbf{u}$, when $\nabla C\cdot\mathbf{u}>0$. The gain g has dimensions of inverse attractant concentration and is proportional to the gain of Eq. (1.3), $\tilde{g}$, which has dimensions of inverse attractant concentration times inverse time. Equation (9.1) is a useful approximation because it yields a tumble rate that depends only upon the cell's position and direction of motion. However, as discussed in Sec. VIII, such an approximation is not valid unless we perform trajectory averages, to account for differences in prior tumble times and rotational Brownian motion. But it will not suffice to blindly replace quantities in (9.1) with their trajectory-averaged value, as in

$$\langle\alpha\rangle=\langle\alpha_0\rangle-gv\nabla C\cdot\mathbf{u}=\langle\alpha_0\rangle+\langle\boldsymbol{\alpha}_1\rangle\cdot\mathbf{u} . \tag{9.2}$$

In Sec. VIII, we required that all the tumble rates being averaged over were of the form (5.6). Equation (1.3) does not meet this requirement, and so we cannot expect the wild-type trajectory-averaged tumble rate to satisfy (9.2). However, we know that the trajectory-averaged tumble rate will be a monotonically increasing function of the angle between the attractant gradient ∇C and the direction of motion $\mathbf{u}$:

$$\langle\alpha\rangle=\langle\alpha_0\rangle-h(\mathbf{u}\cdot\nabla C) , \tag{9.3}$$

where h is a monotonically decreasing function. We now expand (9.3) in a power series in $\mathbf{u}\cdot\nabla C$, and because we are primarily interested in qualitative behavior, we truncate the series after linear order. After factoring out the swimming speed v from the coefficient of the linear term, we find a tumble rate of the form (9.2). The gain g in (9.2) contains all the factors of order unity arising from the average over all the possible trajectories that could bring a cell to position $\mathbf{x}$ with direction of motion $\mathbf{u}$; fortunately, we shall not need to calculate g explicitly. Furthermore, to use results of Sec. VIII, we assume that (9.2) holds even when the cell is moving down the attractant gradient. This last assumption contradicts experimental findings, but we argue that it shall also not upset qualitative predictions. Comparison with experiment shall support this claim, as well as extensive simulation data to be published elsewhere [33].

No experimental evidence exists suggesting $\langle\boldsymbol{\alpha}_1\Theta\rangle\neq\langle\boldsymbol{\alpha}_1\rangle\langle\Theta\rangle$, or that the mean speed is altered in response to attractant stimuli, and so combining (9.2) and (8.9), we approximate the wild-type cell flux as

$$\mathbf{J}=\left[\frac{-\langle v\rangle^2[\nabla\rho-\rho g\nabla C(1-\langle\Theta\rangle)]}{n[\langle\alpha_0\rangle(1-\langle\Theta\rangle)+(n-1)\langle D_{\text{rot}}\rangle]}\right] . \tag{9.4}$$

The drift velocity in (9.4) is

$$v_{\text{drift}}=\frac{\langle v\rangle^2 g\nabla C(1-\langle\Theta\rangle)}{n[\langle\alpha_0\rangle(1-\langle\Theta\rangle)+(n-1)D_{\text{rot}}]} . \tag{9.5}$$

If we assume that ∇C is steep, so that the tumble rate is $\approx 0.5\ \text{s}^{-1}$ when cells head straight up the gradient [9], and use $\langle\alpha_0\rangle\approx 1\ \text{s}^{-1}$, $\langle v\rangle\approx 20\ \mu\text{m/s}$ [9], we find $g\nabla C\approx 0.025\ \mu\text{m}^{-1}$. Taking $\langle D_{\text{rot}}\rangle\approx 0.1\ \text{rad}^2/\text{s}$ [10], and $\langle\Theta\rangle\approx 0.5$ [9], in three dimensions we find $v_{\text{drift}}\approx 5\ \mu\text{m/s}$. This prediction is in reasonable agreement with experimentally observed drift velocities of $\approx 4\ \mu\text{m/s}$ in steep gradients [20]. However, swimming speeds may range up to $30\ \mu\text{m/s}$, so this quantitative agreement is not robust. More importantly, we have predicted that the wild-type chemotactic response is strong: the drift velocity is $\sim 25\%$ of the raw swimming speed.

