



Published in final edited form as:

Am J Phys. 2006 February 1; 74(2): 123–133. doi:10.1119/1.2142789.

Teaching the principles of statistical dynamics

Kingshuk Ghosh,

Department of Biophysics, University of California, San Francisco, California 94143

Ken A. Dill,

Department of Biophysics, University of California, San Francisco, California 94143

Mandar M. Inamdar,

Division of Engineering and Applied Science, California Institute of Technology, Pasadena, California 91125

Effrosyni Seitaridou, and

Division of Engineering and Applied Science, California Institute of Technology, Pasadena, California 91125

Rob Phillips^{a)}

Division of Engineering and Applied Science and Kavli Nanoscience Institute, California Institute of Technology, Pasadena, California 91125

Abstract

We describe a simple framework for teaching the principles that underlie the dynamical laws of transport: Fick's law of diffusion, Fourier's law of heat flow, the Newtonian viscosity law, and the mass-action laws of chemical kinetics. In analogy with the way that the maximization of entropy over microstates leads to the Boltzmann distribution and predictions about equilibria, maximizing a quantity that E. T. Jaynes called "caliber" over all the possible *microtrajectories* leads to these dynamical laws. The principle of maximum caliber also leads to dynamical distribution functions that characterize the relative probabilities of different microtrajectories. A great source of recent interest in statistical dynamics has resulted from a new generation of single-particle and single-molecule experiments that make it possible to observe dynamics one trajectory at a time.

I. INTRODUCTION

We describe an approach for teaching the principles that underlie the dynamical laws of transport of particles (Fick's law of diffusion), energy (Fourier's law of heat flow), momentum (the Newtonian law for viscosity),¹ and mass-action laws of chemical kinetics.² Recent experimental advances now allow for studies of forces and flows at the single-molecule and nanoscale level, representative examples of which may be found in Refs. 3-10. For example, single-molecule methods have explored the packing of DNA inside viruses⁷ and the stretching of DNA and RNA molecules.^{9,10} Similarly, video microscopy now allows for the analysis of trajectories of individual submicron size colloidal particles,¹¹ and the measurement of single-channel currents has enabled the kinetic studies of DNA translocation through nanopores.^{3,6}

One of the next frontiers in biology is to understand the "small numbers" problem: how does a biological cell function given that most of its proteins and nucleotide polymers are present

in numbers much smaller than Avogadro's number?¹² For example, one of the most important molecules, a cell's DNA, occurs in only a single copy. Also, it is the flow of matter and energy through cells that makes it possible for organisms to maintain a relatively stable form.¹³ Hence, cells must be in this stable state far from equilibrium to function. Thus, many problems of current interest involve small systems that are out of equilibrium.

Our interest here is twofold: to teach students a physical foundation for the phenomenological macroscopic laws, which describe the properties of averaged forces and flows, and to teach them about the dynamical fluctuations away from those average values for systems with small numbers of particles.

In this article we describe a simple way to teach these principles. We start from the "principle of maximum caliber" first described by E. T. Jaynes.¹⁴ It aims to provide the same type of foundation for the dynamics of systems with many degrees of freedom that the second law of thermodynamics provides for the equilibria of such systems. To illustrate the principle, we use a slight variant of one of the oldest and simplest models in statistical mechanics, the dog-flea or two urn model.¹⁵⁻¹⁷ Courses in dynamics often introduce Fick's law, Fourier's law, and the Newtonian-fluid model as phenomenological laws, rather than deriving them from some deeper foundation. We describe an unified perspective that we have found useful for teaching these laws from a foundation in statistical dynamics. In analogy with the role of *microstates* as a basis for the properties of equilibria, we focus on *microtrajectories* as the basis for predicting dynamics.

One argument that might be leveled against the kind of framework we present here is that in the cases we consider it is not clear that it leads to anything different from what one obtains using conventional nonequilibrium thinking. On the other hand, restating the same physical result in different language often can provide a better starting point for subsequent reasoning. This point was well articulated by Feynman in his Nobel lecture:¹⁸

"Theories of the known, which are described by different physical ideas may be equivalent in all their predictions and are hence scientifically indistinguishable. However, they are not psychologically identical when trying to move from that base into the unknown. For different views suggest different kinds of modifications which might be made and hence are not equivalent in the hypotheses one generates from them in one's attempt to understand what is not yet understood."

We begin with the main principle embodied in Fick's law. Why do particles and molecules in solution flow from regions of high concentration toward regions of low concentration? To keep it simple, we consider one-dimensional diffusion along a coordinate x . The diffusion is described by Fick's first law of particle transport,^{1,2} which says that the average flux $\langle J \rangle$ is given in terms of the gradient of the average concentration $\langle c \rangle$ by

$$\langle J \rangle = -D \frac{\partial \langle c \rangle}{\partial x}, \quad (1)$$

where D is the diffusion coefficient. To clearly distinguish quantities that are dynamical averages from those that are not, we indicate the averaged quantities explicitly by angular brackets, $\langle \dots \rangle$. Before we describe the nature of this averaging and the nature of the dynamical distribution functions over which the averages are taken, we briefly review the standard derivation of the diffusion equation. We combine Fick's first law with particle conservation,

$$\frac{\partial \langle c \rangle}{\partial t} = - \frac{\partial \langle J \rangle}{\partial x}, \quad (2)$$

and obtain Fick's second law, which is also known as the diffusion equation:

$$\frac{\partial \langle c \rangle}{\partial t} = D \frac{\partial^2 \langle c \rangle}{\partial x^2}. \quad (3)$$

The solution of Eq. (3) subject to two boundary conditions and one initial condition gives $\langle c(x,t) \rangle$, the average concentration in time and space, and the average flux $\langle J(x,t) \rangle$ when no other forces are present. The generalization to situations involving additional applied forces is the Smoluchowski equation.²

A simple experiment shows the distinction between *averaged* quantities and *individual microscopic realizations*. By using a microfluidics chip like that shown in Fig. 1(a), it is possible to create a small fluid chamber divided into two regions by control valves. The chamber is filled on one side with a solution containing a small concentration of micronscale colloidal particles. The other region contains just water. The three control valves on top of that microfluidic chamber serve two purposes: The two outer ones are used for isolation so that no particles can diffuse in and out of the chamber, while the middle control valve provides the partition between the two regions. The evolution of the system is then monitored after the removal of the partition [see Fig. 1(b)]. The time-dependent particle density is determined by dividing the chamber into a number of equal sized boxes in the long direction and by computing histograms of the numbers of particles in each slice as a function of time. This system is a colloidal solution analog of the gas diffusion experiments of classical thermodynamics. The corresponding theoretical model usually used is the diffusion equation. Figure 1(c) shows the solution to the diffusion equation as a function of time for the geometry of the microfluidics chip. The initial condition is a step function in concentration at $x=200 \mu\text{m}$ at time $t=0$.

We use this simple experiment to illustrate one main point. The key distinction is that the theoretical curves are very smooth, while there are very large fluctuations in the experimentally observed dynamics of the particle densities. The fluctuations are large because the number of colloidal particles is small, tens to hundreds. The experimental data show that the particle concentration $c(x,t)$ is a highly fluctuating quantity. It is not well described by the standard smoothed curves that are calculated from the diffusion equation. Of course, when averaged over many trajectories or when particles are at high concentrations, the experimental data should approach the smoothed curves that are predicted by the classical diffusion equation.

