

RAPID COMMUNICATION | AUGUST 06 2015

Communication: Maximum caliber is a general variational principle for nonequilibrium statistical mechanics FREE

Michael J. Hazoglou; Valentin Walther; Purushottam D. Dixit; Ken A. Dill



J. Chem. Phys. 143, 051104 (2015)

<https://doi.org/10.1063/1.4928193>

CHORUS



View
Online



Export
Citation

Articles You May Be Interested In

Maximum caliber inference of nonequilibrium processes

J. Chem. Phys. (July 2010)

Maximum Caliber: A variational approach applied to two-state dynamics

J. Chem. Phys. (May 2008)

Markov processes follow from the principle of maximum caliber

J. Chem. Phys. (February 2012)



The Journal of Chemical Physics

Special Topics Open for Submissions

[Learn More](#)

Communication: Maximum caliber is a general variational principle for nonequilibrium statistical mechanics

Michael J. Hazoglou,^{1,a)} Valentin Walther,^{1,a)} Purushottam D. Dixit,² and Ken A. Dill^{1,3,4}

¹Department of Physics and Astronomy, Stony Brook University, Stony Brook, New York 11794, USA

²Department of Systems Biology, Columbia University, New York, New York 10032, USA

³Laufer Center for Physical and Quantitative Biology, Stony Brook, New York 11794, USA

⁴Department of Chemistry, Stony Brook University, Stony Brook, New York 11794, USA

(Received 20 May 2015; accepted 27 July 2015; published online 6 August 2015)

There has been interest in finding a general variational principle for non-equilibrium statistical mechanics. We give evidence that *Maximum Caliber* (Max Cal) is such a principle. Max Cal, a variant of maximum entropy, predicts dynamical distribution functions by maximizing a path entropy subject to dynamical constraints, such as average fluxes. We first show that Max Cal leads to standard near-equilibrium results—including the Green-Kubo relations, Onsager's reciprocal relations of coupled flows, and Prigogine's principle of minimum entropy production—in a way that is particularly simple. We develop some generalizations of the Onsager and Prigogine results that apply arbitrarily far from equilibrium. Because Max Cal does not require any notion of “local equilibrium,” or any notion of entropy dissipation, or temperature, or even any restriction to material physics, it is more general than many traditional approaches. It also applicable to flows and traffic on networks, for example. © 2015 AIP Publishing LLC. [<http://dx.doi.org/10.1063/1.4928193>]

I. INTRODUCTION

There has been interest in identifying a variational principle basis for non-equilibrium statistical mechanics (NESM). On the one hand, phenomenological dynamics is well established. It consists of (a) phenomenological relationships, such as Ohm's Law of electrical current flow, Fick's Law of diffusion, Fourier's Law of conduction, the Newtonian Law of viscosity, combined with (b) conservation laws, such as Kirchhoff's current relationship, or similar relationships for flows of mass or heat. The combinations of relationships of types (a) and (b) lead to well-known dynamical equations such as the Navier-Stokes and Burgers' equations of hydrodynamics or the diffusion and Smoluchowski equations, as elucidated in standard textbooks.¹

However, the search for a microscopic statistical basis for these relationships has been more challenging; see for example.² While there have been many powerful and important methods for particular calculations, including the Langevin equation, Master equation, Fokker-Planck, Smoluchowski equations, and others, nevertheless so far, there has been no foundational basis for NESM that provides the same power that the Second Law of thermodynamics provides for equilibria.^{3,4} Indeed, multiple context dependent variational principles have been proposed for NESM, such as minimum entropy production,⁴ maximum entropy production,^{5,6} and minimum energy dissipation.⁷ Nevertheless, a general variational quantity remains illusive.

