Entropy as an effective action: general expression for the entropy of Esposito's non-equilibrium polymer model

Tom Weinreich and Eric Smith

August 30, 2011

Abstract

In this summer project, we sought to derive in a principled way an expression for the entropy of a classical system which would be independent of assumptions of equilibrium. The chosen system, as simple model of polymer growth is taken from Esposito et. al [1]. It was chosen for the phenomenological property that in certain regimes of parameter space, energy can be extracted from the environment through the growth of a disordered solid. The model is formulated as a stochastic process, taken through the formal development laid out by Smith in [2], and the resulting expression for the entropy is manipulated appropriately to elucidate its structure. The results are, at this point, so complex as to be enlightening, but we expect that further investigation will yield an understanding of appropriate simplifying assumptions, and hope to eventually derive an expression that is generally valid for the system and has a clear conceptual interpretation.

1 Introduction

Entropy is the central quantity in thermodynamics; from it, all other quantities can be derived. Given thermodynamics' role as a general theory of averaged processes, it is as such one of the most fundamental concepts in science, but despite that it tends to be very poorly understood beyond certain special cases. Classical thermodynamics deals only with conditions of equilibrium, which are an important but exceptionally limited set of the general case of dynamical systems. Its only content, however, beyond the microscopic equations of motion that define the system of interest, is statistics. It seems natural, then, to seek an extension of thermodynamics beyond the domain of equilibrium systems. In classical thermodynamics, the primary statistical property of interest is the stability of the means of probability distributions as the number of samples becomes large. As we move into the nonequilibrium domain, however, we will be more interested in the stability of the entire limiting distribution, and the limited number of functional forms that the distribution takes in the thermodynamic, or large-numbers limit [3]. Given entropy's central role, the way to make this extension will likely be to generalize its definition and then develop methods to calculate it.

The Shannon-Boltzmann entropy of a dynamical system is a sum over states of the probability p of each state times its logarithm. In the limit of large numbers of states, however, the probabilities become very small. When we are calculating macroscopic quantities that involve averaging over a very large number of microstates, we can take the limit of small microstate probabilities. To leading order in this limit, the entropy of a macrostate becomes proportional to the logarithm of the total probability of all microstates consistent with that macrostate. A state, however, is an insufficient amount of information to describe the conditions in a system that is not necessariy equilibrated. The added freedom is time-evolution, and to account for it we will keep track of entire time-dependent trajectories in state space, which we will refer to as "histories". By analogy with classical mechanics, given the prior knowledge that a system is equilibrated, a point in coordinate space uniquely (albeit trivially) specifies its subsequent time evolution. However, away from equilibrium, we need to know something about its history in order to predict its future behavior. Of course, in classical mechanics we neatly tie that information up into momentum, and the Hamiltonian formulation of mechanics elucidates the symmetry between coordinates and momenta in the structure of the theory: the two stand on equal footing, and complete specification of a system requires knowledge of both. As we move away from equilibrium in thermodynamics, we might hope to find a similar such encapsulation of the dynamics in additional state variables, and it turns out that the construction to be described below provides just that, formally converting a representation in terms of histories into a Hamiltonian dynamical system, complete with coordinates that correspond to classical thermodynamical state variables and their conjugate momenta. [4] We note that although the formalism will be identical, the physical interpretations of the

1

momenta that we will derive will be quite different than those in classical mechanics. [2]

With this generalization in hand, we turn to the second part of the prescription for the generalization of thermodynamics and attempt to derive the value of the entropy. Of course, this is easier said than done, and a more realistic goal is to do that for a particular system; the one that we chose is simple enough to be analytically tractable to a low-order approximation, but exhibits the interesting feature that entropy production drives its behavior. The remainder of the paper is devoted to the development of methodology and analysis of the model. In Section II, I will describe so-called Doi-Peliti construction [5, 6, 7, 8] and the resulting Friedlin-Wentzell description of the behavior of stochastic processes in terms of a dynamical system. This discussion will be oriented to those seeking a conceptual understanding; for a more rigorous description, see [2] and references therein. Section III will introduce the details of the Esposito model, and in Section IV I will describe my attempts at elucidation of the expression for the entropy and discuss likely directions of future work.

