

Theory for non-equilibrium statistical mechanics

Phil Attard*

Received 23rd March 2006, Accepted 8th May 2006

First published as an Advance Article on the web 7th July 2006

DOI: 10.1039/b604284h

This paper reviews a new theory for non-equilibrium statistical mechanics. This gives the non-equilibrium analogue of the Boltzmann probability distribution, and the generalization of entropy to dynamic states. It is shown that this so-called second entropy is maximized in the steady state, in contrast to the rate of production of the conventional entropy, which is *not* an extremum. The relationships of the new theory to Onsager's regression hypothesis, Prigogine's minimal entropy production theorem, the Langevin equation, the formula of Green and Kubo, the Kawasaki distribution, and the non-equilibrium fluctuation and work theorems, are discussed. The theory is worked through in full detail for the case of steady heat flow down an imposed temperature gradient. A Monte Carlo algorithm based upon the steady state probability density is summarized, and results for the thermal conductivity of a Lennard-Jones fluid are shown to be in agreement with known values. Also discussed is the generalization to non-equilibrium mechanical work, and to non-equilibrium quantum statistical mechanics. As examples of the new theory two general applications are briefly explored: a non-equilibrium version of the second law of thermodynamics, and the origin and evolution of life.

1. Introduction

Equilibrium statistical mechanics is based upon the Boltzmann distribution,

$$\mathcal{P}_i = \frac{1}{Z} e^{-E_i/k_B T}. \quad (1.1)$$

This gives the probability of finding a sub-system in a microstate i when the sub-system is in thermal equilibrium with a heat reservoir of temperature T , (E_i is the microstate energy, k_B is Boltzmann's constant, and Z is the normalizing partition function). According to Feynman,¹ "This fundamental law is the summit of statistical mechanics, and the entire subject is either the slide-down from this summit, as the principle is applied to various cases, or the climb-up to where the fundamental law is derived."

Few would disagree that the Boltzmann distribution plays such a central role in equilibrium statistical mechanics. However, a fundamental limitation of the Boltzmann distribution is that it is not applicable to non-equilibrium systems, (at least not without making some sort of quasi-equilibrium approximation). There is a pressing need to develop an analogous probability distribution for non-equilibrium systems.

Such dynamic or time-dependent systems comprise the vast majority of biological, technological, and even cosmological systems. One can broadly classify non-equilibrium systems based on whether the phenomenon is transient, harmonic, or steady. Often there is an applied, time-dependent field, which might be mechanical (such as an applied electric or magnetic field, or a mechanical force), or thermodynamic (such as an applied temperature or concentration gradient).

Of the various time-dependent systems, it is the steady state non-equilibrium systems that appear to offer the simplest path to generalizing the Boltzmann distribution. The canonical equilibrium case is a sub-system in equilibrium with a heat reservoir of temperature T . The closest non-equilibrium analogue is steady heat flow, where a sub-system is in thermal contact with two heat reservoirs of different temperatures. Like an equilibrium system, in the steady state the sub-system itself shows no net macroscopic change over time. Unlike an equilibrium system, the surroundings (in this case the thermal reservoirs) do display a net change over time.

This is the essential similarity and essential difference between an equilibrium system and a steady state system. One can immediately conclude two things about the non-equilibrium probability distribution for a steady state: because it describes the sub-system it should not depend explicitly on time, and because it describes a non-equilibrium system it should depend upon the direction of time.

The arrow of time has excited debate ever since the birth of statistical mechanics: "How does the irreversible behavior of the world around us arise from the reversible nature of the microscopic equations of motion?" Nowadays few would find a paradox here, citing in explanation the second law of thermodynamics, namely that over time systems tend to move in the direction of increasing entropy. The absence or presence of an arrow of time distinguishes equilibrium from non-equilibrium systems. This is reflected in the Boltzmann distribution, which is insensitive to the direction of time. In contrast, a non-equilibrium probability distribution, even the steady state distribution that cannot depend explicitly on time, must differ from the Boltzmann distribution in that it depends on the direction of time.

Onsager, in deriving the reciprocal relations for the transport coefficients, focused attention on the essential role that

School of Chemistry F11, University of Sydney, NSW 2006, Australia. E-mail: attard@chem.usyd.edu.au

time reversibility plays in non-equilibrium statistical mechanics.² Inspired by his work, over several years I have been developing a theory for non-equilibrium statistical mechanics that has been presented in a series of papers.^{3–7} This article highlights the more significant aspects of these recent advances, including the notion of second entropy for transitions and its maximization, (sections 3.2, 4.3, and 4.6.4), the non-equilibrium probability distributions, (sections 4.4, 6.1, and 7), the non-equilibrium Monte Carlo algorithms, (section 5), and the microscopic transition theorems, (sections 4.6.3 and 6.2). The purpose of the article is to present a consistent, self-contained account of the theory, and to attempt to provide clearer and more rigorous justification for it than has been possible to date, with perhaps a little more pedagogy than is usual in the research literature. In doing so I have avoided giving a comprehensive review of the literature of non-equilibrium statistical mechanics, and instead have restricted references to a few key papers that intersect with the present results.

The article is organized like this: Section 2 introduces the dynamical reversibility of equilibrium systems. Section 3 deals with thermodynamic fluctuations and the transitions between thermodynamic states. Described in detail is the connection between the so-called second entropy, which describes such transitions, Onsager's regression hypothesis, and the Green–Kubo formulae for the transport coefficients. Section 4 narrows the focus to the particular case of steady heat flow down an imposed temperature gradient. The rate of second entropy production is shown to be maximal in general. Here also is given the non-equilibrium phase space probability density that generalizes the Boltzmann distribution, and utilizes the so-called mirror work, which depends upon the direction of time and which therefore breaks the symmetry characteristic of equilibrium systems. Section 5 develops two Monte Carlo algorithms for heat flow. One is based upon the non-equilibrium phase space probability density and uses the Metropolis scheme with umbrella sampling, and the other is based upon the transition probability and is a type of stochastic molecular dynamics algorithm. Section 6 generalizes the results to systems subjected to non-equilibrium work, and presents the appropriate non-equilibrium probability distribution and certain transitions theorems for these. Section 7 indicates briefly how the approach may be applied to non-equilibrium quantum systems. Section 8 discusses two applications of the non-equilibrium theory: a dynamic version of the second law of thermodynamics, and an account for the origin of life and the direction of evolution. The conclusion, section 9, summarizes the main results and indicates possibilities for future work.

2. Dynamical reversibility

2.1. Hamilton's equations

Let us begin with the time reversible nature of the microscopic equations of motion and of equilibrium states. Consider an isolated sub-system comprising N particles, with \mathbf{q}_i and \mathbf{p}_i being the position and momentum of particle i , respectively.

Hamilton's equations of motion are

$$\dot{\mathbf{q}}_i = \frac{\partial \mathcal{H}}{\partial \mathbf{p}_i}, \quad \dot{\mathbf{p}}_i = -\frac{\partial \mathcal{H}}{\partial \mathbf{q}_i}, \quad (2.1)$$

where \mathcal{H} is the Hamiltonian, and the over-dot denotes the time derivative or velocity. A microstate of the sub-system is a point in its phase space, $\Gamma = \{\mathbf{q}^N, \mathbf{p}^N\}$. The conjugate point has all the momenta reversed,

$$\Gamma^\dagger = \{\mathbf{q}^N, (-\mathbf{p})^N\}. \quad (2.2)$$

This operation of velocity reversal plays a central role in the non-equilibrium theory, and is equivalent to reversing the direction of time. This follows from the equations of motion, at least in the usual case that the Hamiltonian is insensitive to the direction of the velocities, $\mathcal{H}(\Gamma) = \mathcal{H}(\Gamma^\dagger)$. (The theory can be extended to include magnetic fields and the velocity-dependent Lorentz force, but we do not do so here.)

The trajectory of the sub-system may be denoted $\Gamma(t|\Gamma_0)$, which gives the point at a time t after the system was at Γ_0 . If $\Gamma_1 = \Gamma(t|\Gamma_0)$, then the reversibility of the equations of motion show that

$$\Gamma_0 = \Gamma(-t|\Gamma_1), \text{ and } \Gamma_0^\dagger = \Gamma(t|\Gamma_1^\dagger). \quad (2.3)$$

The first says that we can arrive at the original starting point by running time backwards from the original end-point of the trajectory. Equivalently, the second says that the conjugate of the original starting point is the end-point of a trajectory going forward in time from the conjugate of the original end-point.