A common starting point for deriving NESM variational principles is the “state entropy,” $S = -\sum_i p_i \log p_i$ where the p_i 's are the populations of equilibrium states $i = 1, 2, 3, \dots$. Then, a “local equilibrium” assumption is made that S , which is

fundamentally only defined in the context of an extremum principle for predicting equilibrium, can also be a useful predictor near equilibrium. Then, the entropy production is defined as $\sigma = dS/dt$ to determine the rates of approach to equilibrium or of dissipation in steady states. One major drawback of these approaches is that they usually do not provide a natural quantifier for closeness to equilibrium; there is no systematic way to improve the near-equilibrium predictions in terms of an expansion parameter further away from equilibrium. Here, we describe an approach based on path entropies, not state entropies, which does not have these problems, is applicable even far from equilibrium, has expansion parameters, and happens to give very simple procedures for deriving some results of NESM.

A goal of NESM has been to find a variational quantity that can be maximized subject to suitable dynamical constraints. Such an approach goes beyond phenomenological descriptions of only average forces and flows and explains fluctuating quantities and higher moments of dynamical properties as well. Jaynes proposed a candidate NESM variational principle called “Maximum Caliber” (Max Cal).⁸ We and others have explored its applicability as a general foundation for NESM.^{9–15} While the power of Max Cal is its generality (for example, its applicability far from equilibrium), what is new in the present work is simply demonstrating that key well-known results of NESM are readily recovered in a simple way when maximum caliber is applied in a near-equilibrium approximation.

II. THEORY: THE MAXIMUM CALIBER APPROACH

For concreteness, we consider a discrete-time discrete-state system with two types of fluxes: of “stuff” a and “stuff”

^{a)}M. J. Hazoglou and V. Walther contributed equally to this work.

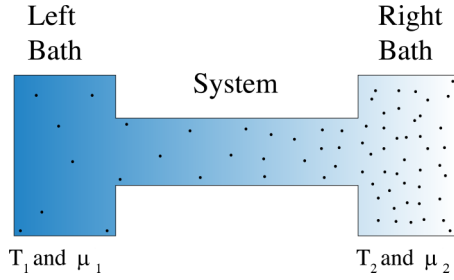


FIG. 1. An illustration of coupled heat and particle flows.

b. Examples of a and b include heat, mass, electrical charge, and momentum. We allow for the coupling between the two fluxes; the flux of a may depend in any way on the flux of b . We assume that the system reaches a macroscopic nonequilibrium stationary state after a long time after it has been coupled to gradients across its boundaries that set up the fluxes. Figure 1 illustrates with an example of coupled heat and particle flows in one dimension.

Suppose the system has N discrete states $\{1, 2, \dots, N\}$. Let us define an ensemble $\{\Gamma\}$ of stationary-state micro-trajectories $\Gamma \equiv \dots \rightarrow i \rightarrow j \rightarrow \dots$ of fixed duration. The duration of the trajectories is not important, so we keep it unspecified. Based on how the system is coupled to the gradients and the internal structure of the system, each micro-trajectory Γ will correspond to a specific value of flux $j_{a\Gamma}(t)$ of type a (and b) at any time t . The ensemble average of the flux $J_a(t)$ of quantity a at time t is then given by

$$J_a(t) = \langle j_{a\Gamma}(t) \rangle = \sum_{\Gamma} p_{\Gamma} j_{a\Gamma}(t), \quad (1)$$

where p_{Γ} is the probability of trajectory Γ . It is clear that different probability distributions p_{Γ} could lead to macroscopic fluxes $J_a(t)$ and $J_b(t)$. The strategy of Max Cal is to seek the particular probability distribution p_{Γ} that maximizes the path entropy and is otherwise consistent with the ensemble averaged fluxes $J_a(t)$ and $J_b(t)$ for all times t . We maximize the path entropy or the caliber,

$$C = - \sum_{\Gamma} p_{\Gamma} \log \frac{p_{\Gamma}}{q_{\Gamma}}, \quad (2)$$

subject to normalization and flux constraints. Here, q_{Γ} is the *reference* probability distribution of trajectories when there are no gradients, i.e., at thermodynamic equilibrium. From Fig. 1, it is clear that there are multiple choices of q_{Γ} , for example, it may correspond to the equilibrium distribution at T_1 and μ_1 or to the distribution at T_2 and μ_2 . As an aside, the conditions for equilibrium (see below) are that (a) there are no net fluxes at equilibrium, i.e.,

$$\langle j_{a\Gamma}(t) \rangle = \sum_{\Gamma} j_{a\Gamma}(t) q_{\Gamma} = 0 \quad (3)$$

and similarly for $\langle j_{b\Gamma}(t) \rangle$, and (b) the equilibrium state satisfies microscopic reversibility: the trajectory ensemble averages at equilibrium are unchanged under path reversal. Under the situation where the flux is odd under time-reversal condition (b) yields (a).

