

A mathematical proof of the zeroth “law” of thermodynamics and the nonlinear Fourier “law” for heat flow

Denis J. Evans, Stephen R. Williams, and Lamberto Rondoni

Citation: *J. Chem. Phys.* **137**, 194109 (2012); doi: 10.1063/1.4766734

View online: <http://dx.doi.org/10.1063/1.4766734>

View Table of Contents: <http://jcp.aip.org/resource/1/JCPSA6/v137/i19>

Published by the [American Institute of Physics](#).

Additional information on *J. Chem. Phys.*

Journal Homepage: <http://jcp.aip.org/>

Journal Information: http://jcp.aip.org/about/about_the_journal

Top downloads: http://jcp.aip.org/features/most_downloaded

Information for Authors: <http://jcp.aip.org/authors>

ADVERTISEMENT



Goodfellow
metals • ceramics • polymers • composites
70,000 products
450 different materials
small quantities fast

www.goodfellowusa.com

A mathematical proof of the zeroth “law” of thermodynamics and the nonlinear Fourier “law” for heat flow

Denis J. Evans,¹ Stephen R. Williams,¹ and Lamberto Rondoni^{2,3}

¹*Research School of Chemistry, Australian National University, Canberra, ACT 0200, Australia*

²*Dipartimento di Scienze Matematiche, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy*

³*INFN, Sezione di Torino, Via P. Giuria 1, 10125 Torino, Italy*

(Received 16 January 2012; accepted 16 October 2012; published online 21 November 2012)

What is now known as the zeroth “law” of thermodynamics was first stated by Maxwell in 1872: at equilibrium, “Bodies whose temperatures are equal to that of the same body have themselves equal temperatures.” In the present paper, we give an explicit mathematical proof of the zeroth “law” for classical, deterministic, T-mixing systems. We show that if a body is initially not isothermal it will in the course of time (subject to some simple conditions) relax to isothermal equilibrium where all parts of the system will have the same temperature in accord with the zeroth “law.” As part of the derivation we give for the first time, an exact expression for the far from equilibrium thermal conductivity. We also give a general proof that the infinite-time integral, of transient and equilibrium autocorrelation functions of fluxes of non-conserved quantities vanish. This constitutes a proof of what was called the “heat death of the Universe” as was widely discussed in the latter half of the 19th century. © 2012 American Institute of Physics. [<http://dx.doi.org/10.1063/1.4766734>]

I. INTRODUCTION

The derivation of a number of fluctuation theorems for classical, time reversible systems leads to a number of macroscopically time irreversible results such as the fluctuation theorems^{1,2} themselves, the second law inequality,³ the Crooks fluctuation theorem,⁴ and the Jarzynski equality.⁵ In 2011, the first dynamical proof⁶ was given of the Clausius inequality. The Clausius inequality (1854) is often taken as the canonical statement of the second “law” of thermodynamics. The century old puzzle of how time reversible microscopic dynamics can lead to time irreversible macroscopic behavior has thus been solved within the context of certain mechanical particulate systems, and in the process the “laws” of thermodynamics have been similarly reduced to mathematical theorems concerning the microscopic dynamics of time reversible particulate systems.

In the present paper, we turn our attention to the proof of the zeroth “law” of thermodynamics. What is now known as the zeroth law of thermodynamics was first stated by Maxwell in 1872. Among numerous equivalent statements, Maxwell said:⁷ “Bodies whose temperatures are equal to that of the same body have themselves equal temperatures.” We have recently given a number of proofs^{8,9,11} of the relaxation of classical particulate systems to thermal equilibrium. This included the relaxation to thermal equilibrium characterized by the canonical distribution, a new result. We also discussed the relaxation of an autonomous Hamiltonian system, to rederive a result that was already known from ergodic theory. Our present derivation is quite different from the derivations given in ergodic theory and gives much greater insight into the relaxation process. To show the reader more precisely how the two derivations differ, we compare the standard ergodic theory derivation with our derivation that is based on the concept

of T-mixing.¹⁰ In T-mixing systems, time integrals of transient time correlation functions (TTCF) of the dissipation function evaluated at time zero with phase functions at a later time t , converge as the integration time goes to infinity. The dissipation function is the argument of the well-known Evans-Searles fluctuation theorem.²

The equilibrium states to which our systems relax are all isothermal, so implicit in these two equilibrium relaxation theorems and the ergodic theory proof for autonomous Hamiltonian systems, is a proof of the zeroth “law” of thermodynamics. In the present paper, we give a mathematical proof of the zeroth “law” for T-mixing, deterministic particulate systems obeying autonomous Hamiltonian dynamics. No external fields are applied to the system. We should add that as in most discussions in physics we only consider inertial coordinate systems since we do not wish to include Coriolis forces, etc.

The derivation also leads to an understanding of how heat flows from hot to cold and how the transport coefficient characterizing this flow is positive and finite when the system is T-mixing. This heat flow gradually equalizes the temperature across the entire system and heat eventually ceases to flow. This is the mechanism by which the zeroth “law” behavior of equilibrium systems is achieved.

The proof gives for the first time, an exact expression (Eq. (23)), for the nonlinear, far from equilibrium thermal conductivity. The proof also explains an integral sum rule for both transient and equilibrium time correlation functions. In the linearized, close to equilibrium case, the sum rule was previously known in certain special cases by Zwanzig and by Berne, Boone and Rice in the 1960s.¹² In the far from equilibrium case, the derivation of the integral sum rule gives a proof of what was known in the latter half of the 19th century, as the “heat death of the Universe.”

II. MATHEMATICAL DERIVATION OF THE ZEROth LAW

A. Background

Consider an ensemble of N -particle systems obeying Newton's or Hamilton's equations of motion. We do not assume that each particle is identical. The particles could differ in masses and interatomic potentials. If the system of particles is isolated, the total energy, linear momentum, and angular momentum are constants of the motion. In our thought experiment, we could imagine the system is composed of two solid three-dimensional boxes so that a left half and a right half of the system are in thermal contact but there is no mass flow between the two sides. These two boxes represent the "bodies" mentioned in Maxwell's statement of the zeroth "law." The two boxes (bodies) contain particles that maybe solid, liquid or gas.