The trajectory may alternatively be written as a conditional transition probability,

$$\Lambda(\Gamma_1|\Gamma_0, t) \equiv \delta(\Gamma_1 - \Gamma(t|\Gamma_0)), \quad (2.4)$$

where a Dirac δ -function appears. Hence the time reversibility of the equations of motion may be written

$$\Lambda(\Gamma_1|\Gamma_0, t) = \Lambda(\Gamma_0|\Gamma_1, -t) = \Lambda(\Gamma_0^\dagger|\Gamma_1^\dagger, t). \quad (2.5)$$

2.2. Macrostates

Statistical mechanics requires the notions of entropy and macrostates.⁸ A macrostate of the system is a particular value of a function of phase space. (Here for notational simplicity a macrostate is taken to be defined by a single variable. Replacing the scalar by a vector allows the following to be applied to an arbitrary set of variables.) Alternatively, any function of phase space signifies a collective of macrostates. If $\mathcal{E}(\Gamma)$ is such a function, then the probability that the isolated sub-system is in the macrostate labelled a is

$$\wp(a|E) = \langle \delta(a - \mathcal{E}(\Gamma)) \rangle_E = W_E^{-1} \int_E d\Gamma \delta(a - \mathcal{E}(\Gamma)), \quad (2.6)$$

where $W_E = \int_E d\Gamma = \int d\Gamma \delta(E - \mathcal{H}(\Gamma))$ is the volume of the energy hypersurface to which the sub-system is confined.

Suppose the collective has a definite parity, $\mathcal{E}(\Gamma) = \sigma_\mathcal{E} \mathcal{E}(\Gamma^\dagger)$. A function that is insensitive to the direction of the velocity is said to have even parity, $\sigma_\mathcal{E} = 1$. A function that changes sign when the velocities are reversed, is said to have odd parity, $\sigma_\mathcal{E} = -1$. With this the unconditional transition probability between macrostates for the isolated system

satisfies

$$\begin{aligned}
 \mathcal{P}(a \leftarrow b|t, E) &= A(a|b, t)\mathcal{P}(b|E) \\
 &= W_E^{-1} \int_E d\Gamma_1 d\Gamma_2 \delta(a - \varepsilon(\Gamma_2)) \\
 &\quad \times \delta(b - \varepsilon(\Gamma_1)) \delta(\Gamma_2 - \Gamma(t|\Gamma_1)) \\
 &= W_E^{-1} \int_E d\Gamma_1^\dagger d\Gamma_2^\dagger \delta(a - \sigma_\varepsilon \varepsilon(\Gamma_2^\dagger)) \\
 &\quad \times \delta(b - \sigma_\varepsilon \varepsilon(\Gamma_1^\dagger)) \delta(\Gamma_1^\dagger - \Gamma(t|\Gamma_2^\dagger)) \\
 &= \mathcal{P}(a_e b \leftarrow \sigma_\varepsilon a|t, E).
 \end{aligned} \tag{2.7}$$

This uses the fact that $d\Gamma = d\Gamma^\dagger$. If the collective has even parity, this says that for an isolated system we shall observe the forward transition $a \rightarrow b$ as frequently as the reverse $b \rightarrow a$. In essence this is what Onsager meant by the principle of dynamical reversibility, which he stated as ‘in the end every type of motion is just as likely to occur as its reverse’. He also said that in consequence for every transition $A \rightarrow B$, ‘the direct reverse transition $B \rightarrow A$ must take place equally often’.²

2.3. Thermal equilibrium

Statistical mechanics depends upon the notion of reservoirs, which are necessarily infinitely larger than the sub-system and unaffected by exchange with the sub-system, which have infinite conductivity, and which are generic in nature and only enter the probability distribution *via* the values of their conjugate field variables. If the sub-system is brought into thermal contact with a heat reservoir of temperature T and allowed to equilibrate, then the probability of the sub-system being in a particular microstate obeys the Boltzmann distribution,⁸

$$\wp_{\text{eq}}(\Gamma|\beta) = Z(\beta)^{-1} e^{-\beta \mathcal{H}(\Gamma)}, \tag{2.8}$$

where $\beta \equiv 1/k_B T$ and Z is the partition function that normalizes the distribution to unity. The exponent of the Boltzmann distribution represents the part of the entropy of the reservoir that depends upon the sub-system.⁸ The subscript ‘eq’ emphasizes that this is an equilibrium result. As such, it must be insensitive to the direction of the velocities,

$$\wp_{\text{eq}}(\Gamma^\dagger|\beta) = \wp_{\text{eq}}(\Gamma|\beta), \tag{2.9}$$

as follows explicitly from the even parity of the hamiltonian.

The microstates Γ of the sub-system are in fact macrostates of the total system (since for each phase point of the sub-system there are many configurations of the molecules of the reservoir). Since the total system is isolated, then the transitions between these macrostates must obey the principle of dynamical reversibility, eqn (2.7). The parity of the collective Γ may be accounted for by noting that the positions have even parity and the momenta have odd parity, and so we may write the conjugate phase point as $\Gamma^\dagger = \{(\sigma_q \mathbf{q})^N, (\sigma_p \mathbf{p})^N\}$. Since the unconditional transition probability for the microstates of the sub-system in the presence of the stochastic perturbations from the thermal reservoir obeys the principle of dynamical

reversibility, it follows that this may be written as

$$\begin{aligned}
 \mathcal{P}_{\text{eq}}(\Gamma_2 \leftarrow \Gamma_1|t) &\equiv A(\Gamma_2|\Gamma_1, t)\mathcal{P}_{\text{eq}}(\Gamma_1|\beta) \\
 &= A(\Gamma_1^\dagger|\Gamma_2^\dagger, t)\mathcal{P}_{\text{eq}}(\Gamma_2^\dagger|\beta) \\
 &\equiv \mathcal{P}_{\text{eq}}(\Gamma_1^\dagger \leftarrow \Gamma_2^\dagger|t).
 \end{aligned} \tag{2.10}$$

The macrostate probability for a sub-system in thermal equilibrium with a reservoir is

$$\begin{aligned}
 \mathcal{P}_{\text{eq}}(a|\beta) &= \langle \delta(a - \varepsilon(\Gamma)) \rangle_{\text{eq}} \\
 &= \int d\Gamma \delta(a - \varepsilon(\Gamma)) \mathcal{P}_{\text{eq}}(\Gamma|\beta).
 \end{aligned} \tag{2.11}$$

The unconditional macrostate transition probability obeys dynamical reversibility,

$$\begin{aligned}
 \mathcal{P}_{\text{eq}}(a \leftarrow b|t) &= \int d\Gamma_1 d\Gamma_2 \delta(a - \varepsilon(\Gamma_2)) \\
 &\quad \times \delta(b - \varepsilon(\Gamma_1)) \mathcal{P}_{\text{eq}}(\Gamma_2 \leftarrow \Gamma_1|t) \\
 &= \int d\Gamma_1^\dagger d\Gamma_2^\dagger \delta(a - \sigma_\varepsilon \varepsilon(\Gamma_2^\dagger)) \\
 &\quad \times \delta(b - \sigma_\varepsilon \varepsilon(\Gamma_1^\dagger)) \mathcal{P}_{\text{eq}}(\Gamma_1^\dagger \leftarrow \Gamma_2^\dagger|t) \\
 &= \mathcal{P}(\sigma_\varepsilon b \leftarrow \sigma_\varepsilon a|t).
 \end{aligned} \tag{2.12}$$

2.4. Time correlation function

The functions of phase space are implicitly dependent upon time, and we may use the short-hand notation $A(t) \equiv A(\Gamma(t|\Gamma_0))$. The equilibrium time correlation function is defined as

$$\begin{aligned}
 C_{AB}(t) &\equiv \langle A(t)B(0) \rangle_{\text{eq}} \\
 &\equiv \langle A(\Gamma(t|\Gamma_0))B(\Gamma_0) \rangle_{\text{eq}} \\
 &= \int d\Gamma_1 d\Gamma_2 A(\Gamma_1)B(\Gamma_2) \mathcal{P}_{\text{eq}}(\Gamma_1 \leftarrow \Gamma_2|t) \\
 &= \sigma_A \sigma_B \int d\Gamma_1^\dagger d\Gamma_2^\dagger A(\Gamma_1^\dagger)B(\Gamma_2^\dagger) \mathcal{P}_{\text{eq}}(\Gamma_2^\dagger \leftarrow \Gamma_1^\dagger|t) \\
 &= \sigma_A \sigma_B \langle A(0)B(t) \rangle_{\text{eq}} \\
 &= \sigma_A \sigma_B C_{BA}(t).
 \end{aligned} \tag{2.13}$$

Of course, by time homogeneity of an equilibrium system we have $C_{AB}(t) = C_{BA}(-t)$. Note that here $\Gamma(t|\Gamma_0)$ is the trajectory perturbed by the stochastic interactions with the heat reservoir; it is *not* the adiabatic trajectory.

