

LA-UR-02-2549

Approved for public release;
distribution is unlimited.

Title:

**What Is the Microscopic Response of a System
Driven Far From Equilibrium?**

Author(s):

C. Jarzynski

Submitted to:

<http://lib-www.lanl.gov/cgi-bin/getfile?00937182.pdf>

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the University of California for the U.S. Department of Energy under contract W-7405-ENG-36. By acceptance of this article, the publisher recognizes that the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

What Is the Microscopic Response of a System Driven Far From Equilibrium?

C. Jarzynski

Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM

chrisj@lanl.gov

Abstract

The central theme of this lecture is that there exists a surprisingly simple and general answer to the question posed in the title, provided that we frame the question statistically. I will present this result along with several derivations, and will discuss some of its implications and generalizations.

INTRODUCTION

It is widely acknowledged that we understand systems in thermal equilibrium far better than those which are out of equilibrium. The statistical foundations laid over a century ago by (above all) Boltzmann and Gibbs, have proven remarkably effective at explaining phenomena as diverse as phase transitions, superconductivity, and blackbody radiation. By contrast, no such universal and elegant framework exists that would simultaneously encompass the mechanics of protein folding, the onset of turbulence, and Fourier's law of heat conduction. In part, the difficulty arises simply because "nonequilibrium phenomena" covers a lot of territory. If the term is understood to include any situation involving a system not in a state of thermal equilibrium, then it is hard to imagine a useful point of reference upon which to build a general theory. By contrast, the very notion of an equilibrium state – toward which a system naturally evolves if left undisturbed – provides the organizing concept at the heart of equilibrium statistical mechanics.

While the field of nonequilibrium physics as a whole might indeed be so broad in scope as to escape a sweeping microscopic foundation (akin to the Boltzmann-Gibbs formalism in the equilibrium case), much progress has been made by considering sub-classes of nonequilibrium phenomena distinguished by a common feature. For instance, the study of systems *near equilibrium* has given us the linear response formalism, whose early successes include the fluctuation-dissipation theorem^{1,2} and Onsager's reciprocity relations³. It also seems reasonable to expect that *nonequilibrium stationary states* will eventually succumb to an elegant statistical formulation with predictive power.

In this lecture I will focus on *systems driven away from an initial state of thermal equilibrium*, where this is accomplished by the variation of an external parameter. An example would be a container filled with gas (the system), closed off at one end by a piston (the parameter), initially prepared in equilibrium. If we rapidly push the piston into the container by some distance, the gas is driven out of equilibrium. At the microscopic level, the

response of the system during this process is described by a single trajectory evolving in the many-body phase space of the gas. This response, however, is typically complicated and non-universal, e.g. shock waves form, propagate, scatter, etc. Therefore we adopt a statistical attitude, and imagine an *ensemble* of realizations, pictured as a swarm of independent trajectories evolving in the many-body phase space. Each trajectory represents one *realization* (possible microscopic outcome) of the piston experiment. The central assertion of this lecture is that *there exists a statistical representation of this ensemble* (defined by Eq.18 below) *in which the response of the system becomes both simple and universal* (Eq.20). By “universal”, I mean that the validity of this result depends neither on the details of the system, nor – perhaps more surprisingly – on how far it is driven out of equilibrium.

The above-mentioned result is, on its own, somewhat abstract. However, by a single integration it leads to another result, the *nonequilibrium work relation* (Eq.46), which relates the work performed during an irreversible process, to an equilibrium free energy difference. This result is not only potentially useful, but has recently been tested experimentally, and is closely related to the Second Law of Thermodynamics.

This lecture is organized as follows. I will first present a very brief review of relevant facts from macroscopic thermodynamics (Section I). Then I will move to the microscopic level of statistical mechanics, where I will again present a brief review, and will establish notation (Section II). I will then present the central result of this lecture in Section III, and work through three derivations of this result in Section IV. Section V will cover the nonequilibrium work relation, and Section VI will relate this result to the Second Law. Finally, in Section VII I will briefly discuss the generalization of these results to nonequilibrium stationary states.

Classical physics is assumed throughout the lecture.

I. THERMODYNAMICS – BRIEF REVIEW OF RELEVANT FACTS

Most of this talk will take place at the microscopic level of statistical mechanics, where one is interested in the behavior of atoms, molecules, and so forth. It is therefore useful at the outset to briefly focus on the bulk behavior of *macroscopic* systems, and to review a few relevant facts of thermodynamics.

Consider a macroscopic system which depends on an external parameter λ , and which is in thermal contact with a heat bath at temperature T . A simple example is a rubber band held stretched between two stiff toothpicks. The rubber band is the system; the distance between the toothpicks is our parameter λ , which we can manipulate as we like; and the surrounding air plays the role of the heat bath. A starting assumption of equilibrium thermodynamics is that for any set of allowable* values (λ, T) , there exists a unique *equilibrium state* to which the system will relax if λ and T are held fixed. Furthermore, there exist state

*By “allowable”, I mean in this case a range of distances λ such that the band is neither limp nor stretched beyond its breaking point, and temperatures T for which the band is not frozen solid, nor melting, burning, or otherwise losing its structural integrity.

functions such as the internal energy E , entropy S , and (most importantly for our purposes) the *free energy*

$$F = E - ST, \quad (1)$$

which take on unique values at equilibrium states, e.g. $F = F(\lambda, T)$.

Suppose that we now prepare our system by holding the distance between toothpicks fixed at a value λ_A , and allowing the rubber band to relax to the corresponding equilibrium state, A . We then very slowly pull the toothpicks apart, until the distance between them is λ_B . For sufficiently slow pulling the rubber band will progress through a continuous sequence of equilibrium states, from $A = (\lambda_A, T)$ to $B = (\lambda_B, T)$. In carrying out such a *reversible* process, the external work which we perform on the system will be equal to the free energy difference between the initial and final equilibrium states A and B :

$$W = \Delta F \equiv F_B - F_A \quad (\text{reversible}). \quad (2)$$

Now consider an *irreversible* process from A to B : after preparing the system in equilibrium state A we pull the toothpicks apart at a rapid rate, and when the distance reaches λ_B we hold the toothpicks fixed and allow the system to relax to state B . We thus have a process in which the system begins and ends in equilibrium states, but at intermediate times, as we rapidly stretch the rubber band, it is driven out of equilibrium. Physically, this excursion away from equilibrium manifests itself in two ways: (1) the rubber band heats up, and (2) its tension increases relative to the room-temperature value. As a result, more work is required to stretch the rubber band than in the case of slow pulling, i.e.

$$W > \Delta F \quad (\text{irreversible}). \quad (3)$$

This is a good point to insert a comment about nomenclature. Throughout this lecture I will be concerned primarily with processes (as the one describe above) in which a system is driven out of equilibrium *in the presence of a single heat reservoir* prepared at a given temperature. I will refer to these as *irreversible-isothermal* processes. It is important to stress that the word “isothermal” here is *not* meant to imply that the temperature of the system remains fixed (if even well-defined!) during the process, but rather to indicate the presence of only one reservoir. Eq.3 applies universally to irreversible-isothermal processes connecting two equilibrium states.

