The Dynamics of Disorder

http://threeplusone.com/disorder

Gavin E. Crooks
March 18, 2015

Contents

Equilibrium and disequilibrium
1 Equilibrium ............................................. 2
   1.1 The canonical ensemble .................. 2
   1.2 Microscopic reversibility ................. 2
2 ...and disequilibrium .......................... 4
   2.1 Experimental protocols ................. 4
   2.2 Work and Heat ............................... 4
   2.3 Microscopic irreversibility ............. 4
3 The fluctuations of dissipation ............... 6
   3.1 Transient protocols ...................... 6
   3.2 Conjugate fluctuation theorem .......... 6
   3.3 The Jarzynski Identity ................. 6
   3.4 Clausius inequality and the second law 7

Appendices
A Miscellaneous Math .............................. 8
1 Equilibrium ...

We must first consider the defining properties of thermodynamic equilibrium. To do: Elaborate

1.1 The canonical ensemble

The central, fundamental law of statistical mechanics is the following: For a system in thermal equilibrium with its environment, then the probability that a system will be found in a particular microstate is given by the canonical distribution of Gibbs [3]

\[ p(x | \beta, \lambda) = \frac{1}{Z} e^{-\beta E(x, \lambda)} \]

\[ = e^{-\beta E(x, \lambda) + \beta F(\lambda)} \]  \hspace{1cm} (1)

Here, x is the state of the system, \( \beta = 1/k_B T \) is the inverse temperature of the environment in natural units (\( k_B \) is the Boltzmann constant and T is temperature), \( E(x, \lambda) \) is the energy of the system, a function of both the internal state x and any experimentally controllable parameters \( \lambda \), and \( Z \) is a normalization constant, called the partition function, and \( F = -k_B T \ln Z \) is the free energy.

This expression is very general, since the details of a particular system enter through the density of energy states, and the dynamics of the system are largely irrelevant. The entire subject of statistical mechanics consists of the derivations and consequences of this relation.

Temperature: The temperatures that occur in the canonical ensemble, and in our studies of non-equilibrium dynamics, are the equilibrium temperatures of a large thermal environment. For these systems, the temperature \( T \) is the partial derivative of the energy with respect to entropy [at equilibrium].

\[ k_B T = \frac{\partial E}{\partial S} \]  \hspace{1cm} (2)

Temperatures are not well defined away from thermal equilibrium [except if the system is in a meta-stable pseudo-equilibrium].

Conventionally, temperature is measured in the arbitrary units of Kelvin, and the thermodynamic entropy is reported in the illogical units of Kelvin per Joule per mole. This is Plank's fault. He knew that the units of temperature, energy and entropy were related, but he placed the conversion factor [which he called Boltzmann's constant, \( k_B \)] in front of the entropy, \( S = -k_B \sum \lambda p_x \ln p_x \). This was a mistake. The conversion factor belongs in front of the temperature. Entropy, with natural or binary logarithms, already has sensible units [nats or bits]. Therefore we will either declare that that entropy and temperature are measured in sensible units so that \( k_B = 1 \) [e.g. nats for entropy and Joules per nat for temperature], or we will always write a Boltzmann constant in front of the temperature \( k_B T \), and never in front of the entropy.

We’ll use \( \beta \) to denote the inverse temperature in natural units.

\[ \beta = \frac{1}{k_B T} = \frac{\partial S}{\partial E} \]  \hspace{1cm} (3)

Inverse temperature often appears in statistical mechanics, and is, from a fundamental viewpoint, a more natural quantity than the temperature itself. It is natural to control the flow of energy into a system, and inquire about the change in entropy. But we can’t directly control the entropy of a system, and ask about the change in energy.

Further reading: For a modern introduction to thermodynamics and statistical mechanics see Callen’s Thermodynamics and an Introduction to Thermostatistics [4] and Chandler’s Green Book [5]. Foundational and classic texts include [6, 3?, 7?].

1.2 Microscopic reversibility

The second law of thermodynamics states that entropy increases in the future. The principle of microscopic reversibility is a stronger statement: In the future, entropy increases. The fundamental microscopic dynamics of physics are all time reversal symmetric. It is only a change in entropy that can pick out a directionality of time. Nothing in the statistics of an assemblage can distinguish a direction of time when entropy fails to distinguish one.[8]

It follows that at thermodynamic equilibrium not only do the properties of the ensemble not change with time, but there is no discernible difference between moving forward in time and moving backwards. In other words, the probability of observing a system move through a particular sequence of states, e.g. \{1 2 3 2 1 2 1 2\} is the same [at equilibrium] whatever the starting time of the observation, and irrespective of whether we read time left to right or right to left. At equilibrium, there are no currents, and time has no origin nor preferred orientation.