II. THE EQUILIBRIUM PRINCIPLE OF MAXIMUM ENTROPY

Because our strategy follows so closely the Jaynes derivation of the Boltzmann distribution law of equilibrium statistical mechanics,^{2,19} we first show the equilibrium treatment. To derive the Boltzmann law, we start with a set of equilibrium microstates $i=1,2,3, \dots, N$ that are relevant to the problem at hand. Our aim is to calculate the probabilities p_i of these microstates in equilibrium. We define the entropy S of the system as

$$S(\{p_i\}) = -k_B \sum_{i=1}^N p_i \ln p_i, \quad (4)$$

where k_B is Boltzmann's constant. The equilibrium probabilities, $p_i=p_i^*$ are those values of p_i that cause the entropy to be a maximum, subject to two constraints:

$$\sum_{i=1}^N p_i = 1, \quad (5)$$

which is a normalization condition that ensures that the probabilities p_i sum to one, and

$$\langle E \rangle = \sum_i p_i E_i, \quad (6)$$

which says that the energies, when averaged over all the microstates, sum to the macroscopically observable average energy.

By using Lagrange multipliers λ and β to enforce the first and second constraints, respectively, we obtain an expression for the values p_i^* that maximize the entropy:^{2,19}

$$\sum_i [-1 - \ln p_i^* - \lambda - \beta E_i] = 0. \quad (7)$$

The result is that

$$p_i^* = \frac{e^{-\beta E_i}}{Z}, \quad (8)$$

where Z is the partition function, defined by

$$Z = \sum_i e^{-\beta E_i}. \quad (9)$$

After a few thermodynamic arguments, the Lagrange multiplier β can be shown to be equal to $1/k_B T$.¹⁹ This derivation, first given in this simple form by Jaynes,¹⁹ identifies the probabilities that are both consistent with the observable average energy and that otherwise maximize the entropy. Jaynes justified this strategy on the grounds that it would be the best prediction that an observer could make, given the observable mean energy, if the observer is ignorant of all else. Although this derivation of the Boltzmann law is now popular, its interpretation as a method of prediction, rather than as a method of physics, is controversial. Nevertheless, for our purposes here, it does not matter whether we regard this approach as a description of physical systems or as a strategy for making predictions.

We switch from the principle of maximum entropy to the principle of maximum caliber.¹⁴ In particular, rather than focusing on the probability distribution $p(E_i)$ for the various microstates, we seek $p[\{\sigma_i(t)\}]$, where $\sigma_i(t)$ is the i th microscopic trajectory of the system. Again we maximize an entropy-like quantity obtained from $p[\{\sigma_i(t)\}]$ to obtain the predicted distribution of microtrajectories. If there are no constraints, this maximization results in the prediction that all the possible microtrajectories are equally likely during the dynamical process. In contrast, certain microtrajectories will be favored if there are dynamical constraints, such as may be specified in terms of the average flux.

III. FICK'S LAW FROM THE DOG-FLEA MODEL

We want to determine the diffusive evolution of particles in a one-dimensional system. The key features of this system are revealed by considering two columns of particles separated by a plane, as shown in Fig. 2. The left-hand column 1 has $N_1(t)$ particles at time t and the right-hand column 2 has $N_2(t)$ particles. This system is a simple variant of the famous “dog-flea” model of Ehrenfest’s introduced in 1907.^{15,17} Column 1 corresponds to dog 1, which

has N_1 fleas on its back at time t , and column 2 corresponds to dog 2, which has N_2 fleas at time t . In the time interval between time t and $t+\Delta t$, a flea can either stay on its current dog or jump to the other dog. This model has been used extensively to study the Boltzmann H-theorem and to understand how the time asymmetry of diffusion processes arises from the underlying time symmetry in the laws of motion.^{15,17,20} We use this model for a slightly different purpose. In particular, our aim is to take a well-characterized problem like diffusion and to reveal how the principle of maximum caliber may be used in a concrete way.

First consider the equilibrium state of the dog-flea model. The total number of ways of partitioning the (N_1+N_2) fleas is

$$W(N_1, N_2) = \frac{(N_1+N_2)!}{N_1!N_2!}. \quad (10)$$

The state of equilibrium is that for which the entropy, $S=k_B \ln W$, is a maximum. A simple calculation shows that the entropy is a maximum when the value $N_1=N_1^*$ is as nearly equal to $N_2=N_2^*$ as possible. In short, at equilibrium, both dogs will have approximately the same number of fleas, in the absence of any bias.

Our focus here is on how the system reaches equilibrium. We discretize time into a series of intervals Δt . We define a dynamical quantity p , which is the probability that a particle (flea) jumps from one column (dog) to the other in any time interval Δt . Thus, the probability that a flea stays on its dog during the time interval is $q=1-p$. We assume that p is independent of the time t and that all the fleas and jumps are independent of each other.

In equilibrium statistical mechanics, the focus is on the microstates. However, for dynamics we focus on *processes*, which at the microscopic level we call the microtrajectory. Characterizing the dynamics requires more than just information about the microstates; we must also consider the processes. Let m_1 represent the number of particles that jump from column 1 to 2 and m_2 represent the number of particles that jump from column 2 to 1 between time t and $t+\Delta t$. There are many possible values of m_1 and m_2 : it is possible that no fleas will jump during the interval Δt , or that all the fleas will jump, or that the number of fleas jumping will be in between these limits. Each one of these different situations corresponds to a distinct microtrajectory of the system in this idealized dynamical model. We need a principle to tell us what number of fleas will jump during the time interval Δt at time t . Because the dynamics of this model is so simple, the implementation of the caliber idea is reduced to a simple exercise in enumeration and counting using the binomial distribution.

A. The dynamical principle of maximum caliber

The probability $W_d(m_1, m_2 | N_1, N_2)$ that m_1 particles jump to the right and that m_2 particles jump to the left in a time interval Δt , given that there are $N_1(t)$ and $N_2(t)$ fleas on the dogs at time t , is

$$W_d(m_1, m_2 | N_1(t), N_2(t)) = \underbrace{\left[\frac{p^{m_1} q^{N_1 - m_1} N_1!}{m_1! (N_1 - m_1)!} \right]}_{W_{d_1}} \times \underbrace{\left[\frac{p^{m_2} q^{N_2 - m_2} N_2!}{m_2! (N_2 - m_2)!} \right]}_{W_{d_2}}. \quad (11)$$

W_d counts microtrajectories in dynamics in the same spirit that W counts microstates for equilibrium. In the same spirit of the second law of thermodynamics, we maximize W_d over all the possible microtrajectories (that is, over m_1 and m_2) to predict the flux of fleas between the dogs. This maximization is the implementation of the principle of maximum caliber for this model. Maximizing W_d over all the possible processes (different values of m_1 and m_2) gives our prediction (right flux $m_1 = m_1^*$ and left flux $m_2 = m_2^*$) for the macroscopic flux that we should observe experimentally. (We follow the usual definition of flux, the number of particles transferred per unit time and per unit area. For simplicity, we take the cross-sectional area to be unity.)

Because the jumps of the fleas from each dog are independent, we find the predicted macroscopic dynamics by maximizing W_{d_1} and W_{d_2} separately, or for convenience their logarithms:

$$\frac{\partial \ln W_{d_i}}{\partial m_i} \Big|_{N, m_i = m_i^*} = 0 \quad (i=1, 2). \quad (12)$$

The application of Stirling's approximation to Eq. (11) gives

$$\ln W_{d_i} = m_i \ln p + (N_i - m_i) \ln q + N_i \ln N_i - m_i \ln m_i - (N_i - m_i) \ln (N_i - m_i). \quad (13)$$

We call $\mathcal{C} = \ln W_d$ the caliber. Maximizing \mathcal{C} with respect to m gives

$$\frac{\partial \ln W_{d_i}}{\partial m_i} = \ln p - \ln q - \ln m_i^* + \ln (N_i - m_i^*) = 0. \quad (14)$$

This result may be simplified to yield

$$\ln \left(\frac{m_i^*}{N_i - m_i^*} \right) = \ln \left(\frac{p}{1 - p} \right), \quad (15)$$

which implies that the most probable jump number is simply given by

$$m_i^* = p N_i. \quad (16)$$

Because our probability distribution W_d is nearly symmetric about the most probable value of flux, the average number and the most probable number are approximately the same. Hence, the average net flux to the right is

$$\langle J(t) \rangle = \frac{m_1^* - m_2^*}{\Delta t} = p \left[\frac{N_1(t) - N_2(t)}{\Delta t} \right] \approx -\frac{p\Delta x^2}{\Delta t} \frac{\Delta c(x,t)}{\Delta x}, \quad (17)$$

which is Fick's law for this simple model with the diffusion coefficient given by $D=p\Delta x^2/\Delta t$. We have rewritten $N_1-N_2=-\Delta c\Delta x$.