Note that we often choose the functions of phase space in such a way that their average value is zero. This can be accomplished by replacing A by $\delta A \equiv A - \langle A \rangle_{\text{eq}}$ and similarly for B . With this replacement the time correlation function goes to zero at long times, which is the formally exact behavior.

Since a velocity variable has opposite parity to the corresponding position variable, $\sigma_A = -\sigma_{\dot{A}}$, we have the result

$$\langle A(t)\dot{A}(0) \rangle_{\text{eq}} = - \langle \dot{A}(t)A(0) \rangle_{\text{eq}}. \tag{2.14}$$

This reveals the very important result that the velocity is instantaneously uncoupled from the position,

$$\langle A(0)\dot{A}(0) \rangle_{\text{eq}} = 0. \tag{2.15}$$

3. Fluctuations and their transitions

In 1931 Onsager derived the reciprocal relations, which show that the matrix of linear coefficients that couples non-equilibrium fluxes to thermodynamic driving forces is symmetric.² (For example, in a system with an applied electric field and temperature gradient, the cross-coefficient relating the resultant heat flux to the electric field is equal to the cross-coefficient relating the induced electric current to the temperature gradient.) His analysis was based on two ideas, namely dynamic reversibility, which was discussed above, and the regression hypothesis. The latter says that the restoring force that returns a spontaneous thermodynamic fluctuation to its equilibrium value is indistinguishable from an externally applied force that causes a flux in that variable. In other words, equilibrium fluctuations encompass all that is required to describe non-equilibrium fluxes.

3.1. First entropy

In ref. 4, the reciprocal relations for the diffusion constant of a Brownian particle were derived, and it was argued that the analysis carried over for thermodynamic fluctuations. Although the Brownian particle yields to a transparent and physically meaningful analysis, it is arguably better to analyze fluctuations directly rather than by analogy. The reader is referred to ref. 4 for the physical motivation for the various mathematical quantities introduced below.

Let \mathbf{q} represent the displacement of a set of thermodynamic observables from their equilibrium values, $\langle \mathbf{q} \rangle = 0$. The first entropy, which is the constrained entropy for a displacement, is to quadratic order,

$$S(\mathbf{q}) = S(0) + \frac{1}{2} \mathbf{q}^T \underline{\underline{S}} \mathbf{q}. \quad (3.1)$$

The exponential of this gives the probability, which is obviously Gaussian, and the correlation matrix is therefore given by

$$\underline{\underline{S}}^{-1} = -\langle \mathbf{q} \mathbf{q}^T \rangle_0 / k_B. \quad (3.2)$$

The following analysis is restricted to small excursions from equilibrium when the quadratic expansion is valid. It is possible to lift this restriction. In that case the full first entropy $S(\mathbf{q})$ may be used, and the first derivative vector is non-zero, and it and the second derivative matrix may be evaluated at the current value of \mathbf{q} rather than at zero.

3.2. Second entropy

Consider a time interval $|\tau|$ and denote the initial position by $\mathbf{q} \equiv \mathbf{q}(t)$ and the final position by $\mathbf{q}' \equiv \mathbf{q}(t + \tau)$. Position here means the value of the thermodynamic variables, and velocity will mean the rate of change of these variables. Define the coarse velocity by $\dot{\mathbf{q}} = (\mathbf{q}' - \mathbf{q})/\tau$. The coarse velocity is the time average of the instantaneous velocity $\dot{\mathbf{q}}$ over the interval.

The coarse velocity represents a collective of macrostates in the sense of section 2, since \mathbf{I} determines uniquely both $\mathbf{q}(t)$ and $\mathbf{q}(t + \tau)$, and we can define an entropy for it, $\sigma(\mathbf{q}', \mathbf{q}|\tau)$. Because of its relationship with dynamical transitions between pairs of macrostates this may be called the *dynamical entropy* or the *second entropy*.

By dynamic reversibility, eqn (2.12), $\wp_{\text{eq}}(\mathbf{q}' \leftarrow \mathbf{q}|\tau) = \wp_{\text{eq}}(\mathbf{q} \leftarrow \mathbf{q}'|\tau)$, and since the probability is the exponential of the entropy, $\wp_{\text{eq}}(\mathbf{q}' \leftarrow \mathbf{q}|\tau) \propto \exp \sigma(\mathbf{q}', \mathbf{q}|\tau)/k_B$, this says that the dynamical entropy is an even function of the coarse velocity and of time. Similarly, the time correlation function $\underline{\underline{C}}(\tau) \equiv \langle \mathbf{q}' \mathbf{q}^T \rangle_{\text{eq}}$ is an even function of τ , assuming here and below that \mathbf{q} has even parity.

The second entropy is just the generalization to dynamic states of the conventional (or first) entropy, which usually describes structure. Like the latter, the second entropy is a concave function of its arguments and so it has similar properties, including the fact that it is maximized by the physical values of any constraints, and that these states are stable. These physical values represent the most likely macrostate. In the following we shall deal with small fluctuations from the most likely state and approximate the second entropy by a quadratic form. We shall further specialize the analysis to the steady state. It should be understood however that the theory is more general than the present illustration of it.

The most general quadratic form for the dynamical entropy is

$$\sigma(\mathbf{q}', \mathbf{q}|\tau) = \text{const.} + \frac{1}{2} \mathbf{q}'^T \underline{\underline{A}}' \mathbf{q}' + \mathbf{q}'^T \underline{\underline{B}} \mathbf{q} + \frac{1}{2} \mathbf{q}^T \underline{\underline{A}} \mathbf{q}. \quad (3.3)$$

The principle of dynamical reversibility implies that this must be symmetric function of \mathbf{q}' and \mathbf{q} , which means that $\underline{\underline{A}}' = \underline{\underline{A}}$, and that $\underline{\underline{B}} = \underline{\underline{B}}^T$. (The matrix $\underline{\underline{A}} \equiv \partial^2 \sigma / \partial \mathbf{q} \partial \mathbf{q}^T$ is also symmetric.) These two matrices must be even functions of τ . In view of the symmetries the second entropy may be rewritten as

$$\begin{aligned} \sigma(\mathbf{q}', \mathbf{q}|\tau) = \text{const.} + \frac{1}{2} [\mathbf{q}' + \underline{\underline{A}}^{-1} \underline{\underline{B}} \mathbf{q}]^T \underline{\underline{A}} [\mathbf{q}' + \underline{\underline{A}}^{-1} \underline{\underline{B}} \mathbf{q}] \\ + \frac{1}{2} \mathbf{q}^T [\underline{\underline{A}} - \underline{\underline{B}} \underline{\underline{A}}^{-1} \underline{\underline{B}}] \mathbf{q}. \end{aligned} \quad (3.4)$$

A very strong condition is imposed by the reduction of the second entropy to the first, which is equivalent to the normalisation of the conditional transition probability. Since the collectives are complete sets of disjoint states, the weight of the macrostate \mathbf{q} is the sum of the weights of the intersections $(\mathbf{q}', \mathbf{q})$,

$$e^{S(\mathbf{q})/k_B} = \int d\mathbf{q}' e^{\sigma(\mathbf{q}', \mathbf{q}|\tau)/k_B}. \quad (3.5)$$

The trick to evaluating these sorts of Gaussian integrals is to complete the squares, which we did in eqn (3.4). The final term in that equation is independent of \mathbf{q}' and so it gives an exponential pre-factor for the integral. The Gaussian integral that remains gives an immaterial logarithmic constant contribution to the first entropy. Hence we obtain the very important condition that

$$\underline{\underline{S}} = \underline{\underline{A}}(\tau) - \underline{\underline{B}}(\tau) \underline{\underline{A}}(\tau)^{-1} \underline{\underline{B}}(\tau). \quad (3.6)$$

The right hand side has to be independent of τ . The solution, for $\underline{\underline{B}}$ in terms of $\underline{\underline{A}}$, is

$$\underline{\underline{B}}(\tau) = -\underline{\underline{A}}(\tau) \sqrt{\underline{\underline{I}} - \underline{\underline{A}}(\tau)^{-1} \underline{\underline{S}}}. \quad (3.7)$$

3.3. Extensivity of the second entropy

Now the τ dependence of the second entropy is elucidated. For large times, there can be no correlation between \mathbf{q} and \mathbf{q}' , and hence

$$\underline{A}(\tau) \rightarrow \underline{S}, \quad \underline{B}(\tau) \rightarrow \underline{0}, \quad |\tau| \rightarrow \infty. \quad (3.8)$$

Conversely, on small time intervals, based upon the analysis of Brownian motion we expect

$$|\underline{A}(\tau)^{-1} \underline{S}| \ll 1. \quad (3.9)$$

(It will turn out that this is equivalent to $|\bar{\mathbf{q}}\tau| \ll |\mathbf{q}|$.) We shall proceed under this assumption to obtain an explicit expression for the τ -dependence of \underline{A} , and we shall use the final result to deduce the limits on τ when the assumption is justified.