We can concisely express the Principle of Microscopic Reversibility at Equilibrium as :

\[ P(X|\lambda) = P(\tilde{X}|\lambda) \]  \hspace{1cm} (4)

The probability of a trajectory \( X \) is the equal to the probability the time reversed trajectory \( \tilde{X} \), given equilibrium initial conditions and an unperturbed system with fixed control parameters, \( \lambda \).

The path probabilities include the probability of sampling the initial state of the trajectory from equilibrium. We can split the trajectory probability into the probability of the initial state and the conditional probability of the path, given the initial state, using the chain rule for conditional probabilities, \( P(A,B) = P(A|B)P(B) \),

\[ P[X, \lambda] = P[X|\lambda] = P(\tilde{X}|\lambda) \]  \hspace{1cm} (5)
and then rearrange this expression as a relation between probability ratios. [Note that whenever we write ratios of probabilities, we implicitly exclude cases where the probabilities are exactly zero.]

\[
\ln \frac{P[X|x_a,\lambda]}{P[X|x_b,\lambda]} = \ln \frac{P(x_b | \lambda)}{P(x_a | \lambda)} \tag{6}
\]

The right hand side is a ratio of equilibrium probabilities. Assuming a canonical equilibrium ensemble, this is equal to

\[
\ln \frac{P[X|x_a,\lambda]}{P[X|x_b,\lambda]} = -\beta \Delta E_{ab} \tag{7}
\]

**Detailed balance:** Microscopic reversibility is essentially synonymous with detailed balance. Detailed balance is expressed with respect to transitions between states, whereas microscopic reversibility is expressed in terms of trajectories. One implies the other.

\[
P(x|\lambda)P(x \rightarrow x'|\lambda, t) = P(x'|\lambda)P(x \rightarrow x|\lambda, t) \tag{8}
\]

To do: [TODO:Expand]

To do: Essentially, can take MR as the defining property of thermodynamic equilibrium. Implies, via Clausius inequality, that entropy is maximized. Also means that Hamiltonian must be even with respect to odd variables, such as velocities.
2 ...and disequilibrium.

With deceptively simple reasoning, we can adapt and extend the principle of microscopic reversibility to driven systems that are not in equilibrium.

2.1 Experimental protocols

We probe the nature of reality by conducting controlled experiments, typically by twiddling, with premeditation, some parameter of the system, which we shall label \( \lambda \). Typically, \( \lambda \) describes a single degree of freedom, such as the trap to trap distance in Fig. ??, although, in principle, \( \lambda \) could control the energy of every single microstate. An experimental protocol \( \Lambda \), depends on the time history of the controllable parameters, \( \lambda \).

We can identify various important classes of experimental protocols. Initially, we’ll consider transient protocols which perturb the system for a finite time. The control parameter \( \lambda \) varies between two times, \( a \) and \( b \), but before time \( a \) the control is fixed at \( \lambda_a \), and after time \( b \) the control is fixed at \( \lambda_b \).

![Graph of \( \lambda \) vs time with \( a \) and \( b \) marked]

2.2 Work and Heat

The first law of thermodynamics states that energy is neither created nor destroyed. Thus the change in energy of a system is equal to the flow of energy between the system and environment. This internal energy change \( \Delta E \) can be split into two distinct components, the work \( W \) and the heat \( Q \).

\[
\Delta E = Q + W \tag{9}
\]

Heat \( Q \) is the spontaneous flow of energy into the system from the environment, whereas work \( W \) is the change in energy due to a deliberate, controlled perturbation of the system. We will use the the sign convention that work and heat measure energy flux into the system. This convention is common in physics, whereas in engineering, work typically has the opposite sign and measures the flow of energy out of the system. To increase confusion, other conventions are also occasionally encountered.

Consider this sketch of the microscopic dynamics of a three state system. The horizontal axis is time, the vertical axis is energy, and the dot represents the current state of the system.

![Graph of energy vs time]

In the first step the system makes a transition between states. This transition could occur even if there were no external perturbation of the system, and therefore no work done on the system. Therefore, this energy change can be associated with an exchange of energy with the heat bath. In the second step the system remains in the same state, but the energy of the system changes because the external perturbation changed the energies of all the states. We associate this change in energy with the work.