This approach gives us a simple explanation for why there is a net flux of particles diffusing across a plane down a concentration gradient: more microscopic trajectories lead downhill than uphill. It shows that the diffusion coefficient D is a measure of the jump rate p . This model does not assume that the system is near-equilibrium, for example, it does not utilize the Boltzmann distribution law, and thus it indicates that Fick's law should also apply far from equilibrium. We might have imagined that for very steep gradients, Fick's law might have been only an approximation and that diffusion is more accurately represented as a series expansion of higher derivatives of the gradient. But at least for the present model, Fick's law is a general result that emerges from counting microtrajectories. On the other hand, we would expect Fick's law to break down when the particle density becomes so high that the particles start interacting with each other, thus spoiling the assumption of independent particle jumps.

B. Fluctuations in diffusion

We have shown that the most probable number of fleas that jump from dog 1 to dog 2 between time t and $t+\Delta t$ is $m_1^*=pN_1(t)$. The model also tells us that sometimes we will have fewer fleas jumping during this time interval and sometimes we will have more fleas. These variations are a reflection of the fluctuations resulting from the system following different microscopic pathways.

We focus now on predicting the fluctuations. To illustrate, we first construct a table of W_d , the different numbers of possible microtrajectories, for all the values of m_1 and m_2 . To keep the illustration simple, we consider the special case $N_1(t)=4$ and $N_2(t)=2$. We also assume $p=q=1/2$. Table I lists the multiplicities of all the possible routes of flea flow. A given entry tells us how many microtrajectories correspond to the value of m_1 and m_2 .

Table I confirms our previous discussion. The dynamical process for which W_d is a maximum (12 microtrajectories in this case) occurs when $m_1^*=pN_1=1/2 \times 4=2$, and $m_2^*=pN_2=1/2 \times 2=1$. You can calculate the probability of this flux by dividing $W_d=12$ by the sum of entries in Table I, which is $2^6=64$, the total number of microtrajectories. The result, which is the fraction of all the possible microtrajectories that have $m_1^*=2$ and $m_2^*=1$, is 0.18. We have chosen an example in which the particle numbers are very small, so the fluctuations are large; they account for more than 80% of the flow. In systems with large numbers of particles, the relative fluctuations are much smaller.

Now look at the top right entry of Table I. This entry says that there is a probability of $1/64$ that both fleas on dog 2 will jump to the left while no fleas will jump to the right, implying that the net flux for this microtrajectory is backward relative to the concentration gradient. We call these "bad actor" microtrajectories. In these cases, particles flow to increase the concentration gradient, not decrease it.²¹ At time t , there are four fleas on the left dog, and two on the right. At the next instant, $t+\Delta t$, all six fleas are on the left dog, and no fleas are on the right-hand dog.

Similarly, if you look at the bottom left entry of Table I, you see a case of superflux: a net flux of four particles to the right, whereas Fick's law predicts a net flow of only two

particles to the right. Table I illustrates that Fick's law is only a description of the average or most probable flow and that Fick's law is not always exactly correct at the microscopic level. However, such violations of Fick's law are of low probability, a point that we will make more quantitative in the following. Such fluctuations have been experimentally measured in small systems.²³

We can elaborate on the nature of the fluctuations by defining the "potencies" of the microtrajectories. We define the potency to be the fraction of all the trajectories that lead to a substantial change in the macrostate. The potencies of trajectories depend on how far the system is from equilibrium. To see this, we continue our consideration of the simple system of six particles. The total number of microscopic trajectories available to this system at each instant in our discrete time picture is $2^6=64$. Suppose that at $t=0$ all six of these particles are on dog 1. The total number of microscopic trajectories available to the system can be classified using m_1 and m_2 , where in this case $m_2=0$ because there are no fleas on dog 2 (see Table II).

What fraction of all microtrajectories changes the occupancies of both dogs by more than some threshold value, say $\Delta N_i > 1$? In this case, we find that 57 of the 64 microtrajectories cause a change greater than this value to the current state. We call these *potent* trajectories.

Now let us look at the potencies of the same system of six particles in a different situation, $N_1=N_2=3$ when the system is in macroscopic equilibrium (see Table III). In this case only the trajectories with (m_1, m_2) pairs given by (0,2), (0,3), (1,3), (2,0), (3,0), and (3,1) satisfy our criterion. Summing over all of these outcomes shows that just 14 of the 64 trajectories are potent in this case.

There are two key observations conveyed by these arguments. For a system far from equilibrium the vast majority of trajectories at that time are potent and move the system significantly away from its current macrostate. Also when the system is near equilibrium, the vast majority of microtrajectories leave the macrostate unchanged. We now generalize from the tables to see when fluctuations will be important.

Fluctuations and potencies—A simple way to characterize the magnitude of the fluctuations is to look at the width of the W_d distribution.² It is shown in standard texts that for a binomial distribution for which the mean and most probable value both equal $m_i^*=N_i p_i$, the variance is $\sigma_i^2=N_i p_i q$. The variance characterizes the width. Moreover, if N_i is sufficiently large, a binomial distribution can be well-approximated by a Gaussian distribution

$$\mathcal{P}(m_i, N_i) = \frac{1}{\sqrt{2\pi N_i p_i q}} \exp\left(-\frac{(m_i - N_i p_i)^2}{2N_i p_i q}\right), \quad (18)$$

a convenient approximation because it leads to simple analytic results. However, this distribution function is not quite the one we want. We are interested in the distribution of flux, $P(J) = P(m_1 - m_2)$, not the distribution of right-jumps m_1 or left-jumps m_2 alone.

Due to a remarkable property of the Gaussian distribution, it is simple to compute the quantity we want. If you have two Gaussian distributions, one with mean $\langle x_1 \rangle$ and variance σ_1^2 and the other with mean $\langle x_2 \rangle$ and variance σ_2^2 , then the distribution function, $P(x_1 - x_2)$ for the difference will also be a Gaussian distribution with mean $\langle x_1 \rangle - \langle x_2 \rangle$ and variance $\sigma^2 = \sigma_1^2 + \sigma_2^2$.

For our binomial distributions, the means are $m_1^* = pN_1$ and $m_2^* = pN_2$ and the variances are $\sigma_1^2 = N_1 pq$ and $\sigma_2^2 = N_2 pq$. Hence the distribution of the net flux, $J = m_1 - m_2$ is

$$P(J) = \frac{1}{\sqrt{2\pi(pqN)}} \exp\left(-\frac{(J - p(N_1 - N_2))^2}{2pqN}\right), \quad (19)$$

where $N = N_1 + N_2$.

Figure 3 shows an example of the distributions of fluxes at different times using $p=0.1$ and starting from $N_1=100$, $N_2=0$. We update each time step using an averaging scheme, $N_1(t + \Delta t) = N_1(t) - N_1(t)p + N_2(t)p$. Figure 3 shows how the mean flux is large at first and decays toward equilibrium, $J=0$. This result could also have been predicted from the diffusion equation. Equally interesting are the wings of the distributions, which show the deviations from the average flux; these deviations are not predictable from the diffusion equation.