Inserting eqn (3.7) and (3.6) into the second entropy expression eqn (3.4), and expanding to linear order we obtain

$$\begin{aligned} \sigma(\mathbf{q}', \mathbf{q}|\tau) &= \frac{1}{2} [\mathbf{q}' - \mathbf{q}]^T \underline{A}(\tau) [\mathbf{q}' - \mathbf{q}] + \frac{1}{2} [S(\mathbf{q}) + S(\mathbf{q}')] \\ &+ \mathcal{O}(\underline{A}(\tau)^{-1} \underline{S})^2. \end{aligned} \quad (3.10)$$

For the final term we have replaced the quadratic form by the explicit entropies, $\mathbf{q}^T \underline{S} \mathbf{q}' = S(\mathbf{q}) + S(\mathbf{q}')$, which is convenient and permissible to leading order, but is not essential for what follows.

In order to find the τ -dependence of \underline{A} , we invoke a reduction condition on the two step transition, $\mathbf{q} \rightarrow \mathbf{q}' \rightarrow \mathbf{q}''$, each step of duration $|\tau|/2$. Consider the ansatz

$$\begin{aligned} \sigma(\mathbf{q}'', \mathbf{q}', \mathbf{q}|\tau/2, \tau/2) &= \frac{1}{2} [\mathbf{q}'' - \mathbf{q}']^T \underline{A}(\tau/2) [\mathbf{q}'' - \mathbf{q}'] \\ &+ \frac{1}{2} [\mathbf{q}' - \mathbf{q}]^T \underline{A}(\tau/2) [\mathbf{q}' - \mathbf{q}] \\ &+ \frac{1}{2} [S(\mathbf{q}) + S(\mathbf{q}'')]. \end{aligned} \quad (3.11)$$

Note that this has the appearance of a Markov transition, as there is no coupling between \mathbf{q}'' and \mathbf{q} . It turns out to be possible to satisfy the reduction conditions with the coefficient of such a coupling set to zero.

The first reduction comes from integrating one of the terminal positions and we must have,

$$e^{\sigma(\mathbf{q}'', \mathbf{q}'|\tau/2)/k_B} = \int d\mathbf{q} e^{\sigma(\mathbf{q}'', \mathbf{q}', \mathbf{q}|\tau/2, \tau/2)/k_B}. \quad (3.12)$$

Completing the squares, a straightforward calculation shows that the ansatz satisfies this condition, with the first neglected term being $\mathcal{O}(\underline{A}^{-1} \underline{S})^2$.

The second reduction arises from integrating the central position,

$$e^{\sigma(\mathbf{q}'', \mathbf{q}'|\tau)/k_B} = \int d\mathbf{q}' e^{\sigma(\mathbf{q}'', \mathbf{q}', \mathbf{q}|\tau/2, \tau/2)/k_B}. \quad (3.13)$$

Again completing the squares we obtain in this case the desired condition on the matrix, namely

$$\underline{A}(\tau) = \frac{1}{2} \underline{A}(\tau/2). \quad (3.14)$$

In view of the fact that the matrix must be an even function of τ we conclude from this that it must have the form

$$\underline{A}(\tau) = \frac{1}{|\tau|} \underline{\sigma}_{ss}. \quad (3.15)$$

The matrix $\underline{\sigma}_{ss}$ will turn out to be the matrix of linear transport coefficients that characterize steady flow.

With this result and the definition of the coarse velocity, the second entropy may be written as

$$\sigma(\mathbf{q}', \mathbf{q}|\tau) = \frac{|\tau|}{2} \bar{\mathbf{q}}^T \underline{\sigma}_{ss} \bar{\mathbf{q}} + \frac{1}{2} [S(\mathbf{q}') + S(\mathbf{q})], \quad (3.16)$$

where the final term may be written as $\mathbf{q}^T \underline{S} \mathbf{q}'/2$. This result shows that the second entropy is an extensive function of the time interval, the coarse velocities being regarded as intensive variables. This extensivity was originally deduced from the behaviour of a Brownian particle using somewhat different arguments.⁴

The condition under which this result for the second entropy was derived, $|\underline{A}(\tau)^{-1} \underline{S}| \ll 1$, becomes

$$|\tau| \ll \tau_{\text{long}} \equiv |\underline{\sigma}_{ss} \underline{S}^{-1}|. \quad (3.17)$$

The right hand side defines the upper limit on the time interval over which these results are valid.

The ansatz (3.11) cannot hold in the limit $|\tau| \rightarrow 0$. In this limit we expect that the velocity will be constant over the brief interval due to the inertia of the thermodynamic variables. Hence $\exp \sigma(\mathbf{q}'', \mathbf{q}', \mathbf{q}|\tau/2, \tau/2)/k_B \propto \delta(\mathbf{q}'' - (\mathbf{q}' + \mathbf{q})/2)$. In this limit then

$$\underline{A}(\tau) = \frac{1}{\tau^2} \underline{\sigma}_i, \quad |\tau| \rightarrow 0. \quad (3.18)$$

We may define the inertial time as

$$\tau_{\text{short}} \equiv \left| \underline{\sigma}_i \underline{\sigma}_{ss}^{-1} \right|. \quad (3.19)$$

We can also define the τ -dependent velocity correlation matrix, $\underline{\sigma}(\tau) = \tau^2 \underline{A}(\tau)$. In view of the above behavior we have

$$\underline{\sigma}(\tau) = \begin{cases} \underline{\sigma}_i, & |\tau| \lesssim \tau_{\text{short}} \\ |\tau| \underline{\sigma}_{ss}, & \tau_{\text{short}} \lesssim |\tau| \lesssim \tau_{\text{long}}. \end{cases} \quad (3.20)$$

It is possible to smoothly interpolate between these two regimes. However, because $|\tau|$ is a non-analytic function, we should not imagine that we can arrive at this result or at a smooth function *via* a finite Taylor expansion. With this definition the second entropy for small velocities is

$$\sigma(\mathbf{q}', \mathbf{q}|\tau) = \frac{1}{2} \bar{\mathbf{q}}^T \underline{\sigma}(\tau) \bar{\mathbf{q}} + \frac{1}{2} [S(\mathbf{q}') + S(\mathbf{q})]. \quad (3.21)$$

Just as in section 3.1 for the first entropy, it is possible to go beyond the fluctuation level in the description of the second entropy. The present second entropy matrix would be replaced by the matrix of second derivatives of the second entropy evaluated at the current values of \mathbf{q} and \mathbf{q}' .

3.4. Regression, fluctuation, and dissipation

The difference between thermodynamics and statistical mechanics is that the former deals with the most likely state, whereas the latter deals with the average state. This distinction

is discussed in detail in ch. 2 of ref. 8. The most likely value of a constraint, (*i.e.* macrostate), is the state that maximizes the entropy constrained with respect to that variable. Here and below we shall use the phrase ‘most likely’ in precisely this mathematical sense.

Equating the derivative of the second entropy with respect to \mathbf{q}' to zero yields the most likely coarse velocity as

$$\underline{\dot{\mathbf{q}}}(\mathbf{q}) = \begin{cases} \frac{-\tau}{2} \underline{\underline{\sigma}}(\tau)^{-1} \underline{\underline{S}}\mathbf{q}, & |\tau| \ll \tau_{\text{long}} \\ \frac{-\tau}{2} \underline{\underline{\sigma}}_{\tau}^{-1} \underline{\underline{S}}\mathbf{q}, & |\tau| \approx \tau_{\text{short}} \\ \frac{-\text{sign}(\tau)}{2} \underline{\underline{\sigma}}_{\text{ss}}^{-1} \underline{\underline{S}}\mathbf{q}, & \tau_{\text{short}} \ll |\tau| \ll \tau_{\text{long}}. \end{cases} \quad (3.22)$$

Here we have invoked the quadratic form so that $\partial S(\mathbf{q})/\partial \mathbf{q} = \underline{\underline{S}}\mathbf{q}$. For positive τ this result gives the most likely velocity going forward from \mathbf{q} , and for negative τ this is the most likely velocity leading up to \mathbf{q} . Due to dynamical reversibility, these are the negative of each other. This says that the most likely velocity is zero at $\tau = 0$, and increases smoothly and linearly from that point. In the intermediate regime the most likely velocity is constant.