In this model system, heat is a change in energy due to a change in the microscopic state of the system, and the work is a change in energy due to a change in the energy of the state that the system currently resides in.

More generally, we consider the energy of the system \( E(x, \lambda) \) to be a function of both the internal state \( x \) and an externally controllable parameter \( \lambda \). An experiential protocol \( \Lambda \) (or schedule) consists of the complete time dependence of the controllable parameters \( \lambda \). If the energy \( E \), coordinates \( x \) and control \( \lambda \) all vary smoothly with time, we can define the work and heat as:

\[
\int_a^b \frac{dE}{dt} = \int_a^b \left( \frac{\partial E}{\partial \lambda} \frac{d\lambda}{dt} + \frac{\partial E}{\partial x} \frac{dx}{dt} \right) dt \tag{10}
\]

The fundamental thermodynamic distinction between heat and work is that heat is associated with a temperature and therefore heat flowing into a system reduces the entropy of the environment. Work, on the other hand, does not have an associated change in environmental entropy of the.

2.3 Microscopic irreversibility

The sum of heat flow for the individual trajectory segments is the total heat of the trajectory \( Q_{ab} \), and we discover the important principle of temporal disorder \([2, 9]\) for an diabatic system:

\[
\ln \frac{P[X|\lambda_a, \Lambda]}{P[X|\lambda_b, \Lambda]} = -\beta Q[X, \Lambda] \tag{11}
\]

To do: [TODO: Check Signs]

Now, we can imagine a driven system where the control parameter changes in a series of discrete jumps, and is
constant between jumps. We can recover a smoothly varying control parameter by taking the limit of many small jumps, very close together.

The total probability of a trajectory can be written as a product of path probability, one for each time segment \((t_i, t_{i+1})\) with fixed \(\lambda_{t_i, t_{i+1}}\).

\[
P[X|\Lambda, x_{t_0}] = \prod_{i=0}^{n-1} P[X_{t_i, t_{i+1}}, \lambda_{t_i, t_{i+1}}, x_{t_i}] \tag{12}
\]

We can decompose the probability of the conjugate time reversed trajectory similarly,

\[
P[\tilde{X}, \tilde{x}_{t_n}] = \prod_{i=0}^{n-1} P[\tilde{X}_{t_i, t_{i+1}}, \tilde{\lambda}_{t_i, t_{i+1}}, \tilde{x}_{t_i, t_{i+1}}] \tag{13}
\]

We now look at the ratio of the trajectory probability to its time reversal and find that

\[
\ln \frac{P[X|\Lambda, x_{t_0}]}{P[\tilde{X}, \tilde{x}_{t_n}]} = \ln \prod_{i=0}^{n-1} \frac{P[X_{t_i, t_{i+1}}, \lambda_{t_i, t_{i+1}}, x_{t_i}]}{P[\tilde{X}_{t_i, t_{i+1}}, \tilde{\lambda}_{t_i, t_{i+1}}, \tilde{x}_{t_i, t_{i+1}}]} \tag{14a}
\]

\[
= \sum_{i=0}^{n-1} \ln \frac{P[X_{t_i, t_{i+1}}, \lambda_{t_i, t_{i+1}}, x_{t_i}]}{P[\tilde{X}_{t_i, t_{i+1}}, \tilde{\lambda}_{t_i, t_{i+1}}, \tilde{x}_{t_i, t_{i+1}}]} \tag{14b}
\]

\[
= -\sum_{i=0}^{n-1} \beta Q[X_{t_i, t_{i+1}}, \lambda_{t_i, t_{i+1}}] \tag{14c}
\]

\[
= -\beta Q[X_{t_0, t_n}, \lambda_{t_0, t_n}] \tag{14d}
\]

We take the log ratio, so that we can conveniently convert the product on line (a) into a sum on line (b). We then recognize that the log probability ratio for each segment is equal to the negative heat (times the inverse temperature), by the principle of microscopic reversibility discussed above. In the final line we sum up the heat for all the individual segments, to conclude that the log probability ratio of trajectories related by time reversal is equal to the negative heat times the inverse temperature of the environment, irrespective of the amount of work performed on the system.

To do: Elaborate on why this works. System doesn’t know if it is in equilibrium or not, since equilibrium is an ensemble property. An individual system has to respond the same way (while control parameters are fixed.)