2 Constructing the system

2.1 The master equation

Every stochastic process can be described mathematically by its corresponding *master equation*, which describes the time evolution of probabilities in its state space. Following [2], I will use as an example the simplest possible stochastic process, which has 2 states and allows transitions back and forth between them; the generalization to arbitrary processes is fairly straight-forward. We will define a vector $\vec{x}$ which contains as its elements the occupation number of each state. ρ will then represent a state of the system, with the interpretation of the probability of being in any given state. Its master equation, in the continuous-time case, is

$$\frac{\partial \rho}{\partial t} = T\rho. \tag{1}$$

T is known as the *transfer matrix*, and it gives the rate of each transition. For the simple two-state process, with states labeled a and b and the rates of transitions given by r_a and r_b, the transfer matrix takes the form

$$(e^{-\mathrm{d}/\mathrm{d}n_a + \mathrm{d}/\mathrm{d}n_b} - 1)r_a + (e^{-\mathrm{d}/\mathrm{d}n_b + \mathrm{d}/\mathrm{d}n_a} - 1)r_b. \tag{2}$$

The exponentials are shift operators, which increment and decrement the particle numbers; the notation, though it may seem unnecessarily obtuse for the current purpose, will be convenient later.

2.2 The generating function

The next step in the construction is to introduce what is known as a *generating function*, which takes as its arguments a complex vector $\vec{z}$ with one component for each state. Defining the dimensionality of the state space (or equivalently, the number of states) as D, the generating function is defined as

$$\psi(\vec{z}) \equiv \sum_{\vec{x}} \prod_{i=1}^{D} z_i^{x_i} \rho. \tag{3}$$

The substitution $z \equiv e^{i\omega}$ reveals that this expression is equivalent to a Fourier transform of ρ. Intuitively, it may not be surprising that expressing the evolution of probability density in a basis of modes of fluctuation rather than of single-particle hops will prove to be more productive in the effort to understand the structure of fluctuations in a system. It can be shown that the z's in fact correspond to exponentials of chemical potential shifts [2], which will come up later.

Taking the time derivative of ψ, we see that its only time-dependence comes from ρ, and define the Liouville operator $\mathcal{L}$ via the equation

$$\frac{\partial \psi(\vec{z})}{\partial t} = -\mathcal{L}\psi. \tag{4}$$

A bit of thought reveals that $\mathcal{L}$ must have a very similar form to T, but in terms of $\vec{z}$ and $\partial/\partial\vec{z}$ rather than $\vec{x}$ and $\mathrm{d}/\mathrm{d}\vec{x}$, another fact to keep in mind for later. Anticipating the next step, we can write a formal expression for ψ, given its value at some time $t = 0$, at any later time $t = T$ as

$$\psi_T = e^{-\int_0^T \mathcal{L}\mathrm{d}t}\psi_0. \tag{5}$$

2.3 Operator algebra

The next step relies on the realization that the arguments of the Liouville operator are algebraically equivalent to the raising and lowering operators of bosonic states in quantum field theory (QFT). Thus, we can substitute our z_i's and $\partial/\partial z_i$'s for $a_i^\dagger$'s and a_i's, respectively. The symbolic substitution, of course, is trivial, but the operators bring with them all the mathematical development of QFT, including an inner product space and a complete set of coherent states. Summing over coherent states yields the identity; by taking the quadrature of (5) and inserting that expression for the identity between each time-slice, we can convert it into a generally-soluble field integral.