During a more general thermodynamic process from one equilibrium state to another, the system might be brought into contact with a sequence of reservoirs at different temperatures. In that case the *Clausius inequality* applies:

$$\int_A^B \frac{dQ}{T} < \Delta S. \quad (4)$$

The left side is the integrated heat absorbed by the system over the course of the process, inversely weighted by the temperature of the reservoir from which that heat is absorbed, and the right side is the entropy difference between the initial and final equilibrium states. As with Eqs.2 and 3, the equality holds if the process is reversible, whereas for irreversible processes we get a strict inequality. Eqs.2 and 3 represent a special case (one reservoir) of the the Clausius inequality.

II. STATISTICAL MECHANICS – NOTATION AND REVIEW

Moving now from the macroscopic level of thermodynamics to the microscopic level of statistical mechanics, consider the following experimental set-up described in a recent paper by Liphardt *et al.*⁴. In room-temperature water, one end of a strand of RNA is attached to a small polystyrene bead, the other to a micromechanical cantilever, and a laser trap is used to capture the bead. Using piezo-electric techniques to move the cantilever back and forth, the experimentalists are able to cause the RNA strand to unravel as it is stretched, or coil up as the end-to-end distance is made smaller; moreover, they can measure the work performed on the system as this micro-manipulation is carried out! This experimental arrangement is a molecular-level analogue of the rubber band between the toothpicks, and serves as a nice example to illustrate the theoretical analysis that will make up the main part of this lecture.

Fig.1 depicts a cartoon version of the above-mentioned experiment, where for simplicity we consider a chain of atoms (rather than nucleotides) and no beads. The N atoms, of equal mass m , are linked by covalent bonds, and we imagine that the two ends of this chain are trapped by confining harmonic potentials. These N atoms constitute the system in this set-up – we will refer to this simply as the “polymer” – the distance between the traps represents an external parameter λ , and surrounding water molecules provide a heat bath at a temperature T . We will use the variable Γ to denote a *microstate* of the system, in this case a point in the $6N$ -dimensional phase space specifying the position and momentum of each of the N atoms:

$$\Gamma = (\mathbf{r}_1, \dots, \mathbf{r}_N; \mathbf{p}_1, \dots, \mathbf{p}_N). \quad (5)$$

Now let $H(\Gamma, \lambda)$ be the parameter-dependent Hamiltonian which gives the internal energy of the system as a function of its microstate, for a given value of λ . According to the central assertion of equilibrium statistical mechanics, if we hold λ fixed and allow our polymer to come to thermal equilibrium then the probability distribution for finding the system in a particular microstate Γ is given by the Boltzmann-Gibbs distribution:

$$p_\lambda^{BG}(\Gamma) = \frac{1}{Z_\lambda} e^{-\beta H(\Gamma, \lambda)}, \quad (6)$$

where $\beta \equiv (k_B T)^{-1}$, and the partition function Z is given by

$$Z_\lambda = \int d\Gamma e^{-\beta H(\Gamma, \lambda)}. \quad (7)$$

The Boltzmann-Gibbs distribution can be viewed as the microscopic counterpart of the equilibrium state of macroscopic thermodynamics. For later reference, it is useful to note here that the free energy associated with such an equilibrium state is given by the relation

$$F_\lambda = -\beta^{-1} \ln Z_\lambda, \quad (8)$$

which (under the appropriate microscopic definition of E and S) can easily be shown to be equivalent to Eq.1.

Since the focus of this lecture is the behavior of systems driven away from an initial state of equilibrium, we will generally be interested in the evolution of our system over some

interval of time, with the system assumed to be in equilibrium at the start of that interval. This evolution is described by a trajectory Γ_t in the $6N$ -dimensional phase space, which can be viewed as a “microscopic history” chronicling the time-dependence of every degree of freedom of the system over the specified interval. From such a trajectory we can compute various quantities of interest. One which will play a central role in the analysis to follow is the *work* performed on the system. Suppose that we externally vary the parameter λ , beginning at time 0 (for specificity), and we simultaneously observe the evolution of the system. In acting thus on the system, we perform external work, and the amount of work performed from the initial time up to some later time t , is given by⁵

$$w_t = \int_0^t dt' \dot{\lambda}_{t'} \frac{\partial H}{\partial \lambda}(\Gamma_{t'}, \lambda_{t'}). \quad (9)$$

Here λ_t denotes the (externally imposed) time-dependence of λ , and Γ_t is the observed trajectory. The dot denotes, as usual, a total time derivative. (I will not here try to justify the expression for work given in Eq.9, beyond observing that the rate of change of the internal energy of the system is given by

$$\dot{H} = \dot{\lambda} \frac{\partial H}{\partial \lambda} + \dot{\Gamma} \frac{\partial H}{\partial \Gamma}. \quad (10)$$

If we identify the two terms on the right as the rate of work performed on the system, and heat absorbed by the system, respectively, then this equation for \dot{H} amounts to the First Law of Thermodynamics.)

In an experiment involving a system such as the polymer described above, we are typically neither able to prepare the system in a precisely specified initial microstate Γ , nor interested in predicting the subsequent evolution of every degree of freedom. Therefore we adopt a statistical attitude when analysing such a situation: instead of posing the question, “How exactly will each atom evolve?”, we ask, “What are the possible microscopic histories consistent with the known preparation and manipulation of the system?” We refer to these potential microscopic histories as *realizations*, and the set of them as a *statistical ensemble*.

We can picture this ensemble of realizations as a swarm of points moving about in $6N$ -dimensional phase space. A natural description of such an evolving swarm is the time-dependent density $f(\Gamma, t)$, which can be defined by the usual box-counting prescription, as follows. Imagine $R \gg 1$ trajectories Γ_t^i ($i = 1, \dots, R$) evolving independently in phase space, each representing a single realization of our N -body system. At any time t , a “snapshot” of this ensemble of trajectories would appear as a set of R points scattered throughout some region of phase space. Now imagine a tiny box \mathcal{B} , of volume $d\Gamma$, centered around a microstate Γ . The density $f(\Gamma, t)$, multiplied by the volume $d\Gamma$, gives us the fraction of the ensemble found inside this box at time t :

$$f(\Gamma, t) d\Gamma \approx \frac{1}{R} \sum_{\Gamma_t^i \in \mathcal{B}} 1. \quad (11)$$

The sum on the right is taken over all realizations which happen to be found inside the box \mathcal{B} at time t , and this sum is equal to the total number of such realizations. The approximation in Eq.11 becomes an equality in the double limit, $R \rightarrow \infty$ and $d\Gamma \rightarrow 0$, taken in that order. We then obtain

$$f(\Gamma, t) = \lim_{d\Gamma \rightarrow 0} \lim_{R \rightarrow \infty} \frac{1}{d\Gamma} \frac{1}{R} \sum_{\Gamma_t^i \in \mathcal{B}} 1 = \langle \delta(\Gamma - \Gamma_t) \rangle, \quad (12)$$

where Γ_t denotes the microstate (at time t) of one particular realization in our infinite ensemble, and the angular brackets denote an average over the ensemble.