The heat flow divided by the temperature of the bath is equal to the change entropy of the environment \(\beta Q_{ab} = \Delta S_{ab}^{\text{env}}\). Thus, we can also write microscopic irreversibility as

\[
\ln \frac{P[X|\Lambda, x_a]}{P[X|\tilde{X}, \tilde{\Lambda}]} = \Delta S^{\text{env}} \tag{15}
\]

An isolated system is always time reversal symmetric. The time-reversal symmetry is broken by the change in entropy of the environment. This is the general relation, applicable to many different situations. But we will leave these generalization to latter.

To do: Heat odd under time reversal. Therefore, main expression is symmetric.
3 The fluctuations of dissipation

The breaking of time reversal symmetry at the level of trajectories is not directly observable since even in the simplest systems there are very large number of possible paths that the system could follow and we can never ascertain the probability of any of them by direct observation.

3.1 Transient protocols

To do: Elaborate

![Diagram](image)

\[ \ln \frac{P[X|\Lambda]}{P[-X|\Lambda]} = -\Delta F + \beta W(X, \Lambda) \] \hspace{1cm} (16)

3.2 Conjugate fluctuation theorem

A direct, observable, consequence of this time reversal symmetry is the work fluctuation theorem\(^1\) [10]: given a transient protocol, the ratio of the probabilities of the forward and reverse paths is the exponential of the observed dissipation along the forward path.

\[ \ln \frac{P(+W|\Lambda)}{P(-W|\Lambda)} = \beta W - \beta \Delta F \] \hspace{1cm} (17)

To do: Step through how symmetric between forward and reverse.

To derive this relation we express the work probability as an average of a delta function across all possible paths, and then take a time reversal.

\[
\begin{align*}
P[+W|\Lambda] &= \sum_X P[X|\Lambda] \delta(W - W[X, \Lambda]) \\
&= \sum_X P[\bar{X} | \bar{\Lambda}] \delta(W - \bar{W}[X, \Lambda]) e^{-\beta W[X, \Lambda] + \beta \Delta F} \\
&= e^{-\beta W + \beta \Delta F} \sum_X P[\bar{X} | \bar{\Lambda}] \delta(W - \bar{W}[X, \Lambda]) \\
&= e^{-\beta W + \beta \Delta F} \sum_X P[\bar{X} | \bar{\Lambda}] \delta(W + W[\bar{X}, \bar{\Lambda}]) \\
&= e^{-\beta W + \beta \Delta F} P[-W|\Lambda]
\end{align*}
\]

We'll define the excess work as \(W_{\text{ex}} = W - \Delta F\) the difference between the actual work and the free energy change, which is the reversible work, the amount of work performed during a quasi-static transformation between the same end points. In much of the literature [including many of my early papers] (e.g. [2, 2]), the excess work is called the dissipative work, since this is the energy irreversibly lost into the environment if we allow the system to relax back to equilibrium. The problem with this terminology is that at the end of the protocol \(\Lambda\) the ensemble is still out-of-equilibrium, and the excess work has not yet been fully dissipated. We'll reserve the term dissipated work for the energy that has already been irreversible dissipated. [See latter.]

To do: Need to talk about transient path fluctuation theorem beforehand

To do: Nomenclature: detailed versus integrated fluctuation theorems.

3.3 The Jarzynski Identity

For a transient protocol, where the system starts in equilibrium and is perturbed for a finite time, the mean Boltzmann weighted work is equal to the Boltzmann weighted free energy change [11, 12].

\[ \langle e^{-\beta W} \rangle_{\Lambda} = e^{-\beta \Delta F} . \] \hspace{1cm} (18)

An equivalently expression statement is that the log mean Boltzmann weighted excess work is zero,

\[ \ln \langle e^{-\beta W_{\text{ex}}} \rangle_{\Lambda} = 0 . \] \hspace{1cm} (19)

The Jarzynski identity follows as a direct directly from the work fluctuation theorem [17].

\[ \langle e^{-\beta W} \rangle_{\Lambda} = \int dW P(+W|\Lambda) e^{-\beta W} \] \hspace{1cm} (20a)
\[ = \int dW P(-W|\Lambda) e^{-\beta \Delta F} \] \hspace{1cm} (20b)
\[ = e^{-\beta \Delta F} \int dW P(-W|\Lambda) \] \hspace{1cm} (20c)
\[ = e^{-\beta \Delta F} \] \hspace{1cm} (20d)

(a) First explicitly write the Jarzynski relation as an average over the work distributions; (b) then apply a time reversal using the work fluctuation theorem [17]; (c) move the free energy change out of the integral (since the free energy change is a constant); (d) finally recognize that the integral is unity due to normalization of probabilities.

