

# The thermodynamic dual structure of linear-dissipative driven systems

Eric Smith

*Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501*

(Dated: June 3, 2005)

It is often assumed that maximization of the statistical entropy cannot account for the spontaneous emergence of currents or compositional heterogeneity associated with them, and that other principles involving entropy production (its maximization or minimization) must be invoked to explain the emergence and robustness of the order in driven dissipative systems. I show for a class of simple models with driving and linear dissipation that the assumption is invalid. For dynamical ensembles, the exact entropy generally becomes a function of currents as well as the familiar equilibrium state variables, and from this richer dependence a full thermodynamic dual structure can be constructed, which predicts both emergence and robustness of nonequilibrium order. In systems of this type, the need for entropy production principles arises only if the exact entropy is replaced with a coarse-grained entropy function of the equilibrium state variables alone. I briefly consider some simple applications to thermal “ratchets”, in which induced cyclic currents display the essential elements of Onsager cycling in chemical reaction graphs.

## I. INTRODUCTION

### A. Statistical principles for emergence of nonequilibrium order

Most current research in basic thermodynamics concerns systems driven away from the classical equilibrium states of Boltzmann and Gibbs. Their macrostates or ensemble descriptions are distinguished from those of equilibria by one or more of the following properties: 1). breaking of time-reversal or time-translation symmetry [1–3] by the macroscopic dynamics of the system; 2). microscopic ergodicity broken by kinetic factors, which cause fine structure of initial conditions or relaxation history to leave persistent imprints [4]; 3). heterogeneous intensive boundary conditions (such as spatially variable externally imposed temperatures) that overconstrain the set of extensive equilibrium state variables [3], and are thus inconsistent with an equilibrium Legendre dual structure. Despite these differences, nonequilibrium steady states often share with equilibria the emergence of unique values for certain order parameters, and robustness of those parameter values under disturbance of the system [5]. An old, central question is thus whether general principles can be found to account for the emergence and stability of nonequilibrium order [6].

The principle of entropy maximization, subject to constraint on ensemble averages of one or more dynamically conserved quantities [7], has essentially explained the uniqueness and robustness of equilibrium macrostates (though the difficulty of computing entropies makes this explanation non-constructive in many practical cases, and leaves significant technical challenges to the study of phase transitions). Searches for informational principles like entropy maximization have thus been standard in nonequilibrium statistical mechanics, though they are often confounded by the difficulty of defining a proper entropy function. An important consequence of this difficulty, especially in systems that resemble equilibria locally in space, is that most macroscopic characterizations

of nonequilibrium systems have used the state variables and entropy functions defined for equilibria [6, 8]. Since the spontaneous emergence of currents and other dynamical structures clearly does not maximize equilibrium entropy subject to any consistent set of boundary conditions, principles advanced using equilibrium-form entropy functionals have typically involved their rates of change in the system or its environment (termed “entropy production”), either absolute [9] or relative to a phenomenological “dissipation function” [10]. In certain compelling applications of nonequilibrium thermodynamics such as biochemistry, it has become a standard assumption that “new physics” is needed to explain both their far-from-equilibrium composition and their dynamics [11].

I will consider here only a narrow class of nonequilibrium systems: those with some kind of extended configuration space to which reservoirs with different values of the same intensive state variable can be attached at different points, generating currents through the system in response. The extent may be in physical space or in a topological space such as a graph of chemical species and reactions, and the spatially variable boundary condition may be temperature, chemical potential, or any other intensive equilibrium state variable (in the examples I will suppose it is temperature). Earlier work on both linear [3] and nonlinear [2] systems of this type, in non-dissipative limits, showed that the spontaneously generated currents (which generally break time-translation symmetry) arise within a formal structure identical to equilibrium finite-temperature statistical mechanics. Here I will introduce dissipation but restrict to the regime of linear response, and obtain broadly similar results, except that with dissipation steady states become possible that break time reversal but not time translation. Many real applications of greatest interest, from convective weather to biochemistry, are both dissipative and nonlinear, and will present further technical and conceptual challenges in combining the present and prior results.

My aim in this paper is to make three general points about entropy and dynamical order, by demonstrating them in a particular class of exactly solvable systems, chosen for simplicity and because the emergence of dissipation has been well-studied in them from first principles. Some of these points have been recognized in other contexts, but taken together they argue that the informational interpretation of thermodynamics, and ordinary entropy maximization, can account for more of the order in nonequilibrium systems than has been appreciated. The three points are:

1. The Shannon/Boltzmann definition of entropy (or its quantum-mechanical generalization) remains exactly applicable to some nonequilibrium systems, but in such cases generally produces different functional forms than it does for equilibria. The full Jaynesian interpretation of maximum entropy subject to constraints on averages can even apply, and with it the Legendre dual structure of thermodynamics, though in general that too takes a richer form than in equilibrium.
2. In particular, in dynamical or dissipative systems, current-valued as well as charge-valued observables become eligible as extensive state variables, and the exact entropy is generally a function of both (noted also in Ref. [12]). By “current-valued” I mean observables that change sign under time-reversal, as opposed to “charge-valued” observables (such as total energy) that are time-reversal symmetric. The increased complement of extensive state variables is needed by a Legendre dual structure compatible with the richer boundary conditions that drive systems away from their equilibria.
3. The spontaneous emergence and stability of currents can in some such cases be explained by entropy maximization subject to nonequilibrium boundary constraints, when the exact entropy rather than an equilibrium-form approximation is used. In such systems *new principles beyond entropy maximization are not needed*; the invocation of entropy production or phenomenological dissipation functions are consequences of using a coarse-grained entropy that is not the actual measure of statistical uncertainty in the ensemble description of the driven system.

## B. Exact versus coarse-grained entropies

The use of equilibrium state variables and functional forms for quantities like the specific entropy has become so standard that, even in nonequilibrium contexts, terms like “the entropy” and “entropy production” are used as if these quantities were the true statistical uncertainty measures of known ensembles [5], rather than approximations too coarse in some cases [3] to resolve the entropies associated with the structures of interest [21]. It

is therefore important to state clearly the relation between equilibrium-form and exact entropies, at least for systems like those considered here, where that relation is well-defined.

Equations of state based on space-local averages corresponding to equilibrium state variables, and equilibrium functions of them like specific entropy, encode approximations: 1). that the projection of the distribution of microstates onto a small number of spatially local degrees of freedom has the same relation to the local average of conserved quantities as holds in equilibrium, and 2). that correlations between local regions are ignorable for the purpose at hand [8]. A distribution in which the exact weight over each global microstate is replaced by the product of the marginal weights on local patches is one instance of a *coarse-graining* [13] of the exact distribution, and the integral of the equilibrium-form specific entropy, which is the proper uncertainty measure for the coarse-graining, is a (generally slack) upper bound on the entropy of the exact distribution (this is the origin of a careful version of Boltzmann’s H theorem).

In the examples below, the marginal distribution for each local degree of freedom will indeed appear exactly thermal, though the implied temperatures will vary in space and generally in time. However, the correlations among regions responsible for macroscopic currents are not ignorable, in the sense that the entropy associated with them is responsible for stability of those currents. The equilibrium-form entropy is obtained by setting the current-valued state variables (a kind of off-diagonal order parameter) to zero and keeping only the charges, leaving the Jaynesian specification of a maximum-entropy distribution that is another (possibly different) coarse-graining of the exact distribution.

The use of coarse-grained rather than exact entropy may not represent a large relative error. However, that error is variable in time, depending on details of the dynamics in relation to the coordinates chosen for projection onto marginals, and can in general be larger than the entropy changes associated with the nonequilibrium order parameter of interest [3]. It is therefore not surprising that maximization of the exact entropy may explain the form of induced currents, while quite different-looking principles are required for application to the coarse-grained entropy, which may depend on phenomenological coefficients like conductivities, and which generalize poorly across different systems. Practically, this may be the best one can do in many cases, and particularly in the chemistry of solutions, which is deeply decoherent, “the entropy” has always referred only to the coarse-grained value. Whether it is possible to identify corrections to the true entropy of chemistry from flows, along the lines of the examples in this paper, is an important question for future work.

### C. Legendre dual structure, currents, and cycling

If principles applicable to the course-grained entropy such as maximum entropy production [10] or minimum entropy production [9] generate approximately correct answers case by case, and the coarse-grained entropy is easier to compute in a wider variety of cases, why bother with the exact entropy ever? I argue here that the interpretive structure of Legendre duality and the equation of state are fundamental enough to our understanding of thermodynamic stability and universality, that their extension to nonequilibrium settings and current-valued state variables can lead to better notions of *dynamical ground states*, self-organization, and universality.

The equilibrium equation of state relates intensive boundary conditions to extensive charges induced in the system. How many such dual pairs exist is determined by the dimensional content of the theory, which may depend on the timescales of interest. Energy/temperature and volume/pressure are universal state variable pairs for systems with temporal and spatial extent. Atomic species numbers, and their dual chemical potentials, may be valid pairs for macroscopic chemical systems over very long times, defining a fundamental dimensional analysis of chemistry. On shorter times, molecular species numbers and chemical potentials, of which there are far more, may carry independent dimensions. On even shorter time scales and coarser levels of abstraction, functional groups or bond types may define state variable pairs. Ordinarily we rely on dimensional analysis to count the number of constraints that can consistently be imposed independently to generate an equilibrium system, but the above points, central to understanding the order of biochemistry, illustrate that such counting can be subtle if multiple timescales are coupled.