This derivation may be regarded as a proof of Onsager’s regression hypothesis.² The term $\partial S(\mathbf{q})/\partial \mathbf{q}$ may be interpreted as the thermodynamic force (divided by T). Hence this result says that the most likely velocity of the fluctuating variable in the intermediate regime is linearly proportional to the force acting on it. In this case the force is the internal thermodynamic restoring force.

The regression hypothesis is closely related to the fluctuation–dissipation theorem.⁹ In the context of the Langevin equation, the theorem relates the strength of the stochastic or fluctuating force to the magnitude of the friction or dissipation.¹⁰ In the present context the theorem says that the coefficient that describes the rate of return of a thermodynamic fluctuation to equilibrium is the same as the coefficient that describes the dissipative flux due to an externally imposed thermodynamic force or gradient. (The term ‘dissipative flux’ originally meant the transfer of heat from a hot body to a cold body, but now more generally means an increase in entropy.) In the case of a fluctuation an internal restoring force arises from the departure from equilibrium, and this creates an internal flux back to equilibrium. In the other case an external force maintains a departure from equilibrium. The consequent internal flux seeking to return the system to equilibrium is exactly cancelled by an external flux from the reservoirs that increases their entropy. The coefficient of interest, which may be called a susceptibility, or a transport coefficient, can be obtained from the velocity correlations, as is done next, or from the ratio of the flux to the applied field, as is done for the case of heat flow in section 4.3.

The second entropy matrix is negative because it describes fluctuations from equilibrium. Hence a state with non-zero velocity has a reduced sub-system second entropy and it therefore represents a more ordered state of the sub-system. An external driving force or applied thermodynamic gradient makes the velocity non-zero, (and hence orders the sub-system), and we shall show that this corresponds to a dissipative flux that continually increases the entropy of the reservoirs, and the second entropy of the total system. In other words, order is created in the sub-system by the increas-

ing disorder of the surroundings. We shall show that this sub-system order maximizes the total second entropy, and hence it is a more probable state of the universe than is a disordered sub-system. This is typical non-equilibrium behavior, and it reconciles the order that we observe in our biosphere with the second law of thermodynamics, (see section 8 below).

3.5. Velocity correlations and Green–Kubo

Because the transition probability is proportional to the exponential of the dynamical entropy, the transition matrix may be extracted from the equilibrium fluctuations. From the Gaussian integrals we obtain

$$\begin{aligned} \langle \mathbf{q}(t+\tau)\mathbf{q}(t)^T \rangle_0 &= \frac{1}{Z} \int d\mathbf{q}' d\mathbf{q} \mathbf{q}'^T \mathbf{q}^T e^{\sigma(\mathbf{q}',\mathbf{q}(t))/k_B} \\ &= k_B \underline{\underline{A}}(\tau)^{-1} \underline{\underline{B}}(\tau) \underline{\underline{S}}^{-1} \\ &= -k_B \sqrt{\underline{\underline{I}} - \underline{\underline{A}}(\tau)^{-1} \underline{\underline{S}} \underline{\underline{S}}^{-1}}. \end{aligned} \quad (3.23)$$

On not too long and on intermediate time scales this reduces to

$$\begin{aligned} \langle \mathbf{q}(t+\tau)\mathbf{q}(t)^T \rangle_0 / k_B &= - \left[\underline{\underline{S}}^{-1} - \frac{\tau^2}{2} \underline{\underline{\sigma}}(\tau)^{-1} \right], |\tau| \ll \tau_{\text{long}} \\ &= - \left[\underline{\underline{S}}^{-1} - \frac{|\tau|}{2} \underline{\underline{\sigma}}_{\text{ss}}^{-1} \right], \tau_{\text{short}} \ll |\tau| \ll \tau_{\text{long}}. \end{aligned} \quad (3.24)$$

Using the fact that $\langle \mathbf{q}\mathbf{q}^T \rangle_0 = -k_B \underline{\underline{S}}^{-1}$, this may be written in terms of the coarse velocity,

$$\begin{aligned} \langle \dot{\mathbf{q}}(t; \tau)\mathbf{q}(t)^T \rangle_0 &= \frac{k_B \tau}{2} \underline{\underline{\sigma}}(\tau)^{-1}, |\tau| \ll \tau_{\text{long}} \\ &= \frac{k_B \text{sign}(\tau)}{2} \underline{\underline{\sigma}}_{\text{ss}}^{-1}, \tau_{\text{short}} \ll |\tau| \ll \tau_{\text{long}}. \end{aligned} \quad (3.25)$$

Differentiating both sides of the first equality with respect to τ and taking the limit that τ goes to zero shows that

$$\begin{aligned} k_B \underline{\underline{\sigma}}_{\tau}^{-1} &= \langle \ddot{\mathbf{q}}(t)\mathbf{q}(t)^T \rangle_0 \\ &= \lim_{\tau \rightarrow 0} \frac{1}{\tau} \langle \dot{\mathbf{q}}(t+\tau)\mathbf{q}(t)^T \rangle_0 \\ &= - \langle \dot{\mathbf{q}}(t)\dot{\mathbf{q}}(t)^T \rangle_0. \end{aligned} \quad (3.26)$$

(These results use the fact that $\langle \dot{\mathbf{q}}(t)\mathbf{q}(t)^T \rangle_0 = \underline{\underline{0}}$, as follows from dynamical reversibility.) This enables the inertial mass to be determined from the velocity fluctuations.

On intermediate time scales, eqn (3.25) must be independent of τ . The derivative of the left hand vanishes if, and only if,

$$\begin{aligned} \langle \dot{\mathbf{q}}(t; \tau)\mathbf{q}(t)^T \rangle_0 &= \langle \dot{\mathbf{q}}(t+\tau)\mathbf{q}(t)^T \rangle_0, \\ \tau_{\text{short}} \ll |\tau| \ll \tau_{\text{long}}. \end{aligned} \quad (3.27)$$

It also follows that the instantaneous velocity is constant in the intermediate regime,

$$\langle \dot{\mathbf{q}}(t+\tau)\mathbf{q}(t)^T \rangle_0 = \underline{\underline{0}}, \tau_{\text{short}} \ll |\tau| \ll \tau_{\text{long}}. \quad (3.28)$$

Using these results, in the intermediate regime the transport coefficient may be obtained from a number of equivalent

expressions,

$$\begin{aligned}\underline{\underline{\sigma}}_{\text{ss}}^{-1} &= \frac{2}{k_{\text{B}}} \langle \dot{\mathbf{q}}(t; \tau) \mathbf{q}(t)^{\text{T}} \rangle_0 \\ &= \frac{-\tau}{k_{\text{B}}} \langle \dot{\mathbf{q}}(t; \tau) \dot{\mathbf{q}}(t; \tau)^{\text{T}} \rangle_0 \\ &= \frac{2}{k_{\text{B}}} \langle \mathbf{q}(t + \tau) \mathbf{q}(t)^{\text{T}} \rangle_0 \\ &= \frac{2}{k_{\text{B}}} \int_0^\tau dt' \langle \dot{\mathbf{q}}(0) \dot{\mathbf{q}}(t')^{\text{T}} \rangle_0, \\ \tau_{\text{short}} &\ll \tau \ll \tau_{\text{long}}.\end{aligned}\quad (3.29)$$

The final form may be recognized as a Green–Kubo expression.^{9–13}

The Onsager matrix for the transport coefficients is related to the second entropy matrix by

$$\underline{\underline{L}} = \frac{-1}{2T} \underline{\underline{\sigma}}_{\text{ss}}^{-1}. \quad (3.30)$$

Since $\underline{\underline{\sigma}}_{\text{ss}}$ is derived from the symmetric $\underline{\underline{A}}(\tau)$, this shows that the Onsager matrix is symmetric. This is called the reciprocal relation.

In general $\langle \dot{\mathbf{q}}(t; \tau) \mathbf{q}(t)^{\text{T}} \rangle_0$ will yield the transport function, $\underline{\underline{g}}(\tau)$, and it may be possible to extract the linear transport coefficient $\underline{\underline{\sigma}}_{\text{ss}}$ by fitting. If the system is driven by an external thermodynamic force, then it achieves a steady state in which $\underline{\underline{g}}_{\pm}$ is identically zero, (and $\tau_{\text{short}} = 0$). The flux in this steady state and at this state point is characterized by $\underline{\underline{\sigma}}_{\text{ss}}$.