2.4 Classical mechanics

The evaluation of that integral, after a change of variables, yields an expression for ψ in terms of $\vec{x}$

and another set of variables, $\vec{\eta}$. Formally, these turn out to be conjugate momenta to the $\vec{x}$ in the Hamiltonian sense, with $\mathcal{L}$ acting as their Hamiltonian. Other terms familiar from Hamiltonian mechanics also show up: for example, the action is defined in classical mechanics in terms of the Hamiltonian $\mathcal{H}$ and the canonical coordinates and momenta $\vec{q}$ and $\vec{p}$ as $\int \mathrm{d}t[\vec{p}\delta_t\vec{q} - \mathcal{H}]$. In the stochastic-process formalism, we encounter an *effective action*, which shows up in the functionally identical form

$$S_{\text{eff}} \equiv \int \mathrm{d}t[-\vec{\eta}\delta_t\vec{x} + \mathcal{L}]. \tag{6}$$

We can take this further and write down "stochastic effective" versions of almost any quantity we are interested in from classical mechanics [**?**]; for an example, see the discussion of effective potentials in [2]. This formalism is known as the Friedlin-Wentzell construction. It is noteworthy that the Liouville operator, as we noted earlier, is related to the transfer matrix, but in these variables it is almost identical: generally, it takes exactly the same form but with a minus sign in front and all $\mathrm{d}/\mathrm{d}n_i$ replaced by $-\eta_i$. Thus, after defining a system, one can skip directly to the Friedlin-Wentzell representation of the stochastic process, writing down the Liouville operator by inspection from the transfer matrix.

2.5 Probabilities and the entropy

The evaluation of the field integral discussed above yields

$$\psi = \int \mathrm{d}\vec{x} e^{\vec{\eta} \cdot \vec{x} - S_{\text{eff}}}, \tag{7}$$

with the effective action as defined above. It will be instructive to compare this expression for ψ to our original definition, keeping in mind the correspondence between the z_i and exponentials of chemical potential shifts. The sum and integral that begin the two expressions are taken over the same argument; the $\vec{\eta}$, although I will not show it here, are similar to chemical potential shifts, and thus the remaining terms, ρ and $e^{-S_{\text{eff}}}$, must be equivalent. Remembering our definition of entropy as a log-probability, we realize that the effective action is the entropy that we set out after. With this result in hand, we turn to our model.

3 The model

3.1 Motivation

In order to motivate the selection of this model, I will quote some results of its solution. They are easily reproduced by investigating the mean-field equations of motion, which will be described below. The most interesting feature of the model is contained in the title of [1], the relevant part of which I will quote: "Extracting chemical energy by growing disorder". The phenomenology referred to here is that when the monomers' binding energy is below a certain positive maximum value, a sufficient asymmetry in the rate constants that define the rates of attachment and detachment for each type of monomer will induce growth of the polymer at concentrations that would favor its decay if only one monomer were present. An intuitive explanation for this is that most of the events involve attachments or detachments of the faster monomer, but every so often a slow one attaches. This tends to "lock in" the fast monomers behind it, the number of which is typically greater than zero, because only the last monomer in the chain can detach and the fast monomer tends to quickly move in and occupy that position. The thermodynamic justification is that a positive energy change is more than offset by a positive entropy change, keeping the change in *free* energy, the minimized quantity for the ensemble, below zero. Additionally, this phenomenon is one that can occur arbitrarily far from equilibrium, and it is this combination of independence of equilibrium with nontrivial dependence on entropy that makes this model an attractive candidate, and a quantitative demonstration of the above-mentioned entropy growth is one of our principal goals.

3.2 Definition

We will describe a simple model of polymer growth lifted from an article by Esposito et al. [1], in which a kinetic analysis primarily focused on the maximum extractable power was done. The basic units of the model are two types of monomers, which we will label m_i, $i \in a, b$. The defining features of the dynamics are that detachment and attachment events can only occur at one end of the polymer chain, and that their rates are dependent only on the identity of the attaching or detaching monomer. That is, given a polymer $\pi_{\vec{v}_k}$ where $\vec{v}_k \equiv (i_k i_{k-1} \ldots i_1)$ is a sequence whose elements are the ordered indices of the monomers in the polymer, the allowed processes are given by the chemical equation

$$m_i + \pi_{\vec{v}_k} \rightleftharpoons \pi_{i\vec{v}_k}, \tag{8}$$
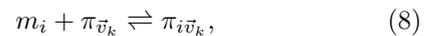
with the rate of the process k_i only a function of i. It is assumed that attachment rates are linear in the concentration of the appropriate monomer in solution, and that the detachment rate is a constant function of the end monomer. We will make some further assumptions for the remainder of the analysis in order to simplify the algebra. They are that the binding energy for both monomers is zero, and that equilibrium for a system that contains only one

type of monomer occurs at unit concentration of that monomer in solution.