This paper does not attempt a proper statistical mechanics of chemistry, but it will display the role of dimensions and the way a dynamical generalization of Legendre duality can address the subtleties of counting constraints in their simplest instantiation: a system with multiple internal degrees of freedom coupled to multiple thermal reservoirs. If the system degrees of freedom are independent (the reservoir coupling commutes with the system Hamiltonian), each reservoir temperature is dual to an average energy for its system degrees of freedom. The system ensemble becomes a tensor product of independent subsystem ensembles, and Legendre duality and dimension-counting are those of equilibrium in each subsystem.

In the examples below, the reservoir coupling will not commute with the Hamiltonian, and the system will mediate energy and particle flows between the reservoirs. In this case, only one extensive energy characterizes the system, and it must be Legendre-dual to some average of the reservoir temperatures. The temperature differences among reservoirs, however, continue to be independent boundary control parameters with the same dimensions; what has become of their duals as we introduced off-

diagonal couplings in the system Hamiltonian? In the dissipative examples here their duals are *energy currents*, and the dissipation rate provides the necessary time dimension to relate energy currents, rather than energy charges, to temperature-valued boundary controls.

Thermal currents become particularly interesting when the system Hamiltonian is periodic (on systems either of infinite extent or of finite extent, periodically identified). The directional symmetry breaking of thermal currents has close relations to particle flows in Brownian ratchets [14], which will be illustrated in one example below. Cyclic thermal currents capture many properties of the cycles in chemical reaction graphs first considered by Onsager [10], and widely recognized as central to the relation between energy flow and material cycling in biology and biochemistry [16]. It is in order to understand how to think about these systems that I emphasize here the role of thermal currents as state variables and order parameters of statistically defined dynamical ground states.

### D. Open or closed?

Driven systems of the type considered here are usually called “open”, to distinguish them from equilibrium systems that do not support net flows of energy or particles [5]. Because of the extensive formal equivalence I will demonstrate between particular driven systems and equilibrium ensembles, a more useful distinction between open and closed is the one that already exists within equilibrium thermodynamics.

Closed systems are those for which the extremized thermodynamic potential is the entropy [22]. Open systems are those for which any reservoirs are represented only through their intensive state variables, which act as boundary conditions for the system. The extremized potentials for open systems are various Legendre transforms of the entropy, such as the free energies. The same distinction will be salient for the systems here, and it will be of secondary importance that in these systems net currents flow through the interfaces with the reservoirs, while in equilibrium they do not.

As in the equilibrium theory of adiabatic transformations, to derive the flows generated in open systems from first principles, it will be straightforward to convert them into closed systems by including the reservoir degrees of freedom, and starting the total system away from equilibrium. While the exact system  $\times$  reservoir entropy cannot change in the resulting closed system, the coarse-grained entropy which is the sum of system and reservoir marginal entropies will increase, and will reproduce as one component the entropy for the open-system representation.

### E. Outline of the paper

The presentation is arranged as follows:

I introduce the model class of linear-dissipative oscillator systems in Sec. II, and construct formally all the objects needed to understand their evolution and to demonstrate the claims about entropy summarized here in the introduction. A little material on  $P$ -representations and path integral Green's functions is included in Sec. IID, which is not strictly needed to understand the main claims of the paper, but which makes interesting connections to field theoretic and effective-potential methods for studying nonequilibrium systems. This subsection may be skipped without loss of continuity.

The physics behind the entropy of induced currents is more intuitively grasped in the simplest non-trivial model of a two-dimensional oscillator, which is developed in Sec. III. Both open-system and exact closed-system solutions are developed, to demonstrate the emergence of dissipation in a linear system.

Sec. IV presents a somewhat richer example motivated by Brownian ratchets. This steady-state ratchet, driven by a temperature heterogeneity, exhibits directional symmetry breaking by currents in response to reflection asymmetry of the system potential, and provides a very simple example of Onsager's cycling of internal system currents in response to flow between external reservoirs.

To close the derivation Sec. V returns to general forms, to check that the results shown for the exact entropy lead to distributions that reproduce Onsager's relation for production of the coarse-grained entropy, when the exact distribution is replaced by a product of its marginals. Sec. VI then offers summary and conclusions, and a brief appendix derives the exact entropy state function and the entropy differences relevant to coarse graining, by a variant on the replica trick.

## II. LINEAR-DISSIPATIVE DRIVEN SYSTEMS

### A. Choosing a model class

I wish to consider systems in the so-called "Onsager regime" of linear current response [10] to heterogeneous boundary potentials. Examples are Ohm's law for electrical conduction (one-dimensional) and Fourier's law for heat flow (diffusion in arbitrary dimension). The simplest models capable of capturing the essential features of dissipation and linear response, while abstracting away all other details of Fermion conduction, spatial versus graph topology, etc., are systems of linear quantum harmonic oscillators, coupled linearly to reservoirs which are themselves systems of linear quantum harmonic oscillators. It is easy to give such models significant structure, they are exactly solvable, and they have been extensively studied in the quantum optics literature [15], as systems in which dissipation by quantum decoherence can be derived from first principles.

The basic structure in the quantum optics models is a Fock space of kinematically independent bosonic excitations, which in a second-quantized framework have the

interpretation of particles, or in a first-quantized framework, of energy excitations of a particle in a potential well. I will draw from either representation freely in descriptions. My interest is not in quantum optics per se, and the structures I will use as examples are different from those considered in standard models of optical cavities. Engineered quantum optical systems would, however, provide an easy testing ground for these ideas, where their range of validity is well understood.

The use of quantum models offers the advantage that currents have a definition as phase-coherent structures independent from charges. This will allow us to recognize current-valued state variables as independent entities, not defined merely through linear functions of charge differences as in the conventional Onsager relations, or through conservation laws from time derivatives of charges, which appear as operator relations in the quantum theory. Though none of the macroscopic observables is itself a quantum-coherent object, this raises a caution that these simple models could be inadequate for application to chemistry (in particular in solution), where the fundamental excitations in any similarly abstract description would be required to represent complex quasiparticles among which phase coherence could be irrelevant.

I am thus assuming that, more than just a phenomenological category, Onsager linear response defines a statistical universality class and that the algebra of discrete linear oscillators provides a simplest instance in which to study it. That linear dissipation defines a universality class within these models, independent of most details of reservoir coupling or dimension, is already established [15]. I will work only in the high-temperature regime where Bose statistics are indistinguishable from Boltzmann to whatever approximation I require, leaving unaddressed whether a map to the same limit can be made for driven Fermi liquids, mechanical Brownian ratchets, chemical transformations, or other specific applications.

### B. Fock space for extended systems

The physical state space of a *system* with  $D$  internal degrees of freedom then results from the action of  $D$  orthogonal creation and annihilation operators starting from a ground (ket) state  $|0\rangle$ . They have commutation relations  $[a^\mu, a_\nu^\dagger] = \delta_\nu^\mu$ ,  $\mu, \nu \in 1, \dots, D$ . The system has a "spatial" basis indexed  $i$ , in which reservoir coupling is diagonal, and a (generally different) basis indexed  $\mu$  in which the Hamiltonian is diagonal.

In any basis, the diagonal elements of the dyadic matrix operator  $\hat{n} \equiv a^\dagger a$  (i.e.,  $[\hat{n}_\nu^\dagger] \equiv [a_\nu^\dagger a^\nu]$ ) constitute a set of the independent number components in the Fock space. The Hamiltonian (up to constants) is written  $\hat{H} \equiv \text{Tr}[E\hat{n}]$ , with eigenvalues  $E_\mu^\nu \equiv \delta_\mu^\nu E(\mu)$ . Eigenstates

of the number operators take the form

$$|\vec{n}\rangle \equiv \prod_{\mu=1}^D \frac{(a_{\mu}^{\dagger})^{n_{\mu}}}{\sqrt{n_{\mu}!}} |0\rangle, \quad (1)$$

where  $\vec{n}$  denotes the vector of components  $[n_{\mu}]$ .

For a column vector  $\xi \equiv [\xi^{\mu}]$  of complex scalars, a general coherent state for the system is compactly written

$$|\xi\rangle \equiv e^{-(\xi^{\dagger}\xi)/2} \sum_{N=0}^{\infty} \frac{(a^{\dagger} \cdot \xi)^N}{N!} |0\rangle. \quad (2)$$

Density matrices diagonal in the (overcomplete) set of coherent states constitute a sufficient representation for general well-behaved finite-temperature systems [15], and a particularly simple subset of them describe locally-thermal dissipative dynamical systems.