B. The dissipation function

We denote the phase space vector describing the coordinates and momenta of the particles ($\mathbf{q}_i, \mathbf{p}_i; i = 1, \dots, N$), as $\mathbf{\Gamma} = (\mathbf{q}_1, \dots, \mathbf{q}_N, \mathbf{p}_1, \dots, \mathbf{p}_N)$. The N -particle phase space density at time t is denoted as $f(\mathbf{\Gamma}; t)$. The Evans-Searles fluctuation theorem,¹ second law inequality,³ dissipation theorem,¹³ and relaxation theorems^{8,9,11} each refer in some way to the dissipation function. Regardless of whether the system is thermostatted or adiabatic or whether there are external fields applied, the integral of the dissipation over a time t , starting from a given initial probability distribution $f(\mathbf{\Gamma}; 0)$ is defined as²

$$\Omega_t(\mathbf{\Gamma}) \equiv \ln \left[\frac{f(\mathbf{\Gamma}; 0) \exp \left(- \int_0^t ds \Lambda(S^t \mathbf{\Gamma}) \right)}{f(\tau S^t \mathbf{\Gamma}; 0)} \right], \quad (1)$$

where the operator τ denotes the time reversal map:¹⁴ $\tau(\mathbf{q}, \mathbf{p}; i = 1, \dots, N) \equiv (\mathbf{q}, -\mathbf{p}; i = 1, \dots, N)$, S^t is the time evolution operator for the phase space vector, so that $S^t \mathbf{\Gamma}$ denotes the position in phase space reached by a phase space trajectory that started at $\mathbf{\Gamma}$ time $t = 0$. Λ is the divergence of the equations of motion. This is the rate at which phase space volumes expand: for equations of motion expressed by $\dot{\mathbf{\Gamma}} = \mathbf{G}(\mathbf{\Gamma})$, one has $\Lambda = \text{div}(\mathbf{G}(\mathbf{\Gamma}))$.

By definition the number of ensemble members inside any volume element is conserved (conservation of ensemble members). In general, their density varies in time, as determined by the factor $\exp(\int_0^t ds \Lambda(S^t \mathbf{\Gamma}))$. Indeed, if the system gains or loses heat to its surroundings, the magnitude of the phase space volume changes in time. With the loss of heat to thermostatted surroundings, phase space volumes contract (i.e., $\int_0^t ds \Lambda(S^t \mathbf{\Gamma})$ averages to a negative quantity).

Taking the time derivative of $\Omega_t(\mathbf{\Gamma})$ we obtain the instantaneous dissipation function $\Omega(\mathbf{\Gamma})$. Because the microscopic dynamics is time-reversal symmetric, the denominator of (1) denotes the probability density of observing at time zero [i.e., with respect to the initial distribution $f(\mathbf{\Gamma}; 0)$], the set of trajectories that are the conjugate time reversed trajectories to those appearing in the numerator.¹⁴ Along the trajectories in the denominator, the average of the function Ω takes values opposite to those that it takes along the trajectories in the nu-

merator. This time reversal conjugacy is inherent in the definition of the dissipation function. It leads immediately to the Evans-Searles transient fluctuation relation

$$\frac{\Pr(\Omega_t \approx At \pm \delta At)}{\Pr(\Omega_t \approx -At \pm \delta At)} = \exp(At + \text{small error}), \quad (2)$$

where the small error is $O(\delta At)$.¹⁵

Now we assume that the system is isolated and simply obeys Newton's or Hamilton's equations for N interacting particles. We assume throughout this paper that the Hamiltonian is time independent (i.e., autonomous) and is translationally and rotationally invariant. The energy, momentum, and angular momentum are thus constants of the motion and there is no phase space volume expansion or contraction: $\int_0^t ds \Lambda(S^t \mathbf{\Gamma}) = 0$ for all $\mathbf{\Gamma}$ and all t . Without loss of generality we choose a co-moving coordinate frame in which the total linear momentum is zero, $\mathbf{P} = \mathbf{0}$. We assume the angular momentum is also zero, $\mathbf{L} = \mathbf{0}$. As we will see later, if the total angular momentum is nonzero the system cannot be mixing or T-mixing.

From the definition of the dissipation function, it is trivial to see that if the states are distributed as

$$f_{\mu C}(\mathbf{\Gamma}) = \frac{\delta(H_0(\mathbf{\Gamma}) - E)\delta(\mathbf{P})(\mathbf{L})}{\int d\mathbf{\Gamma} \delta(H_0(\mathbf{\Gamma}) - E)\delta(\mathbf{P})(\mathbf{L})} \\ = \frac{1}{\int d\mathbf{\Gamma}} \text{ if } \mathbf{\Gamma} \in D, \quad = 0 \text{ if } \mathbf{\Gamma} \notin D, \quad (3)$$

the dissipation function is identically zero, everywhere in ostensible phase space, D . The distribution function in Eq. (3) is therefore an equilibrium distribution function. It is referred to as the equilibrium microcanonical distribution ($f_{\mu C}(\mathbf{\Gamma})$). In (3), D is the ostensible phase space domain where the particle coordinates range over some physical volume on the energy, zero linear and zero angular momentum hypersurface ($\delta(H_0(\mathbf{\Gamma}) - E)\delta(\mathbf{P})\delta(\mathbf{L})$). Within this ostensible domain D , T-mixing systems have no non-trivial constants of the motion. Later, we will prove this statement from the T-mixing definition.¹⁰ Of course if the particular Hamiltonian we are dealing with contains more symmetries than those discussed here, there will be additional non-trivial constants of the motion. These should be handled by inserting additional delta functions into the microcanonical distribution (3) so that the ostensible phase space is constrained to a fixed value for these additional constants of the motion. [We make the distinction of the ostensible phase space domain because in nonequilibrium steady states in particular, the region of phase space that has nonzero phase space density is of lower dimension than the ostensible phase space. Not all the ostensible phase space has nonzero density.]