III. CENTRAL RESULT – A PREDICTION

We now have in place the basic theoretical elements required to address the central issue of this talk, namely, the response of a system driven (arbitrarily far!) away from equilibrium. Consider the following thought experiment. We prepare the system depicted in Fig.1 in thermal equilibrium with a reservoir at temperature T , with the parameter fixed at some initial value. Then, beginning at $t = 0$, we vary the parameter according to some pre-determined “protocol”, λ_t , and we observe the resulting microscopic trajectory Γ_t . For instance, we start with the traps separated by a distance λ_A , then we stretch our polymer by pulling the traps apart at a constant rate $\dot{\lambda}$ while observing the motion of each atom. Here, the protocol for varying the parameter, λ_t , represents “what we do to the system”, whereas the trajectory Γ_t specifies “how the system responds” to this external manipulation.

Since we want to analyze this situation statistically, let us imagine that we repeat the above experiment R times, always starting in equilibrium, and always subsequently varying the external parameter according to the same protocol, λ_t . We will then observe R different trajectories

$$\Gamma_t^1, \Gamma_t^2, \dots, \Gamma_t^R. \quad (13)$$

In the limit of infinitely many realizations ($R \rightarrow \infty$), we can construct the time-dependent density $f(\Gamma, t)$ describing the statistical response of the system (Eqs.11,12). By the assumption of equilibrium, the initial density is

$$f(\Gamma, t = 0) = p_A^{BG}(\Gamma) = \frac{1}{Z_A} e^{-\beta H(\Gamma, A)}. \quad (14)$$

(I will often use the symbol A as compact notation for λ_A , and similarly B for λ_B .) Now, if we happen to vary λ very slowly (quasistatically), then we expect the system to remain in equilibrium. More precisely, the ensemble of realizations continuously responds so as to “keep up” with the slowly changing value of λ :

$$f(\Gamma, t) = \frac{1}{Z_{\lambda_t}} e^{-\beta H(\Gamma, \lambda_t)} \quad (\text{quasistatic}). \quad (15)$$

However, if we vary λ at a finite rate – possibly very rapidly – then we will drive the system out of equilibrium, and for $t > 0$ the density $f(\Gamma, t)$ will *not* be the Boltzmann-Gibbs distribution corresponding to the current value of λ .

Thus, as long as our system remains in equilibrium, we have simple, universal expressions for its statistical state (Eqs.14,15.) Can we write down a comparably simple and general formula for $f(\Gamma, t)$ when the system is driven out of equilibrium by the rapid variation of λ ? Probably not: the phase space density can evolve in a horribly complicated way, moreover

this evolution will likely depend on details of the environment, and of the coupling between system and environment. In other words, if we insist on using $f(\Gamma, t)$ to represent the evolving state of our system, then we are stuck with a complicated, non-universal description. However, there exists another statistical representation of the state of our system – which I will call $g(\Gamma, t)$ – which does indeed evolve according to a simple, general formula.

Given our ensemble of realizations, we can use the phase space trajectories (Eq.13) to obtain a set of evolving values,

$$w_t^1, w_t^2, \dots, w_t^R, \quad (16)$$

by Eq.9. Let us think of w_t^i as an evolving auxiliary variable, recording of the work performed on the system during the i 'th realization. By transforming these auxiliary variables into statistical weights, $\Omega_t^1, \Omega_t^2, \dots, \Omega_t^R$, defined by

$$\Omega_t^i \equiv e^{-\beta w_t^i}, \quad (17)$$

we can construct the following *weighted* phase space density:

$$g(\Gamma, t) \equiv \lim_{d\Gamma \rightarrow 0} \lim_{R \rightarrow \infty} \frac{1}{d\Gamma} \frac{1}{R} \sum_{\Gamma_t^i \in \mathcal{B}} \Omega_t^i = \langle e^{-\beta w_t} \delta(\Gamma - \Gamma_t) \rangle. \quad (18)$$

As in Eq.11, the above sum is taken over all realizations which happen to be found within a box \mathcal{B} of volume $d\Gamma$ around the phase space point Γ , at time t ; the difference is that the unity in the sum in Eq.11 has been replaced by the weight Ω_t^i . The function $g(\Gamma, t)$ is a kind of “undemocratic” phase space density, in which the contribution of each realization at time t depends, by Eq.17, on the work performed on the system up to that time. [We can also interpret $g(\Gamma, t)$ as a “mass density” in phase space, by assigning to each realization a fictitious and time-dependent mass $m_t^i = \Omega_t^i$.] By contrast, $f(\Gamma, t)$ is a “democratic” phase space density, as each realization contributes equally (Eq.11). Note that at $t = 0$ the two statistical descriptions are the same, since $\Omega_0^i = 1$:

$$g(\Gamma, 0) = f(\Gamma, 0) = \frac{1}{Z_A} e^{-\beta H(\Gamma, A)}. \quad (19)$$

Recall that $f(\Gamma, t)$ was introduced as a natural description of the evolving state of our ensemble. We can view $g(\Gamma, t)$ as simply an alternative statistical representation (which happens to incorporate information about the work performed during each realization) of the same ensemble. What makes this alternative description interesting is the following prediction. *Even if the system is driven far from equilibrium by rapidly varying λ , the weighted density $g(\Gamma, t)$ evolves in a very simple way:*

$$g(\Gamma, t) = \frac{1}{Z_A} e^{-\beta H(\Gamma, \lambda_t)} \quad t > 0. \quad (20)$$

This statement is the central result around which this lecture is organized, and represents an answer to the question posed by the title.

IV. DERIVATIONS

A number of derivations of Eq.20 (and closely related results) have by now appeared in the literature^{6–15}. These differ from one another in the underlying assumptions that are made regarding the dynamics of the system. Here I will go through the details of three such derivations. The first of these, based on Hamiltonian evolution, applies to the special case of isolated systems. In the second derivation, the presence of a heat reservoir is modeled with Langevin dynamics (white noise plus dissipation). The third derivation applies to arbitrary thermal Markov processes.