The subset, which I will call *Gaussian-coherent ensembles*, are conveniently written in the Glauber-Sudarshan “ $P$ -representation” [15],

$$\rho = \int \frac{d\xi^{\dagger}d\xi}{\pi^D} P(\xi^{\dagger}, \xi) |\xi\rangle \langle \xi|. \quad (3)$$

with the kernel  $P(\xi^{\dagger}, \xi)$  taken to be the Gaussian form

$$P(\xi^{\dagger}, \xi) \equiv \text{Det}(K) e^{-\xi^{\dagger} K \xi}, \quad (4)$$

with  $K$  a  $D \times D$  Hermitian matrix with only positive eigenvalues. The marginal distributions for any single degree of freedom  $P_{\alpha}(n_{\alpha}) \equiv \text{Tr}(\rho |n_{\alpha}\rangle \langle n_{\alpha}|)$  in a Gaussian-coherent ensemble are exponential in  $n_{\alpha}$  in any basis [3]. They include the equilibrium thermal distributions ( $K_{\nu}^{\mu} \equiv \delta_{\nu}^{\mu} (e^{\beta E(\omega)} - 1)$ ), but also time-dependent distributions created by linear dissipation, which exactly preserves Gaussian-coherent form. Thus these ensembles satisfy the usual assumption [8] that any single degree of freedom is distributed according to an internal energy state variable and a Gibbs distribution equivalent

to those of an equilibrium. However, phase correlations *between* degrees of freedom responsible for time dependence will not be ignorable. Whether or not the ensemble is thermal, The matrix  $K$  is the inverse expected number operator

$$n \equiv \text{Tr}(\rho \hat{n}) = K^{-1}. \quad (5)$$

I will show below that the components of  $n$ , or equivalently of  $K$ , are the state variables of the system, and that the asymmetric off-diagonal elements measure macroscopic currents.

### C. Linear-dissipative evolution

At first I will not represent the reservoirs explicitly, but rather introduce linear dissipation in this class of models through the phenomenological equation for evolution of the classical number matrix

$$\frac{dn}{dt} = i[E, n] + r(n_R - n). \quad (6)$$

The system loses particles in proportion to the number present at rate  $r$ , and gains them from the separate components of the reservoir in proportion to the classical number matrix  $n_R$ , which for convenience I will take to be time-independent (a reservoir stability assumption). Eq. (6) arises as a universal form from decoherence of exact models [15], contributing to the view of linear response as a universality class.

Through Eq. (5), linear evolution of  $n$  implies

$$\frac{dK}{dt} = i[E, K] + rK - rKn_RK, \quad (7)$$

which may be converted to a Fokker-Planck equation on  $P$  of the form

$$\frac{\partial P}{\partial t} = \left[ rD + i \left( \xi_{\mu}^* E_{\nu}^{\mu} \frac{\partial}{\partial \xi_{\nu}^*} - \xi^{\mu} E_{\mu}^{\nu} \frac{\partial}{\partial \xi^{\nu}} \right) + \frac{r}{2} \left( \xi_{\mu}^* \frac{\partial}{\partial \xi_{\mu}^*} + \xi^{\mu} \frac{\partial}{\partial \xi^{\mu}} \right) + r(n_R)_{\mu}^{\nu} \frac{\partial}{\partial \xi_{\mu}^*} \frac{\partial}{\partial \xi^{\nu}} \right] P. \quad (8)$$

Integration by parts produces an equation for the density in which the Fokker-Planck operator acts on states,

$$\frac{\partial \rho}{\partial t} = \int \frac{d\xi^{\dagger}d\xi}{\pi^D} P(\xi^{\dagger}, \xi) \left[ i \left( \xi^{\mu} E_{\mu}^{\nu} \frac{\partial}{\partial \xi^{\nu}} - \xi_{\mu}^* E_{\nu}^{\mu} \frac{\partial}{\partial \xi_{\nu}^*} \right) - \frac{r}{2} \left( \xi_{\mu}^* \frac{\partial}{\partial \xi_{\mu}^*} + \xi^{\mu} \frac{\partial}{\partial \xi^{\mu}} \right) + r(n_R)_{\mu}^{\nu} \frac{\partial}{\partial \xi_{\mu}^*} \frac{\partial}{\partial \xi^{\nu}} \right] |\xi\rangle \langle \xi|, \quad (9)$$

and that in turn is readily converted to the operator form

$$\frac{\partial \rho}{\partial t} = i [\hat{H}, \rho]$$

$$\begin{aligned} & - \frac{r}{2} \left\{ \hat{N}, \rho \right\} + r a^{\mu} \rho a_{\mu}^{\dagger} \\ & - r \left\{ \hat{\Sigma}_R, \rho \right\} - r \text{Tr}(n_R) \rho + r(n_R)_{\mu}^{\nu} [a_{\nu}^{\dagger} \rho a^{\mu} + a^{\mu} \rho a_{\nu}^{\dagger}]. \end{aligned}$$

(10)

$\hat{N} \equiv \text{Tr}[\hat{n}] = a_\mu^\dagger a^\mu$ , is the total number operator, and  $\hat{\Sigma}_R \equiv \text{Tr}(\hat{n}n_R) = a_\nu^\dagger (n_R)_\mu^\nu a^\mu$  is a source of particles from the reservoirs. In the so-called ‘‘rotating wave approximation’’ (RWA) derivation [15] of Eq. (10) from the Born approximation on reservoirs that are also systems of oscillators,  $n_R$  arises as the expectation of the number operators for those reservoir components coupled to the system.

The first line of Eq. (10) is standard unitary evolution within the system, the second line contains all terms for particle loss to the reservoirs, and the third line contains terms for gain from the reservoirs. It is straightforward to check that each line individually is trace-preserving.

#### D. Path integral evolution of the kernel

The  $P$ -representation is so named because  $P$  resembles a probability density on classical phase space, except that

$(\xi^\dagger, \xi)$  are overcomplete and  $P$  need not be positive everywhere (though for Gaussian-coherent ensembles it is). Thus it is not surprising that dual to the operator evolution equation (10), one can construct a path integral Green’s function, in which the decoherence terms have relations to the Onsager-Machlup effective action [17], and possibly to the time-loop  $S$ -matrix as it would be inserted in a finite-temperature field theory [1, 3].

One begins with a Gaussian integral relating  $P_{t+\delta t}$  to  $P_t$  for some short time interval  $\delta t$ . Denote dummy variables of integration at  $t + \delta t$  by  $\xi_1^\dagger, \xi_1$ , and those at  $t$  by  $\xi_0^\dagger, \xi_0$ . For a general (nondegenerate) matrix  $a$ , not a function of the variables of integration, completing the square shows that

$$P_a(\xi_1^\dagger, \xi_1) \equiv \text{Det}(a^{-1}) \int \frac{d\xi_0^\dagger d\xi_0}{\pi^D} \exp \left\{ - \left( \xi_1 - e^{(iE-r/2)\delta t} \xi_0 \right)^\dagger a^{-1} \left( \xi_1 - e^{(iE-r/2)\delta t} \xi_0 \right) \right\} P_t(\xi_0^\dagger, \xi_0) \quad (11)$$

has the structure of Eq. (4), with kernel

$$K_a^{-1} = e^{(iE-r/2)\delta t} K_t^{-1} e^{-(iE-r/2)\delta t} + a. \quad (12)$$

If we take  $a = r \delta t n_R$ , we may identify  $K_a$  with  $K_{t+\delta t}$ , with inverse  $n_{t+\delta t}$ . Taking  $(n_{t+\delta t} - n_t)/\delta t \rightarrow dn/dt$  since  $n$  evolves smoothly, recovers Eq. (6). With this choice, if we also denote  $(\xi_1 - \xi_0)/\delta t \rightarrow d\xi/dt \equiv \dot{\xi}$  (as long as we remember that  $\xi$  fluctuates as a Brownian and not a smoothly differentiable field in defining the measure), and drop the subscript 0, we can write

$$\left( \xi_1 - e^{(iE-r/2)\delta t} \xi_0 \right)^\dagger \frac{(n_R)^{-1}}{r \delta t} \left( \xi_1 - e^{(iE-r/2)\delta t} \xi_0 \right) \rightarrow \frac{\delta t}{r} \left( \dot{\xi} - (iE - r/2) \xi \right)^\dagger (n_R)^{-1} \left( \dot{\xi} - (iE - r/2) \xi \right). \quad (13)$$

Define a skeletonized measure for C-number paths  $\xi_t^\dagger, \xi_t$  then as follows: For  $T \equiv M \delta t$ , denote

$$\int^{\xi_T^\dagger, \xi_T} \mathcal{D}_R \xi^\dagger \mathcal{D}_R \xi \equiv \prod_{m=0}^{M-1} \text{Det} \left( \frac{(n_R)^{-1}}{r \delta t} \right) \int \frac{d\xi_m^\dagger d\xi_m \delta t}{\pi^D}. \quad (14)$$

By iterating Eq. (11), we can write

$$P_T(\xi_T^\dagger, \xi_T) = \int^{\xi_T^\dagger, \xi_T} \mathcal{D}_R \xi^\dagger \mathcal{D}_R \xi \exp \left\{ - \frac{1}{r} \int dt \left( \dot{\xi} - (iE - r/2) \xi \right)^\dagger (n_R)^{-1} \left( \dot{\xi} - (iE - r/2) \xi \right) \right\} P_0(\xi_0^\dagger, \xi_0) \quad (15)$$

The time-dependent density follows from Eq. (15) as

$$\rho_T = \int \frac{d\xi_T^\dagger d\xi_T}{\pi^D} P_T(\xi_T^\dagger, \xi_T) |\xi_T\rangle \langle \xi_T|. \quad (16)$$

The three lines of Eq. (10) re-appear in Eq. (15) as

three distinct kinds of transformations on  $\xi$ . The imaginary rotation  $iE\xi$  is the familiar action of Hamiltonian evolution on coherent states.  $-r\xi/2$  contracts the number insertion (dyadic matrix)  $\xi\xi^\dagger$  at rate  $r\xi\xi^\dagger$  in the state relative to the kernel, while Gaussian integration over

kernel  $(rn_R)^{-1}$  introduces particles by admitting fluctuations of  $\xi\xi^\dagger$  at rate  $rn_R$ . All of these are weight-preserving (because the coherent states are normalized), so they preserve trace.