Further reading: The Jarzynski identity was discovered by the following reasoning [2]: In computer simulations (pre Jarzynski) one can measure free energy changes either by perturbing the system very slowly, so that the work is very close to the reversible work, or we can perturb the system infinity fast, in a single jump, and analyze the work performed using thermodynamic perturbation. Since we can extract free engines with both infinitely fast
and infinitely slow perturbations, should we not be able to interpolate between the two extremes and extract free energies from finite time perturbations? This line of reasoning, along with a consistent, microscopic definition of thermodynamic work (??) (plus 135 years of intense scientific effort), lead to this remarkable simple, but profound, modification of the second law. The Jarzynski identity has been called the simplest relation in statistical mechanics that isn’t in Landa and Lifshits [2].

It was soon realized that the Jarzynski identity could be derived from microscopic reversibility [9], which in turn lead to the connection between the Jarzynski relation and fluctuation theorems [13]. The next key advance was when Hummer and Szabo observed that the Jarzynski relation was not just a theoretical construct of interest only to simulations, but directly applicable to real world, single molecule experiments [14, 15]. The definitive experiment was subsequently carried out by Liphardt et al. in 2002 [16].

The Jarzynski identity has been explicitly derived for many different model dynamics, including Hamiltonian dynamics [11, 12], Monte Carlo Markov chain [12, 9, 17], Langevin dynamics [12, 18], isothermal molecular dynamics [11, 12, 19, 20], and quantum dynamics [2, 2]. Hummer and Szabo derived the Jarzynski identity from the Feynman-Kac theorem [14].

### 3.4 Clausius inequality and the second law

Application of the Jensen inequality \( \ln(\exp(-x)) \geq \langle x \rangle \) to the Jarzynski identity (18) gives the Clausius inequality [11, 12]

\[
\beta \langle W \rangle \geq \beta \Delta F \quad (21a)
\]

\[
\langle \Delta S^{\text{total}} \rangle \geq 0 \quad (21b)
\]

The second law is sometimes erroneously understood to mean that entropy always increases, or almost always increases, or at least increases more often than decreases. But the total entropy can, in fact, decrease on occasion. In fact, such transient violations of the naive understanding of the second law are inevitable. The actual constraint of the Second Law is that the average entropy increase is positive, where the average is performed over multiple realizations of the protocol.

To do: Need to actually demonstrate that violations are inevitable. To do: Comment on how pioneers understood second law was statistical, but didn’t seem to know the correct way of expressing it.
A Miscellaneous Math

Delta function The defining property of the Dirac delta function is that
\[ \int_{-\infty}^{\infty} f(x) \delta(x - a) \, dx = f(a) \tag{22} \]
Informally, one can think of the delta function as being zero everywhere except for an infinitely tall and narrow spike, with unit area, located at \( x = 0 \).

A similar function is the Kronecker delta
\[ \delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \tag{23} \]
The Kronecker delta is a discrete function, so its amplitude at equality is one (rather than area as for the Dirac delta function.)

A not-uncommon convention is to use the same notation \( \delta(i-j) \) for both Dirac and Kronecker delta functions, and distinguish which is which by context: Dirac delta inside integrals, Kronecker delta inside sums.


Jensen’s inequality [21] states that the mean of a convex function is equal, or greater than, the convex function of the mean.
\[ \langle f(x) \rangle \geq f(\langle x \rangle) \tag{24} \]
Here, angled brackets indicate an average, and the function \( f \) is convex (The function always lies below the straight line connecting any two points of the function.)

For no apparently good reason, the special case \( \langle \exp(x) \rangle \geq \exp(\langle x \rangle) \) is often referred to by some long string of names in the statistical mechanics literature, e.g. Gibbs-Bogoliubov-Feynman-Jensen-Peierls inequality [5].

To do: Check

Logistic function The logistic (or Fermi) function is defined as
\[ f(x) = \frac{1}{1 + e^{-x}}. \tag{25} \]

Relative entropy The relative entropy (or Kullback-Leibler divergence) between two probability distribution \( p \) and \( q \) is
\[ D(p||q) = \sum_i p_i \ln \frac{p_i}{q_i} \tag{26} \]
Relative entropy is non-negative, and zero if, and only if, every probability is the same, \( p_i = q_i \) for all \( i \).

I prefer the nomenclature “relative entropy” since its shorter and more descriptive.

References
[2] [citation needed].