The caliber^{8,9} is maximized with the constraints on the macroscopic fluxes of a and b at time t for all t , which are

enforced by the set of Lagrange multipliers $\{\lambda_a(t)\}$ and $\{\lambda_b(t)\}$,

$$- \sum_{\Gamma} p_{\Gamma} \log \frac{p_{\Gamma}}{q_{\Gamma}} + \sum_t \lambda_a(t) \left(\sum_{\Gamma} p_{\Gamma} j_{a\Gamma}(t) - J_a(t) \right) + \sum_t \lambda_b(t) \left(\sum_{\Gamma} p_{\Gamma} j_{b\Gamma}(t) - J_b(t) \right) + \alpha \left(\sum_{\Gamma} p_{\Gamma} - 1 \right). \quad (4)$$

Maximizing the Caliber with respect to the trajectory probability p_{Γ} gives

$$p_{\Gamma} = \frac{q_{\Gamma}}{Z} \exp \left(\sum_t [\lambda_a(t) j_{a\Gamma}(t) + \lambda_b(t) j_{b\Gamma}(t)] \right), \quad (5)$$

with the dynamical partition function Z ,

$$Z = \sum_{\Gamma} q_{\Gamma} \exp \left(\sum_t [\lambda_a(t) j_{a\Gamma}(t) + \lambda_b(t) j_{b\Gamma}(t)] \right). \quad (6)$$

Equations (5) and (6) are the expressions of the principle of maximum caliber for two types of flows. Like maximum entropy for equilibrium statistical mechanics, maximum caliber for nonequilibrium computes macroscopic quantities as derivatives of a partition-function-like quantity (in this case, a sum of weights over the different pathways). For example, average flux quantities are first derivatives of the logarithm of the *dynamical partition function*,

$$J_a(t) = \langle j_{a\Gamma}(t) \rangle = \sum_{\Gamma} p_{\Gamma} j_{a\Gamma}(t) = \frac{\partial \log Z}{\partial \lambda_a(t)}. \quad (7)$$

Identical equations follow for $J_b(t)$. Equations (7) allow the calculation of $\lambda_a(t)$ and $\lambda_b(t)$ from the knowledge of the functional form of $J_a[\lambda_a(t), \lambda_b(t)]$ and $J_b[\lambda_a(t), \lambda_b(t)]$ as well as the constrained values. Higher moments of the dynamical distribution function can be calculated by taking higher derivatives of $\log Z$. For example, the second order cumulants are

$$\begin{aligned} \langle j_{a\Gamma}(t) j_{a\Gamma}(\tau) \rangle - \langle j_{a\Gamma}(t) \rangle \langle j_{a\Gamma}(\tau) \rangle &= \frac{\partial \langle j_{a\Gamma}(t) \rangle}{\partial \lambda_a(\tau)} \\ &= \frac{\partial^2 \log Z}{\partial \lambda_a(t) \partial \lambda_a(\tau)}, \end{aligned} \quad (8)$$

$$\begin{aligned} \langle j_{a\Gamma}(t) j_{b\Gamma}(\tau) \rangle - \langle j_{a\Gamma}(t) \rangle \langle j_{b\Gamma}(\tau) \rangle &= \frac{\partial \langle j_{a\Gamma}(t) \rangle}{\partial \lambda_b(\tau)} = \frac{\partial \langle j_{b\Gamma}(\tau) \rangle}{\partial \lambda_a(t)} \\ &= \frac{\partial^2 \log Z}{\partial \lambda_a(t) \partial \lambda_b(\tau)}. \end{aligned} \quad (9)$$

Identical expressions follow for b .