The quadratic form in  $\dot{\xi}$  and  $\xi$  in Eq. (15) resembles a quantum version of the Onsager-Machlup effective potential [17], in which  $rn_R$  is the correlation function of a Langevin field. At the same time,  $\xi$  as the eigenvector of the annihilation operator gives the path for  $\xi_t$  the interpretation of a forward time-evolved state, and the reversed path for  $\xi_t^\dagger$  the interpretation of a backward-evolving Hermitian conjugate. The resulting path integral over  $\xi_t$  and  $\xi_t^\dagger$  resembles a time-loop  $S$ -matrix for the computation of time-dependent correlations at finite temperature [3], except that the insertion  $rn_R$  creates a partial trace between the field and its conjugate within each interval of time, which is absent from coherent quantum evolution. Eq. (15) thus suggests that at least some linear dissipative systems may have a compact path integral description that systematically generalizes from the Matsubara construction for equilibrium density matrices [18].

### E. The generation of currents

The phenomenon of “spontaneous emergence of currents” occurs if the reservoir components do not couple directly to current-valued excitations in the system, but only to charges, yet the combination of system dynamics with dissipative coupling to the reservoirs leads to the presence of macroscopic currents in  $\rho$ . Formally this is achieved if  $[E, n_R] \neq 0$ , but  $n_R$  is real-valued and symmetric when written in the eigenbasis of  $E$ . The off-diagonal components of  $n_R$  measure charge asymmetries the reservoirs attempt to impose on the system, or equivalently in the quantum language, phase correlations the reservoir coupling attempts to maintain between system eigenvectors of different frequency.

The linear evolution equation (6) is readily solved over any time interval with arbitrary  $n_R$ , but it suffices here to consider constant  $n_R$  and late times, where  $n$  approaches an asymptotic value  $\bar{n}$  given by

$$\bar{n} = \int_0^\infty r dt e^{-rt} e^{iEt} n_R e^{-iEt}. \quad (17)$$

In components in the Eigenbasis of the Hamiltonian,

$$(\bar{n})_\mu^\nu = \frac{(n_R)_\mu^\nu}{1 - i(E_{(\mu)} - E_{(\nu)})/r}. \quad (18)$$

If  $n_R$  is real and symmetric,  $\bar{n}$  takes on imaginary, asymmetric components proportional to the off-diagonal charge components of  $n_R$ . The imaginary components vanish at  $r \rightarrow \infty$ , and all off-diagonal components vanish at  $r \rightarrow 0$ . Whereas real, symmetric off-diagonal components of  $\bar{n}$  represent standing-wave combinations of eigenvectors, imaginary, off-diagonal components represent traveling waves.

### F. Entropy, partition function, and free energy

The Gaussian-coherent ensembles are minimal generalizations of Gibbs equilibria. Not only are the single degree-of-freedom marginal distributions exponential, but the definition of entropy in terms of the quantum density matrix is the same as that for an equilibrium ensemble, and has an elementary evaluation in terms of the classical state variables.

App. A shows that the exact entropy has the von Neumann form

$$\begin{aligned} S(\rho) &\equiv -\text{Tr}(\rho \log \rho) \\ &= \text{Tr}[(I + n) \log(I + n) - n \log n], \end{aligned} \quad (19)$$

(in which parentheses denote density matrix trace, and square brackets the scalar matrix trace over index  $\mu$ ). Gaussian-coherent ensembles are specified among general density matrices  $\rho$  by standard Jaynesian entropy maximization [7],

$$\delta_{\rho, \lambda} \{S(\rho) - \text{Tr}[\lambda(\text{Tr}(\rho \hat{n}) - K^{-1})]\} = 0, \quad (20)$$

where  $\lambda$  is a matrix of Lagrange multipliers enforcing a constraint on the *trace* of the number operator  $\hat{n}$ . It is the role of  $K^{-1}$  as the trace constraint value that identifies its components as the extensive state variables of the system. The partition function of the distribution (3,4) is indistinguishable from that for a tensor product of thermal states (diagonalize  $K$ ), and evaluates to

$$\log Z = \text{Tr}[\log(I + n)] = S(n) - \text{Tr}\left[n \frac{\delta S(n)}{\delta n}\right]. \quad (21)$$

$-\log Z$  is the Legendre transform of the entropy to a function of the intensive state variables  $\delta S(n)/\delta n$ , though I will show below that a subset of these variables, carrying the boundary temperatures, are the actual control parameters on the system.

The entropy (19) evolves under Eq. (6) as

$$\begin{aligned} \frac{d}{dt} S(\rho) &= -r \text{Tr}\left[(n - n_R) \frac{\delta S}{\delta n}\right] = r \text{Tr}[(\rho - \rho_R) \log \rho] \\ &= r[\Delta(\rho; \rho_R) - S(\rho) + S(\rho_R)], \end{aligned} \quad (22)$$

where  $\Delta(\rho; \rho_R)$  is the Kullback-Leibler divergence [19] (also derived using App. A)

$$\begin{aligned} \Delta(\rho; \rho_R) &\equiv \text{Tr}[\rho_R(\log \rho_R - \log \rho)] \\ &= \text{Tr}\left[(I + n_R) \log(I + n)(I + n_R)^{-1} - n_R \log n n_R^{-1}\right], \end{aligned} \quad (23)$$

and  $\rho_R$  is to be understood as the Gaussian-coherent ensemble the reservoirs would impose through  $n_R$  if  $[E, n_R]$  were zero. The steady state condition  $dS/dt = 0$  is equivalent to the condition  $S(\rho) - S(\rho_R) = \Delta(\rho; \rho_R)$  that  $\rho$  be a coarse-graining of  $\rho_R$  [13].

**From Eq. (18) and Eq. (22) we obtain the first two main claims of the paper.** First, the spontaneous

generation of currents can *increase* the exact entropy of a well-defined ensemble description of the system, relative to the reference entropy associated with the boundary charge distribution  $n_R$ . Second, though the entropy of  $\bar{n}$  is greater than that of  $n_R$ , it is still less than that of an equilibrium ensemble with the same average energy for each eigenvector, as would be obtained from the diagonal components of  $\bar{n}$  (or  $n_R$ , which are the same), acting alone. This intuitive result says that the imposition of more constraints than just average energy by the boundary conditions can only decrease our uncertainty relative to an equilibrium ensemble, but that the extensive state variable on which the entropy reduction depends generally has both charge and current components.

Geometrically, we may understand how  $\rho$  can be a coarse-graining of  $\rho_R$  by observing that  $n_R = \bar{n} - i[E, \bar{n}]/r$  is in the tangent plane to the surface  $\delta S = 0$  at  $\bar{n}$ .  $n_R - \bar{n}$  is proportional to the vector  $i[E, \bar{n}]$ , tangent to  $\delta S = 0$  because unitary evolution exactly preserves  $S$ . I will illustrate this graphically in the specific coordinates of the two-dimensional oscillator example of Sec III.

### G. Dual structure of the driven system

The ensemble description in which the reservoirs are treated phenomenologically, through the extensive state variables  $n_R$  or their dual intensive variables  $\delta S(n_R)/\delta n_R$ , is an open-system description, in which the system entropy is maximized subject to a normal-derivative constraint. The thermodynamic potential whose extremum determines the stationary response should thus be the formal equivalent of a Helmholtz potential, rather than the entropy alone. To understand what happens to the dual structure when the reservoir coupling does not commute with the Hamiltonian, we return to the question of how many Lagrange multipliers are actually needed to specify a distribution  $\rho$ , from the full set introduced in Eq. (20).

For a general solution  $n_\tau$  of Eq. (6), define

$$\Lambda_\tau \equiv \frac{\delta S(n_\tau)}{\delta n_\tau} = \log(1 + n_\tau^{-1}), \quad (24)$$

the outward normal to the surface  $\delta S = 0$  at  $n_\tau$ . All moments  $\text{Tr}(\rho \hat{n})$  of the maximum-entropy distribution with intensive state  $\Lambda_\tau$  are fixed by maximizing  $S(\rho)$  subject to the single constraint  $\text{Tr}[(\text{Tr}(\rho \hat{n}) - n_\tau) \Lambda_\tau] = 0$ , because the set  $\delta S \geq 0$  from any  $n_\tau$  is convex. Entropy maximization with one Lagrange multiplier for an extensive trace constraint,

$$\delta_{\rho, \lambda} \{S(\rho) - \lambda_\tau [\text{Tr}(\text{Tr}[\hat{n} \Lambda_\tau] \rho) - \text{Tr}[n_\tau \Lambda_\tau]]\} = 0, \quad (25)$$

is thus a *sufficient* principle to extract  $\rho(n_\tau)$ , but fails to capture the physical role of the temperatures in the reservoir as independently specified intensive state variables.