As suggested by (8), the number of possible transitions, given an arbitrary number of monomers to work with, is infinite. This is because transitions must be accounted for between polymers with any arbitrary sequence of any finite length. Due to the impossibility of considering all such transitions, we are forced to make an approximation. Before doing that, we increase the number of polymers from 1 to some $M > 0$, and denote by p_i the number of polymers terminated by a monomer of type i. The approximation we then make is that rather than keeping track of the entire sequence of each polymer, we store only these values, and take the probability of "uncovering" a monomer of type i after a detachment event as proportional to p_i. It is trivially shown that $p_a + p_b \equiv M$, so we may eliminate the dependent variable by defining $p \equiv (p_a - p_b)/2$. However, for convenience, we will keep the formally independent notation for now. We complete the mathematical specification by defining the remaining state variables, the number of monomers in solution, as n_i for $i \in a, b$.

3.3 The master equation

Let $\vec{x} \equiv (n_a, n_b, p_a, p_b)$ define the set of state variables. We also introduce a solution volume, V. On-rates are proportional to the product of the concentration n_i/V with the fraction of polymers of the appropriate type p_j/M, and off-rates are proportional only to p_j/M. Then the master equation, under the approximation outlined above, is

$$\frac{\partial \rho_{\vec{n}}}{\partial t} = \sum_{i,j=a,b} k_i [(e^{-\partial/\partial n_i} - 1)\frac{n_i p_i}{V} +$$

$$(e^{-\partial/\partial n_i + \partial/\partial p_i} - 1)\frac{n_i p_j}{V} + (e^{\partial/\partial n_i} - 1)\frac{p_i^2}{M} +$$

$$(e^{\partial/\partial n_i - \partial/\partial p_i} - 1)\frac{p_i p_j}{M}]\rho_{\vec{n}}. \quad (9)$$

.

3.4 The mean-field approximation

In this section, we write down the mean-field equations, ignoring all fluctuations in the state variables beyond their average behavior. They can be derived from the master equation, or simply written down by inspection from the definition of the model. We take the limits as V and M become large, and rewrite the state variables as real-valued "concentrations". Let

$$u_i = \frac{p_i}{M} \qquad \nu_i = \frac{n_i}{V}. \quad (10)$$

We will denote by $\bar{\nu}_i$ and $\bar{u}_i$ the average of ν_i and u_i, respectively. The mean-field equations of motion,

following usual methods of construction, then become

$$\frac{d\bar{\nu}_i}{dt} = \frac{k_i M}{V}(\bar{u}_i - \bar{\nu}_i) \quad (11)$$

$$\frac{d\bar{u}_i}{dt} = k_i[\bar{\nu}_i \bar{u}_j - \bar{u}_i \bar{u}_j] + k_j[-\bar{\nu}_j \bar{u}_i + \bar{u}_j \bar{u}_i]. \quad (12)$$

It is worth noting that there are actually only 3 independent equations here, because $u_a + u_b \equiv 1$ and thus $d\bar{u}_a/dt \equiv -d\bar{u}_b/dt$.

We should check to confirm that the results here are consistent with a simple kinetic analysis of the problem, and indeed, the predictions agree. One example we can quickly check is the growth velocity per polymer at unit concentrations, $v = -V/M(d\bar{\nu}_a/dt + d\bar{\nu}_b/dt$. We find the self-consistent instantaneous value of u_a by setting $d\bar{u}_a/dt = 0$ with $n_a = n_b = 1$ and plugging it into Eqs. (4); the result is $v = \sqrt{k_a k_b}$, as expected from the kinetic analysis.