One measure of the importance of the fluctuations is the ratio of the standard deviation σ to the mean, $\sqrt{\sigma^2}/J$. In the limit of large N_1 , $\sqrt{\sigma^2}/J$ reduces to

$$\sqrt{\frac{\sigma^2}{J}} = \sqrt{\frac{Npq}{(N_1 - N_2)p}} N^{-1/2}. \quad (20)$$

In a typical bulk experiment, the particle numbers are large, of the order of Avogadro's number 10^{23} . In such cases, the width of the flux distribution is exceedingly small, and it becomes overwhelmingly probable that the mean flux will be governed by Fick's law. However, within biological cells and in applications involving small numbers of particles, the variance of the flux can become significant. It has been observed that both rotary and translational single motor proteins sometimes transiently step backward relative to their main direction of motion.²⁴

As a measure of the fluctuations, we now calculate the variance in the flux. It follows from Eq. (19) that $\langle J^2 \rangle = Npq$, where $N = N_1 + N_2$. Thus, we can represent the magnitude of the fluctuations as

$$\delta = \sqrt{\frac{\langle (\Delta J)^2 \rangle}{\langle J \rangle^2}} = \frac{\sqrt{Npq}}{pfN} \propto \frac{1}{f} \sqrt{\frac{q}{p}} N^{-1}, \quad (21)$$

where $N = N_1 + N_2$ is the total number of fleas and $f = (N_1 - N_2)/N$ is the normalized concentration difference. The quantity δ is also a measure of the degree of backflux. In the limit of large N , δ goes to zero. That is, the noise diminishes with system size. However, even when N is large, δ can still be large (indicating the possibility of backflux) if the concentration gradient, $N_1 - N_2$, is small.

Another measure of fluctuations is the potency. Trajectories that are not potent should have $|m_1 - m_2| \approx 0$, which corresponds to a negligible change in the current state of the system as a result of a given microtrajectory. In Fig. 4 the impotent microtrajectories are shown as the shaded band for which $m_1 \approx m_2$. We define impotent trajectories as those for which $|m_1 - m_2| < h$ ($h \ll N$). In the Gaussian model, the fraction of impotent trajectories is

$$\Phi_{\text{impotent}} \approx \int_{-h}^h dJ \frac{1}{\sqrt{2\pi Npq}} \exp\left(-\frac{(J - (N_1 - N_2)p)^2}{2Npq}\right) \quad (22)$$

$$= \frac{1}{2} \left(\operatorname{erf} \left[\frac{h + (N_1 - N_2) p}{\sqrt{2Npq}} \right] + \operatorname{erf} \left[\frac{h - (N_1 - N_2) p}{\sqrt{2Npq}} \right] \right) \quad (23)$$

and corresponds to summing over the subset of trajectories that have a small flux. To keep it simple, we take $p=q=1/2$ for which the probability distribution for the microscopic flux m_1 - m_2 is given by Eq. (19). The choice of h is arbitrary and we choose h to be one standard deviation, $\sqrt{N}/4$. Figure 5 shows the potencies for various values of N_1 and N_2 . When the concentration gradient is large, most trajectories are potent, leading to a statistically significant change of the macrostate, whereas when the concentration gradient is small, most trajectories have little effect on the macrostate.

As another measure of the fluctuations, let us now consider the “bad actors” (see Fig. 6). If the average flux is in the direction from dog 1 to dog 2, what is the probability we will observe flux in the opposite direction (bad actors)? We use Eq. (19) for $P(J)$ to obtain

$$\Phi_{\text{bad actors}} \approx \int_{-\infty}^0 \frac{1}{\sqrt{2\pi Npq}} \exp \left(\frac{-(J - (N_1 - N_2) p)^2}{2Npq} \right) \quad (24)$$

$$= \frac{1}{2} \left(1 - \operatorname{erf} \left[\frac{(N_1 - N_2) p}{\sqrt{2Npq}} \right] \right) \quad (N_2 > N_1), \quad (25)$$

which amounts to summing up the fraction of trajectories for which $J \leq 0$. Figure 7 shows the fraction of bad actors for $p=q=1/2$. Bad actors are rare when the concentration gradient is large, and highest when the gradient is small. The discontinuity in the slope of the curve in Fig. 7 at $N_1/N_2=1/2$ is a reflection of the fact that the mean flux abruptly changes sign at that value.

IV. FOURIER'S LAW OF HEAT FLOW

Although particle flow is driven by concentration gradients according to Fick's law, $\langle J \rangle = -D \, c' / x$, energy flow is driven by temperature gradients according to Fourier's law:¹

$$\langle J_q \rangle = -\kappa \frac{\partial T}{\partial x}. \quad (26)$$

Here, J_q is the energy transferred per unit time and per unit cross-sectional area and $\partial T / \partial x$ is the temperature gradient that drives it. The thermal conductivity¹ κ plays the role that the diffusion coefficient plays in Fick's law.

To explore a simplified version of Fourier's law that depends only on particle transport, we return to the dog-flea model as described in Sec. III. Now columns 1 and 2 can differ not only in their particle numbers, $N_1(t)$ and $N_2(t)$, but also in their temperatures, $T_1(t)$ and $T_2(t)$. For simplicity, we assume that each column is at thermal equilibrium and that each particle that jumps carries with it the average energy $\langle m v^2 / 2 \rangle = k_B T / 2$ from the column it left. We again take the cross-sectional area to be unity. In this simple model all energy is transported by hot or cold molecules switching dogs. The average heat flow at time t is

$$\langle J_q \rangle = \frac{m_1^*}{\Delta t} (k_B T_1 / 2) - \frac{m_2^*}{\Delta t} (k_B T_2 / 2) = \frac{p k_B}{\Delta t} [N_1 T_1 - N_2 T_2], \quad (27)$$

where m_1^* and m_2^* are as defined in Sec. III A the numbers of particles jumping from each column at time t . If the particle numbers are identical, $N_1=N_2=N/2$, then

$$\langle J_q \rangle = \frac{pk_B N}{\Delta t} (T_1 - T_2) = -\kappa \frac{\Delta T}{\Delta x}, \quad (28)$$

which is Fourier's law for the average heat flux for this two-column model. The model predicts that the thermal conductivity is $\kappa=(pk_B N \Delta x)/(\Delta t)$, which can be expressed in a more canonical form as $\kappa=pk_B n v_{av} \Delta x$ in terms of the particle density $n=N/\Delta x$ and the average velocity, $v_{av}=\Delta x/\Delta t$. Our model gives a value for the thermal conductivity similar to that found in the kinetic theory of gases,¹ $\kappa=(1/3)k_B n v_{av} l$, if Δx in our model corresponds to l , the mean free path.

The factors of p and $1/3$ in our model and kinetic theory, respectively, can be reconciled by the following observation. Kinetic theory deals with the motion of particles in three dimensions so that each particle can move in six possible directions ($\pm x, \pm y, \pm z$). As a result, only $1/6$ of the particles will contribute to the heat flux in our direction of interest. Also, on the plane of interest there are particles coming in from the positive and negative directions. So the factor of $1/6$ is increased by two to $1/3$. In contrast, in our model, the particles move with probability p or stay with probability $1-p$. So the contribution to the flux comes from a fraction p of the particles. The origin of the numerical factors is thus clear.

The numerical factors of $1/3$ or p are not really important to the point we are trying to make. It is also not fair to read too much into this simple but illustrative model. The key point is that the simple model captures the main physical features of heat flow by appealing to the idea of summing over the weighted microtrajectories available to the system.