So far, this development is general, allowing for time-dependent fluxes. However, for our purpose below of touching base with three well-known results of NESM, we now restrict consideration to stationary flows, $\langle j_{a\Gamma}(t) \rangle = J_a$ and $\langle j_{b\Gamma}(t) \rangle = J_b$, for all times t . See section S1 of the supplementary material for a proof that Lagrange multipliers λ_a, λ_b corresponding to fluxes J_a and J_b are also independent of time.²⁷ Now, we show how the Green-Kubo relations, Onsager's reciprocal relations, and Prigogine's minimum entropy production theorem are derived quite simply from Equations (5) and (6).

A. Deriving the Green-Kubo relations from Max Cal

The Green-Kubo relations are well-known relationships between various transport coefficients, on the one hand, and

time correlation functions at equilibrium, on the other. Here, we show that they can be derived quite directly from Max Cal. Consider a coupled flow system in a stationary state (time invariant). Consider the fluxes at some particular time, call it $t = 0$. Now, expand around $\lambda \approx 0$ (small driving forces, see below). That is, expand $\langle j_{a\Gamma}(t) \rangle = J_a(0)$ at $t = 0$ to first order around $\lambda_a(\tau), \lambda_b(\tau) = 0$ for all τ ,

$$\begin{aligned} J_a(0) &\approx \sum_{\tau} \left[\frac{\partial \langle j_{a\Gamma}(0) \rangle}{\partial \lambda_a(\tau)} \lambda_a(\tau) + \frac{\partial \langle j_{a\Gamma}(0) \rangle}{\partial \lambda_b(\tau)} \lambda_b(\tau) \right] \\ &= \lambda_a \sum_{\tau} \langle j_{a\Gamma}(0) j_{a\Gamma}(\tau) \rangle_{\lambda=0} \\ &\quad + \lambda_b \sum_{\tau} \langle j_{a\Gamma}(0) j_{b\Gamma}(\tau) \rangle_{\lambda=0}. \end{aligned} \quad (10)$$

In Eq. (10), when $\lambda = 0$, note that p_{Γ} is equal to the equilibrium distribution q_{Γ} and note that $\langle j_{a\Gamma}(t) \rangle = \langle j_{b\Gamma}(t) \rangle = 0$. When the system is in steady state, any time dependence of the Lagrange multipliers λ_a and λ_b will vanish (see section S1 proof of the supplementary material²⁷). Eq. (10) is simply the Green-Kubo relations^{16–18} between the linear transport coefficient and the flux autocorrelations. And, up to a constant factor, the Lagrange multipliers can be seen as the driving forces.

B. Deriving the Onsager reciprocal relationships from Max Cal

Onsager considered the near-equilibrium situation in which multiple flows are linearly proportional to their corresponding forces,⁷

$$J_a = L_{aa}\lambda_a + L_{ab}\lambda_b, \quad (11)$$

$$J_b = L_{ba}\lambda_a + L_{bb}\lambda_b. \quad (12)$$

For this situation, Onsager derived the reciprocal relationship that $L_{ab} = L_{ba}$.^{19,20} This result can be derived straightforwardly from maximum caliber. From Eqs. (9) and (10) above, we have

$$\begin{aligned} L_{ab} &= \sum_{\tau} \frac{\partial^2 \log Z}{\partial \lambda_a(0) \partial \lambda_b(\tau)} \Big|_{\lambda=0} = \sum_{\tau} \langle j_a(0) j_b(\tau) \rangle_{\lambda=0} \\ &= \sum_{\tau} \langle j_a(\tau) j_b(0) \rangle_{\lambda=0} \\ &= L_{ba}. \end{aligned} \quad (13)$$

The last equality follows from microscopic reversibility of the equilibrium state. We have assumed that both fluxes have the same parity under time reversal. Summing (or integrating) over τ will not affect the symmetry of the matrix L .

Eq. (13) shows that Max Cal gives a quantitative measure of how far away a system is from equilibrium, in terms of the expansion quantity λ . Operationally, since the λ 's represent the forces, “near equilibrium” refers to any situation in which λ is small enough that fluxes are linearly proportional to forces.

