

PHYSICAL REVIEW LETTERS

VOLUME 48

7 JUNE 1982

NUMBER 23

Generalization of the Onsager Reciprocity Theorem

James Hurley and Claude Garrod

Department of Physics, University of California, Davis, California 95616

(Received 22 February 1982)

The Onsager reciprocity theorem has been since its discovery in 1931 one of the most prominent successes of nonequilibrium statistical mechanics. Its interpretation and application suffer from several limiting assumptions made in its derivation. The most notable of these is the linear relation between the currents and the affinities. A generalization of the Onsager theorem has been obtained, which is comparatively free of *ad hoc* assumptions.

PACS numbers: 05.70.Ln

It has long been customary in the field of nonequilibrium thermodynamics to employ linear constitutive relations between fluxes and gradients of affinities. Perhaps the most familiar of these is the proportionality assumed between the heat flux and the temperature gradient. The phenomenological coefficient of proportionality is the thermal conductivity.

Besides heat conduction we may also have particle diffusion and/or electron conduction. One finds that the flux of thermal energy depends not only on the temperature gradient but upon the concentration gradient and electric potential gradient (if the particles are charged). Similarly the particle diffusion depends not only on the concentration gradient but on the temperature and electric potential gradients as well. In the early development of the subject it was found that if one assumed linear constitutive relations between the fluxes and the affinity gradients that certain relations existed between the coefficients. Among these are the well-known Kelvin relations.

Although models were proposed from which it was possible to derive many of these relationships it was the grand synthesis of Lars Onsager that showed that all such interrelationships are exam-

ples of a single fundamental theorem. This theorem now bears his name and is referred to as the "Onsager Reciprocity Theorem." Onsager demonstrated that the reciprocity among the constitutive coefficients depends fundamentally on the time-reversal invariance of the microscopic equations of motion.

To obtain the reciprocal relations Onsager assumed a linear proportionality between fluxes and affinities. The coefficients of proportionality were assumed to be constant, independent of the thermodynamic state. We shall show that it is possible to remove this assumption and thereby construct a generalization of Onsager's original result.

We begin by considering a macroscopic system with observables a_1, a_2, \dots, a_K which we collectively denote by \underline{a} . These observables are functions of the microscopic state of the system. The observables might, for example, be the numbers of K chemical constituents or the energies and particle number in given subdomains of the system.

The microscopic state of the system is represented by the generalized coordinates $(q_1, \dots, q_N, p_1, \dots, p_N)$ in a $2N$ -dimensional phase space Γ . We shall find it convenient to introduce a vector

z with components $(z_1, \dots, z_{2N}) = (q_1, \dots, q_N, p_1, \dots, p_N)$. We may now succinctly express the functional dependence of the observables \underline{a} on the phase point by the relation $\underline{a} = \underline{a}(z)$.

Within Γ we define two point transformations: (1) the time evolution operator $T_t z_0 = z(t, z_0)$, where $z(t, z_0)$ is the solution of Hamilton's equations with initial value z_0 , and (2) the time-reversal operator $Rz = (\underline{q}, -\underline{p})$. A function $f(z)$ is called *time reversible* if $f(Rz) = f(z)$ and *time invariant* if $f(T_t z) = f(z)$ for all t . The Hamiltonian is assumed to have both properties. This implies that $T_t R T_t z = Rz$ or, equivalently, that

$$RT_t = T_{-t} R. \quad (1)$$

We consider an ensemble with a time-reversible and time-invariant density $\rho(z)$. For any t , the $K \times K$ correlation matrix $\underline{C}(t)$ is defined as

$$\underline{C} = \int_{\Gamma} \underline{a}(z) \underline{a}(T_t z) \rho(z) d^{2N} z. \quad (2)$$

Introducing into Eq. (2) the transformation of variables $z' = RT_t z$ and using the facts that $\rho(z') = \rho(z)$, $d^{2N} z' = d^{2N} z$, and $RT_t = T_{-t} R$ we write $\underline{C}(t)$ as

$$\underline{C} = \int_{\Gamma} \underline{a}(RT_t z') \underline{a}(Rz') \rho(z') d^{2N} z'. \quad (3)$$