From (9.2), the condition (8.7) is satisfied with

$$\nabla\phi=-g\nabla C(1-\langle\Theta\rangle) , \tag{9.6}$$

and so equilibrium results when

$$\rho(\mathbf{x})=\rho(\mathbf{x}_0)\exp\{g[C(\mathbf{x})-C(\mathbf{x}_0)](1-\langle\Theta\rangle)\} . \tag{9.7}$$

If $C(\mathbf{x})-C(\mathbf{x}_0)>0$, then the cells have an exponentially higher equilibrium density at position $\mathbf{x}$ than at $\mathbf{x}_0$. By all measures, the wild-type strategy for modulating the tumble rate is very successful for seeking out food.

Given that the chemotactic behavior of wild-type *E. coli* evolved over millions of years, it is hardly surprising that its strategy for chemotaxis is successful. We now study cells that cannot adapt. The impulse response of such cells (Fig. 2) has only a positive lobe of duration $\Delta t\sim 1$ s [16,17], and so the cell cannot perform temporal comparisons; instead it merely decreases the tumble rate by an amount proportional to the average of the attractant concentration experienced over the last ~ 1 s. Assuming that the attractant concentration experienced over the last Δt seconds is relatively constant, $\Delta t v\nabla C \ll C$, we approximate (1.3) as

$$\alpha(t)=\alpha_0-\tilde{g}C[\mathbf{x}(t)] , \tag{9.8}$$

for a cell at position $\mathbf{x}$ at time t. However, because the determination of $C(\mathbf{x})$ requires Δt seconds, the measured value will depend somewhat on the cell's direction of motion. Cells moving down (up) an attractant gradient come from a region of slightly higher (lower) attractant concentration; consequently, cells moving down the gradient will have a lower tumble rate than those moving up. In contrast, cells that perform temporal comparisons have the lower tumble when moving up the gradient, so

we expect a *reverse* taxis for cells that do not adapt. To add direction dependence to (9.8), we take trajectory averages and make the same approximations discussed above, yielding

$$\langle \alpha(\mathbf{x},\mathbf{u})\rangle = \langle \alpha_0\rangle - \tilde{g}[C(\mathbf{x}) - v\langle \Delta t\rangle \nabla C\cdot\mathbf{u}] , \quad (9.9)$$

where $\langle \Delta t\rangle$ is the effective measurement duration after trajectory averaging. Equation (9.9) has the form (5.6) with

$$\mathbf{J}(\mathbf{x}) = -\left[\frac{\langle v\rangle^2[\nabla\rho + \rho\tilde{g}\langle \Delta t\rangle \nabla C(1-\langle\Theta\rangle)]}{n\{[\langle\alpha_0\rangle - \tilde{g}C(\mathbf{x})](1-\langle\Theta\rangle) + (n-1)\langle D_{\rm rot}\rangle\}}\right] , \quad (9.12)$$

and so equilibrium results when

$$\rho(\mathbf{x}) = \rho(\mathbf{x}_0)\exp\{-\tilde{g}\langle \Delta t\rangle [C(\mathbf{x}) - C(\mathbf{x}_0)](1-\langle\Theta\rangle)\} . \quad (9.13)$$

If $C(\mathbf{x}) - C(\mathbf{x}_0) > 0$, cells that do not adapt have an exponentially *lower* equilibrium density at position $\mathbf{x}$ than at $\mathbf{x}_0$. This strategy is clearly unsuccessful and indeed represents reverse chemotaxis. This effect has been noted in computer simulations [8], but has yet to be observed in free swimming cells.