In each case the analysis will begin with an evolution equation for the ordinary phase space distribution $f(\Gamma, t)$, e.g. the Liouville equation for Hamiltonian dynamics (Eq.21), the Smoluchowski equation for Langevin dynamics (Eq.34), etc. Following that, we will consider $h(\Gamma, w, t)$, the joint probability distribution for observing a microstate Γ and work value w at time t , and the evolution equation for h will be obtained from that for f by the simple addition of a continuity term $-\dot{w}\partial h/\partial w$. Finally, the “undemocratic” distribution $g(\Gamma, t)$ will be expressed as a weighted projection of $h(\Gamma, w, t)$ (Eq.26), and an explicit evolution equation for g will be obtained. In all three cases we will find that Eq.20 is an exact solution of the evolution equation for g , thus proving our central assertion for the dynamics considered.

A. Hamiltonian evolution

Let us modify the thought experiment described above, by assuming that the polymer in Fig.1 is isolated from any thermal environments as we pull apart the laser traps. Thus, we first equilibrate the system with a reservoir at temperature T , holding $\lambda = A$ fixed, and then *we remove the reservoir*, start the clock running, and implement the protocol λ_t . The trajectory Γ_t then evolves under Hamilton’s equations – as appropriate for an isolated classical system – with a time-dependent Hamiltonian $H(\Gamma, \lambda_t)$. Given a statistical ensemble of realizations, the phase space density $f(\Gamma, t)$ satisfies the Liouville equation,

$$\frac{\partial f}{\partial t} = \{H, f\} = -\dot{\mathbf{q}}\frac{\partial f}{\partial \mathbf{q}} - \dot{\mathbf{p}}\frac{\partial f}{\partial \mathbf{p}}, \quad (21)$$

with initial conditions given by the Boltzmann-Gibbs distribution, Eq.14. In the above equation, $\{\cdot, \cdot\}$ represents the Poisson bracket; $H = H(\Gamma, \lambda_t)$; and $\Gamma = (\mathbf{q}, \mathbf{p})$, where

$$\mathbf{q} = (\mathbf{r}_1, \dots, \mathbf{r}_N) \quad \text{and} \quad \mathbf{p} = (\mathbf{p}_1, \dots, \mathbf{p}_N) \quad (22)$$

specify the configurations and corresponding momenta, satisfying Hamilton’s equations:

$$\dot{\mathbf{q}} = \partial H/\partial \mathbf{p} \quad , \quad \dot{\mathbf{p}} = -\partial H/\partial \mathbf{q}. \quad (23)$$

If we keep track of the work w_t performed during each realization, then we can represent the evolution of the system by a trajectory (Γ_t, w_t) evolving in an “expanded phase space” which includes the auxiliary variable w . A statistical ensemble of such trajectories is described by joint probability distribution $h(\Gamma, w, t)$, defined as

$$h(\Gamma, w, t) = \langle \delta(\Gamma - \Gamma_t) \delta(w - w_t) \rangle, \quad (24)$$

by analogy with Eq.12. Making use of the fact that $\dot{w}_t = \dot{\lambda} \partial H / \partial \lambda$ (from Eq.9), we can immediately write down the evolution equation

$$\frac{\partial h}{\partial t} = \{H, h\} - \dot{\lambda} \frac{\partial H}{\partial \lambda} \frac{\partial h}{\partial w} = -\dot{\mathbf{q}} \frac{\partial h}{\partial \mathbf{q}} - \dot{\mathbf{p}} \frac{\partial h}{\partial \mathbf{p}} - \dot{w} \frac{\partial h}{\partial w}. \quad (25)$$

Just as Eq.21 is simply a continuity equation in phase space – describing the deterministic flow generated by Hamilton’s equations – so Eq.25 is a continuity equation in the expanded phase space.

Generically, we will not be able to obtain simple expressions for either $f(\Gamma, t)$ or $h(\Gamma, w, t)$. Indeed, if the underlying classical dynamics shows evidence of chaos, then these distributions will become stretched and folded with time in a very complicated manner. However, let us now consider the following function:

$$g(w, t) = \int dw h(\Gamma, w, t) e^{-\beta w}. \quad (26)$$

This is exactly the weighted distribution defined earlier (see Eqs.18 and 24). With this expression, the equation of motion for g follows after a single integration by parts from Eq.25:

$$\frac{\partial g}{\partial t} = \{H, g\} - \beta \dot{\lambda} \frac{\partial H}{\partial \lambda} g. \quad (27)$$

In sharp contrast with the equations of motion for f and h , Eq.27 affords a simple solution, namely,

$$g(\Gamma, t) = \frac{1}{Z_A} e^{-\beta H(\Gamma, \lambda_t)}, \quad (28)$$

as predicted. Here we have assumed initial equilibrium (Eq.19), and have made use of the identity $\{H, e^{-\beta H}\} = 0$, which follows from the chain rule for Poisson brackets:

$$\{H, e^{-\beta H}\} = -\beta e^{-\beta H} \{H, H\} = 0. \quad (29)$$

Eq.20 is thus satisfied for isolated Hamiltonian systems.

B. Langevin evolution

Let us now assume that the system of interest is in thermal contact with a heat reservoir while we vary λ , and let us model the evolution of the system as a Langevin process¹⁶. In other words, instead of explicitly including the water molecules in the analysis, we mock up their presence by incorporating randomness into the evolution of the polymer. Specifically, we assume that the Hamiltonian for the polymer itself takes has the usual kinetic + potential form,

$$H(\mathbf{q}, \mathbf{p}, \lambda) = \frac{\mathbf{p}^2}{2m} + V(\mathbf{q}, \lambda) \quad (30)$$

(where $\mathbf{p}^2 = \sum_{i=1}^N p_i^2$, and the potential energy V includes the confining harmonic traps for atoms 1 and N) and we model the evolution of the polymer with the following equations of motion:

$$\dot{q}_\mu = \frac{p_\mu}{m} \quad , \quad \dot{p}_\mu = -\frac{\partial V}{\partial q_\mu} - \gamma p_\mu + \xi_\mu. \quad (31)$$

Here, q_μ and p_μ are the μ 'th components of the $3N$ -dimensional vectors \mathbf{q} and \mathbf{p} , γ is a friction coefficient, and ξ_μ represents *white noise*, with an autocorrelation function

$$\overline{\xi_\mu(t_1) \xi_\nu(t_2)} = D \delta_{\mu\nu} \delta(t_2 - t_1). \quad (32)$$

These evolution equations (Eq.31) are just Hamilton's equations, modified by the addition of friction ($-\gamma p_\mu$) and noise (ξ_μ). These terms represent a straightforward method for modeling the "random kicks" received by the polymer from the surrounding water molecules, and are related by a fluctuation-dissipation relation,

$$\gamma = \beta D / 2m. \quad (33)$$

A statistical ensemble of trajectories evolving under Eq.31 is described by the Smoluchowski equation for the time-dependent phase space distribution¹⁶:

$$\frac{\partial f}{\partial t} = \{H, f\} + \gamma \frac{\partial}{\partial \mathbf{p}} \cdot (\mathbf{p}f) + \frac{D}{2} \nabla_p^2 f, \quad (34)$$

where $\nabla_p^2 \equiv \sum_\mu \partial^2 / \partial p_\mu^2$, and $H = H(\mathbf{q}, \mathbf{p}, \lambda)$, as always.