We shall now assume that each of the observables is either even or odd under time reversal:

$$a_i(Rz) = \epsilon_i a_i(z), \quad \epsilon_i = \pm 1. \quad (4)$$

We may also write Eq. (4) in terms of a $K \times K$ matrix $\epsilon_{ij} = \epsilon_i \delta_{ij}$, as $\underline{a}(Rz) = \underline{\epsilon} \cdot \underline{a}(z)$. Using Eq. (4) in Eq. (3) gives the result

$$C_{ij}(t) = \epsilon_i \epsilon_j C_{ji}(t). \quad (5)$$

At time zero we assume that the observables $\underline{a}(z)$ take on specified values $\underline{\alpha}$ defined in a K -dimensional space γ . In γ we define an ensemble entropy function $S(\underline{\alpha})$ and a time evolution operator T_t by the equations¹

$$e^{S(\underline{\alpha})} = \int_{\Gamma} \delta(\underline{a}(z) - \underline{\alpha}) \rho(z) d^{2N} z \quad (6)$$

and

$$T_t \underline{\alpha} = \int_{\Gamma} \underline{a}(T_t z) \delta(\underline{a}(z) - \underline{\alpha}) \rho(z) d^{2N} z / e^{S(\underline{\alpha})}, \quad (7)$$

where $\delta(\underline{\alpha})$ is a K -dimensional Dirac δ function. Without loss of generality we may assume that the maximum of $S(\underline{\alpha})$ (i.e., the equilibrium state) occurs at $\underline{\alpha} = 0$. With these definitions we obtain the identity

$$\begin{aligned} \underline{C} &= \int_{\Gamma} \underline{a}(z) \underline{a}(T_t z) \rho(z) d^{2N} z \\ &= \int_{\gamma} \underline{\alpha} \int_{\Gamma} \underline{a}(T_t z) \delta(\underline{a}(z) - \underline{\alpha}) d^{2N} z d^K \alpha \\ &= \int_{\gamma} \underline{\alpha} T_t \underline{\alpha} e^{S(\underline{\alpha})} d^K \alpha. \end{aligned} \quad (8)$$

The gradient of S defines a vector field in γ , $\underline{\psi}(\underline{\alpha})$, with components $\psi_k(\underline{\alpha}) = \partial S(\underline{\alpha}) / \partial \alpha_k$, which, for reasonable thermodynamic ensembles, is zero only at the origin. The conventional Onsager reciprocity relation can be easily obtained if we assume that there exists some interval $\tau < t < T$ and some neighborhood of the origin in γ within which one can find a constant $K \times K$ matrix, \underline{P} satisfying

$$T_t \underline{\alpha} = \underline{\alpha} + t \underline{\psi} \cdot \underline{P}. \quad (9)$$

It is then possible to use Eq. (5) to show that $\underline{\epsilon} \cdot \underline{P}$ is a symmetric matrix and this is the traditional Onsager relation.²⁻⁴ For typical thermodynamic systems the lower time limit, τ , is on the order of a microscopic collision time and is thus macroscopically indistinguishable from zero. The upper limit, T , need only be much less than any macroscopic relaxation time of the system so that terms of order t^2 and higher in $T_t \underline{\alpha} - \underline{\alpha}$ can be neglected.

We shall now show that these restrictions are unnecessary and obtain a more general consequence of time reversibility. We define $\underline{P}(\underline{\alpha}, t)$ as any matrix function satisfying Eq. (9). That at least one such function exists is guaranteed by the fact that the dyadic $\underline{P} = [\underline{\psi}(T_t \underline{\alpha} - \underline{\alpha}) / t] \underline{\psi}^2$ is a solution of Eq. (9). However, once we no longer require that \underline{P} be a constant matrix it is clear that Eq. (9) is not sufficient to define a unique matrix function. We shall return to the question of the nonuniqueness of $\underline{P}(\underline{\alpha}, t)$ as defined by Eq. (9) later. Combining Eqs. (5), (8), and (9) we obtain

$$(1 - \epsilon_i \epsilon_j) \int \alpha_i \alpha_j e^{S(\underline{\alpha})} d^K \alpha + t \int \sum_k (\alpha_i P_{kj} - \epsilon_i \epsilon_j \alpha_j P_{ki}) \psi_k e^{S(\underline{\alpha})} d^K \alpha = 0. \quad (10)$$

Since this equation is an identity in time, the first term must be independently equal to zero. To verify this we observe that

$$(1 - \epsilon_i \epsilon_j) \int \alpha_i \alpha_j e^{S(\underline{\alpha})} d^K \alpha = (1 - \epsilon_i \epsilon_j) \int_{\Gamma} a_i(\underline{q}, \underline{p}) a_j(\underline{q}, \underline{p}) \rho(\underline{q}, \underline{p}) d^{2N} z. \quad (11)$$

If a_i and a_j are of the same parity with respect to time reversal, then $(1 - \epsilon_i \epsilon_j)$ is zero. If they are of opposite parity, the integral is zero since $\rho(\underline{q}, \underline{p})$ is of even parity. In either event the term is zero.