Langevin dynamics, and its probabilistic variant the Fokker–Plank equation, have long provided a phenomenological approach to non-equilibrium problems.^{9–11,14} It may be shown that the present second entropy expression for the transition probability in the intermediate regime corresponds to Langevin dynamics in the high friction limit.⁴ It may also be shown by direct evaluation that the second entropy transition probability inserted into the probability evolution equation, (equivalently Liouville’s theorem), yields Smoluchowski’s equation.

4. Heat flow

The analysis in the preceding section 3 dealt with the fluctuations and rates of change of thermodynamic variables. The analysis may now be made more concrete, and specific non-equilibrium results may be derived, by application to the problem of heat flow. In this case a thermal gradient is imposed on a sub-system separating two heat reservoirs of different temperatures. There is consequently a steady flow of heat from the hot reservoir to the cold reservoir through the sub-system. First we have to identify explicitly the thermodynamic variables appropriate for this problem.

4.1. Thermodynamic variables

We consider a sub-system connected to thermal reservoirs of temperatures T_{\pm} and located at $z = \pm L/2$. (To simplify the notation we treat only this scalar one dimensional case.) We expect that the imposed temperature gradient, $(T_{+} - T_{-})/L$, will induce a corresponding temperature gradient in the sub-system, and also a gradient in the energy density.

To see this quantitatively we require the first entropy. We let the energies of the respective reservoirs be $E_{r\pm}$. We identify a fixed region of the sub-system adjacent to each boundary and denote the energy of these regions by $E_{s\pm}$. We impose the energy conservation laws

$$\Delta E_{s+} = -\Delta E_{r+}, \text{ and } \Delta E_{s-} = -\Delta E_{r-}. \quad (4.1)$$

This constraint is not exact because the boundary regions can also exchange energy with the interior of the sub-system. However, we assume that the rate of this exchange is much slower than that with the reservoir. We address this point further below.

Using the definition of temperature, $T^{-1} = \partial S/\partial E$, the conservation laws, and a Taylor expansion about some fixed state, we may write the reservoirs’ entropy as

$$S_{r\pm}(E_{r\pm}) = S_{r\pm}(E_{r0\pm}) - \frac{E_{s\pm}}{T_{\pm}}. \quad (4.2)$$

The first term is independent of the sub-system and may be neglected. With this, the entropy of the total system constrained to be in the macrostate $E_{s\pm}$ is the sum of that of the isolated sub-system in that macrostate and that of the reservoirs,

$$S_{\text{total}}(E_{s+}, E_{s-} | T_{+}, T_{-}) = S(E_{s+}, E_{s-}) - \frac{E_{s+}}{T_{+}} - \frac{E_{s-}}{T_{-}}. \quad (4.3)$$

Here $S(E_{s+}, E_{s-})$ is the entropy of the isolated sub-system. Differentiating this with respect to the constraints we have

$$\frac{\partial S_{\text{total}}}{\partial E_{s\pm}} = \frac{1}{T_{s\pm}} - \frac{1}{T_{\pm}}. \quad (4.4)$$

The *most likely* state, which is denoted throughout by an overline, is the one that maximizes the entropy. In this case the entropy derivative vanishes when the boundary temperatures of the sub-system equal that of the respective reservoirs,

$$\bar{T}_{s\pm} = T_{\pm}. \quad (4.5)$$

This result is intuitively appealing and not unexpected.

We now treat the same problem from a slightly different perspective. Motivated by the fact that it is the inverse temperature that is thermodynamically conjugate to the energy, we define the zeroth temperature,^{3,5}

$$\frac{1}{T_0} \equiv \frac{1}{2} \left[\frac{1}{T_{+}} + \frac{1}{T_{-}} \right], \quad (4.6)$$

and the first temperature,

$$\frac{1}{T_1} \equiv \frac{1}{L} \left[\frac{1}{T_{+}} - \frac{1}{T_{-}} \right]. \quad (4.7)$$

The zeroth temperature is essentially the average temperature of the two reservoirs, and the first temperature is essentially the applied temperature gradient, $T_1^{-1} \equiv \nabla T^{-1} = -\nabla T/T_0^2$. The sub-system temperatures T_{s0} and T_{s1} can be defined identically in terms of T_{s+} and T_{s-} .

The sub-system energies used above can be rearranged in the form of energy moments,

$$E_0 = [E_{s+} + E_{s-}], \quad (4.8)$$

and

$$E_1 = (L/2)[E_{s+} - E_{s-}]. \quad (4.9)$$

It is not hard to show that with these definitions the zeroth and first temperatures of the sub-system are thermodynamically conjugate to the zeroth and first energy moments,

$$\frac{1}{T_{s0}} = \frac{\partial S(E_0, E_1)}{\partial E_0}, \text{ and } \frac{1}{T_{s1}} = \frac{\partial S(E_0, E_1)}{\partial E_1}. \quad (4.10)$$

Actually this is the more fundamental definition of the temperatures, but it is entirely consistent with the preceding expressions.

With the above definitions it is a brief exercise to rearrange the constrained total entropy as

$$S_{\text{total}}(E_0, E_1|T_0, T_1) = S(E_0, E_1) - \frac{E_0}{T_0} - \frac{E_1}{T_1}. \quad (4.11)$$

Maximizing the total entropy with respect to its arguments shows that the most likely state then is one in which these sub-system temperatures equal their reservoir counterparts, $\bar{T}_{s0} = T_0$, and $\bar{T}_{s1} = T_1$.

The point of this second formulation is its generality. The energy moments are always well defined, for example

$$E_x = \int_V dr z^\alpha \varepsilon(\mathbf{r}), \quad (4.12)$$

where $\varepsilon(\mathbf{r})$ is the energy density. In this formulation there is no need to invoke an arbitrary boundary region. It will turn out that the rate of change of the first energy moment is related to the heat flux through the sub-system. We *define* the zeroth and first temperatures to be conjugate to the respective energy moments *via* the derivatives of the entropy, and again we can avoid having to invoke a boundary region. A formal derivation of the conjugate relations may be found in ref. 3. The idea that moments and gradients are conjugate is due to Onsager.²

4.2. Gaussian approximation

The first energy moment of the isolated system is not conserved and it fluctuates about zero. The general analysis of section 3.1 holds. Accordingly the entropy of the isolated system may be written as a quadratic form,

$$S(E_0, E_1) = S_0(E_0) + \frac{1}{2} E_1 S E_1, \quad (4.13)$$

where the correlation matrix satisfies

$$S^{-1} = \frac{-1}{k_B} \langle E_1 E_1 \rangle. \quad (4.14)$$

In this approximation the first temperature is given by

$$T_{s1}^{-1} = S E_1. \quad (4.15)$$

In view of this, an applied temperature gradient induces an energy moment in the sub-system that is given by

$$\bar{E}_1 = S^{-1} T_1^{-1}. \quad (4.16)$$

4.3. Second entropy

4.3.1 Isolated system. We now consider the transition $E_1 \rightarrow E_1'$ in time $\tau > 0$. For the case of an isolated system,

the results of section 3.2 apply. The second entropy in the intermediate regime is

$$\sigma(E_1', E_1|\tau) = \frac{|\tau|}{2} \overset{\circ}{E}_1^0 \sigma_{ss} \overset{\circ}{E}_1^0 + \frac{1}{2} [S(E_1') + S(E_1)]. \quad (4.17)$$

Here the coarse velocity is $\overset{\circ}{E}_1^0 \equiv (E_1' - E_1)/\tau$; the zero is appended to emphasize that this is the internal or adiabatic rate of change of moment (*i.e.* the system is isolated from any reservoirs).

The derivative with respect to E_1' is

$$\begin{aligned} \frac{\partial \sigma(E_1', E_1|\tau)}{\partial E_1'} &= \text{sign}(\tau) \sigma_{ss} \overset{\circ}{E}_1^0 + \frac{\partial S(E_1')}{2 \partial E_1'} \\ &= \text{sign}(\tau) \sigma_{ss} \overset{\circ}{E}_1^0 + \frac{1}{2 T_{s1}'}. \end{aligned} \quad (4.18)$$

Hence the most likely rate of change of moment due to internal processes is

$$\overset{\circ}{E}_1^0 = -\text{sign}(\tau) \sigma_{ss}^{-1} / 2 T_{s1}' = -\text{sign}(\tau) \sigma_{ss}^{-1} S E_1' / 2, \quad (4.19)$$

the final line holding in quadratic approximation. (For $|\tau \overset{\circ}{E}_1^0| \ll E_1$, we can replace E_1' by E_1 on the right hand side.)