In general, that role is nonlinear and difficult to extract, but it becomes simple in the Onsager regime of

high temperature and weak perturbation away from equilibrium. Suppose, for example, that the trace constraint on  $n_R$  comes from average energy and some other (non-diagonal) charge matrix  $Q$ . Suppose also a high temperature  $\bar{\beta} E_{(\mu)} \ll 1$ , and weak perturbation  $\beta' Q_{(\nu)}/\bar{\beta} E_{(\mu)} \ll 1$ , for all eigenvalues  $E_{(\mu)}$  of  $E$  and  $Q_{(\nu)}$  of  $Q$ . With these trace constraints, we may expand  $n_R$  to leading order

$$n_R = \left(e^{\bar{\beta} E + \beta' Q} - 1\right)^{-1} \approx \frac{E^{-1}}{\bar{\beta}} - \frac{E^{-1}}{\bar{\beta}} \beta' Q \frac{E^{-1}}{\bar{\beta}}. \quad (26)$$

Under Eq. (6) we may then write

$$n \approx \frac{E^{-1}}{\bar{\beta}} - \frac{E^{-1}}{\bar{\beta}} \beta' J \frac{E^{-1}}{\bar{\beta}} \approx \left(e^{\bar{\beta} E + \beta' J} - 1\right)^{-1}, \quad (27)$$

as long as

$$\frac{dJ}{dt} = i[E, J] + r(Q - J). \quad (28)$$

If we think of  $Q$  as denominated in energy units,  $1/(\bar{\beta} \mp \beta') \equiv kT_\pm$  are the two temperature parameters represented in the environment. Dual to the average temperature  $1/\bar{\beta}$  is an average energy

$$\text{Tr}[\bar{n} E] = \text{Tr}[n_R E] \equiv \bar{\mathcal{E}}, \quad (29)$$

(equal in  $\bar{n}$  and  $n_R$ ), and dual to  $\beta'/\bar{\beta}^2$  is a new trace

$$\text{Tr}[n J] \equiv \mathcal{J}. \quad (30)$$

Denoting the Hermitian operators formed from  $Q$  and  $J$   $\hat{Q} \equiv \text{Tr}[Q \hat{n}]$  and  $\hat{J} \equiv \text{Tr}[J \hat{n}]$  respectively, the steady state distribution solves the maximization problem

$$\delta_{\rho, \lambda, \lambda'} \left\{ S(\rho) - \lambda \left[ \text{Tr}(\hat{H} \rho) - \bar{\mathcal{E}} \right] - \lambda' \left[ \text{Tr}(\hat{J} \rho) - \mathcal{J} \right] \right\} = 0. \quad (31)$$

Dual to entropy maximization (31) under extensive state variable constraints is the minimization with intensive constraints,

$$\delta_\rho \left\{ \bar{\beta} \text{Tr}(\hat{H} \rho) + \beta' \text{Tr}(\hat{J} \rho) - S(\rho) \right\} = 0. \quad (32)$$

Reducing to the residual dependence on  $n$ , we recognize that this is the minimization of the log inverse partition function (21)

$$\delta_n \{-\log Z\} = \delta_n \left\{ \text{Tr}[(\bar{\beta} E + \beta' J) n] - S(n) \right\} = 0, \quad (33)$$

equivalent to minimization of the Helmholtz free energy for the equilibrium canonical ensemble, **which was the third main claim of the paper.**

The relations (31 - 33), while derived for a carefully selected model class, are macroscopic relations among the classical state variables of a system, with standard thermodynamic form. Moreover they preserve the definitions of the charge-valued state variables from equilibrium, and their relations to the local distributions of charge-valued

excitations as usually assumed in nonequilibrium thermodynamics. To these they add only independently defined current-valued state variables and a refined derivation of the entropy, both features approximated away but not ruled out by the commonly made coarse-grainings. These may thus be general relations within the Onsager linear-response universality class, or they may define a subclass within the larger set of systems considered equivalent by Onsager.

### III. THE TWO-DIMENSIONAL OSCILLATOR

#### A. The minimal nontrivial case

In a linear system all currents are ultimately generated through coupling of pairs of eigenvalues, per Eq. (18) or its time-dependent generalizations. Therefore they are most simply seen in a system with  $D = 2$ , and Hamiltonian defined from

$$E = \begin{bmatrix} E_x & \\ & E_y \end{bmatrix}, \quad (34)$$

with  $E_x \neq E_y$ . General Hermitian  $n$  take the form

$$n = \begin{bmatrix} n_0 + n_3 & n_1 + in_2 \\ n_1 - in_2 & n_0 - n_3 \end{bmatrix}, \quad (35)$$

with coefficients related to the physical state basis of Fig. 1 by  $2n_3 = n_x - n_y$ ,  $2n_1 = n_u - n_v$ ,  $2n_2 = n_+ - n_-$ . The transformation of annihilation operators from basis  $x, y$  to basis  $u, v$  is

$$\begin{bmatrix} a_u \\ a_v \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} a_x \\ a_y \end{bmatrix}. \quad (36)$$

The transformation to basis  $\pm$  is

$$\begin{bmatrix} a_+ \\ a_- \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix} \begin{bmatrix} a_x \\ a_y \end{bmatrix}, \quad (37)$$

and the creation operators are defined by conjugation.

If by convention  $n_{(x,y)}$  and  $n_{(u,v)}$  refer to standing waves (whose densities are charges),  $n_{\pm}$  are traveling waves (whose densities are currents). By symmetry the entropy (19) is preserved under arbitrary similarity transformation of  $n$  by elements of  $SU(2)$ , illustrating most clearly why, if the equilibrium entropy is a function of charge asymmetries, its dynamical generalization must be the properly symmetrized function of both charge and current asymmetries.

Hamiltonian evolution preserves  $n_0$  and  $n_3$ , and cycles net charge and current excesses in time as

$$(n_1 + in_2)_t = e^{i(E_x - E_y)t} (n_1 + in_2)_0. \quad (38)$$

A picture of  $\delta S = 0$  and a Hamiltonian orbit within it is given in Fig. 2, for an arbitrary fixed value of  $n_3$ .

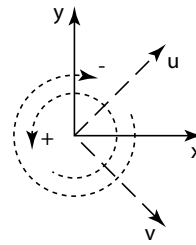


FIG. 1: Three bases for the two-dimensional oscillator.  $xy$  (solid) is the eigenbasis of the Hamiltonian.  $uv$  (dash) is the “spatial” basis in which a thermal asymmetry is imposed by  $n_R$ .  $\pm$  (dotted) is the current basis excited spontaneously, and switched under time reversal.

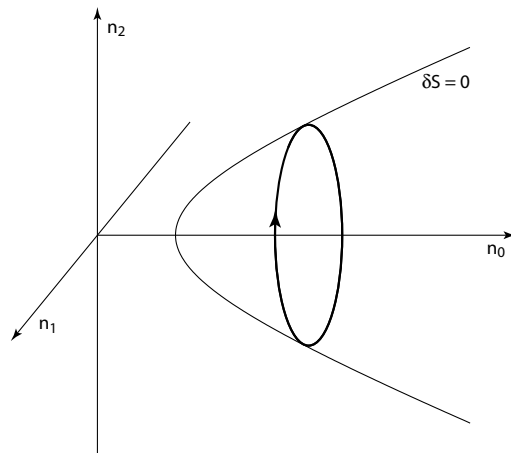


FIG. 2: A characteristic surface  $\delta S = 0$ , (the surface obtained by rotating the thin curve about its axis of symmetry) and within it a Hamiltonian orbit (heavy line; the direction of the arrow corresponds to  $E_x > E_y$ ). The fixed points of Hamiltonian evolution are the axis  $n_1, n_2 = 0$ .

The set of coarse-grainings of real  $n_R$  preserving  $n_0$  and  $n_3$  is given by  $(n_1 + in_2)_\lambda = (n_R)_1 / (1 - i\lambda)$  for real  $\lambda$ , and shown in Fig. 3. By Eq. (18) for  $\bar{n}$ , we solve for  $\lambda = (E_x - E_y)/r$ . The current magnitude  $|\bar{n}_2|$  is maximized at  $r^2 = (E_x - E_y)^2$ , reproducing a common feature of driven dissipative systems, that currents are maximized when the characteristic timescales of coherent transport and dissipation are comparable.

#### B. Exact solutions

As with the adiabatic transformations considered in equilibrium thermodynamics, it is possible to replace the classical boundary conditions and open system with an explicit treatment of the system and reservoirs as a larger closed system, with non-equilibrium initial conditions and finite-rate relaxation. For  $D = 2$  the exact solution is simple enough that, rather than use a Born approximation, we may generate dissipation by first evolving the

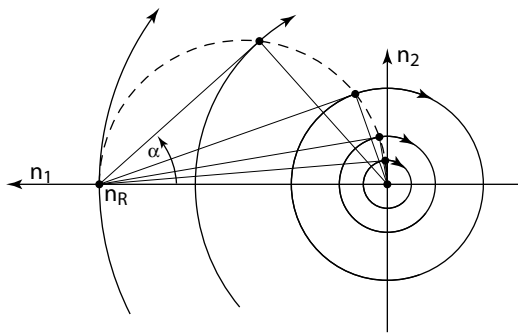


FIG. 3: End view corresponding to Fig. 2, showing coarse-grainings from charges  $n_1$  to linear combinations  $n_1 + in_2$  of charges and currents. Concentric circles are Hamiltonian orbits, and dashed circle is the set of coarse grainings from given  $(n_R)_1$ . The angle  $\alpha$  obeys  $\tan \alpha = 1/\lambda = r/(E_x - E_y)$ .

whole system and then tracing over all of the reservoir degrees of freedom.