V. NEWTONIAN VISCOSITY

Another phenomenological law of gradient-driven transport is that of Newtonian viscosity,¹

$$\tau = \eta \frac{dv_y}{dx}, \quad (29)$$

where τ is the shear stress that is applied to a fluid, dv_y/dx is the resultant shear rate, and the proportionality coefficient η is the viscosity of a Newtonian fluid. Whereas Fick's law describes particle transport and Fourier's law describes energy transport, Eq. (29) describes the transport (in the x direction, from the top moving plate toward the bottom fixed plate) of linear momentum in the y direction (parallel to the plates)(see Fig. 8). In the same spirit as our simplified treatment of Fourier's law, the dog-flea model can be used as the basis of a particle transfer version of momentum transport. Suppose each particle in column 1 of Table I carries momentum $m v_{y1}$ along the y axis, and m_1^* particles hop from column 1 to 2 at time t carrying with them some linear momentum. As before, we consider the simplest model for which every particle carries the same average momentum from the column it leaves to its destination column.

The flux J_p is the amount of y -axis momentum that is transported from one plane to the next in the x direction per unit area:

$$\langle J_p \rangle = \frac{m_1^*}{\Delta t} (m v_{y1}) - \frac{m_2^*}{\Delta t} (m v_{y2}) = \frac{pm}{\Delta t} [N_1 v_{y1} - N_2 v_{y2}]. \quad (30)$$

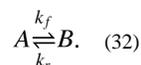
If the number of particles is the same in both columns, $N/2=N_1=N_2$, Eq. (30) simplifies to

$$\langle J_p \rangle = \frac{pmN}{\Delta t} [v_{y1} - v_{y2}] = \eta \frac{\Delta v_y}{\Delta x}, \quad (31)$$

which is the Newtonian law of viscosity for this two-column model. The viscosity is predicted by this model to be $\eta = (pmN\Delta x)/(\Delta t)$. If we convert this result to the more canonical form, we have $\eta = pmnv_{av}\Delta x$, where $n = N/\Delta x$ is the particle density, and $v_{av} = \Delta x/\Delta t$ is the average velocity. This form is equivalent to the value given by the kinetic theory of gases,¹ $\eta = (1/3)mnv_{av}$ from our model equals l . The numerical factors of $1/3$ and p in the kinetic theory and our model, respectively, have the same origin as discussed in the context of heat flow. Note that this simple model based on molecular motions will clearly not be applicable to complex fluids where the underlying molecules possess internal structure.

VI. CHEMICAL KINETICS WITHIN THE DOG-FLEA MODEL

Let us now look at chemical reactions using the dog-flea model. Chemical kinetics can be modeled using the dog-flea model when the fleas have preference for one dog over the other. Consider the reaction



The time-dependent average concentrations, $[A](t)$ and $[B](t)$ are often described by chemical rate equations,²

$$\frac{d[A]}{dt} = -k_f[A] + k_r[B], \quad (33a)$$

$$\frac{d[B]}{dt} = k_f[A] - k_r[B], \quad (33b)$$

where k_f is the average conversion rate of an A to a B , and k_r is the average conversion rate of a B to an A . These rate expressions describe only average rates; they do not give the distribution of rates. Some A 's will convert to B 's faster than the average rate $k_f[A]$ predicts, and some will convert more slowly. Again we use the dog-flea model and consider the average concentrations and the fluctuations in concentrations. A particularly fruitful area for applications of the $A \rightleftharpoons B$ dynamics considered here is to problems involving molecular motors and ion channels.

We assume that dog 1 represents chemical species A and dog 2 represents chemical species B . The net chemical flux from 1 to 2 is given by $J_c = m_1^* - m_2^*$. What is different about our model for these chemical processes than in our previous situations is that now the intrinsic jump rate from column 1 (species A), p_1 , is different than the jump rate from column 2, p_2 . This difference reflects the fact that a forward rate can differ from a backward rate in a chemical reaction. We assume that the fleas have a different escape rate from each dog. Fleas escape from dog 1 at rate p_1 and fleas escape from dog 2 at rate p_2 . Maximizing W_d gives $m_1^* = N_1 p_1$ and $m_2^* = N_2 p_2$, so the average flux (which is almost the same as the most probable flux because of the approximately symmetric nature of the binomial distribution) at time t is

$$\langle J \rangle = N_1 p_1 - N_2 p_2 = k_f[A] - k_r[B], \quad (34)$$

which is just the standard mass-action rate law, expressed in terms of the mean concentrations. The mean values satisfy detailed balance at equilibrium ($\langle J \rangle = 0$) implies that $N_2/N_1 = p_1/p_2 = k_f/k_r$

More interesting than the behavior of the mean chemical reaction rate is the behavior of the fluctuations. For example, if the number of particles is small, then even when $k_f[A] - k_r[B] > 0$, indicating an average conversion of A 's to B 's, the reverse can happen occasionally. When will these fluctuations be large? As in Sec. III, we first determine the probability distribution of the flux J . In this case, the probability distribution becomes

$$p(J) = \frac{1}{\sqrt{2\pi(p_1q_1N_1 + p_2q_2N_2)}} \times \exp\left(-\frac{(J - (N_1p_1 - N_2p_2))^2}{2(p_1q_1N_1 + p_2q_2N_2)}\right). \quad (35)$$

We use this flux distribution function to consider the fluctuations in the chemical reaction. The relative variance in the flux is

$$\frac{\langle (\Delta J)^2 \rangle}{\langle J \rangle^2} = \frac{\sqrt{N_1p_1q_1 + N_2p_2q_2}}{N_1p_1 - N_2p_2}. \quad (36)$$

As before, the main message is that when the system is not yet at equilibrium (that is, the denominator is nonzero), macroscopically large systems will have negligibly small fluctuations. The relative magnitude of the fluctuations scales approximately as $N^{-1/2}$. Let us also look at the potencies of microtrajectories as another window into the fluctuations. If we use Eq. (23) with p_1 and p_2 , we find that the fraction of trajectories that are impotent is

$$\Phi_{\text{impotent}} \approx \int_{-h}^h dJ \frac{1}{\sqrt{2\pi(N_1p_1q_1 + N_2p_2q_2)}} \times \exp\left(\frac{-(J - (N_1p_1 - N_2p_2))^2}{2(N_1p_1q_1 + N_2p_2q_2)}\right) \quad (37)$$

$$= \frac{1}{2} \left(\text{erf} \left[\frac{h + (N_1p_1 - N_2p_2)}{\sqrt{2(N_1p_1q_1 + N_2p_2q_2)}} \right] + \text{erf} \left[\frac{h - (N_1p_1 - N_2p_2)}{\sqrt{2(N_1p_1q_1 + N_2p_2q_2)}} \right] \right). \quad (38)$$

For $N_1 + N_2 = N = 100$ and $p_1 = 0.1$ and $p_2 = 0.2$, $\Phi_{\text{potent}} = 1 - \Phi_{\text{impotent}}$ is shown in Fig. 9 as a function of N_1/N .

VII. DERIVATION OF THE DYNAMICAL DISTRIBUTION FUNCTION FROM MAXIMUM CALIBER

We have used the binomial distribution function W_d as the basis for our treatment of stochastic dynamics. The maximum caliber assumption is that if we find the value of W_d that is a maximum with respect to the microscopic trajectories, this value will give the macroscopically observable flux. We now restate this assumption more generally in terms of the probabilities of the trajectories.