4.3.2 Reservoirs. We now place the sub-system in thermal contact with the two reservoirs discussed in the preceding section. In this case the energy moment can change by the internal processes just discussed, or by exchange with the reservoirs,

$$E_1' = E_1 + \Delta^0 E_1 - \Delta E_{1r}. \quad (4.20)$$

The second entropy is now the sum of that of the isolated system and that due to the reservoirs. Taking into account the extra degree of freedom it is

$$\begin{aligned} \sigma_{\text{total}}(\Delta^0 E_1, \Delta E_{1r}, E_1|\tau) &= \frac{\tau}{2} \overset{\circ}{E}_1^0 \sigma_{ss} \overset{\circ}{E}_1^0 + \frac{1}{2} [S(E_1') + S(E_1) \\ &\quad - \frac{2E_1 - \Delta E_{1r}}{T_1}]. \end{aligned} \quad (4.21)$$

Here we have used $\tau > 0$ because we are dealing with the established steady state. The final term in the brackets is the sum of the sub-system-dependent reservoir entropy in the initial and final states of the transition. The derivative with respect to the external change yields

$$\begin{aligned} \frac{\partial \sigma_{\text{total}}}{\partial \Delta E_{1r}} &= \frac{\partial S(E_1')}{2 \partial E_1'} \frac{\partial E_1'}{\partial \Delta E_{1r}} + \frac{1}{2 T_1} \\ &= \frac{-1}{2 T_{s1}'} + \frac{1}{2 T_1}. \end{aligned} \quad (4.22)$$

This vanishes when $\bar{T}_{s1}' = T_1$, which is the most likely result obtained from the first entropy.

Maximising the total second entropy with respect to E_1 yields essentially the same result, $1/\bar{T}_{s1}' + 1/\bar{T}_{s1} = 2/T_1$, which is to say that $T_{s1} = T_1$, and hence $\bar{E}_1' = \bar{E}_1$. That is $\overline{\Delta E_{1r}} = \overline{\Delta^0 E_1}$, which is to say that in the steady state the change in the sub-system moment due to energy flow from the reservoir is exactly cancelled by the internal energy flows within the sub-system, and so the sub-system structure remains unchanged.

The derivative with respect to the internal change yields

$$\begin{aligned}\frac{\partial \sigma_{\text{total}}}{\partial \Delta^0 E_1} &= \sigma_{\text{ss}} \overset{\circ}{E}_1^0 + \frac{\partial S(E_1')}{2\partial E_1'} \frac{\partial E_1'}{\partial \Delta^0 E_1} \\ &= \sigma_{\text{ss}} \overset{\circ}{E}_1^0 + \frac{1}{2T_{s1}'}.\end{aligned}\quad (4.23)$$

This vanishes at the most likely velocity, which is

$$\overset{\circ}{E}_1^0 = -\sigma_{\text{ss}}^{-1}/2T_{s1}' = -\sigma_{\text{ss}}^{-1}/2T_1. \quad (4.24)$$

This says that the most likely rate of change of the moment due to internal processes is linearly proportional to the imposed temperature gradient. This is a particular form of the transport law, eqn (3.22), with the imposed temperature gradient providing the thermodynamic driving force for the flux. Note again that in this case the sign of τ has been taken as positive because we are assuming that the system has been in a steady state for some time already, (*i.e.* the system is *not* time reversible).

This result confirms Onsager's regression hypothesis. The most likely velocity in an isolated system following a fluctuation from equilibrium, eqn (4.19), is equal to the most likely velocity due to an externally imposed force, eqn (4.24), when the internal force is equal to the external force, $T_{s1}' = T_1$.

This result for the most likely change in moment is equivalent to Fourier's law of heat conduction. To see this we note that in the steady state the rate of change of moment is zero, $\dot{E}_1 = 0$, so that the internal change is cancelled by that due to the reservoirs, $\dot{E}_1^0 = \dot{E}_{r1}$. The rate of change of the reservoir moment is just the energy flux through the system,

$$J = \frac{\pm \dot{E}_{r\pm}}{A} = \frac{\dot{E}_{r1}}{V} = \frac{\dot{E}_1^0}{V}, \quad (4.25)$$

where A is the cross-sectional area of the sub-system and $V = AL$ is the sub-system volume. Now Fourier's law says

$$J = -\lambda \nabla T = \lambda T_0^2/T_1, \quad (4.26)$$

where λ is the thermal conductivity. Accordingly we have

$$\overset{\circ}{E}_1^0 = \frac{\lambda VT_0^2}{T_1}. \quad (4.27)$$

This means that thermal conductivity and second entropy transport coefficient are related by $\lambda = -1/2VT_0^2\sigma_{\text{ss}}$. (In the steady state, $\sigma_i = 0$ and $\overset{\circ}{E}_1^0 = E_1^0$.)

4.3.3 Rate of entropy production. These results give the rate of entropy production. The rate of change of the first entropy of the universe, which, since the sub-system's structure does not change with time, is the same as that of the reservoirs, is given by

$$\dot{S}^{(1)} = \frac{\overline{\Delta E_{1r}}}{T_1 \tau} = \frac{\overset{\circ}{E}_1^0}{T_1} = \frac{-\sigma_{\text{ss}}^{-1}}{2T_1^2}. \quad (4.28)$$

This is evidently positive, (since the second entropy is concave down). The second equality follows because $T_{1s}(\overset{\circ}{E}_1') = T_1$, and $T_{1s}(\bar{E}_1) = T_1$, which implies that $\bar{E}_1' = \bar{E}_1$, and hence $\overline{\Delta E_{1r}} = \Delta^0 E_1$.

It should be clear that the most likely or physical rate of first entropy production is neither minimal nor maximal; these would correspond to values of the heat flux of $\pm\infty$. The conventional first entropy does not provide any variational principle for heat flow, or for non-equilibrium dynamics more generally. In the literature one finds claims for both theorems: some claim that the rate of entropy production is minimal and others claim that it is maximal, whereas the present results indicate that it is neither. One has to distinguish the first or structural entropy, which conventionally is what is meant by the word 'entropy' in the literature, from the second or dynamic entropy, which was introduced in ref. 4 and which provides the basis for the present non-equilibrium theory. It is the second entropy which obeys a variational principle that allows the physical non-equilibrium state to be obtained by maximizing it with respect to constraints.

The most likely value of the total second entropy, which is of course its maximum value, is readily shown to be

$$\begin{aligned}\bar{\sigma}_{\text{total}} &= \frac{\tau}{2} \sigma_{\text{ss}} \left(\overset{\circ}{E}_1^0 \right)^2 + \frac{\overline{\Delta^0 E_1} - \overline{\Delta E_{1r}}}{2\bar{T}_{1s}} + \frac{\overline{\Delta E_{1r}}}{2T_1} + S(\bar{E}_1) - \frac{\bar{E}_1}{T_1} \\ &= \tau \left[\frac{\sigma_{\text{ss}}^{-1}}{8T_1^2} - \frac{\sigma_{\text{ss}}^{-1}}{4T_1^2} \right] + S(\bar{E}_1) - \frac{\bar{E}_1}{T_1}.\end{aligned}\quad (4.29)$$

The final, constant, term is the static entropy of the total system, (sub-system plus sub-system-dependent part of the reservoir), due to establishing the structure. The first term in the brackets is the constant rate of second entropy *decrease* due to the dynamic order of the sub-system, and the second, larger, term in the brackets is the constant rate of entropy *increase* of the reservoirs. Consequently the most likely rate of change of second entropy is

$$\dot{\bar{\sigma}}_{\text{total}} = -\frac{\sigma_{\text{ss}}^{-1}}{8T_1^2}, \quad (4.30)$$

which is greater than zero. This result shows that for a steady state system, the dynamic order induced in the intervening sub-system consumes second entropy at a certain rate, but that this ordering enables the external reservoirs that impose the thermodynamic gradient to produce second entropy at a greater rate so that the second entropy of the universe increases. (Because the conductivity of the reservoirs is by definition much greater than that of the intervening sub-system, the reservoir contribution to the total second entropy is just half their first entropy.)

Since the second entropy scales linearly with time, maximizing the total second entropy to find the most likely flux is the same as maximizing the rate of entropy production of the universe. If one were to increase the flux \dot{E}_1^0 beyond its optimum value, one would increase the rate of entropy consumption by the sub-system due to its increased dynamic order by a greater amount than one would increase the entropy production of the reservoirs due to the faster transfer of heat. The converse holds for a less than optimum flux. In both cases the total rate of second entropy production would fall from its maximum value.

This result must be stressed. The present analysis of non-equilibrium systems is equivalent to a variational procedure that *maximizes* the rate of second entropy production of the universe. One should not confuse this with the principle of *minimum* rate of (first) entropy production due to Prigogine.^{15,16} As we now show, the latter refers to the fact that when external forces are applied to a sub-system to create a steady state, additional forces and fluxes would increase the rate of production of (first) entropy.