Denote a Hamiltonian for the (system  $\times$  reservoir) degrees of freedom

$$\hat{\Xi} \equiv \begin{bmatrix} \hat{H} & \hat{g} \\ \hat{g}^\dagger & \hat{H}_R \end{bmatrix}, \quad (39)$$

where  $\hat{H}_R$  is the (Hermitian) Hamiltonian over the reservoir degrees of freedom (block diagonal if the reservoir has separated components), and  $\hat{g}$  is the vector of system/reservoir exchange terms. The Hamiltonian (39) comes from the coupling matrix

$$\Xi \equiv \begin{bmatrix} E & g \\ g^\dagger & E_R \end{bmatrix}. \quad (40)$$

Similarly, decompose the Hermitian matrix  $\mathcal{K}$  for an arbitrary Gaussian-coherent ensemble for the system  $\times$  reservoir as

$$\mathcal{K} \equiv \begin{bmatrix} K & b \\ b^\dagger & C \end{bmatrix}. \quad (41)$$

For the dynamics here we imagine  $C$  block-diagonal, with the two blocks initially representing thermal distributions at different temperatures.  $\mathcal{K}$  evolves under

$$\frac{d\mathcal{K}}{dt} = i[\Xi, \mathcal{K}]. \quad (42)$$

If we trace over reservoir degrees of freedom, the effective matrix for the system is not simply  $K$ , but the result of completing the square in the integrated reservoir coherent-state variables:

$$K_{\text{eff}} = K - bC^{-1}b^\dagger, \quad (43)$$

at each instant of time.

Fig. 4 shows the topology of a sample system, in which the  $u, v$  system oscillators are coupled randomly to two

reservoir blocks, each of which has random internal couplings to produce a dense band of states of average energy  $(E_x + E_y)/2$ . Fig 5 shows the three independent components of the system  $n$  inverse to  $K_{\text{eff}}$  from Eq. (43), and the average particle number in each reservoir (total number divided by number of reservoir sites), averaged over 50 random instantiations of the couplings.

The system rapidly converges on a fixed ratio of  $n_+ - n_-$  to  $n_u - n_v$ , and of both relative to  $(n_R)_u - (n_R)_v$ , which shows the slow decay of the reservoir temperature difference due to particle exchange. On longer timescales, the numbers  $n_x$  and  $n_y$ , initially set equal, approach their unequal thermal values at the mean temperature, satisfying  $E_x n_x = E_y n_y$ .

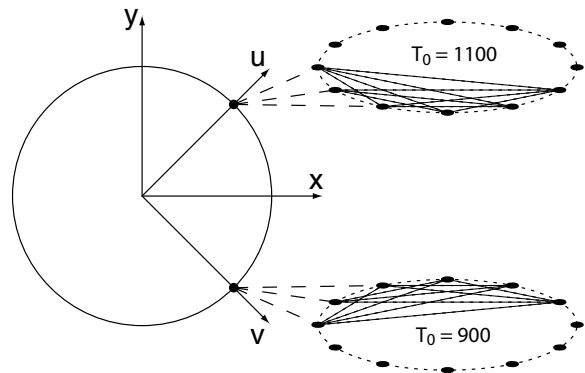


FIG. 4: A  $D = 2$  system coupled randomly to two reservoir blocks. Dashed lines represent random couplings  $g$  between modes  $u$  or  $v$  and all sites in their respective reservoirs (only a subset shown). Dotted lines represent random couplings off-diagonal in  $E_R$  spanning all pairs of sites in each reservoir block (subset shown). In the numerical example each reservoir has 36 sites, and both the cross-site couplings in  $E_R$  and the system/reservoir exchange  $g$  are drawn from uniform distributions in magnitude and phase, over a range adjusted to “impedance match” the system to the reservoirs for rapid phase decoherence and reservoir equalization. Reservoir blocks are initially given equilibrium thermal distributions at the temperatures  $T_0$  indicated, in units where  $(E_x + E_y)/2 = 1$ .

## IV. A SPATIALLY EXTENDED RATCHET

### A. Driving with temperature rather than mechanics

The Brownian ratchet, originally conceived as a model of biological symmetry breaking in systems like muscle fibers [14], converts mechanical energy to an asymmetric particle current. An asymmetric sawtooth potential for a gas of particles is turned on and off, and the particles congregate asymmetrically in wells when the potential is on, and diffuse symmetrically when it is off, generating net current flow down the shallower gradient of the sawtooth.

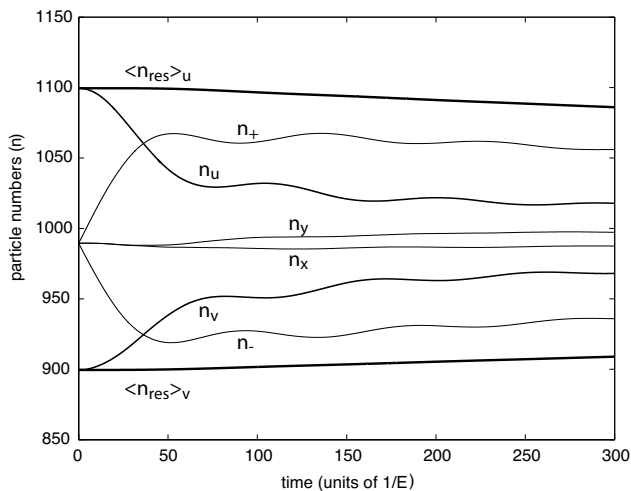


FIG. 5: Particle numbers in the three bases of the two-dimensional oscillator, and in the two reservoir components.  $(E_x + E_y)/2 = 1$ , and  $E_x - E_y = 0.025$ , so that the period of free oscillation in the system is  $2\pi/0.025 \approx 251$ . Curves represent an average over 50 random instantiations of reservoir-internal and system-reservoir couplings.

A related notion of ratchet may be used to drive an asymmetric and even cyclic current in response to a steady-state but spatially inhomogeneous temperature field. The ratchet consists of  $D$  sites uniformly distributed in a spatial interval  $x \in [-\pi, \pi]$ , which may be periodically identified or a single period in an extended space. The system Hamiltonian, given by

$$\hat{H} \equiv \sum_{i=1}^D \left( 1 - \frac{V}{2} x_i \right) a_i^\dagger a^i - \frac{h}{2} \left( a_{i+1}^\dagger a^i + a_{i-1}^\dagger a^i \right), \quad (44)$$

has a site-diagonal energy profile shown in Fig. 6, and uniform nearest-neighbor hopping of strength  $h$ .

The lattice sites are excited by an external temperature profile

$$T_i = \bar{T} + \delta T \sin x_i. \quad (45)$$

The parameters  $\bar{T}$  and  $\delta T$  are the intensive state variables used to control the system. Roughly, the high-temperature region excites symmetric particle hopping. Particles that hop to the nearby energy peak are quenched, and diffuse preferentially down the gradient to make one cycle of the potential well, generating a net positive-sense current. Though the source of energy in the thermal and mechanical ratchets is different, the combined elevation of a subset of particles from the well to the peak, and their subsequent adiabatic drift under dissipation, are the common elements responsible for directional symmetry breaking.

In this ratchet, currents are generated by the traveling waves whose creation and annihilation operators are the Fourier transforms of the site-diagonal operators. We can resolve the net particle current into components, each proportional to the mean excitation number

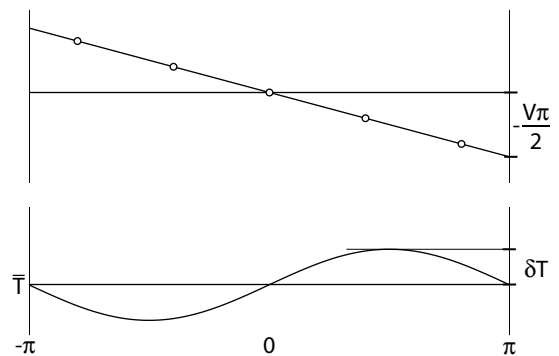


FIG. 6: Site-diagonal potential (top) and temperature profile (bottom) of a thermal ratchet. Abscissa is the space coordinate;  $D$  (here 5) sites are the little circles, drawn at position  $x_i$  and height representing the energy  $(1 - Vx_i/2)$ . The two intensive boundary constraints on this quasithermal ensemble are  $\bar{T}$  and  $\delta T$ . Temperatures, like the potential, are only evaluated at the  $x_i$ .

of its associated Fourier mode. Excitation numbers are graphed versus wavenumber index and dissipation rate  $r$  in Fig. 7. The profile of excitation is roughly sinusoidal in wavenumber, and the current asymmetry in all modes is maximized when  $r$  is comparable to the difference in the eigenvalues of  $\hat{H}$ .