We now proceed exactly as we did in Section IV A. That is, given an ensemble of trajectories Γ_t governed by this stochastic process, we imagine that we keep track of the work w_t for each realization, and we construct the joint probability distribution $h(\Gamma, w, t)$ defined by Eq.24. This distribution satisfies the evolution equation

$$\frac{\partial h}{\partial t} = \{H, h\} + \gamma \frac{\partial}{\partial \mathbf{p}} \cdot (\mathbf{p}h) + \frac{D}{2} \nabla_p^2 h - \dot{w} \frac{\partial h}{\partial w}. \quad (35)$$

As before, we have simply added a continuity term, $-\dot{w} \partial h / \partial w$, to account for the auxiliary variable w_t .

Since neither the evolution of f nor that of h is simple[†], we consider $g = \int dw h e^{-\beta w}$ (Eq.26), and we use Eq.35 to obtain:

$$\frac{\partial g}{\partial t} = \{H, g\} + \gamma \frac{\partial}{\partial \mathbf{p}} \cdot (\mathbf{p}g) + \frac{D}{2} \nabla_p^2 g - \beta \dot{\lambda} \frac{\partial H}{\partial \lambda} g. \quad (36)$$

Simple inspection reveals that a solution of this equation is given by $g = Z_A^{-1} \exp -\beta H(\Gamma, \lambda_t)$.

[†]except in the quasistatic limit, where we get $f = p_{\lambda_t}^{BG}(\Gamma)$ and $h = p_{\lambda_t}^{BG}(\Gamma) \delta(w - F_{\lambda_t})$.

C. Thermal Markov dynamics

The Langevin evolution considered above is an example of a *thermal Markov process*. In this section I briefly discuss the defining features of Markov processes in general¹⁷, and then establish the validity of Eq.20 for arbitrary *thermal* Markov processes. An alternative approach in this context has been taken by Hummer and Szabo¹⁴, who have recently pointed out that Eq.20 follows directly from the Feynman-Kac theorem¹⁸.

A *stochastic* process is simply one which contains an element of randomness. When modeling a system in contact with a thermal environment, this randomness acts as a substitute for explicitly including the degrees of freedom of the reservoir. Thus, acknowledging that the atoms in Fig.1 are continuously buffeted by water molecules, we can model these collisions as random “kicks” and attempt to construct a reasonable prescription for generating them. Langevin dynamics (Eq.31) represents one approach, with the kicks modeled as a combination of white noise and friction. Another approach would be to occasionally – with probability-per-unit-time α – select one of the atoms (i) at random and replace its momentum with one sampled randomly from the Maxwell-Boltzmann distribution ($\propto \exp -\beta p_i^2/2m$); between such “thermalizing collisions”, the evolution obeys Hamilton’s equations. This scheme is known as the *Andersen thermostat*¹⁹.

The Markov assumption is easily illustrated in the context of a numerical simulation, which inherently involves some prescription for obtaining a new microstate Γ at time $t + dt$, from an old microstate Γ' at time t , where dt is the small but finite numerical time step. Typically this prescription utilizes some deterministic integrator (such as the Runge-Kutta algorithm²⁰), but if the process being simulated is stochastic, then there will be additional terms requiring the generation of random numbers. *The process is Markovian if the new microstate Γ is uniquely determined from the old microstate Γ' , along with whatever random numbers are generated at time t .* In particular, the random numbers generated at previous time steps must not enter directly into the determination of Γ . For instance, the Andersen thermostat might be implemented as follows. At time t we first integrate Hamilton’s equations for a time step dt , and then we generate a random number ζ_t between 0 and 1. If $\zeta_t \leq \alpha dt$, then this signifies a thermalizing collision, in which case we generate further random numbers to choose a specific atom and assign its new momentum. However, if we were to impose a “dead time” after every collision, e.g. by prohibiting two collisions at successive time steps, then this would violate the Markov assumption, since the probability for generating a collision at time t would depend explicitly on whether or not $\zeta_{t-dt} \leq \alpha dt$.

Formally, a Markov process occurring in discrete time steps dt is completely specified by a function $P(\Gamma, t + dt|\Gamma', t)$, which gives the probability distribution of obtaining Γ at time $t + dt$, given Γ' at time t . A continuous-time Markov process is obtained in the limiting case of infinitesimal time steps, and is specified by a *transition function*

$$\mathcal{L}_t(\Gamma' \rightarrow \Gamma) = \lim_{dt \rightarrow 0} \frac{1}{dt} [P(\Gamma, t + dt|\Gamma', t) - P(\Gamma, t|\Gamma', t)], \quad (37)$$

which gives the instantaneous rate for making transitions to microstate Γ , starting from Γ' at time t . [Note that $P(\Gamma, t|\Gamma', t) = \delta(\Gamma - \Gamma')$.] The evolution of an ensemble of trajectories governed by such a Markov process is described by the following equation for the time-dependent phase space density:

$$\frac{\partial f}{\partial t}(\Gamma, t) = \int d\Gamma' f(\Gamma', t) \mathcal{L}_t(\Gamma' \rightarrow \Gamma), \quad (38)$$

or more compactly,

$$\frac{\partial f}{\partial t} = \hat{\mathcal{L}}_t f, \quad (39)$$

where $\hat{\mathcal{L}}_t$ is a time-dependent linear operator acting on the space of distributions f .

In the context of the problem we are considering, Eq.39 represents – statistically – the evolution of our system as we vary the parameter λ in the presence of a heat reservoir. Assuming the characteristics of the reservoir itself (e.g. its temperature, density, etc.) are constant, the time-dependence of the transition operator $\hat{\mathcal{L}}_t$ enters only through the value of λ . Hence, we will change notation slightly and write

$$\frac{\partial f}{\partial t} = \hat{\mathcal{L}}_{\lambda_t} f, \quad (40)$$

where $\hat{\mathcal{L}}_{\lambda}$ is a *parameter*-dependent transition operator. Eq.40 is a *master equation*: given an initial distribution $f(\Gamma, 0)$ and a protocol λ_t for varying the external parameter, Eq.40 determines the subsequent evolution of the distribution $f(\Gamma, t)$.