C. Ergodic theory approach

It is known from ergodic theory that for a finite, autonomous, Hamiltonian system that is mixing, an arbitrary initial state described by an initial phase space distribution $f(\mathbf{\Gamma}; 0)$ will eventually relax to microcanonical equilibrium. A system is said to be mixing if for integrable phase functions, time correlation functions computed with respect to a stationary distribution factorize into products of averages computed

with respect to the same distribution

$$\lim_{t \rightarrow \infty} \langle A(\Gamma)B(S^t \Gamma) \rangle_\infty - \langle A(\Gamma) \rangle_\infty \langle B(\Gamma) \rangle_\infty = 0. \quad (4)$$

Here, the brackets denote an ensemble average with respect to an invariant (i.e., stationary) probability distribution μ_∞ . In case μ_∞ has density $f(\Gamma; \infty)$, one may write

$$\langle A \rangle_\infty = \int d\mu_\infty(\Gamma) A(\Gamma) = \int d\Gamma f(\Gamma; \infty) A(\Gamma),$$

where $d\mu_\infty = d\Gamma f(\Gamma; \infty)$ is a (dimensionless and normalized) distribution.

If $f(\Gamma; \infty)$ is singular, one would write only the first equality $\langle A \rangle_\infty = \int d\mu_\infty(\Gamma) A(\Gamma)$, where $d\mu_\infty(\Gamma)$ is dimensionless and normalized.

We note that if the system has nonzero angular momentum no stationary long-time measure is possible (unless we transform to a non-inertial, co-rotating coordinate frame where Hamiltonian dynamics breaks down). So zero angular momentum as a constraint on the ostensible phase space not only guarantees the possibility of Hamiltonian dynamics but also allows for the possibility of having a stationary measure.

The mixing property is a property of the stationary state of interest, in which observables take the average values denoted by $\langle \cdot \rangle_\infty$. It represents the fact that, in the macroscopically stationary state, correlations among time evolving microscopic properties (such as the local and instantaneous values of the phase functions) decay in time. Therefore, in general the mixing condition does not guarantee relaxation to an invariant state. Mixing already assumes stationarity of the macrostate, whether it is reached asymptotically in time, as implied by our notation, or it is initially prepared in that state. Only in the special case of autonomous Hamiltonian systems, does mixing actually imply relaxation to the (microcanonical) stationary state (cf. below). This is indeed also a case, for which the density $f(\Gamma; \infty)$ exists.

D. T-mixing conditions

Mixing is closely related to, but subtly different from, the T-mixing condition, which we recently introduced in slightly different versions to obtain correspondingly different results.¹⁵ In particular, to prove steady state fluctuation relations, only a weak and hence very general form of T-mixing is required. This proves how general the steady state fluctuation relation for the dissipation function^{10,15} actually is. On the other hand, the relaxation theorems require sufficiently fast rates of correlation decay. Below this is made part of the definition of T-mixing itself.¹⁰ The T-mixing condition assumes that for a real sufficiently smooth phase function $A(\Gamma)$

$$\int_0^\infty ds \langle \Omega(\Gamma)A(S^s \Gamma) \rangle_0 = L_0 \in \mathbf{R}, \quad (5)$$

i.e., L_0 is real and finite, where $\Omega(\Gamma)$ is the instantaneous dissipation at the phase Γ and $A(S^s \Gamma)$ is the phase function $A(\Gamma)$ evaluated at the time evolved phase $S^s \Gamma$. In contradistinction to the well-known mixing condition of ergodic theory, the T-mixing condition considers time correlation functions re-

ferred to the initial state, here denoted by $\langle \cdot \rangle_0$, where the distribution of phases is usually known.

The weak T-mixing condition that looks very similar to the mixing condition is that

$$\lim_{t \rightarrow \infty} [\langle A(\Gamma)B(S^t \Gamma) \rangle_0 - \langle A(\Gamma) \rangle_0 \langle B(S^t \Gamma) \rangle_0] = 0, \quad (6)$$

where $A(\Gamma)$, $B(\Gamma)$ are any integrable phase functions. The main difference between weak T-mixing (6) and standard mixing (4) lies in the fact that the second factor in the second term inside square brackets in (6) is not time independent. It takes the form $\langle B(S^t \Gamma) \rangle_0 = \langle B(\Gamma) \rangle_t$, hence it cannot be taken out of the limit. This time dependence is a reflection of the fact that the ensemble averages in (6) are taken with respect to the initial distribution rather than an invariant long-time distribution.

For weak T-mixing (6) and T-mixing (5), the relevant probability distribution in (5) and (6) is not the invariant one; it is the initial ensemble $d\mu_0(\Gamma) = d\Gamma f(\Gamma; 0)$, whose averages are denoted by $\langle \cdot \rangle_0$. Mixing (4) and weak T-mixing (6) do not say anything about the rate of convergence to a stationary state or even whether such convergence actually occurs.

Throughout the rest of this paper if we use the term ‘‘T-mixing’’ we are referring to the condition given in Eq. (5). If we refer to ‘‘mixing,’’ we are discussing mixing given by Eq. (4). If we discuss ‘‘weak T-mixing’’ we are referring only to the condition given in Eq. (6).

We obviously exclude the constants of the motion inherent in the Hamiltonian symmetries from being possible phase functions in (5) (i.e., $A(\Gamma)$, $B(\Gamma) \notin H_0(\Gamma)$, $P_\alpha(\Gamma)$, $L_\alpha(\Gamma)$, $\alpha = x, y, z$), since each of these variables is obviously a constant of the motion. So our ostensible phase space domain D is some specified physical volume on a zero linear and angular momentum energy hypersurface. The zero linear momentum condition could be relaxed but the total angular momentum must be fixed at zero.