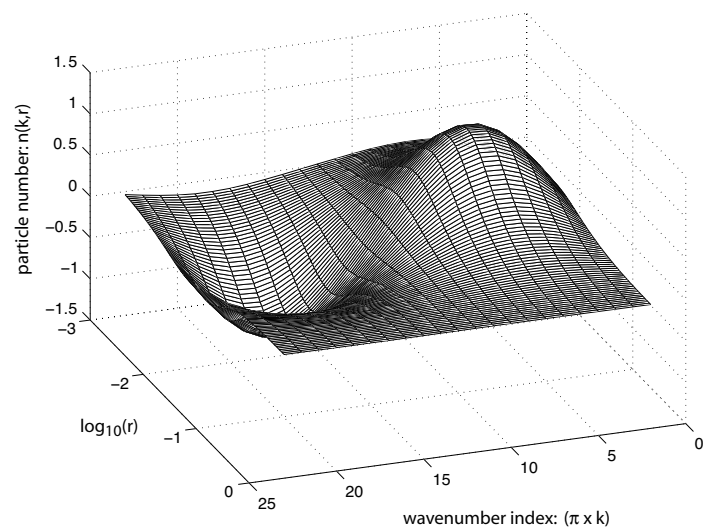


FIG. 7: The diagonal of  $\bar{n}$  in a wavenumber basis, versus dissipation rate  $r$ . Typical eigenvalue differences  $(E_{(\mu)} - E_{(\nu)}) \sim 0.1$  for these parameters, and maximal currents occur around  $r \sim 10^{-1.7}$ . The profile in wavenumber is roughly sinusoidal.

Though it is unnecessary for a linear system, it is relevant to more complex cases to note that the steady-state currents in a system of this type can be estimated perturbatively. The integral solution (17) for  $\bar{n}$  has a geometric

power-series expansion in commutators

$$\bar{n} \approx n_R + \left[ \frac{iE}{r}, n_R \right] + \left[ \frac{iE}{r}, \left[ \frac{iE}{r}, n_R \right] \right] + \dots, \quad (46)$$

which converges for sufficiently large  $r$ . No current appears before the third order, corresponding to a vertex in which the mode-one temperature profile and mode-one component of the potential couple to the mode-two component of the potential. (Thus wavenumbers sum to zero in the three-point vertex, leaving a mode-zero net current.) The leading Fourier expansion of the potential

$$\frac{V}{2} x_i \approx V \left( \sin x_i - \frac{1}{2} \sin 2x_i \right) \quad (47)$$

therefore produces the leading term in the perturbative solution for the currents.

Nothing about the thermal ratchet model implies mathematically that the sites are spatial positions. It therefore serves also as a poor-man's model of Onsager cycling in a chemical reaction graph [10]. In this interpretation, sites are chemical species that may contain some functional group or atomic cluster of interest, and particles are the groups or clusters that move among sites through chemical reactions. Thermal excitation is a proxy (often poor) for activation of functional groups that enables them to participate in reactions, though it might be a more appropriate model for light-induced charge separation, as occurs in recyclable photosynthetic pigments [20]. More realistic models of chemical cycling can readily be built field-theoretically, but their solution and the computation of appropriate entropies is no longer simple.

## V. THE PHENOMENOLOGICAL PRINCIPLES FOR COARSE-GRAINED ENTROPY

To emphasize that these results for the exact entropy in no way contradict results of phenomenological theories like Onsager's maximum (coarse-grained) entropy production, we may check that the latter is recovered if we coarse-grain the exact distribution for the Gaussian-coherent ensemble. Because the operation of projecting the exact distribution onto the product of its marginals causes system and reservoir components to appear to each other only through their intensive state variables, the Onsager relation in fact follows from minimization of the generalized exact free energies (33) in the linear-current regime.

Full coarse-graining in these models would be achieved by projecting  $\rho$  onto its diagonal components  $\tilde{\rho}$  in a Fock space over independent number excitations. The resulting coarse-grained entropy would then be function of the charge-valued state variables only. As we have only introduced notation and derived linear particle exchange for factoring the (system  $\otimes$  reservoir) distribution into independent marginals, that partition will be used as

an example. It simplifies the presentation to consider the reservoir a unified system in its own right (rather than a set of independent components), with state variable  $n_R$ , and distribution  $\rho_R$  extremized as in Eq. (33), except with trace constraint  $\bar{\beta}E + \beta'Q$ , for consistency with Eq. (26). Label the reservoir partition function  $Z_R$ , and its entropy  $S_R$  computed from  $\rho_R$  as for  $S(\rho)$ . Let overdot denote time differentiation. Then consider the potential

$$\begin{aligned} \frac{d}{dt} \log(ZZ_R) &= \dot{S} - \text{Tr} \left[ (\bar{\beta}E + \beta'J) \dot{n} - \beta' \dot{J} n \right] \\ &+ \dot{S}_R - \text{Tr} \left[ (\bar{\beta}E + \beta'Q) \dot{n}_R - \beta' \dot{Q} n_R \right]. \end{aligned} \quad (48)$$

Use Eq. (28) to evaluate  $\dot{J}$ , and by symmetric treatment of system and reservoir let  $\dot{Q} = r(J - Q)$  (taking the reservoir energy diagonal in  $n_R$  purely for convenience). In terms of currents defined phenomenologically from  $n$  and  $n_R$ ,

$$j \equiv \dot{n} - i[E, n], \quad (49)$$

and  $j_R \equiv \dot{n}_R$ , we may then write Eq. (48) as

$$\begin{aligned} \frac{d}{dt} \log(ZZ_R) &= \text{Tr} \left[ \left( \frac{\delta S}{\delta n} - \bar{\beta}E - \beta'J \right) j \right. \\ &+ \left. \left( \frac{\delta S_R}{\delta n_R} - \bar{\beta}E - \beta'Q \right) j_R - \beta'(J - Q)(n - n_R) \right]. \end{aligned} \quad (50)$$

Maximization of  $\log Z$  and  $\log Z_R$  at each time is equivalent to saddle-point extremization of Eq. (50), maximizing in  $j$  and  $j_R$ , minimizing in  $n$  and  $n_R$ .

For high temperatures and linear perturbations it is convenient to write  $n = (\bar{\beta}E)^{-1} + \delta n$ ,  $n_R = (\bar{\beta}E)^{-1} + \delta n_R$ , and Eq. (50) is readily expressed to leading order as a difference of quadratic forms in these variables. Variation with  $\delta n + \delta n_R$  then sets  $j + j_R = 0$ , and variation with  $j + j_R$  recovers the sum of forms (26,27). Variation with  $\delta n - \delta n_R$  gives  $\delta n - \delta n_R$  in terms of  $j - j_R$ , recovering the linear dissipation rule. Evaluating Eq. (50) on these three extrema leaves the function of  $(j - j_R)/2 \rightarrow j$ :

$$\begin{aligned} \frac{d}{dt} \log(ZZ_R) &= 4r \text{Tr} \left[ \beta'(Q - J)j - \frac{1}{2r} (\bar{\beta}E)j (\bar{\beta}E)j \right] \\ &+ r(2r - 1) \text{Tr} \left[ (\bar{\beta}E) \beta'(Q - J) (\bar{\beta}E) \beta'(Q - J) \right]. \end{aligned} \quad (51)$$

The linear form  $\text{Tr} [\beta'(Q - J)j]$  is Onsager's "entropy production rate", with the difference of inverse temperatures  $\beta'(Q - J)$  regarded as constants under variation, while

$$\phi(j, j) \equiv \frac{1}{2r} \text{Tr} \left[ (\bar{\beta}E)j (\bar{\beta}E)j \right] \quad (52)$$

is the phenomenological "dissipation function" [10]. Under complete coarse-graining of both system and reservoirs, the single function (51) would expand into the

structure of physical diffusion currents determined by the Hamiltonian, though the resulting coarse-grained entropy, even within the system, would progressively diverge from the exact value (19).

## VI. CONCLUDING REMARKS

This paper has demonstrated some features of entropy, and some roles for it in explaining the order in dissipative systems, which appear not to have been recognized before, due to a heavy reliance on equilibrium state variables and state relations in most descriptions of systems whose micro-statistics locally resemble those of equilibria. Quantum-mechanical examples have been ideal to illustrate these roles, because they provide principled definitions of currents and charges as independent excitations, and because the entropy of the quantum density matrix remains well-defined whether or not the system in question sustains macroscopic currents. This has enabled exact computation of the dependence of a well-defined entropy measure on currents, which is supported by physical intuition because of the intrinsic symmetry between charge and current excitations in quantum state representations.

The particular choice to study Gaussian-coherent ensembles provided us with systems whose local excitations satisfy exactly the equilibrium-like state relations usually assumed, allowing us at the same time to quantify the error made by ignoring the correlations between local regions as is done by restricting to the use of equilibrium state variables alone. We have therefore seen that these correlations, which in quantum theory are precisely the source of macroscopic currents, also carry the entropy whose maximization explains the emergence and stability of the currents, within a standard Jaynesian formulation of statistical mechanics.

The choice to consider linear quantum models has made possible exact solutions and first-principles derivations of dissipation, but it has left two important questions about the range of applicability of these results as general principles. One is whether the methods for deriving exact entropies used here, or those in previous work on nonlinear but non-dissipative systems [2] (for which entropy is implicitly given by the partition function), can be extended to nonlinear dissipative systems. These results are probably a good description, with minor corrections, for Fourier heat laws or Ohm's law. At the same time, nonlinear feedbacks, dynamical phase transitions, and far-from-equilibrium steady states (however treated) are certainly at the heart of the applications that aren't trivial to understand with phenomenological laws, and the test of these methods will ultimately be their ability

to predict the properties of such systems.