Why has reverse chemotaxis not been observed? Competing pseudochemotactic effects could conceivably be to blame. Unlike (9.4), (9.12) exhibits a diffusion coefficient and a drift velocity that vary spatially even in a uniform gradient. In the limit that $\langle \Delta t\rangle$ is very large, reverse chemotaxis outweighs the pseudochemotactic flux. In the opposite limit, drift down the gradient becomes less significant than the pseudochemotactic flux toward regions of high diffusion coefficient, i.e., high attractant concentration. In theory, this small $\langle \Delta t\rangle$ limit might be the relevant one for *cheRcheB* cells, commonly designated as adaption defective. However, the analysis below suggests that *cheRcheB* cells may still have partial adaption abilities, preventing reverse chemotaxis from being observed.

X. PARTIAL ADAPTION IN METHYLATION-DEFECTIVE CELLS?

The question of whether *cheRcheB* cells are chemotactic has attracted interest [20–22] because these cells lack methylating and demethylating enzymes thought to play a significant role in *E. coli*'s chemosensory system (for a review, see Ref. [34]). If these cells are found to be nonchemotactic, this result would strengthen the view that the absent enzymes are necessary for chemotaxis. Over time scales of a few seconds, *cheRcheB* cells appear adaption defection in tethering experiments [16,17], but over a minute or so, a fraction of the *cheRcheB* population appears partially adaptive [17]. Experiments with swimming cells show that *cheRcheB* tend to weakly head up attractant gradients [21,22]; however, such results cannot distinguish nonadaptive pseudochemotaxis from true adaptive chemotaxis.

Experiments performed by Berg and Turner [20] were specifically designed to separate pseudochemotactic effects from true chemotactic ones, and are highly amenable to theoretical analysis. With a population of *cheR-cheB* cells placed initially at the top (bottom) of a linear attractant gradient, Berg and Turner measured the steady-state flux of cells moving down (up) the gradient, Berg and Turner measured the steady-state flux of cells moving down (up) the gradient, and found that the ratio of the flux up the gradient to the flux down the gradient, J_+/J_-, was ≈ 2. However, if a second attractant was added uniformly, causing a global reduction in α_0, $J_+/J_- \approx 1$. What do these results imply about *cheRcheB* cells?

$$\alpha_1 = \tilde{g}v\langle \Delta t\rangle \nabla C , \quad (9.10)$$

and an effective α_0,

$$\langle \alpha_0^{\rm eff}(\mathbf{x})\rangle = \langle \alpha_0\rangle - \tilde{g}C(\mathbf{x}) . \quad (9.11)$$

From (8.9), (9.10), and (9.11), we are led to the flux

As we now show, in the absence of any drift term in the flux, cells undergoing pseudochemotaxis alone, (2.7), should exhibit a J_+/J_- ratio of 1. At steady state in a linear attractant gradient, the flux in the direction of the gradient is constant and perpendicular fluxes are zero, so we use one-dimensional theory. If x_1 is the length of the gradient, integrating (2.7) from $x=0$ to $x=x_1$ yields a steady-state flux

$$J = \frac{\rho_0 - \rho_1}{\int_0^{x_1} \frac{dx}{D(x)}} , \quad (10.1)$$

where ρ_0 and ρ_1 are the cell densities at $x=0$ and $x=x_1$, respectively. If the attractant gradient is now reversed in sign, $D(x)$ is replaced in the integral by $D(x_1-x)$, because regions previously characterized by a low (high) diffusion coefficient are now characterized by a high (low) one. But this switch leaves the integral unchanged, through the change of integration variables $x' = x_1 - x$ in (10.1). Thus pseudochemotaxis alone yields a J_+/J_- ratio of 1.