If the space is orientationally isotropic and the total angular momentum is a nonzero constant of the motion, for reasons that are rather obvious the system cannot possibly be T-mixing (5) or mixing (4). When viewed from an inertial coordinate frame, the measure required for mixing (4) cannot be time invariant but rather will be periodic. Likewise the integrals required for the T-mixing property (5) will not in general converge but may also be periodic functions of the integration time. Rotating systems may however be weak T-mixing (6).

In a T-mixing system there can be no non-trivial constants of the motion other than those inherent in the Hamiltonian symmetries. If there were such constants we could form transient time correlation functions that violated Eqs. (5) and (6). The fixed values of the various constants of the motion must be chosen to provide an inertial coordinate frame, within which we can construct a Hamiltonian dynamical system.

All T-mixing systems are ergodic over the ostensible phase space because if the phase space broke up into non-intersecting phase space subdomains characterized by different macroscopic averages for smooth phase functions, we could form constants of the motion depending on whether a system was on one subdomain or another. These subdomain

occupation numbers could then be substituted as $A(\Gamma)$ in (5) thereby violating the T-mixing condition.

If the relevant time correlation functions (5) decay asymptotically as t^{-1} or more slowly, the system may be weak T-mixing (6) but cannot be T-mixing (5). In contradistinction to mixing (4), if a system is T-mixing (5), it must relax to a stationary state at long times, whether this state is characterized by a smooth probability density $f(\Gamma; \infty)$ or not. If a system is weak T-mixing, but not T-mixing, relaxation to an invariant state from a non-invariant initial state, will not occur.

In general, it is exceedingly difficult to prove that a given system is mixing and perhaps even harder to prove whether it is T-mixing. However, because of the many properties of T-mixing systems it is easy to perform numerical/experimental tests of whether a system is T-mixing.

E. Relaxation to microcanonical equilibrium

The standard proof of relaxation for autonomous Hamiltonian systems from ergodic theory begins by computing the time dependent average of an integrable phase function $A(\Gamma)$

$$\begin{aligned} \langle A \rangle_t &= \int d\Gamma A(\Gamma) f(\Gamma; t) \\ &= \int d\Gamma A(\Gamma) f(S^{-t}\Gamma; 0) \\ &= \int d\Gamma A(S^t\Gamma) f(\Gamma; 0), \end{aligned} \quad (7)$$

where the second and third line follow from the fact that the dynamics is Hamiltonian, hence coordinate changes from Γ to $S^t\Gamma$ or to $S^{-t}\Gamma$ have unitary Jacobian determinants. In (7), stationarity is not assumed.

Now we multiply and divide the last expression in (7) by the (necessarily finite) volume of the phase space. This casts the first line in a form for which the mixing property can be used

$$\begin{aligned} \langle A \rangle_t &= \frac{1}{\int_D d\Gamma} \cdot \int_D d\Gamma A(S^t\Gamma) f(\Gamma; 0) \cdot \int_D d\Gamma \\ &= \langle A(S^t\Gamma) f(\Gamma; 0) \rangle_{\mu C} \cdot \int_D d\Gamma. \end{aligned} \quad (8)$$

In order to derive (8), the ostensible phase space volume needs to be finite.

A few words need to be said about $\langle A(S^t\Gamma) f(\Gamma; 0) \rangle_{\mu C}$. This function is an equilibrium microcanonical cross time correlation function. It results from the fact that for Hamiltonian dynamics, any time dependent nonequilibrium ensemble average, say $\langle A \rangle_t$, equals a time dependent nonequilibrium average $\langle A(S^t\Gamma) \rangle_0$ computed with respect to the initial distribution $f(\Gamma; 0)$. We now take the long time limit

$$\begin{aligned} \lim_{t \rightarrow \infty} \langle A \rangle_t &= \langle A(\Gamma) \rangle_{\mu C} \langle f(\Gamma; 0) \rangle_{\mu C} \cdot \int_D d\Gamma \\ &= \langle A(\Gamma) \rangle_{\mu C} \frac{1}{\int_D d\Gamma} \int_D d\Gamma f(\Gamma; 0) \cdot \int_D d\Gamma \\ &= \langle A(\Gamma) \rangle_{\mu C} \cdot 1 = \langle A(\Gamma) \rangle_{\mu C}, \end{aligned} \quad (9)$$

where we have used the mixing assumption (4), which allows us to factorize the invariant (microcanonical) time correlation function into a product of two invariant (microcanonical) averages. Finally, we use the normalization of the initial distribution function. Here, we do not need to assume the existence of a stationary state, since the microcanonical distribution is indeed invariant for Hamiltonian dynamics.

So $\langle A \rangle_t$ tends to a microcanonical average, whatever phase function, A , or initial probability density $f(\Gamma; 0)$ one considers. By definition, this amounts to a proof of relaxation to the microcanonical equilibrium state denoted by $\langle \cdot \rangle_{\mu C}$.

There is one crucial requirement in the proof above: the proof cannot be extended to thermostatted dissipative systems because the asymptotic steady state would be singular, and have no smooth density. If the system is subject to a time independent dissipative external field and no thermostat is applied, then the total Hamiltonian is constant and if the system is mixing the derivation above still applies. Finite mixing Hamiltonian systems, regardless of whether external or only internal fields are applied, ultimately relax to microcanonical equilibrium.

The mixing condition (4) explicitly requires stationary measures. In Eq. (8), stationarity of the microcanonical ensemble is indeed guaranteed by the Hamiltonian dynamics, provided the ostensible phase space has a finite volume. In that case, stationarity is therefore not a separate logical assumption, as is invoked in (4) above. The fact is that the microcanonical ensemble constitutes an invariant probability density for bounded Hamiltonian dynamics, as can be easily checked by substituting the microcanonical distribution into the phase continuity equation (usually incorrectly attributed to Liouville). It constitutes a stationary density even if the system is not mixing or ergodic and relaxation to stationarity does not occur. If an autonomous Hamiltonian system is prepared in a microcanonical state, it will remain in that state forever.