The other question is whether the notion of a current-dependent dynamical entropy remains useful for complex processes like chemical reactions in solution, where the fundamental particles may not be quantum coherent at the level of abstraction appropriate to these arguments, and the definitions of currents as independent excitations is more problematic. I am optimistic that chemistry remains within this regime, because chemistry lies between two regimes in which currents and their entropies are well defined. One is the quantum-coherent regime studied in this paper. The other is the regime of classical mechanics, where probability densities on phase space have well-defined Shannon entropies if the phase-space measure is given [12]. While the Gibbs distribution has no net currents, it is sensible classically to define distributions with other constrained boundaries, which do. Classical Hamiltonian phase space dynamics is a fully-decohered regime of long-range inertial correlation, rather than short-range quantum phase correlation. The  $P$ -representation in some sense bridges these two regimes, mapping phase correlation in quantum mechanics onto phase-space correlations in  $P(\xi^\dagger, \xi)$ . The question to answer for diffusive chemistry is whether, perhaps lacking both quantum phase coherence and long-range phase-space coherence, it still lies under this bridge between them.

## Acknowledgments

I am grateful to Insight Venture Partners for support of this work, and to Anita Goel for many helpful discussions, including the suggestion to use Brownian ratchets as systems in which to study symmetry breaking by currents.

## APPENDIX A: CALCULATION OF ENTROPIES AND KULLBACK-LEIBLER MEASURES

Analytic continuation in powers of densities, akin to replica methods [4] for handling quenched randomness, provide a simple derivation of the von Neumann formula for the entropy in Gaussian-coherent ensembles, and the equivalent form for the Kullback-Leibler divergence.

### 1. Integer-number products of densities

Begin by computing traces of powers of density matrices:

$$\text{Tr} [\rho_1 \dots \rho_m] = \prod_{i=1}^m \text{Det} (K_i) \int \frac{d\xi_i^\dagger d\xi_i}{\pi^D} \langle \xi_0 | \xi_1 \rangle \dots \langle \xi_m | \xi_0 \rangle e^{-\sum_{i=1}^m \xi_i^\dagger K_i \xi_i}. \quad (\text{A1})$$

From the commutation relations of the raising and lowering operators, it follows that

$$\langle \xi_i | \xi_j \rangle = e^{\xi_i^\dagger \xi_j - (\xi_i^\dagger \xi_i + \xi_j^\dagger \xi_j)/2}, \quad (\text{A2})$$

so that

$$\text{Tr} [\rho_1 \dots \rho_m] = \frac{\prod_{i=1}^m \text{Det} (K_i)}{\text{Det} (A)}, \quad (\text{A3})$$

where

$$A \equiv \begin{bmatrix} A_1 & -I & & & & \\ & A_2 & -I & & & \\ & & \ddots & \ddots & & \\ & & & A_{m-1} & -I & \\ -I & & & & & A_m \end{bmatrix}, \quad (\text{A4})$$

in which the diagonal blocks are  $A_i \equiv K_i + I$ , and  $I$  is the  $D \times D$  identity matrix. Decomposing any vector transformed by  $A$  into block form

$$v \equiv \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix}, \quad (\text{A5})$$

the definition of an eigenvector of  $A$  may be put into  $m$  alternate forms equivalent to

$$Av = \lambda v \Leftrightarrow (A_m - \lambda) \dots (A_1 - \lambda) v_1 = v_1 \quad (\text{A6})$$

Thus the characteristic polynomial has the equivalent representations

$$\text{Det} [A - \lambda] = \text{Det} [(A_m - \lambda) \dots (A_1 - \lambda) - I] = 0 \quad (\text{A7})$$

Since the coefficient of the highest-order term  $\lambda^{mD}$  is one in both expressions in Eq. (A7), the polynomials have the same scale, and therefore the same constant term

$$\text{Det} [A] = \text{Det} [A_m \dots A_1 - I]. \quad (\text{A8})$$

The trace (A3) is then written

$$\text{Tr} [\rho_1 \dots \rho_m] = \frac{\prod_{i=1}^m \text{Det} [K_i]}{\text{Det} [(K_m + I) \dots (K_1 + I) - I]}. \quad (\text{A9})$$

## 2. Replica calculation of entropies and Kullback-Leibler measures

Next represent the logarithm as the analytic continuation of a function of the power  $p$  to which a density is raised:

$$\log \rho = \lim_{p \rightarrow 0} \frac{1}{p} (\rho^p - 1). \quad (\text{A10})$$

Similarly, for any normalized  $\rho_i$  not necessarily equal to  $\rho_j$ , represent

$$\text{Tr} [\rho_i \log \rho_j] = \lim_{p \rightarrow 0} \frac{1}{p} (\text{Tr} [\rho_i \rho_j^p] - 1). \quad (\text{A11})$$

Using Eq. (A9), evaluate

$$\begin{aligned} \text{Tr} [\rho_i \rho_j^p] &= \frac{\text{Det} [K_i] \text{Det} [K_j]^p}{\text{Det} [(K_i + I) (K_j + I)^p - I]} \\ &= \exp \{ \text{Tr} [\log K_i + p \log K_j - \log ((K_i + I) (K_j + I)^p - I)] \} \\ &= \exp \{ \text{Tr} [p \log K_j - \log ((I + K_i^{-1}) (I + p \log (K_j + I) + \dots) - K_i^{-1})] \} \\ &= \exp \{ p \text{Tr} [\log K_j - (I + K_i^{-1}) \log (K_j + I)] + \dots \} \\ &= \exp \{ -p \text{Tr} [(I + K_i^{-1}) \log (I + K_j^{-1}) - K_i^{-1} \log K_j^{-1}] + \dots \}, \end{aligned} \quad (\text{A12})$$

where  $\dots$  denotes terms of  $\mathcal{O}(p^2)$  that are dropped in the replica  $p \rightarrow 0$  limit. As  $n \equiv K^{-1}$  from Eq. (5), Eq. (A11)

becomes

$$\text{Tr} [\rho_i \log \rho_j] = -\text{Tr} [(I + n_i) \log (I + n_j) - n_i \log n_j]. \quad (\text{A13})$$

When  $\rho_i \equiv \rho_j$  Eq. (A13) produces the form (19) for the entropy of a Gaussian-coherent ensemble in terms of its expected occupation number matrix. More gener-

ally, Eq. (A13) is used to obtain the expression (23) for the Kullback-Leibler divergence [19].

- 
- [1] E. Smith, *Phys. Rev. E* **58** (1998) 2818-2832.  
 [2] E. Smith, *Phys. Rev. E* **60** (1999) 3633-3645.  
 [3] E. Smith, *Phys. Rev. E* **68**, 046114 (2003).  
 [4] M. Mezard, G. Parisi, M. A. Virasoro, *Spin glass theory and beyond*, (World Scientific, Teaneck, New Jersey, 1987).  
 [5] G. Nicolis and I. Prigogine, *Self-organization in nonequilibrium systems: from dissipative structures to order through fluctuations* (Wiley, New York, 1977).  
 [6] S. R. de Groot and P. Mazur, *Non-equilibrium thermodynamics* (Dover, New York, 1984).  
 [7] E. T. Jaynes, *Papers on probability, statistics and statistical physics* (D. Reidel Publishing Co., Dordrecht, Holland, 1983) R. D. Rosenkrantz, ed.  
 [8] P. Glansdorff, and I. Prigogine, *Thermodynamics of Structure, Stability and Fluctuations* (Wiley-Interscience, New York, 1971); Ref. [5], Ch. 3.  
 [9] I. Prigogine, *Bull. Acad. Roy. Belg. C. Sci.* **31** (1945) 600; *Etude Thermodynamique des Processus Irreversibles*, (Desoer, Liège, 1947); Ref. [5], Sec. 3.5.  
 [10] L. Onsager, *Phys. Rev.* **37**, 405 (1931); *Phys. Rev.* **38**, 2265 (1931).  
 [11] E. F. Schrödinger, *What is Life?: the physical aspect of the living cell* (Cambridge U. Press, New York, c1992).  
 [12] A. B. Corbet and H. J. Morowitz, *Phys. Rev. A* **6** (1972) 2298-2302; A. B. Corbet, *Phys. Rev. A* **9** (1974) 1371-1384.  
 [13] M. Gell-Mann and S. Lloyd, *Complexity* **2** (1996) 44-52.  
 [14] R. D. Astumian and P. Hänggi, *Physics Today* (Nov. 2002) 33-39.  
 [15] H. J. Carmichael, *Statistical Methods in Quantum Optics 1: Master Equations and Fokker-Planck Equations* (Springer-Verlag, New York, 1999), Ch. 1.  
 [16] H. J. Morowitz, *Energy Flow in Biology* (Ox Bow Press, Woodbridge, Conn., 1979), Ch. II, p. 29,33.; H. J. Morowitz, *J. Theor. Biol.* **13**, 60 (1966).  
 [17] L. Onsager and S. Machlup, *Phys. Rev.* **91** (1953) 1505.  
 [18] G. D. Mahan, *Many-Particle Physics*, (Plenum, New York, 1990), 2<sup>nd</sup> Ed., sec. 2.9  
 [19] T. Cover, and J. Thomas, *Elements of Information Theory* (Wiley, New York, 1991) p.18.  
 [20] R. E. Blankenship, *Molecular mechanisms of photosynthesis*, (Blackwell Science, London, 2002).  
 [21] A constructive review of an earlier version of this paper pointed out that computations based on exact entropy are even likely to be read as “an idiosyncratic approach to nonequilibrium thermodynamics”.  
 [22] This is a slightly more restrictive criterion than sometimes used, as it implies flow of neither energy nor particles across the system boundary. “Closed” is sometimes used only to refer to absence of particle exchange [5], but the former usage is logically more appropriate to this discussion, where energy is transferred only in the form of particle excitations.