It is clear that the symmetries in transport coefficients arise from microscopic reversibility of fluxes at equilibrium. Are there other such symmetries amongst the higher-order transport coefficients for systems not near equilibrium? While

there has been a considerable effort to discover such symmetries, no clear general results have been obtained.^{21–24} Our development shows that n^{th} order expansion of the flux in terms of the thermodynamic gradients will involve $(n+1)^{\text{st}}$ order cumulants functions among fluxes at equilibrium. These cumulants do have some symmetry properties owing to microscopic reversibility and translational invariance with respect to time. See section S2 of the supplementary material where we show that there are no simple relationships between higher-order transport coefficients.²⁷

C. Deriving Prigogine’s principle of minimum entropy production from Max Cal

Prigogine developed a variational principle for two coupled flows near equilibrium. If one of those flows a is driven by a given force, and if the other b is unconstrained, then the flux of b is predicted to be that which has the minimum rate of entropy production.^{25,26} First, here is the standard development in terms of entropy production. If the state entropy is S , then the rate of entropy production in a system carrying two fluxes J_a and J_b is given by

$$\sigma = \frac{dS}{dt} = J_a \lambda_a + J_b \lambda_b, \quad (14)$$

where λ_a and λ_b are driving gradients. Using the Onsager relationships near equilibrium, we have

$$\sigma = L_{aa}\lambda_a^2 + 2L_{ab}\lambda_a\lambda_b + L_{bb}\lambda_b^2. \quad (15)$$

Prigogine’s principle then seeks the entropy production rate that is minimal with respect to variations in λ_b ,

$$\frac{\partial \sigma}{\partial \lambda_b} = 2(L_{ab}\lambda_a + L_{bb}\lambda_b) = 2J_b = 0, \quad (16)$$

which correspondingly also predicts that $J_b = 0$.⁴

Now, here instead is the same principle derived from Max Cal. First, express the caliber as

$$\begin{aligned} C &= - \sum_{\Gamma} p_{\Gamma} \ln \left(\frac{p_{\Gamma}}{q_{\Gamma}} \right) \\ &= \ln Z - \sum_t [\lambda_a(t) J_a(t) + \lambda_b(t) J_b(t)]. \end{aligned} \quad (17)$$

Now, maximizing the caliber gives

$$\begin{aligned} \frac{\partial C}{\partial \lambda_b(\tau)} &= - \sum_t \left[\lambda_a(t) \frac{\partial J_a(t)}{\partial \lambda_b(\tau)} + \lambda_b(t) \frac{\partial J_b(t)}{\partial \lambda_b(\tau)} \right] \\ &\approx -\lambda_a L_{ab} - \lambda_b L_{bb} + O(\lambda^2) = -J_b = 0, \end{aligned} \quad (18)$$

which gives the a force-flux relationship near equilibrium that is identical to the one Prigogine derives using the entropy-production argument. However, in addition, Max Cal makes a useful prediction beyond the linear regime. The caliber is maximized when

$$\sum_t \left[\lambda_a(t) \frac{\partial J_a(t)}{\partial \lambda_b(\tau)} + \lambda_b(t) \frac{\partial J_b(t)}{\partial \lambda_b(\tau)} \right] = 0. \quad (19)$$

So, if we are given how J_a and J_b depend on the imposed thermodynamic gradients λ_a and λ_b , then solving Eq. (19)

gives the gradient λ_b to which the system adjusts itself when it is not constrained.

III. CONCLUSIONS

The principle of maximum caliber is a putative variational principle for nonequilibrium statistical mechanics. We show here that Max Cal provides a natural and simple route to deriving several key results of NESM, including the Green-Kubo relations, the Onsager reciprocal relations, and Prigogine's minimum entropy production principle. It gives a way to explore higher-order generalizations of Onsager relationships and Prigogine's principle for situations not near equilibrium. The advantages of Max Cal over other variational schemes are that it is not limited to near equilibrium, or to "local equilibrium assumptions," has a natural "order parameter" for defining a distance from equilibrium, and has a sounder basis in principle than quantities like entropy production rates. The power of the Max Cal method derives from its focus on path entropies, not state entropies.