F. Relaxation under strong T-mixing

We now give a proof of relaxation to the stationary state based on the strong T-mixing condition (5). From the T-mixing assumption, there can be no constants of the motion other than the trivial ones, the internal energy, H_0 and the linear and angular momenta, \mathbf{P}, \mathbf{L} which are assumed to take on fixed values of $E, \mathbf{0}, \mathbf{0}$, respectively. As mentioned above, if the Hamiltonian has further symmetries we can extend the theory in the obvious way by fixing any of those additional constants of the motion at particular values. As will soon become obvious, this derivation reveals many additional details about the relaxation process. Within our ostensible phase space domain D (see Eq. (3)), there are no constants of the motion. If there were other constants of the motion transient time correlation functions with these variables would never decay, thereby violating the T-mixing condition.

If we consider any deviation from the microcanonical form (3) generated by a real-valued smooth deviation

function, $g(\Gamma)$ that is even in the momenta and differentiable,

$$f_g(\Gamma) = \frac{\exp[-g(\Gamma)]\delta(H_0(\Gamma) - E)\delta(\mathbf{P})\delta(\mathbf{L})}{\int d\Gamma \exp[-g(\Gamma)]\delta(H_0(\Gamma) - E)\delta(\mathbf{P})\delta(\mathbf{L})}, \quad (10)$$

the dissipation function will not vanish and we would have

$$\Omega(\Gamma) = \dot{g}(\Gamma), \quad (11)$$

where $\dot{g}(\Gamma) \equiv \dot{\Gamma} \cdot \partial g(\Gamma)/\partial \Gamma$ denotes the time derivative. The second law inequality³ states that the ensemble average of the time integral of the dissipation from 0 to some time t is non-negative for all values of t . It is only equal to zero if the system is at equilibrium. Thus, for finite values of the deviation function g we have

$$\langle \Omega_t \rangle_0 = \langle g(S^t \Gamma) - g(\Gamma) \rangle_0 > 0, \quad g(\Gamma) \neq 0, \quad \forall t > 0. \quad (12)$$

Thus, if there is any deviation from the equilibrium distribution (3), the dissipation function will not vanish (because there are no other constants of the motion) and further, the ensemble average of the time integrated dissipation function must be positive. It is impossible for it to vanish because of cancellation, with some areas of phase space having positive $\Omega_t(\Gamma)$, and other areas negative. This is obviously impossible

because of the Evans-Searles transient fluctuation theorem. This means that for T-mixing systems, the equilibrium distribution function is unique and given by Eq. (3).

One can prove that the system must relax to equilibrium by using the T-mixing property and the dissipation theorem for the dissipation function itself. We do not give those general arguments here as they have been given before.⁸ Also below we repeat these arguments for a particular form of initial distribution rather than the general form in Ref. 8. One key point about the relaxation process is that it is not necessarily monotonic. This is very different from the Boltzmann H-theorem applied to uniform ideal gases.

G. Proof of the zeroth “law”

In order to prove the zeroth “law” of thermodynamics consider a system with different temperatures in its left and right sides. We let $\Delta\beta$ denote the difference in the reciprocal absolute temperatures of the two bodies left L , and right R , divided by Boltzmann’s constant. The absolute temperature of each body, T_L, T_R is not known. Only the reciprocal difference: $\Delta\beta \equiv 1/k_B T_R - 1/k_B T_L$ is known. The absolute temperature of each body requires knowledge of the equation of state for the system. Thus, we consider an initial distribution of the form

$$\begin{aligned} f(\Gamma; 0) &= \frac{\exp[\Delta\beta H_L(\Gamma)]\delta(H_0(\Gamma) - E)\delta(\mathbf{P})\delta(\mathbf{L})}{\int d\Gamma \exp[\Delta\beta H_L(\Gamma)]\delta(H_0(\Gamma) - E)\delta(\mathbf{P})\delta(\mathbf{L})}, \quad \Delta\beta \neq 0 \\ &= \frac{\exp[-\Delta\beta H_R(\Gamma)]\delta(H_0(\Gamma) - E)\delta(\mathbf{P})\delta(\mathbf{L})}{\int d\Gamma \exp[-\Delta\beta H_R(\Gamma)]\delta(H_0(\Gamma) - E)\delta(\mathbf{P})\delta(\mathbf{L})}, \quad \Delta\beta \neq 0, \end{aligned} \quad (13)$$

where

$$H_C = \sum_{i \in C} \left[\frac{p_i^2}{2m} + \frac{1}{2} \sum_j \phi_{i,j} \right], \quad C = L, R, \quad (14)$$

where L, R denote the left or right bodies in Maxwell’s statement of the zeroth “law” and $\phi_{i,j}$ is the potential energy of particles i, j . Clearly, $H_0(\Gamma) = H_L(\Gamma) + H_R(\Gamma)$ and, in contradistinction to common notation, the interaction energy between the two bodies, is accounted for within the two sub-Hamiltonians of our body. We have assumed that there are only pair interactions. We could extend the theory to include many body interactions but this would only increase the complexity of the argument without revealing any more physics.

If there were no interactions between the two subsystems L, R (i.e., $\phi_{i,j} = 0, \forall i \in L, j \in R$) these two parts would remain in separate microcanonical equilibria indefinitely. Such a system would not as a whole be T-mixing. Switching on the interactions between the two subsystems means that the initial system is not in thermodynamic equilibrium and on average

generates future states with positive, time averaged dissipation function.

From (13) the deviation function is

$$g(\Gamma) = -\Delta\beta H_L(\Gamma) = \Delta\beta H_R(\Gamma). \quad (15)$$

If the reciprocal difference is zero the system is isothermal and is at equilibrium. So Eq. (13) provides a convenient mathematical model to study thermal relaxation and hence give a proof of Maxwell’s zeroth “law.”