4 Stochastic formulation

4.1 Approximation methods

From the master equation for the model system, we can immediately write down the Liouville operator and therefore an integral expression for the effective action. However, it is not immediately useful, as the interpretation of the $\vec{\eta}$ variables is not clear, and the integral is not soluble in that form, other than numerically. To solve these problems, we write $\mathcal{L}$ as a sum of symmetric and antisymmetric functions of $\vec{\eta}$, in the form $\mathcal{L} = \sum_i f_i(1 - \cosh\eta_i) - g_i\sinh\eta_i$. The f_i and g_i are functions only of $\vec{x}$. In the current case, the summation index will run over linear combinations of the components of $\vec{\eta}$. We then expand the Liouville operator to second order in $\vec{\eta}$, and collapse the f_i into a square matrix $\mathbf{f}$ and the g_i into a column vector $\vec{g}$, multiplied by the appropriate number of powers of $\vec{\eta}$. The resulting formula for the effective action is the time integral of a quadratic expression in $\vec{\eta}$, of the form

$$S_{\text{eff}} = \int dt[-\frac{1}{2}\vec{\eta}\mathbf{f}\vec{\eta} - (\delta_t \vec{x} + \vec{g})\vec{\eta}]. \quad (13)$$

This is profitable because we can then complete the square in the above equation. The result of that, apart from a boundary term that turns out to exactly cancel, is

$$S_{\text{eff}} = -\frac{1}{2}\int dt[\mathbf{f}^{-1}(\delta_t \vec{x} - \vec{g})]^T \mathbf{f}[\mathbf{f}^{-1}(\delta_t \vec{x} - \vec{g})]. \quad (14)$$

The second approximation that we will make is that given a stable state, whether it is equilibrium or a non-equilibrium steady state, the most-likely trajectory to arrive at some nearby fluctuation state is

the time-reverse of the classical diffusion path. This result is generally true in one dimension, but in three it is not. For the sake of convenience, we make this approximation for now, as it greatly simplifies the expression for the effective action. We have not done so, but it is a fairly straightforward task to numerically solve for the most-likely trajectories, and in the future this will provide a test for the amount of error that the approximation introduces. We expect that it will prove quite accurate, as the dynamical system is fairly simple, lacking cycling behavior or other features that often significantly affect escape trajectories. With this approximation, because we know that the effective action for the classical diffusion path must be zero, $g \equiv -\delta_t \vec{x}$ on that path. The escape trajectory is then just the opposite, so the coefficient of the linear term in $\vec{\eta}$ collapses to a single quantity.

We can now collapse the first $\delta_t \vec{x} - \vec{g}$ term into simply $2\delta_t \vec{x}$, and the second into $2\vec{g}$. The result, noting that $dt\delta_t \vec{x} = d\vec{x}$, is that

$$S_{\text{eff}} = 2 \int_{\vec{\bar{x}}}^{\vec{x}} d\vec{x'}(\mathbf{f}^{-1}\vec{g}). \qquad (15)$$

The limits of integration, in practice, will be taken from some known steady state to a fluctuation state, and the evaluation of the integral will yield, within the approximations just given, the log-probability of the fluctuation state at some late time in the ensemble where the steady state is most probable. Thus we can use this result to get at the structure of fluctuations in the system around any steady state.

4.2 Limitations

Unfortunately, there are still several problems with this approach, the most troubling of which is that the algebra simply does not reduce enough. At equilibrium, the expression that results from the evaluation of the above integral is almost succinct enough to attempt to derive some sort of meaning, but this is not very useful for two reasons. The first is the obvious fact that we set out to derive an expression for the nonequilibrum entropy, so an expression for the equilibrium version is hardly a success at all. However, such an expression could allow us to make educated guesses at generalizations, so we should not discard it entirely.