When we include reverse chemotaxis effects in addition, we intuitively expect adaption-defective cells to exhibit a J_+/J_- ratio less than 1. Using (9.12), an explicit calculation for a linear gradient with $|\nabla C| = C'$ yields

$$\frac{J_+}{J_-} = \frac{[(1-e^{-x_1/l})(Al+Bl^2) - Blx_1]}{[(e^{x_1/l}-1)(Al+Bl^2) - Blx_1 e^{x_1/l}]} , \quad (10.2)$$

where

$$A = n[\langle \alpha_0 \rangle (1-\langle \Theta \rangle) + (n-1)\langle D_{\rm rot} \rangle] , \quad (10.3a)$$

$$B = n\tilde{g}C'(1-\langle \Theta \rangle) , \quad (10.3b)$$

$$l^{-1} = \tilde{g}\langle \Delta t \rangle C'(1-\langle \Theta \rangle) . \quad (10.3c)$$

The aspartate attractant gradients used by Berg and Turner rose from 0 to 100 μM over 500 μm; *cheRcheB* cells typically tumbled at the rate $\approx 3\ {\rm s}^{-1}$ at the bottom of the gradient, but suppressed tumbles significantly less near the top [35]. For the following rough calculation, we assume no tumbles occur at the top of the gradient; this behavior gives $\tilde{g}C' \approx 0.006\ \mu{\rm m}^{-1}\,{\rm s}^{-1}$. Using the full width at half maximum of Fig. 2 to approximate $\langle \Delta t \rangle \approx 0.5\ {\rm s}^{-1}$, and taking other values from above, yields $J_+/J_- \approx 0.4$. The finding $J_+/J_- < 1$ is robust when pseudochemotaxis is combined with negative drift (9.12) and is insensitive to the values chosen. Because $J_+/J_- \geq 1$ for *cheRcheB* cells, it is likely that at least a fraction of these cells partially adapt. However, this adaption ability is certainly small compared to that of wild-type cells, which exhibited $J_+/J_- \approx 350$ in aspartate gradients ten times gentler than the gradients used in the *cheRcheB* experiments.

ACKNOWLEDGMENTS

I am grateful to Professor H. Stone for his help with Appendix B, to Professor H. J. Schnitzer for supplying the identity (A2), to Professor J. Stock for bringing very tumbly mutant *E. coli* strains to my attention, and to L. Turner for discussion about Ref. [20]. I am also particularly grateful to my teachers Dr. S. M. Block and Professor H. C. Berg for introducing me to the subjects of random walks and bacterial chemotaxis, for innumerable helpful discussions, and for constant encouragement. Without the help of these two scientists, this paper would never have been written. This research has been supported by the Rowland Fund of the Harvard Physics Department and by the Rowland Institute for Science.

APPENDIX A

In this appendix we show that equilibrium is unattainable in the general case (5.4). We emphasize that this proof does not exclude the possibility of steady-state solutions with nonvanishing particle flux, or even equilibrium solutions not obeying (5.6).

It is easier to work with indices and the Einstein summation convention than with boldfaced tensors. We shall also need to know the following integrals over the $(n-1)$-dimensional hypersphere:

$$\oint u_{i_1} u_{i_2} \cdots u_{i_{2k+1}} d^{n-1}\mathbf{u} = 0 \quad (A1)$$

and

$$\oint u_{i_1} u_{i_2} \cdots u_{i_{2k}} d^{n-1}\mathbf{u} = C_{2k}\Omega(n)[\delta_{i_1,i_2}\delta_{i_3,i_4}\cdots\delta_{i_{2k-1},i_{2k}} + (\text{all possible pairings})] . \quad (A2)$$

In (A2), (all possible pairings) stands for the $(2k)!/k!2^k$ distinct products of k Kronecker deltas involving the indices i_1 to i_{2k}. The coefficient C_{2k} is given by the recursion formula

$$C_{2k+2} = \frac{C_{2k}}{n-1+\left[\dfrac{(2k+2)!}{(k+1)!2^{k+1}}\right]} , \quad (A3a)$$

$$C_0 = 1 . \quad (A3b)$$

Equations (4.6) and (4.11) will be recognized as special cases of (A1) and (A2).