The two quantities in (15) differ by a constant $\Delta\beta H_0$ but this constant has no physical relevance. The instantaneous dissipation function, $\dot{g}(\Gamma)$, is

$$\Omega(\Gamma) = -\Delta\beta \dot{H}_L(\Gamma) = \Delta\beta \dot{H}_R(\Gamma). \quad (16)$$

In deriving this equation we have used the fact that the energy $H_0 = H_L + H_R$ is a constant of the motion. From the dissipation theorem the time evolution of the phase space distribution function is given by¹³

$$f(\Gamma; t) = f(\Gamma; 0) \exp\left(-\int_0^t ds \Omega(S^s \Gamma)\right). \quad (17)$$

Therefore, if the system is at equilibrium, where the phase space density does not evolve, $f(\mathbf{\Gamma}; t) = f(\mathbf{\Gamma}; 0)$, and it follows that $\Omega(\mathbf{\Gamma}) = 0$. Indeed, the latter equality is our mathematical definition of an equilibrium distribution.

The second law inequality shows that the time integral of the dissipation function satisfies the inequality

$$\begin{aligned} \langle \Omega_t(\mathbf{\Gamma}) \rangle_0 &= -\langle \Delta\beta \Delta H_L(S^t \mathbf{\Gamma}) \rangle_0 \\ &= \langle \Delta\beta \Delta H_R(S^t \mathbf{\Gamma}) \rangle_0 > 0, \quad \forall t > 0, \end{aligned} \quad (18)$$

where $\Delta H_L(S^t \mathbf{\Gamma}) \equiv H_L(S^t \mathbf{\Gamma}) - H_L(\mathbf{\Gamma})$. If the left side of the system is hotter than the right $T_L > T_R$, and $\Delta\beta > 0$ in order for Eq. (18) to be satisfied, $\langle \Delta H_L(S^t \mathbf{\Gamma}) \rangle_0$ must be negative meaning that the left hand side loses heat energy to the right. This is in accord with the second “law” of thermodynamics. The hotter side of the system loses heat to the cooler.

Substituting into the dissipation theorem for averages gives¹³

$$\langle H_L(S^t \mathbf{\Gamma}) \rangle_0 = \langle H_L(\mathbf{\Gamma}) \rangle_0 - \Delta\beta \int_0^t ds \langle \dot{H}_L(\mathbf{\Gamma}) H_L(S^s \mathbf{\Gamma}) \rangle_0. \quad (19)$$

Since the system is T-mixing, in the long time limit the integral of the correlation function converges to a finite value and the average energy of the left and right sides of the system become constant in time. If the average energy of the left hand body is constant the instantaneous dissipation must have a zero average value,

$$\lim_{t \rightarrow \infty} \Delta\beta \langle \dot{H}_L(S^t \mathbf{\Gamma}) \rangle_0 = 0. \quad (20)$$

The whole system must have relaxed to its unique equilibrium distribution (3). This is because, as we have already seen, any deviation from this distribution must produce a positive value for the time integral of the ensemble averaged dissipation function (12) and (16). The temperatures of the left and right hand sides of the system must be equal because the unique equilibrium distribution, Eq. (3), is spatially isotropic. This completes our proof of the zeroth “law” of thermodynamics.¹⁶

H. Thermal conduction

Since there is no flux of particles between the two regions or bodies and if the boundary between the two bodies has a cross sectional area σ , the energy change is simply related to a heat flux $J_{Q,R}$, from the appropriate side of the system

$$\dot{H}_L \equiv J_{Q,R} \sigma \quad (21)$$

(outward normal convention is used) and we can write the dissipation theorem for the heat flux as

$$\langle J_{Q,R}(S^t \mathbf{\Gamma}) \rangle_0 = -\Delta\beta \sigma \int_0^t ds \langle J_{Q,R}(\mathbf{\Gamma}) J_{Q,R}(S^s \mathbf{\Gamma}) \rangle_0. \quad (22)$$

Note that $\langle J_{Q,R}(\mathbf{\Gamma}) \rangle_0 = 0$, because of the form of (8) which ensures that all functions that are odd in the momentum (including all fluxes of non-conserved quantities) are zero at time zero. Equation (22) is obviously a form of Fourier’s

“law” for heat flow. In fact it gives, for the first time, an exact expression for the nonlinear far from equilibrium, thermal conductivity. Indeed, the magnitude of $\Delta\beta$ has not been specified, in our derivation, and can be arbitrarily large or small. Fourier’s “law” for heat flow only relates to the linear response regime close to equilibrium. It is a linear constitutive relation. Previous¹⁷ time correlation expressions for the thermal conductivity were also limited to the linear response regime close to equilibrium. In (20) above, the transient time correlation integral is dependent on the size of the temperature gradient (difference). In the linear response regime, we take the weak gradient limit of the transient time correlation function, namely, the corresponding equilibrium time correlation function.

If the temperature difference is converted into a temperature gradient and if the heat capacity of the systems is large relative to the heat fluxes, this equation gives an expression for the nonlinear thermal conductivity $J_{Q,L} \equiv \lambda(\Delta\beta)\Delta\beta$ of the pseudo steady state that develops initially,

$$\lambda(\Delta\beta) = \sigma \int_0^{t_c} ds \langle J_Q(\mathbf{\Gamma}) J_Q(S^s \mathbf{\Gamma}) \rangle_0, \quad (23)$$

where t_c is the convergence time for the pseudo steady state. In the weak gradient limit, Eq. (23) is consistent with the Green-Kubo relations for thermal conductivity. [Note that in our system the heat flux appearing in our correlation functions is defined in terms of the energy flux across a plane whereas the usual heat flux appearing in Green-Kubo expressions is defined over a volume. Also the J_Q appearing in (23) can be either the left or right fluxes as the formula is symmetric.] However for our system, because the total energy is conserved but the energies of each of the two regions are not separately conserved, and our T-mixing system eventually relaxes to equilibrium, the heat flux eventually goes to zero

$$\lim_{t \rightarrow \infty} \langle J_{Q,R}(S^t \mathbf{\Gamma}) \rangle_0 = -\Delta\beta \sigma \int_0^\infty ds \langle J_{Q,R}(\mathbf{\Gamma}) J_{Q,R}(S^s \mathbf{\Gamma}) \rangle_0 = 0. \quad (24)$$

In this equation all dynamics is Newtonian and the initial distribution is the initial nonequilibrium distribution (13). It is valid when the initial state is arbitrarily far from equilibrium. In the far from equilibrium regime, the time correlation function is not an equilibrium time correlation function but rather is a TTCF (nonequilibrium) that is dependent on the magnitude of the initial temperature difference.