Using the fact that $\psi_k e^S = \partial e^S / \partial \alpha_k$ and integrating Eq. (10) by parts, we obtain

$$\int [P_{ij} - \epsilon_i \epsilon_j P_{ji} + \sum_k (\alpha_i \partial P_{kj} / \partial \alpha_k - \epsilon_i \epsilon_j \alpha_j \partial P_{ki} / \partial \alpha_k)] e^{S(\underline{\alpha})} d^K \alpha = 0. \quad (12)$$

That Eq. (12) is a generalization of the Onsager reciprocity relation may be easily seen because the assumption that \underline{P} is a constant matrix (with respect to $\underline{\alpha}$ but not necessarily with respect to t) leads immediately to the conventional Onsager relation $P_{ij} = \epsilon_i \epsilon_j P_{ji}$.

We may also see from Eq. (12) that the Onsager relation remains valid to first order when P_{ij} is not a constant matrix. Let us choose the $\underline{\alpha}$'s to be the *deviation* of the macroscopic observable from their equilibrium values. We may then expand P in a power series in $\underline{\alpha}$, which for small $\underline{\alpha}$ we may approximate by

$$P_{ij}(\underline{\alpha}, t) = P_{ij}^0(t) + \sum_k P_{ijk}(t) \alpha_k, \quad (13)$$

and substituting it into Eq. (12) we find that

$$P_{ij}^0(t) = \epsilon_i \epsilon_j P_{ji}^0(t). \quad (14)$$

The contribution from the sum in Eq. (12) vanishes, for

$$\int \alpha_i e^{S(\underline{\alpha})} d^K \alpha = 0 \quad (15)$$

since the mean value of $\underline{\alpha}$ is zero by construction. Thus to first order we recover the Onsager relation for the constant part of \underline{P} .

A form of the Onsager relation may be retained to all orders by taking advantage of the fact that Eq. (9) does not define a unique matrix \underline{P} . Let us add the subsidiary condition $\nabla \cdot \underline{P}(\underline{\alpha}, t) = 0$. (Note that this condition is trivially satisfied by a constant matrix.) With this added condition Eq. (12) takes the simpler form

$$\bar{P}_{ij} = \epsilon_i \epsilon_j \bar{P}_{ji}, \quad (16)$$

where \bar{P} is the ensemble average of $\underline{P}(\underline{\alpha}, t)$,

$$\bar{P}(t) = \int \underline{P}(\underline{\alpha}, t) e^{S(\underline{\alpha})} d^K \alpha. \quad (17)$$

In summary we have in Eq. (12) a mathemati-

cal identity. The only assumptions made in its derivation are that the $\underline{\alpha}$'s are either even or odd functions of the momenta, the dynamics of a point in phase space is time invariant and the Hamiltonian is an even function of the momenta. It is unnecessary to make any assumptions regarding the magnitude of t , or that the system is close to its equilibrium state. The entropy need not be a quadratic function of the thermodynamic observables. The dimension K of the thermodynamic state space is arbitrary. One need not attempt to construct a maximal set of independent observables. Last, and perhaps most important, the Onsager matrix \underline{P} need not be constant with respect to time or the thermodynamic observables.

We have restricted ourselves to discrete observables $\underline{\alpha}$. In a forthcoming paper these results will be extended to the treatment of continuous systems. We will also consider the relation between the traditional entropy production formalism and the Onsager reciprocity relation.

¹What we have defined as $S(\underline{\alpha})$ is the standard entropy only if $\rho(z)$ describes a microcanonical ensemble [$\rho = \delta(H(z) - E)$]. If ρ is a canonical ensemble density then our $S(\underline{\alpha})$ is the usual canonical potential $S - \beta E$. With a slight generalization of the formalism one could also treat the grand canonical ensemble. All of our later results remain valid for these other ensembles.

²L. Onsager, Phys. Rev. **38**, 2265 (1931).

³S. R. deGroot and P. Mazur, *Non-Equilibrium Thermodynamics* (North-Holland, Amsterdam, 1962).

⁴*Foundations of Continuum Mechanics*, edited by J. J. D. Domingos, M. N. R. Nina, and J. H. Whitelaw (Wiley, New York, 1973).