To see if equilibrium can be attained, we substitute (5.4) and (5.5) into (5.3), require that the time derivative in (5.3) vanish, and use (A1)–(A3) to find

$$\frac{-\partial_j(v\rho)}{n} = \rho \sum_{k=1}^{\infty} \{C_{2k}(\alpha_{2k-1})_{i_1 i_2 \cdots i_{2k-1}} \times [\delta_{i_1,i_2}\delta_{i_3,i_4}\cdots\delta_{i_{2k-3}i_{2k-2}}\delta_{i_{2k-1},j} + (\text{all possible pairings})]\} . \quad (A4)$$

For equilibrium to be possible, putting (A4) and (5.6) back into (5.1) must yield zero. Instead, we find

$$\frac{\partial P}{\partial t} = -\rho \left\{ \sum_{k=1}^{\infty} (\alpha_{2k})_{i_1 i_2 \cdots i_{2k}} \{u_{i_1}u_{i_2}\cdots u_{i_{2k}} - C_{2k}[\delta_{i_1,i_2}\delta_{i_3,i_4}\cdots\delta_{i_{2k-1},i_{2k}} + (\text{all possible pairings})]\} - \sum_{k=2}^{\infty} (\alpha_{2k-1})_{i_1 i_2 \cdots i_{2k-1}} \{u_{i_1}u_{i_2}\cdots u_{i_{2k-1}} - C_{2k}u_j[\delta_{i_1,i_2}\delta_{i_3,i_4}\cdots\delta_{i_{2k-3},i_{2k-2}}\delta_{i_{2k-1},j} + (\text{all possible pairings})]\} \right\} . \quad (A5)$$

Thus, if (5.6) does not hold, then (5.4) and (A4) do not describe equilibrium.

APPENDIX B

To show

$$\nabla_{\mathbf{u}}^2 \mathbf{u} = (1-n)\mathbf{u} , \quad (B1)$$

it is easier to use indices and the Einstein summation convention than to use vector notation. Consider a vector on the $(n-1)$-dimensional unit hypersphere. Representing this vector in Cartesian coordinates in the n-dimensional imbedding space, the ith component is

$$\frac{u_i}{(u_k^2)^{1/2}} . \quad (B2)$$

The denominator of (B2) manifestly contrains the vector to have length 1, so differentiating (B2) will not involve changes with respect to the vector's length, only its direction. Taking the Laplacian of (B2) yields

$$\frac{\partial^2}{\partial u_i^2}\left[\frac{u_j}{(u_k^2)^{1/2}}\right]=\frac{\partial}{\partial u_i}\left[\frac{\delta_{ij}}{(u_k^2)^{1/2}}-\frac{u_j u_i}{(u_k^2)^{3/2}}\right]$$
$$=\frac{-(n+2)u_i\delta_{ij}}{(u_k^2)^{3/2}}+\frac{3u_j u_i u_i}{(u_k^2)^{5/2}}$$
$$=(1-n)\frac{u_j}{(u_k^2)^{3/2}}\ . \tag{B3}$$

Having differentiated, we explicitly set $(u_i^2)^{1/2}$ to 1 in (B3), and so reach (B1).

As an example, if direction space is the two-sphere parametrized by the polar angle θ and the azimuthal angle φ, then,

$$\nabla_{\mathbf{u}}^2=\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left[\sin\theta\frac{\partial}{\partial\theta}\right]+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\ . \tag{B4}$$

The unit vector $\mathbf{u}$ is represented as

$$\mathbf{u}=\begin{bmatrix}\sin\theta\cos\varphi\\ \sin\theta\sin\varphi\\ \cos\theta\end{bmatrix}, \tag{B5}$$

and so $\nabla_{\mathbf{u}}^2\mathbf{u}=-2\mathbf{u}$.

*Present address: Joseph Henry Laboratories, Princeton University, Princeton, NJ 08544.