III. CONCLUSION

We have given a direct mathematical proof of the zeroth “law” of thermodynamics. The form of the zeroth “law” given by Prigogine¹⁸ is closest to that used in our proof: “If a system A is in equilibrium with system B and if system B is in equilibrium with system C, then it follows that system A is in equilibrium with system C. This transitivity of the state of equilibrium is sometimes called the zeroth “law.” Thus, equilibrium systems have one uniform temperature. . . .”

Our proof is constructed using an array of previous results. Our proof of the zeroth “law” is far more informative

than the corresponding derivation using ergodic theory. Combining the present proof with the observation that for an isolated mechanical system the energy is constant and our recent proof⁶ of the Clausius inequality, we see that all the so-called “laws” of classical thermodynamics are mathematical results provable from the laws of mechanics, supplemented by the axiom of causality,² and by the T-mixing condition. This might be regarded as changing the logical status of thermodynamics.

A second less obvious result of our work is that for nonequilibrium systems entropy seems to play no role at all! Its place is taken by dissipation. The idea that one could use the Gibbs entropy in proving relaxation to equilibrium is obviously erroneous, as discussed in a range of papers.¹⁹ In the first place, away from equilibrium, the quantity known as the Gibbs entropy does not represent any physical entropy at all. In fact, as was known to Gibbs,²⁰ the fine grained Gibbs entropy is a constant of the motion for autonomous Hamiltonian systems.^{20,21} It therefore cannot increase to a maximum in microcanonical equilibrium. It is just a constant of the motion. For thermostatted driven systems in nonequilibrium steady states, the fine grained Gibbs entropy diverges to negative infinity for any nonzero value of the dissipative field no matter how small. The usual coarse-grained Gibbs entropy has further unphysical properties.²¹

The present work points out, however, that entropy is not really necessary away from equilibrium. It is only at, or very near to, equilibrium when dissipation is identically zero or so small that local thermodynamic equilibrium can be assumed, that entropy may be useful. One may resort to other dynamical notions of entropy, cf. the Boltzmann entropy, or the one defined in the third paper of Ref. 21 to try to avoid these problems.

Our proof also provides a simple derivation of the transient time correlation function expression for the nonlinear thermal conductivity as well as the limiting Green-Kubo form for the linear response.¹⁷ Further, since the total energy of the left half of our system is not a constant of the motion, we have the now familiar result that the integral of the left or right hand heat flux autocorrelation function integrates to zero.^{12,22,23} This shows the deep connection between this, Zwanzig and Berne, Boon and Rice (ZBBR) integration rule^{12,22,23} (such as (24)) and the equilibrium relaxation theorem.

To give a simple generalized derivation of the ZBBR sum rule, let $g(\Gamma)$ be some non-conserved phase function, that is, like the Hamiltonian, an even function of the momenta.²⁴ We assume the dynamics is Newtonian and the total linear and angular momentum are zero. We know that among the possible initial distributions, Eq. (3) gives the unique equilibrium distribution. If at time zero the initial distribution is given by

$$f_{\lambda g}(\Gamma) = \frac{\exp[-\lambda g(\Gamma)]\delta(H_0(\Gamma) - E)\delta(\mathbf{P})\delta(\mathbf{L})}{\int d\Gamma \exp[-\lambda g(\Gamma)]\delta(H_0(\Gamma) - E)\delta(\mathbf{P})\delta(\mathbf{L})}, \quad (25)$$

where $\lambda > 0$ is a scaling parameter used to control the amplitude of the deviation from equilibrium. We do not assume we are close to equilibrium; λ is arbitrary in magnitude. The instantaneous dissipation function is given by (11). Since the

system is T-mixing we know that the system will relax to equilibrium and therefore from the T-mixing property and the dissipation theorem

$$\lim_{t \rightarrow \infty} \langle g(S^t \Gamma) \rangle_{\lambda g} = \text{const}, \quad (26)$$

where $\langle \rangle_{\lambda g}$ is the ensemble average evaluated using the distribution function (22). This in turn means that $\lim_{t \rightarrow \infty} \langle \dot{g}(S^t \Gamma) \rangle_{\lambda g} = 0$. Applying the dissipation theorem to the instantaneous dissipation function itself yields

$$\lim_{t \rightarrow \infty} \langle \dot{g}(S^t \Gamma) \rangle_{\lambda g} = \lambda \int_0^\infty ds \langle \dot{g}(\Gamma) \dot{g}(S^s \Gamma) \rangle_{\lambda g} = 0, \quad \forall \lambda > 0. \quad (27)$$

The equation represents the heat death of the Universe as discussed in the late 19th century by Lord Kelvin and others. Any flux $\dot{g}(\Gamma)$, of a non-conserved quantity $g(\Gamma)$ in an autonomous, T-mixing, Hamiltonian system will eventually average to zero. In 1852, Kelvin wrote²⁵ a paper with the title “On the universal tendency in Nature to the dissipation of mechanical energy.” Equation (27) and the discussion above gives a mathematical proof of precisely the same proposition: a mathematical proof of heat death in autonomous Hamiltonian systems assuming the laws of mechanics, the axiom of causality, and the T-mixing property.