ACKNOWLEDGMENTS

We thank Kinghuk Ghosh and Steve Presse for many deeply engaging initial discussions that led to this work. And, we thank Amos Maritan and Jayanth Banavar for the insightful derivation of the results in Appendix I (supplementary material).²⁷ We appreciate the support of NSF Grant No. PHY-1205881 and of the Laufer Center.

- ¹R. Bird, W. Stewart, and E. Lightfoot, *Transport Phenomena* (Wiley, 2007).
- ²F. Bonetto, J. L. Lebowitz, and L. Rey-Bellet, "Fourier's law: A challenge to theorists," preprint [arXiv:math-ph/0002052](https://arxiv.org/abs/math-ph/0002052) (2000).
- ³W. Grandy, Jr., *Entropy and the Time Evolution of Macroscopic Systems*, International Series of Monographs on Physics (OUP, Oxford, 2008).
- ⁴D. Kondepudi and I. Prigogine, *From Heat Engines to Dissipative Structures* (John Wiley & Son, 1998).
- ⁵R. Dewar, *J. Phys. A: Math. Gen.* **36**, 631 (2003).
- ⁶R. C. Dewar, *J. Phys. A: Math. Gen.* **38**, L371 (2005).
- ⁷L. Onsager and S. Machlup, *Phys. Rev.* **91**, 1505 (1953).
- ⁸E. T. Jaynes, in *Complex Systems—Operational Approaches*, edited by H. Haken (Springer-Verlag, Berlin, 1985), p. 254.
- ⁹S. Pressé, K. Ghosh, J. Lee, and K. A. Dill, *Rev. Mod. Phys.* **85**, 1115 (2013).
- ¹⁰P. D. Dixit and K. A. Dill, *J. Chem. Theory Comput.* **10**, 3002 (2014).
- ¹¹G. Stock, K. Ghosh, and K. A. Dill, *J. Chem. Phys.* **128**, 194102 (2008).
- ¹²Q. A. Wang, *Chaos, Solitons Fractals* **23**, 1253 (2005).
- ¹³H. Ge, S. Pressé, K. Ghosh, and K. A. Dill, *J. Chem. Phys.* **136**, 064108 (2012).
- ¹⁴J. Lee and S. Pressé, *J. Chem. Phys.* **137**, 074103 (2012).
- ¹⁵P. D. Dixit, J. Abhinav, S. Gerhard, and A. Dill Ken (unpublished).
- ¹⁶M. S. Green, *J. Chem. Phys.* **20**, 1281 (1952).
- ¹⁷M. S. Green, *J. Chem. Phys.* **22**, 398 (1954).
- ¹⁸R. Kubo, *J. Phys. Soc. Jpn.* **12**, 570 (1957).
- ¹⁹L. Onsager, *Phys. Rev.* **37**, 405 (1931).
- ²⁰L. Onsager, *Phys. Rev.* **38**, 2265 (1931).
- ²¹G. Bochkov and Y. E. Kuzovlev, *Zh. Eksp. Teor. Fiz.* **72**, 238 (1977).
- ²²G. Bochkov and Y. E. Kuzovlev, *Zh. Eksp. Teor. Fiz.* **76**, 1071 (1979).
- ²³R. D. Astumian, *Phys. Rev. Lett.* **101**, 046802 (2008).
- ²⁴M. N. Artyomov, *Phys. Rev. Lett.* **102**, 149701 (2009).
- ²⁵U. Seifert, *Eur. Phys. J. B* **64**, 423 (2008).
- ²⁶T. Tomé and M. J. de Oliveira, *Phys. Rev. Lett.* **108**, 020601 (2012).
- ²⁷See supplementary material at <http://dx.doi.org/10.1063/1.4928193> for the time independence of Lagrange multipliers when fluxes are time independent, lack of higher order reciprocal relationships, and modified reciprocal relationships when fluxes have different parities.