Let $P(i)$ be the probability of a microtrajectory i during the interval from time t to $t + \Delta t$. A microtrajectory is a specific set of fleas that jump; for example microtrajectory $i=27$ might be the case for which fleas 4, 8, and 23 jump from dog 1 to 2. We take as a constraint the average number of jumps, $\langle m \rangle$, the macroscopic observable. The quantity $m_F=3$ in this case indicates that trajectory i involves three fleas jumping. We express the caliber \mathcal{C} as

$$\mathcal{C} = \sum_i P(i) \ln P(i) - \lambda \sum_i m_i P(i) - \alpha \sum_i P(i), \quad (39)$$

where λ is the Lagrange multiplier that enforces the constraint of the average flux and α is the Lagrange multiplier that enforces the normalization condition that the $P(i)$'s sum to one. Maximizing the caliber gives the populations of the microtrajectories,

$$P(i) = e^{-\alpha - \lambda m_i}. \quad (40)$$

Note that the probability $P(i)$ of the i th trajectory depends only on the total number m_i of the jumping fleas. Also, all trajectories with the same m_i have the same probabilities. In the same way that it is sometimes useful in equilibrium statistical mechanics to switch from microstates to energy levels, we now express the population $P(i)$ of a given microtrajectory in terms of $\rho(m)$, the fraction of all the microtrajectories that involve m jumps during this time interval,

$$\rho(m) = g(m) Q(m), \quad (41)$$

where $g(m) = N/[m!(N-m)!]$ is the density of trajectories with flux m (in analogy with the density of states for equilibrium systems), and $Q(m)$ is the probability $P(i)$ of microtrajectory i with $m_i = m$. In other words, i denotes a microtrajectory (a specific set of fleas jumping) and m denotes a microprocess (the number of fleas jumping). The total number of i 's associated with a given m is $g(m)$. It can also be easily seen that

$$\sum_i P(i) = \sum_{m=0}^N g(m) Q(m) = \sum_{m=0}^N \rho(m) = 1, \quad (42)$$

$$\langle m \rangle = \sum_i m_i P(i) = \sum_{m=0}^N m g(m) Q(m) = \sum_{m=0}^N m \rho(m). \quad (43)$$

Thus the distribution of jump-processes written in terms of the jump number m is

$$\rho(m) = \frac{N!}{m!(N-m)!} e^{-\alpha} e^{-\lambda m}. \quad (44)$$

The Lagrange multiplier α can be eliminated by summing over all trajectories and requiring that $\sum_{m_i=0}^N \rho(m_i) = 1$, that is,

$$e^{\alpha} = \sum_m^N g(m) e^{-\lambda m} = \sum_m^N \frac{N!}{m!(N-m)!} e^{-\lambda m} = (1 + e^{-\lambda})^N. \quad (45)$$

We combine Eqs. (44) and (45) to obtain

$$\rho(m) = \frac{N!}{m!(N-m)!} \frac{e^{-\lambda}}{1 + e^{-\lambda}}. \quad (46)$$

If we now let

$$p = \frac{e^{-\lambda}}{1 + e^{-\lambda}}, \quad (47)$$

we obtain

$$p^m = \frac{e^{-\lambda m}}{[1 + e^{-\lambda}]^m}, \quad (48)$$

and

$$(1 - p)^{N-m} = \frac{1}{[1 + e^{-\lambda}]^{N-m}}. \quad (49)$$

If we combine Eqs. (46), (48), and (49), we find the simple form

$$\rho(m) = \frac{N!}{m!(N-m)!} p^m (1-p)^{N-m}, \quad (50)$$

which appears in Eq. (11) and which we have used throughout this paper.

VIII. SUMMARY AND COMMENTS

We have shown how to derive the phenomenological laws of nonequilibrium transport, including Fick's law of diffusion, Fourier's law of heat conduction, the Newtonian law of viscosity, and the mass-action laws of chemical kinetics from a simple physical foundation that can be taught in undergraduate courses. We used the dog-flea model for describing how particles, energy, or momentum can be transported across a plane. We combined this model with the principle of maximum caliber, a dynamical analog of the principle of maximum entropy for the laws of equilibrium. For dynamics we focus on microtrajectories rather than microstates and maximize a dynamical entropy-like quantity, subject to an average flux constraint. In this way maximizing the caliber is the dynamical equivalent of minimizing a free energy for predicting equilibria. A particular value of this approach is that it also gives us fluctuation information, not just averages. In diffusion, for example, sometimes the flux can be a little higher or lower than the average value expected from Fick's law. These fluctuations can be important for biology and nanotechnology, where the numbers of particles can be very small, and therefore where there can be significant fluctuations in rates, around the average.

Acknowledgments

It is a pleasure to acknowledge the helpful comments and discussions with Dave Drabold, Mike Geller, Jané Kondev, Stefan Müller, Hong Qian, Darren Segall, Pierre Sens, Jim Sethna, Ron Siegel, Andrew Spakowitz, Zhen-Gang Wang, and Paul Wiggins. We would also like to thank Sarina Bromberg for help with the figures. K.A.D. and M.M.I. would like to acknowledge support from NIH Grant No. R01 GM034993. R.P. acknowledges support from NSF Grant No. CMS-0301657, CIMMS, the Keck Foundation, NSF NIRT Grant No. CMS-0404031, and NIH Director's Pioneer Award Grant No. DP1 OD000217.

References

1. Reif, F. *Fundamentals of Statistical and Thermal Physics*. McGraw-Hill; New York: 1965.
2. Dill, K.; Bromberg, S. *Molecular Driving Forces: Statistical Thermodynamics in Chemistry and Biology*. Garland Science; New York: 2003.

3. Kasianowicz J, Brandin E, Branton D, Deamer D. Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl. Acad. Sci. U.S.A.* 1996; 93:13770–13773. [PubMed: 8943010]
4. Lu HP, Xun L, Xie XS. Single molecule enzymatic dynamics. *Science*. 1998; 282:1877–1882. [PubMed: 9836635]
5. Reif M, Rock RS, Mehta AD, Mooseker MS, Cheney RE, Spudich JA. Myosin-V stepping kinetics: A molecular model for processivity. *Proc. Natl. Acad. Sci. U.S.A.* 2000; 97:9482–9486. [PubMed: 10944217]
6. Meller A, Nivon L, Branton D. Voltage-driven DNA translocations through a nano-pore. *Phys. Rev. Lett.* 2001; 86:3435–3438. [PubMed: 11327989]
7. Smith DE, Tans SJ, Smith SB, Grimes S, Anderson DL, Bustamante C. The bacteriophage [H9278]29 portal motor can package DNA against a large internal force. *Nature (London)*. 2001; 413:748–752. [PubMed: 11607035]
8. Li H, Linke WA, Oberhauser AF, Carrion-Vazquez M, Kerkvliet JG, Lu H, Marszalek P, Fernandez JM. Reverse engineering of the giant muscle protein. *Nature (London)*. 2002; 418:998–1002. [PubMed: 12198551]
9. Liphardt J, Dumont S, Smith SB, Tinocho I, Bustamante C. Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality. *Science*. 2002; 296:1832–1835. [PubMed: 12052949]
10. Bustamante C, Bryant Z, Smith SB. Ten Years of tension: single-molecule DNA mechanics. *Nature (London)*. 2003; 421:423–427. [PubMed: 12540915]
11. Dufresne ER, Altman D, Grier DG. Brownian dynamics of sphere in a slit pore. *Europhys. Lett.* 2001; 53:264–270.
12. Alberts, B.; Bray, D.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. *Essential Cell Biology: An Introduction to the Molecular Biology of the Cell*. Garland; New York: 1997.
13. Kondepudi, D.; Prigogine, I. *Modern Thermodynamics: From Heat Engines to Dissipative Structures*. Wiley; New York: 1998.
14. Jaynes, ET.; Jaynes, ET.; Rosenkrantz, RD. *Papers on Probability, Statistics and Statistical Physics*. Kluwer Academic; 1980. Chap. 14
15. Klein M. Entropy and the Ehrenfest urn model. *Physica (Amsterdam)*. 1956; 22:569–575.
16. Ambegaokar V, Clerk A. Entropy and time. *Am. J. Phys.* 1999; 67:1068–1073.
17. Emch, GG.; Liu, C. *The Logic of Thermostatistical Physics*. Springer-Verlag; New York: 2002.
18. Feynman, RP. *Nobel Lectures in Physics: 1901-1995*. World Scientific; Singapore: 1998.
19. Reference 14, Chap. 1.
20. Kac M. Random walk and the theory of Brownian motion. *Am. Math. Monthly*. 1947; 54:369–391.
21. The Maxwell demon is an imaginary microscopic object that was invoked in similar situations in heat flow processes, that is, where heat flows from a colder object to a hotter one, albeit with low probability (Ref. 22). In particular, the demon is supposed to capture the bad actor microtrajectories.
22. Gamow, G. *One Two Three ... Infinity*. Dover; New York: 1988.
23. Wang GM, Sevcik EM, Mittag E, Searles DJ, Evans DJ. Experimental verification of the fluctuation theorem. *Phys. Rev. Lett.* 2002; 89:050601–1. [PubMed: 12144431]
24. Howard, J. *Mechanics of Motor Proteins and the Cytoskeleton*. Sinauer Associates; 2001.