The ZBBR integration rule is obtained in the weak deviation limit where we employ Hamiltonian dynamics and an equilibrium microcanonical initial distribution

$$\int_0^\infty ds \langle \dot{g}(\Gamma) \dot{g}(S^s \Gamma) \rangle_0 = 0. \quad (28)$$

It is the final relaxation to equilibrium that means that the long-time flux goes to zero. This in turn means that the integral of the correlation function must be zero.^{12,22,23} This provides an almost trivial derivation of the ZBBR integration rule. If the system is not T-mixing, the integration rule (27) and (28) cannot be expected to hold for all non-conserved phase functions.

Finally, we should make another observation. If a system satisfies T-mixing, then by definition, infinite-time integrals of the ensemble average of the dissipation function at time zero with any integrable phase function at some later time, are convergent and therefore long-time averages of those phase functions are guaranteed to be stationary. T-mixing systems always evolve to stationary states at long times.

ACKNOWLEDGMENTS

We would like to thank Professor D. J. Searles, Ms. Charlotte F. Petersen, and Mr. Richard Skelton for useful suggestions. We would also like to thank the Australian Research Council for support of this research. L.R. thanks the European Research Council, for funding under the European Community’s (EC) Seventh Framework Programme (FP7/2007-2013)/ERC Grant Agreement No. 202680. The EC is not liable for any use that can be made on the information contained herein.

¹D. J. Evans, E. G. D. Cohen, and G. P. Morriss, *Phys. Rev. Lett.* **71**, 2401 (1993); **71**, 3616 (1993); D. J. Evans and D. J. Searles, *Phys. Rev. E* **50**, 1645 (1994).

- ²D. J. Evans and D. J. Searles, *Adv. in Phys.* **51**, 1529 (2002).
- ³D. J. Searles and D. J. Evans, *Aust. J. Chem.* **57**, 1119 (2004).
- ⁴G. E. Crooks, *J. Stat. Phys.* **90**, 2690 (1998).
- ⁵C. Jarzynski, *Phys. Rev. Lett.* **78**, 2690 (1997).
- ⁶D. J. Evans, S. R. Williams, and D. J. Searles, *J. Chem. Phys.* **134**, 204113 (2011).
- ⁷J. C. Maxwell, *Theory of Heat*, 3rd ed. (Longmans, Green, London, 1872), p. 32.
- ⁸D. J. Evans, D. J. Searles, and S. R. Williams, "A simple mathematical proof of Boltzmann's equal *a priori* probability hypothesis," in *Diffusion Fundamentals III*, edited by C. Chmelik, N. Kanellopoulos, J. Karger, and D. Theodorou (Leipziger Universitätsverlag, Leipzig, 2009), pp. 367–374.
- ⁹D. J. Evans, D. J. Searles, and S. R. Williams, "Fluctuation relations applied to an optical trapping system" in *Nonequilibrium Statistical Physics of Small Systems*, Nonlinear Dynamics of Nanosystems Vol. 4, edited by R. Klages, W. Just, and C. Jarzynski (Wiley-VCH, Weinheim, to be published).
- ¹⁰The term T-mixing was first defined in S. R. Williams and D. J. Evans, *J. Chem. Phys.* **132**, 184105 (2010).
- ¹¹D. J. Evans, D. J. Searles, and S. R. Williams, *J. Stat. Mech.: Theory Exp.* **2009**, P07029.
- ¹²R. Zwanzig, *J. Chem. Phys.* **40**, 2527 (1963); B. J. Berne, J. P. Boon, and S. A. Rice, *ibid.* **45**, 1086 (1966).
- ¹³D. J. Evans, D. J. Searles, and S. R. Williams, *J. Chem. Phys.* **128**, 014504 (2008); **128**, 249901 (2008).
- ¹⁴For some nonequilibrium systems slightly different time reversal maps may need to be used (e.g., shear flow).
- ¹⁵D. J. Searles, L. Rondoni, and D. J. Evans, *J. Stat. Phys.* **128**, 1337 (2007); L. Rondoni and C. Mejia-Monasterio, *Nonlinearity* **20**, R1 (2007).
- ¹⁶The term "zerth law" was first introduced by R. Fowler and E. A. Guggenheim, *Statistical Thermodynamics: A Version of Statistical Mechanics for Students of Physics and Chemistry* (Cambridge University Press, Cambridge, England, 1939), p. 56.
- ¹⁷R. Zwanzig, *Ann. Rev. Phys. Chem.* **16**, 67 (1965).
- ¹⁸D. Kondepudi and I. Prigogine, *Modern Thermodynamics: From Heat Engines to Dissipative Structures* (Wiley, Chichester, 1999), p. 6.
- ¹⁹J. L. Lebowitz, *Phys. Today* **46**(9), 32 (1993); L. Rondoni and E. G. D. Cohen, *Physica D* **168–169**, 341 (2002).
- ²⁰J. W. Gibbs, *Elementary Principles in Statistical Mechanics* (Yale University Press, New Haven, CT, 1902); see also P. Ehrenfest and T. Ehrenfest, *The Conceptual Foundations of the Statistical Approach in Mechanics* (Dover, New York, 1990).
- ²¹D. J. Evans, S. R. Williams, and D. J. Searles, *J. Chem. Phys.* **135**, 194107 (2011); see also M. Falcioni, L. Palatella, S. Pigolotti, L. Rondoni, and A. Vulpiani, *Physica A* **385**, 170 (2007); and for driven systems: D. J. Evans and L. Rondoni, *J. Stat. Phys.* **109**, 895 (2002).
- ²²D. J. Evans and H. J. M. Hanley, *Phys. Rev. A* **25**, 1771 (1981).
- ²³D. J. Evans, *Phys. Rev. A* **23**, 2622 (1981).
- ²⁴Note that it is certainly possible to treat distribution functions that are not even functions of the momenta. If the system moves relative to the observer care needs to be exercised on defining thermodynamic quantities. If the distribution function is even in the momenta these matters are much simpler.
- ²⁵W. T. Thompson (Lord Kelvin), Proc. R. Soc. Edinburgh article 2, April 19, page 139 (1852).