The transition operator formalism is quite generally applicable to Markov processes. However, in order to obtain any meaningful results from the application of this formalism to the problem at hand, we must incorporate the specifically *thermal* nature of the random forces acting on the system. We do so by imposing the following constraint on our transition operator:

$$\hat{\mathcal{L}}_{\lambda} e^{-\beta H(\Gamma, \lambda)} = 0, \quad (41)$$

or, equivalently, $\hat{\mathcal{L}}_{\lambda} p_{\lambda}^{BG} = 0$. Combined with Eq.40, this simply says that if we begin with our ensemble in the Boltzmann-Gibbs distribution corresponding to a given value of λ , and then we keep the external parameter fixed at that value, then the ensemble will not change with time. In other words, *the Boltzmann-Gibbs distribution is stationary when λ is held fixed*. If a Markov process satisfies Eq.41, then we will call it a *thermal* Markov process.

With these definitions in place, we now proceed to derive Eq.20 for arbitrary thermal Markov dynamics. As in Sections IV A and IV B, for an ensemble of trajectories Γ_t evolving under this stochastic process, we write down an evolution equation for the joint probability distribution $h(\Gamma, w, t)$:

$$\frac{\partial h}{\partial t} = \hat{\mathcal{L}}_{\lambda_t} h - \dot{\lambda} \frac{\partial H}{\partial \lambda} \frac{\partial h}{\partial w}. \quad (42)$$

The first term on the right describes the evolution of the ensemble of trajectories Γ_t , the second is a continuity term accounting for the auxiliary variable w_t .[‡]

[‡]Integrating both sides of Eq.42, and then performing an integration by parts on the last term, we recover Eq.39.

Given Eq.42, the weighted distribution $g = \int dw h e^{-\beta w}$ satisfies

$$\frac{\partial g}{\partial t} = \int dw \frac{\partial h}{\partial t} e^{-\beta w} \quad (43)$$

$$= \int dw \left(\hat{\mathcal{L}}_{\lambda_t} h - \dot{\lambda} \frac{\partial H}{\partial \lambda} \frac{\partial h}{\partial w} \right) e^{-\beta w} \quad (44)$$

$$= \left(\hat{\mathcal{L}}_{\lambda_t} - \beta \dot{\lambda} \frac{\partial H}{\partial \lambda} \right) g. \quad (45)$$

Now making use of Eq.41, we once again find that a solution of this evolution equation is given by $g(\Gamma, t) = Z_A^{-1} \exp -\beta H(\Gamma, \lambda_t)$.

D. Other derivations

In addition to the derivations presented above, two other approaches to establishing the validity of Eq.20 deserve mention, although details of the analyses will not be provide here.

In the dynamics considered above, time was taken to be a continuous variable. However, it is common to model thermal processes as occurring in discrete time steps. The discreteness of time is not to be viewed here as a necessary evil related to practical issues of numerical simulation, but rather as intrinsic to the process. Assuming that this evolution is Markovian, a given realization of the process is specified by a *Markov chain* (an ordered sequence of microstates, $\Gamma_0 \rightarrow \Gamma_1 \rightarrow \dots$). For such discrete-time Markov processes, the validity of Eq.20 is easily established^{7,12,15}, given an appropriate definition of work, along with a thermal assumption akin to Eq.41.

As mentioned earlier, stochastic processes (Markovian or otherwise) are ultimately convenient models which allow us to avoid explicitly dealing with the degrees of freedom of the heat reservoir. As such, these models are only approximations of physical reality, and it would be desirable to have a derivation of Eq.20 based on more fundamental equations of motion. An obvious approach is to proceed along the lines taken in Section IV A, but to drop the assumption that the system is isolated, and instead include *all* microscopic degrees of freedom – e.g. of both the polymer and the surrounding water – in the analysis. This is in fact not very difficult: treating the system and reservoir together as a large, isolated system, one can simply apply the analysis of Section IV A to this larger system, and then project out the reservoir degrees of freedom²¹. A certain amount of care is needed if the physical interaction between the system and reservoir is non-negligible in magnitude (in effect, the Hamiltonian of the polymer needs to be “renormalized” to include a free energy of solvation²²), but at the end of the day one again obtains Eq.20. Moreover, this approach can be extended to more general (non-isothermal) irreversible processes¹¹.

V. THE NONEQUILIBRIUM WORK RELATION

The central prediction of this lecture, Eq.20, is admittedly somewhat abstract, not the sort of result which typically inspires experimentalists to rush off to the laboratory to test its validity! Let us now assume (for specificity) that from $t = 0$ to $t = \tau$ we vary our parameter from the initial value λ_A to a final value λ_B , and let $W = w_\tau$ denote the total

work we perform during this process. By setting $t = \tau$ in Eq.20, integrating both sides of the equation over all of phase space, and invoking Eq.8, we obtain the *nonequilibrium work relation*:

$$\langle e^{-\beta W} \rangle = e^{-\beta \Delta F}, \quad (46)$$

where $\Delta F = F_B - F_A$. This result relates irreversible work (l.h.s.) to an equilibrium free energy difference (r.h.s.), and – as with Eq.20 – *remains valid no matter how quickly or slowly we vary λ* .

As formulated above, the nonequilibrium work relation applies regardless of whether or not the system itself ever reaches equilibrium state B : if we simply pull the laser traps apart at a constant, finite rate, then at the moment we reach λ_B the polymer will be in some nonequilibrium state.[§] However, we are perfectly within our rights to consider a process whereby λ is first varied rapidly from λ_A to λ_B , and then held fixed until the system relaxes to equilibrium state B , over a total time interval from $t = 0$ to $t = \tau$. In this scenario we have an irreversible process from one equilibrium state to another ($A \rightarrow B$), and Eq.46 relates the work performed during this process to the free energy difference between the initial and final states of the system.

The nonequilibrium work relation is both easier to state and experimentally more accessible than the more abstract result from which it was derived. Fundamentally, while Eq.20 makes a prediction regarding a (weighted) distribution in a $6N$ -dimensional phase space, Eq.46 involves a distribution on the one-dimensional work axis, in the following sense. Given infinitely many repetitions of our pulling experiment with the polymer, let $\rho(W)$ denote the distribution of values of work: $\rho(W) dW$ is the fraction of realizations for which the work value fell between W and $W + dW$. Then the nonequilibrium work relation predicts that

$$\int dW \rho(W) e^{-\beta W} = e^{-\beta \Delta F} \quad (47)$$

even if the system is driven far from equilibrium as we vary the parameter from λ_A to λ_B .

To test this prediction experimentally, of course, one does not need to actually construct $\rho(W)$. Rather, given R repetitions of the pulling experiment, we expect

$$\frac{1}{R} \sum_{i=1}^R e^{-\beta W^i} \approx e^{-\beta \Delta F}, \quad (48)$$

where W^i is the work performed during the i 'th realization. There is a catch, however: because the quantity whose average is being taken is highly nonlinear, very many repetitions might be needed in order for this approximation to be a good one. Specifically, if $\rho(W)$ is much wider than β^{-1} , then the above average will be dominated by values of work which are in the far left tail of the distribution⁶. In that case R must be extremely large in order to sample that region of the W axis with good statistics, making Eq.46 experimentally inaccessible, though formally correct.