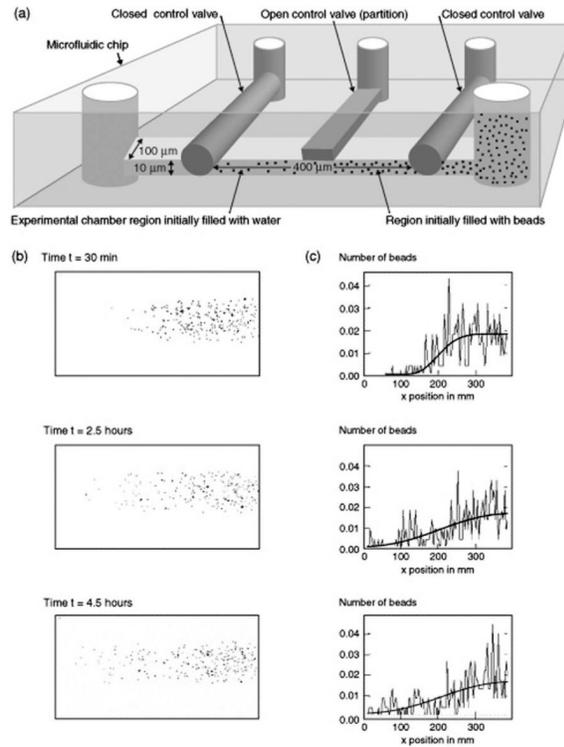


Fig. 1. Colloidal free expansion setup to illustrate diffusion involving small numbers of particles. (a) Schematic of experimental setup. (b) Several snapshots from the experiment. (c) Normalized histogram of particle positions during the experiment. The solution to the diffusion equation for the microfluidic “free expansion” experiment is superposed for comparison.

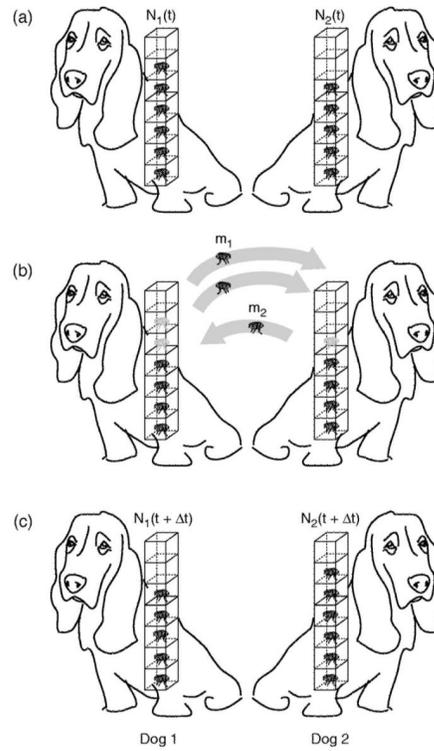


Fig. 2. Schematic of the simple dog-flea model. (a) State of the system at time t ; (b) a particular microtrajectory in which two fleas jump from the dog on the left and one flea jumps from the dog on the right; (c) occupancies of the dogs at time $t + \Delta t$.

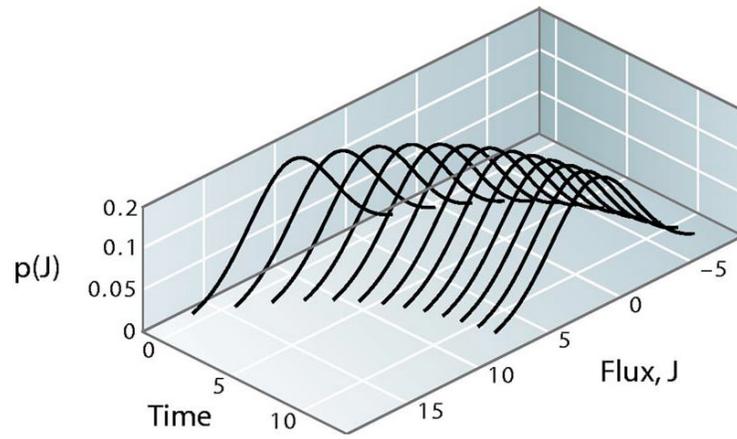


Fig. 3. Schematic of the distribution of fluxes for different times as the system approaches equilibrium.

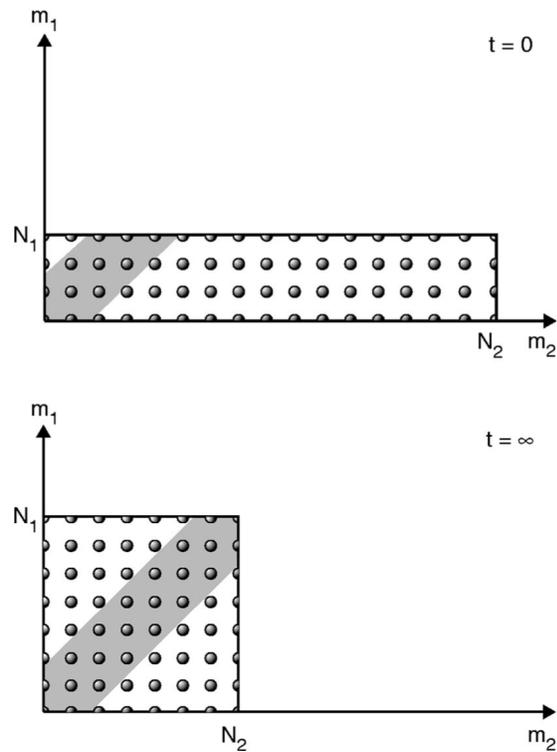


Fig. 4. Schematic of which trajectories are potent and which are impotent. The shaded region corresponds to the impotent trajectories for which m_1 and m_2 are either equal or approximately equal and hence make relatively small change in the macrostate. The unshaded region corresponds to potent trajectories.