[1] S. Goldstein, Q. J. Appl. Math. **4**, 2 (1951).
[2] G. H. Weis, J. Stat. Phys. **15**, 157 (1976); J. Masoliver, K. Lindenberg, and G. H. Weis, Physica A **157**, 891 (1989); M. Araujo, S. Havlin, G. H. Weis, and H. E. Stanley, Phys. Rev. A **43**, 5207 (1991).
[3] P. J. Flory, *Principles of Polymer Chemistry* (Cornell University Press, Ithaca, 1953).
[4] P. S. Lovely and F. W. Dahlquist, J. Theor. Biol. **50**, 477 (1975).
[5] G. I. Taylor, Proc. London Math. Soc. **20**, 196 (1920).
[6] M. Kac, Rocky Mount. J. Math. **4**, 497 (1974).
[7] R. K. Clayton, Arch. Mikrobiol. **27**, 311 (1957).
[8] M. J. Schnitzer, S. M. Block, H. C. Berg, and E. M. Purcell, Symp. Soc. Gen. Microbiol. **46**, 15 (1990).
[9] H. C. Berg and D. A. Brown, Nature **239**, 500 (1972).
[10] H. C. Berg, *Random Walks in Biology* (Princeton University Press, Princeton, NJ, 1983).
[11] R. Nossal, J. Stat. Phys. **30**, 392 (1983).
[12] J. Adler, Science **166**, 1588 (1969).
[13] D. A. Brown and H. C. Berg, Proc. Natl. Acad. Sci. USA **71**, 1388 (1974).
[14] H. C. Berg and D. A. Brown, Antibiot. Chemother. **19**, 55 (1974).
[15] H. C. Berg and E. M. Purcell, Biophys. J. **20**, 193 (1977).
[16] S. M. Block, J. E. Segall, and H. C. Berg, Cell **31**, 215 (1982).
[17] J. E. Segall, S. M. Block, and H. C. Berg, Proc. Natl. Acad. Sci. USA **83**, 8987 (1986).
[18] S. M. Block, J. E. Segall, and H. C. Berg, J. Bacteriol. **154**, 312 (1983).
[19] A. Ishihara, J. E. Segall, S. M. Block, and H. C. Berg, J. Bacteriol. **155**, 228 (1983).
[20] H. C. Berg and L. Turner, Biophys. J. **58**, 919 (1990).
[21] J. Stock, A. Borczuk, F. Chiou, and J. E. B. Burchenal, Proc. Natl. Acad. Sci. USA **82**, 8364 (1985).
[22] R. M. Weis and D. E. Koshland, Jr., Proc. Natl. Acad. Sci. USA **85**, 83 (1988).
[23] D. A. Mcquarrie, *Statistical Mechanics* (Harper & Row, New York, 1976).
[24] B. Gaveau, T. Jacobson, M. Kac, and L. S. Schulman, Phys. Rev. Lett. **53**, 419 (1984).
[25] D. G. C. Mckeon and G. N. Ord, Phys. Rev. Lett. **96**, 3 (1992).
[26] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953).
[27] I. R. Lapidus, J. Theor. Biol. **86**, 91 (1980).
[28] M. A. Rivero, R. T. Tranquillo, H. M. Buettner, and D. A. Lauffenburger, Chem. Eng. Sci. **44**, 2881 (1989).
[29] J. Crank, *The Mathematics of Diffusion* (Oxford University Press, Oxford, 1975).
[30] W. Jost, *Diffusion in Solids, Liquids, Gases* (Academic, New York, 1969).
[31] D. W. Stroock, Z. Wahrsch. Ver. Geb. **28**, 305 (1974).
[32] R. M. Ford and P. T. Cummings, SIAM J. Appl. Math. **52**, 1426 (1992).
[33] M. J. Schnitzer, S. M. Block, and H. C. Berg (unpublished).
[34] R. M. Macnab, in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology,* edited by F. C. Neidhart *et al.* (American Society for Microbiology, Washington, DC, 1987), p. 732.
[35] H. C. Berg and L. Turner (personal communication).