[§]We should in this case interpret F_B as a free energy associated with the final value of λ (Eqs.7 and 8), rather than as “the final free energy of the system”.

While these considerations seem to rule out laboratory tests using macroscopic systems (roughly speaking, we can expect the dispersion of work values to grow with the size of the system), microscopic systems are a different story. In recent experiments using a laser trap and a micromechanical cantilever to pull on the two ends of an RNA strand, as described briefly in Section II, Liphardt *et al*⁴ have found good agreement with Eq.46. These results have provided the first experimental evidence of the validity of the nonequilibrium work relation, and by extension of Eq.20, suggesting that these theoretical predictions might ultimately become offer useful tools in the analysis of experimental data. Indeed, Hummer and Szabo¹⁴ had earlier proposed a scheme, based on Eq.20, for extracting equilibrium free energy information – specifically a potential of mean force – from micromanipulation experiments.

The nonequilibrium work relation might also provide a useful tool for estimating free energy differences from numerical simulations. A traditional method of computing ΔF involves simulating the system under consideration, at fixed temperature, as an external parameter λ is varied from an initial value (A) to a final value (B). If the simulation proceeds sufficiently slowly that the system remains very close to equilibrium from start to finish, then the work performed on the system gives a good estimate of ΔF (Eq.2). However, for many systems of practical interest – such as large biomolecules, where the estimation of free energy differences plays a central role in rational drug design²³ – such nearly-reversible simulations are prohibitively time-consuming. The nonequilibrium work relation offers a potential method for getting around this problem, by using a number of irreversible simulations and then either directly taking the “exponential average” of the work, $\langle e^{-\beta W} \rangle$, to obtain ΔF , or else using an approximation derived from Eq.46. Recent results have suggested that this may indeed be a practical approach to numerical free energy estimation^{24–28}.

VI. RELATION TO SECOND LAW OF THERMODYNAMICS

Both the nonequilibrium work relation and the Second Law of Thermodynamics have something to say about irreversible-isothermal processes connecting two equilibrium states A and B . The former claims that $\langle e^{-\beta W} \rangle = e^{-\beta \Delta F}$, the latter that $W > \Delta F$ (Eq.3). What is the relationship between these two statements? Indeed, are they even compatible? On the face of it, the answer to the second question seems to be negative: if $W > \Delta F$ for *every* realization, then the $\langle e^{-\beta W} \rangle$ would necessarily be less than $e^{-\beta \Delta F}$. However, we know that the Second Law must ultimately be interpreted *statistically*, thus allowing for the possibility that occasionally (though extremely rarely, for macroscopic systems) we will observe a realization for which $W < \Delta F$. With this in mind we now show that the nonequilibrium work relation implies two inequalities which are closely related to the Second Law.

*Jensen’s inequality*²⁹ states that $\overline{\exp x} \geq \exp \bar{x}$, where the overbar denotes an average over any set of values of a real variable x . If we combine this mathematical result with Eq.46, we get $\exp -\beta \Delta F = \langle \exp -\beta W \rangle \geq \exp -\beta \langle W \rangle$, i.e.

$$\langle W \rangle \geq \Delta F. \tag{49}$$

In other words, the nonequilibrium work relation immediately and rigorously implies that the average work performed, over an ensemble of realizations of an irreversible-isothermal process, is no less than ΔF . This is certainly consistent with the Second Law, but in fact we can say more, as we now show.

Given a particular irreversible-isothermal process, suppose we observe a realization for which $W = \Delta F - \epsilon$, where $\epsilon > 0$ has units of energy. The value of ϵ is a measure of the degree to which the Second Law was “violated” for this particular realization. Using the nonequilibrium work relation, we now rigorously derive a prediction regarding the probability distribution of such putatively illegal events. With $\rho(W)$ as defined as in Section V, let

$$\text{Prob}(W < \Delta F - \epsilon) = \int_{-\infty}^{\Delta F - \epsilon} dW \rho(W) \quad (50)$$

denote the probability of observing a work value less than $\Delta F - \epsilon$. Using the chain of inequalities,

$$\int_{-\infty}^{\Delta F - \epsilon} dW \rho(W) \leq \int_{-\infty}^{\Delta F - \epsilon} dW \rho(W) e^{\beta(\Delta F - \epsilon - W)} \leq e^{\beta(\Delta F - \epsilon)} \int_{-\infty}^{+\infty} dW \rho(W) e^{-\beta W}, \quad (51)$$

and then invoking Eq.47, we get

$$\text{Prob}(W < \Delta F - \epsilon) \leq e^{-\beta\epsilon}. \quad (52)$$

This result states that the probability of observing a violation of at least magnitude ϵ , decays exponentially (or faster) with $\epsilon/k_B T$. This in turn implies that the chance of observing a macroscopic violation ($\epsilon \gg k_B T$) is fantastically small, in complete agreement with the empirical evidence.

VII. NONEQUILIBRIUM STATIONARY STATES

Up to this point, this lecture has concerned systems driven away from equilibrium by the external variation of some parameter. The underlying assumption has been that, if the parameter is held fixed, the system relaxes to and then remains in a state of thermal equilibrium. However, recent work by Hatano and Sasa^{30,31} has revealed that much of the above formalism can be generalized to arbitrary stationary states, as I now briefly discuss.

Suppose we have a system whose dynamics is a Markov process specified by a parameter-dependent transition function $\mathcal{L}_\alpha(\Gamma' \rightarrow \Gamma)$. If the parameter α is held fixed, then a statistical ensemble of realizations of this process obeys the master equation

$$\frac{\partial f}{\partial t} = \hat{\mathcal{L}}_\alpha f. \quad (53)$$

Let us assume that this Markov process has a unique stationary distribution, $f_\alpha^S(\Gamma)$, for any value of α . This stationary distribution satisfies

$$\hat{\mathcal{L}}_\alpha f_\alpha^S = 0, \quad (54)$$

but we do *not* assume that this distribution is a Boltzmann-Gibbs distribution. Rather, $f_\alpha^S(\Gamma)$ represents a *nonequilibrium stationary state* of the system.

Let us now define $\phi(\Gamma, \alpha)$ by

$$f_{\alpha}^S(\Gamma) = e^{-\phi(\Gamma, \alpha)}, \quad (55)$$

and introduce an auxiliary variable

$$y_t = \int_0^t dt' \dot{\alpha}_{t'} \frac{\partial \phi}{\partial \alpha}(\Gamma_{t'}, \alpha_{t'}), \quad (56)$$

by analogy with our earlier definition of work (Eq.9). Here we have assumed some protocol α_t for varying the external parameter.