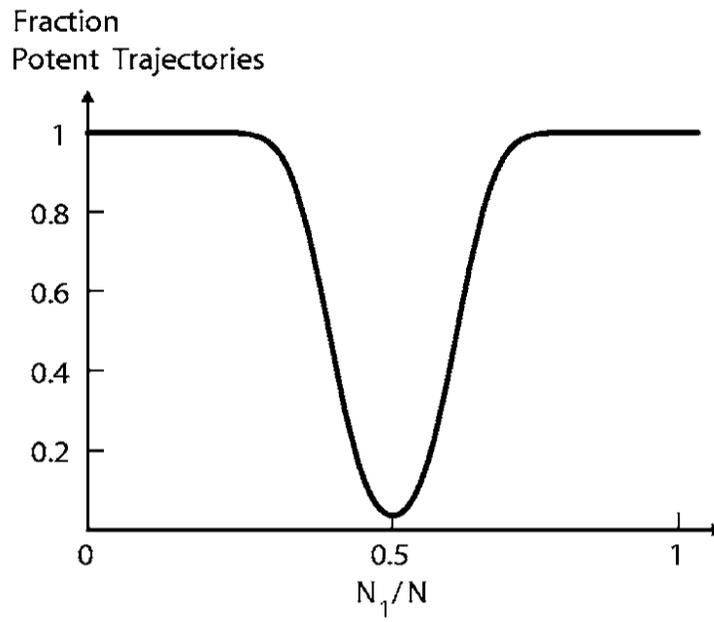


Fig. 5. Illustration of the potency of the microtrajectories associated with different distributions of N particles on the two dogs. The total number of particles $N_1 + N_2 = N = 100$.

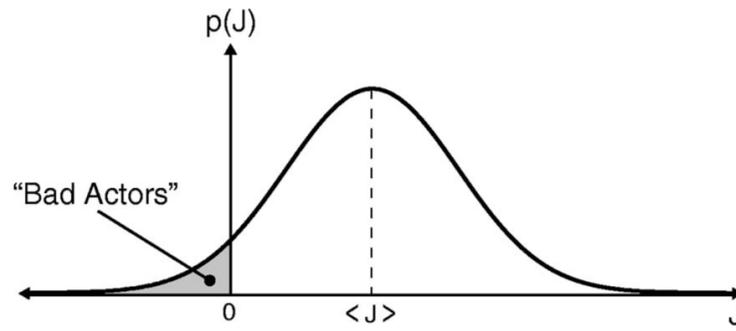


Fig. 6. Illustration of the notion of bad actors. Bad actors are the microtrajectories that contribute net particle motion that has the opposite sign from the macroflux.

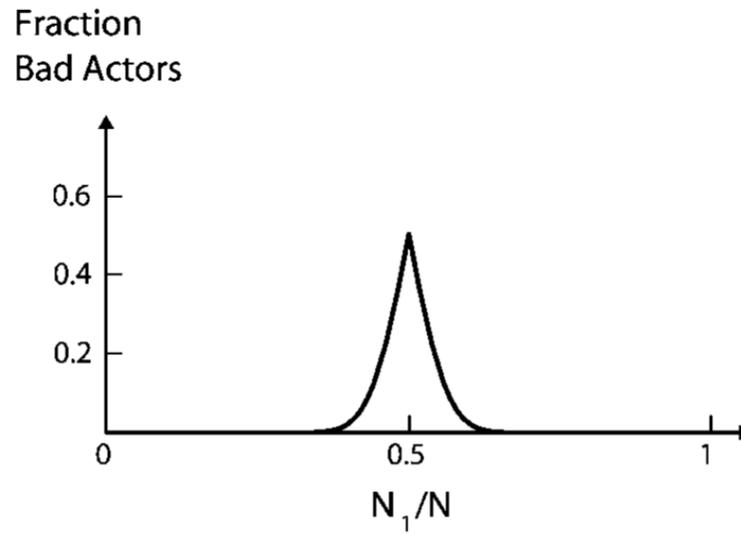


Fig. 7. The fraction of all possible trajectories that go against the direction of the macroflux for $N=100$. The fraction of bad actors is highest at $N_1=N/2=50$.

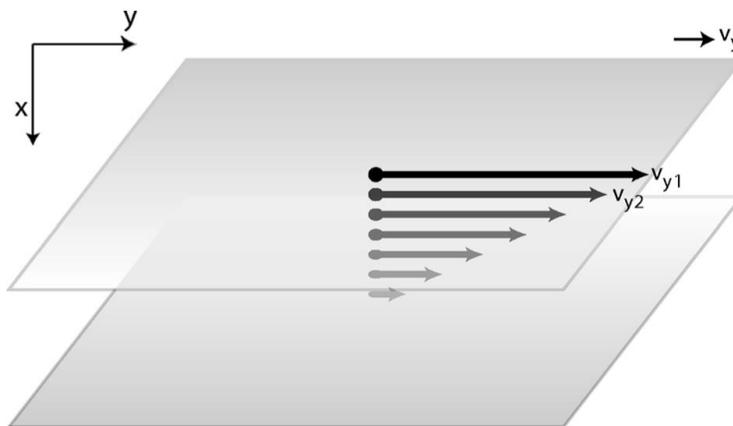


Fig. 8. Illustration of Newton's law of viscosity. The fluid is sheared with a constant stress. The fluid velocity decreases continuously from its maximum value at the top of the fluid to zero at the bottom. There is thus a gradient in the velocity which can be related to the shear stress in the fluid.

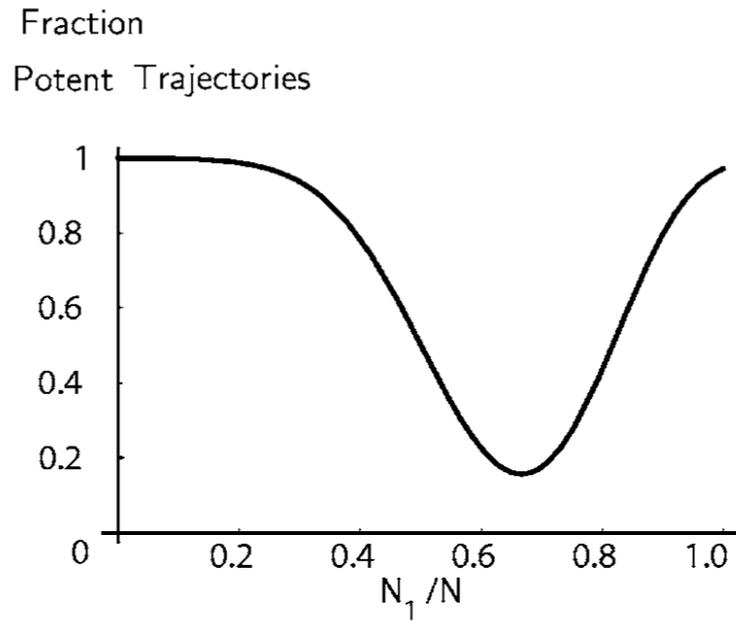


Fig. 9. The fraction of potent trajectories Φ_{potent} as a function of N_1/N for $N_1+N_2=N=100$, and $p_1=0.1$ and $p_2=0.2$. The minimum value of the potency does not occur at $N_1/N=0.5$, but at $N_1/N=0.66$. This value of N_1/N also corresponds to its equilibrium value given by $p_2/(p_1+p_2)$.

Table I

Trajectory multiplicity for the case where $N_1(t)=4$ and $N_2(t)=2$. Each entry corresponds to the total number of trajectories for the particular values of m_1 and m_2 .

m_1	$m_2=0$	$m_2=1$	$m_2=2$
0	1	2	1
1	4	8	4
2	6	12	6
3	4	8	4
4	1	2	1

Table II

Trajectory multiplicity for $N_1(t)=6$ and $N_2(t)=0$ when the system is far from macroscopic equilibrium. In this case $m_2=0$.

m_1	$m_2=0$
0	1
1	6
2	15
3	20
4	15
5	6
6	1

Table III

Trajectory multiplicity for $N_1(t)=3$ and $N_2(t)=3$ when the system is at macroscopic equilibrium.

m_1	$m_2=0$	$m_2=1$	$m_2=2$	$m_2=3$
0	1	3	3	1
1	3	9	9	3
2	3	9	9	3
3	1	3	3	1