The more serious problem is that a linearization of the mean-field equations of motion about equilibrum reveal a zero eigenvalue. Its associated eigenvector is a linear shift in u; further analysis of the full problem reveals that the true associated symmetry is a combination of the shift in u with a spatial variation of polymer composition. These variations can be decomposed into an infinite series of allowed sine and cosine fluctuations, which means that the polymer composition can take almost any function of position. The problem that this creates is simpler than that, though: the zero eigenvalue means that mean-regression is nonexistent to first order along the direction of that eigenvector, which makes the time-reverse assumption outlined above seriously flawed. To solve this problem, and also because the resulting states are interesting in their own right, we introduce constant currents j_a and j_b to the system, which add monomers of their respective types to the solution at a constant rate. We will then seek to determine the fluctuation structure of the nonequilibrium steady state that results.

Relaxing the assumptions of equilibrium in the expression for the effective action, however, introduces new issues. The expression that was already difficult to interpret at equilibrium becomes a real "spaghetti bowl", with no conceptual understanding forthcoming. This issue remains in symmetrized variables, where n_a and n_b are replaced by their independent sum and difference, and becomes significantly worse when the coordinates are first diagonalized with respect to the principal axes of the classical trajectory which we assume the escape takes. We have thus arrived at a small impasse, as this is essentially the simplest description of the entropy that we can reasonably assume in the current framework, and it appears to be impenetrable.

4.3 Future directions

However, there are several directions in which we can take this that may solve this problem. The simplest is that, as mentioned in the case of equilibrium, fairly specific expressions that are taken in certain limits or to low orders may prove helpful. If they can be interpreted conceptually in the context of entropic forms we are already familiar with, we can make educated guesses at generalizations and test them. If this doesn't work, we can make a further simplification in the master equation, decoupling the transitions in the n_i and u but leaving the rates the same. This would bring the model further away from the original physical problem of interest, but the mean-field equations would remain unchanged and this would remove some terms from the off-diagonal elements in $\mathbf{f}$. Lastly, the equilibrium system's phenomenology has been written down mathematically but not investigated thoroughly. A deeper understanding of that regime is interesting in its own right, because the property of a driving force having a stabilizing effect is shared by many types of systems, and could potentially suggest productive pathways to move forward in the analysis of the nonequilibrium case.

5   Conclusion

We sought to generalize the notion of entropy to the domain of nonequilibrium systems. By formulating the dynamics of the Esposito model in terms of a stochastic process, and converting it to a Hamiltonian-like system, we were able to derive an integral expression for the quantity we sought. Unfortunately, the algebra has so far proved so complex as to be impenetrable. However, we are optimistic that future efforts will result in an expression for the entropy in terms of the state variables and their conjugates discussed here, which will reduce to the appropriate equilibrium form and have predictive power over the structure of fluctuations in nonequilibrium steady states.

References

[1] M. Esposito, K. Lindenberg, C. Van den Broeck. Extracting chemical energy by growing disorder: efficiency at maximum power. *J. Stat. Mech.* (2010) P01008

[2] D. E. Smith. Large-deviation principles, stochastic effective actions, path entropies, and the structure and meaning of thermodynamic descriptions. *arXiv:1102.3938v1*

[3] Richard S. Ellis. *Entropy, Large Deviations, and Statistical Mechanics.* Springer-Verlag, New York, 1985.

[4] M. I. Freidlin and A. D. Wentzell. *Random perturbations in dynamical systems.* Springer, New York, second edition, 1998.

[5] M. Doi. Second quantization representation for classical many-particle system. *J. Phys. A*, 9:1465–1478, 1976.

[6] M. Doi. Stochastic theory of diffusion-controlled reaction. *J. Phys. A*, 9:1479–, 1976.

[7] L. Peliti. Path-integral approach to birth-death processes on a lattice. *J. Physique*, 46:1469, 1985.

[8] L. Peliti. Renormalization of fluctuation effects in $a + a \rightarrow a$ reaction. *J. Phys. A*, 19:L365, 1986.