Now consider the following situation. Beginning with the system in some stationary state $f_{\alpha_A}^S$, we change the parameter from α_A to a new value α_B according to a protocol α_t ; we observe the evolution of the system, Γ_t ; and from this trajectory we construct the evolution of the auxiliary variable, y_t . Then we repeat this experiment infinitely many times, always employing the same protocol α_t . From this data, assigning a time-dependent statistical weight e^{-y_t} to each trajectory, we finally construct the weighted distribution

$$g(\Gamma, t) = \langle e^{-y_t} \delta(\Gamma - \Gamma_t) \rangle. \quad (57)$$

You can probably guess what I will now tell you: no matter how slowly or quickly we vary α , the evolution of this weighted distribution obeys a very simple equation, namely,

$$g(\Gamma, t) = e^{-\phi(\Gamma, \alpha_t)}. \quad (58)$$

(Since the derivation of this result is essentially identical to that presented in Section IV C, I won't reproduce it here.)

From this result we can obtain predictions entirely analogous to the nonequilibrium work relation³¹, as well as the inequalities given by Eqs.49 and 52. Thus, the entire mathematical structure of Sections IV C, V, and VI carries over to the context considered here. Unfortunately, a serious obstacle currently prevents us from transforming this abstract analysis into predictions which might be tested experimentally: *there exists no general microscopic statistical theory of nonequilibrium stationary states*. In other words, for a specific physical situation, we will not generally have an explicit expression for $f_{\alpha}^S(\Gamma)$, hence we will not know $\phi(\Gamma, \alpha)$, and will not be able to deduce the evolution of the auxiliary variable y_t from knowledge of the trajectory Γ_t .

VIII. CONCLUSIONS

In this lecture I have tried to summarize recent research related to the statistical physics of systems driven away from an initial state of thermal equilibrium. The central point which I have stressed is that there exists a particular statistical representation of such processes which is simple, largely independent of the details of the process, and valid even far from equilibrium. Moreover, this formalism leads to predictions which are both potentially useful and experimentally testable.

The extension of these results to the regime of quantum mechanics represents an obvious potential avenue for future research. Yukawa¹³ has made progress in this direction with

respect to the nonequilibrium relation, but the existence of a quantal version of Eq.20 remains an open question.

Finally, the generalization discussed in Section VII is, in my opinion, very intriguing. It tells us that if some day a statistical theory of nonequilibrium stationary states (analogous to the Boltzmann-Gibbs formalism for thermal equilibrium) will be constructed, then we will automatically be able to make concrete predictions about systems driven away from such states. In particular, the generalization of Eqs.49 and 52 suggests that transitions between nonequilibrium stationary states obey an inequality analogous to the Clausius inequality (Eq.4) for transitions between equilibrium states. Indeed, Hatano and Sasa³¹ have related their work in along these lines to the phenomenological, macroscopic framework introduced by Oono and Paniconi³² for the study of nonequilibrium stationary states. However, the story will not really be complete until, as suggested above, a genuine microscopic understanding of such states is developed.

REFERENCES

- ¹ A.Einstein, Ann. Phys. (Leipzig) **17**, 549 (1905).
- ² M.Smoluchowski, Ann. Phys. (Leipzig) **21**, 756 (1906).
- ³ L.Onsager, Phys.Rev. **37**, 405 (1931); Phys.Rev. **38**, 2265 (1931).
- ⁴ J.Liphardt *et al*, submitted for publication.
- ⁵ The essence of this interpretation seems to date back to Erwin Schrödinger's lectures: *Statistical Thermodynamics* (Cambridge, 1962); see the paragraphs found between Eqs.2.13 and 2.14. For a more recent reference, see K.Sekimoto, Prog.Theor.Phys.Suppl. **130**, 17 (1998).
- ⁶ C.Jarzynski, Phys.Rev.Lett. **78**, 2690 (1997).
- ⁷ C.Jarzynski, Phys.Rev.E **56**, 5018 (1997).
- ⁸ G.E.Crooks, J.Stat.Phys. **90**, 1481 (1998).
- ⁹ C.Jarzynski, Acta Phys.Polonica B **29**, 1609 (1998).
- ¹⁰ G.E.Crooks, Phys.Rev.E **60**, 2721 (1999).
- ¹¹ C.Jarzynski, J.Stat.Phys. **96**, 415 (1999).
- ¹² G.E.Crooks, Phys.Rev.E **61**, 2361 (2000).
- ¹³ S.Yukawa, J.Phys.Soc.Japan **69**, 2367 (2000).
- ¹⁴ G.Hummer and A.Szabo, Proc.Natl.Acad.Sci. (USA) **98**, 3658 (2001).
- ¹⁵ R.M.Neal, Statistics and Computing **11**, 125 (2001).
- ¹⁶ H. Risken, *The Fokker-Planck Equation*, Springer Verlag, 1989.
- ¹⁷ C.W.Gardiner, *Handbook of Stochastic Methods*, Springer Verlag, 2001.
- ¹⁸ Z.Schuss, *Theory and Applications of Stochastic Differential Equations*, Wiley, 1980.
- ¹⁹ H.C.Andersen, J.Chem.Phys. **72**, 2384 (1980).
- ²⁰ W.H.Press *et al*, *Numerical recipes in C : the art of scientific computing*, Cambridge University Press, 1988.
- ²¹ Ref.⁶ discusses this approach in the context of the nonequilibrium work relation.
- ²² C.Jarzynski, unpublished.
- ²³ M.R.Reddy and M.D.Erion, eds, *Free Energy Calculations in Rational Drug Design*, Kluwer Publishing, 2001.
- ²⁴ D.A.Hendrix and C.Jarzynski, J.Chem.Phys. **114**, 5974 (2001).
- ²⁵ G.Hummer, J.Chem.Phys. **114**, 7330 (2001).
- ²⁶ H.Nanda, D.M.Zuckerman, T.B.Woolf, Biophys.J. **82**, 328A (2002).
- ²⁷ D.M.Zuckerman and T.B.Woolf, Chem.Phys.Lett. **351**, 445 (2002).
- ²⁸ M.O.Jensen *et al*, submitted for publication.
- ²⁹ D.Chandler, *Intoduction to Modern Statistical Mechanics*, Oxford University Press, 1987, Section 5.5.
- ³⁰ T.Hatano, Phys.Rev.E **60**, R5017 (1999).
- ³¹ T.Hatano and S.Sasa, Phys.Rev.Lett. **86**, 3463 (2001).
- ³² Y. Oono and M. Paniconi, Prog. Theor. Phys. Suppl. **130**, 29 (1998).

FIGURES

FIG. 1. A toy model of a system (polymer) in contact with a heat reservoir (water). See text for details.