In the present language, let x_i be the values of the sub-system variables that are exchangeable with a reservoir, and let the imposed, conjugate thermodynamic forces be X_i , $i = 1, 2, \dots$, (e.g. $x = E_1$ and $X = 1/T_1$). The total constrained first entropy is

$$S_{\text{total}}(\mathbf{x}|\mathbf{X}) = S(\mathbf{x}) - \mathbf{x}^T \mathbf{X}, \quad (4.31)$$

and the rate of first entropy production of the reservoirs is $S_r = \dot{\mathbf{x}}^T \mathbf{X}$. In the steady state, $\bar{\mathbf{x}} = 0$, and $\bar{\dot{\mathbf{x}}} = \bar{\mathbf{x}}^T$, and in this case the rate of first entropy production of the universe is the same as that of the reservoirs because the structure of the sub-system is constant.

In the steady state the total constrained second entropy is

$$\begin{aligned} \sigma_{\text{total}}(\Delta^0 \mathbf{x}, \Delta \mathbf{x}_r, \mathbf{x}|\mathbf{X}, \tau) &= \frac{\tau}{2} \mathbf{x}^{0T} \underline{\underline{\sigma}}_{\text{ss}}^0 \mathbf{x}^0 \\ &+ \frac{1}{2} [S(\mathbf{x}') + S(\mathbf{x}) - 2 \mathbf{x}^T \mathbf{X} + \Delta \mathbf{x}_r^T \mathbf{X}], \end{aligned} \quad (4.32)$$

where $\mathbf{x}' = \mathbf{x} + \Delta^0 \mathbf{x} - \Delta \mathbf{x}_r$. Maximizing the constrained second entropy yields $\mathbf{X}_s(\bar{\mathbf{x}}) = \mathbf{X}_s(\bar{\mathbf{x}}') = \mathbf{X}$, and hence $\overline{\Delta^0 \mathbf{x}} = \overline{\Delta \mathbf{x}_r}$. Also, the most likely flux is $\bar{\dot{\mathbf{x}}}^0 = -\underline{\underline{\sigma}}_{\text{ss}}^{-1} \mathbf{X}/2$. In the case of the non-exchangeable variables that don't appear explicitly here, it may be shown that their internal flux and their internal conjugate thermodynamic force are both zero, but the values of the non-exchangeable variables themselves are non-zero. (See for example ref. 3 for the coupling of energy and density gradients.)

Using these results, the most likely rate of change of the first entropy of the reservoirs is

$$\bar{S}_r = \bar{\dot{\mathbf{x}}}^{0T} \mathbf{X} = \frac{-1}{2} \mathbf{X}^T \underline{\underline{\sigma}}_{\text{ss}}^{-1} \mathbf{X}. \quad (4.33)$$

This is a positive definite quadratic form that is minimized when the forces are zero. (The second entropy matrix is negative due to its concavity.) This result was given by Prigogine, and from it he gave the principle of minimum rate of (first) entropy production, which in essence says that allowing additional fluxes across the boundary, (*i.e.* applying additional forces), increases the rate of (first) entropy production.^{15,16}

Similarly, the maximum value of the total second entropy is

$$\begin{aligned} \bar{\sigma}_{\text{total}} &= \frac{\tau}{8} \mathbf{X}^T \underline{\underline{\sigma}}_{\text{ss}}^{-1} \mathbf{X} + S(\bar{\mathbf{x}}(\mathbf{X})) - \bar{\mathbf{x}}(\mathbf{X})^T \mathbf{X} - \frac{\tau}{4} \mathbf{X}^T \underline{\underline{\sigma}}_{\text{ss}}^{-1} \mathbf{X} \\ &= S(\bar{\mathbf{x}}(\mathbf{X})) - \bar{\mathbf{x}}(\mathbf{X})^T \mathbf{X} - \frac{\tau}{8} \mathbf{X}^T \underline{\underline{\sigma}}_{\text{ss}}^{-1} \mathbf{X}. \end{aligned} \quad (4.34)$$

The first two, constant, terms are the total (first) entropy increase due to establishing the structure, and the final term

is the ongoing production of (second) entropy. Hence the most likely rate of second entropy production of the universe is

$$\bar{\sigma}_{\text{total}} = \frac{-1}{8} \mathbf{X}^T \underline{\underline{\sigma}}_{\text{ss}}^{-1} \mathbf{X}. \quad (4.35)$$

This result in terms of the applied forces only holds for the optimized second entropy. That is, it is only valid for the physical or most likely steady state, and it is less general than the expression in terms of the constrained fluxes used throughout the text. It does not obey a variational principle, except for the trivial one in that it is a positive definite quadratic form. (It is the constrained second entropy that obeys a variational principle for the fluxes.) The most likely flux predicted by the present second entropy analysis is of course stable because the sub-system second entropy is concave down as a function of its arguments.

4.4. Phase space probability distribution

We now seek the steady state probability distribution for a system with an imposed temperature gradient. This is the microstate probability density for the phase space of the sub-system. The reservoirs enter by the quantities defined in the preceding section, namely the zeroth, $\beta_0 \equiv 1/k_B T_0$, and the first, $\beta_1 \equiv 1/k_B T_1$ temperatures. We denote this steady state probability density as $\varphi_{\text{ss}}(\Gamma|\beta_0, \beta_1)$.

From the preceding section it is evident that the zeroth and first energy moments of the sub-system are going to enter the probability distribution. The zeroth energy moment is the ordinary Hamiltonian,

$$E_0(\Gamma) = \mathcal{H}(\Gamma) = \sum_{i=1}^N \varepsilon_i, \quad (4.36)$$

and the first energy moment in the z -direction is just

$$E_1(\Gamma) = \sum_{i=1}^N \varepsilon_i z_i, \quad (4.37)$$

where ε_i is the total energy of particle i . The adiabatic rate of change of the energy moment, which is the natural or Hamiltonian motion where no heat flows to the isolated system, is denoted $\dot{E}_1^0(\Gamma) = \Gamma \cdot \nabla E_1(\Gamma)$. Both moments have even phase space parity, $E_0(\Gamma) = E_0(\Gamma^\dagger)$ and $E_1(\Gamma) = E_1(\Gamma^\dagger)$, since it is assumed that there are no velocity dependent forces in the Hamiltonian. The rate of change of the energy moment necessarily has odd parity, $\dot{E}_1^0(\Gamma^\dagger) = -\dot{E}_1^0(\Gamma)$.

As mentioned in the Introduction, the steady state probability density cannot depend explicitly on time, but it must depend upon the direction of time. Accordingly it cannot have even phase space parity,

$$\varphi_{\text{ss}}(\Gamma|\beta_0, \beta_1) \neq \varphi_{\text{ss}}(\Gamma^\dagger|\beta_0, \beta_1). \quad (4.38)$$

Since both E_0 and E_1 have even parity, the probability density cannot be a function of them alone. Nevertheless the fact that these are conjugate to β_0 and β_1 suggests that they will occur in the probability density. It seems likely that the flux \dot{E}_1^0 will also occur in order to break the even parity.

In section 3.2 we elucidated the nature of fluctuations in the thermodynamic variables of an isolated system. We concluded that on short time scales, $t \lesssim \tau_{\text{short}}$, the system behaved

inertially, and that on intermediate time scales, $\tau_{\text{short}} \lesssim t \lesssim \tau_{\text{long}}$, the system was in a steady state. In the inertial regime the system was insensitive to the direction of time, (*i.e.* it had even phase space parity), whereas in the steady state it moved in a definite direction in time. Hence our strategy in constructing the steady state probability density is to begin with an even probability density appropriate for a fluctuation, and to propagate it forward in time into the intermediate regime where it becomes the steady state probability density.

This is done as follows. Suppose the sub-system is at Γ at $t = 0$, and let $\Gamma_0(t|\Gamma)$ be the adiabatic trajectory, (*i.e.* the Hamiltonian evolution if the sub-system were isolated). Denote the past and future energy moments by $E_1^\pm(\Gamma) \equiv E_1(\Gamma_0(\pm\tau|\Gamma))$, for some time interval $\tau > 0$. From the time reversible nature of the equations of motion, $\Gamma_0(t|\Gamma) = [\Gamma_0(-t|\Gamma^\dagger)]^\dagger$, it follows that $E_1^\pm(\Gamma^\dagger) = E_1^\mp(\Gamma)$. On the trajectory the zeroth moment is conserved, $E_0^\pm(\Gamma) = E_0(\Gamma)$.

We shall tentatively consider the steady state probability as that which adiabatically evolved to the current point Γ from a probability density of even parity induced by the reservoirs a time $-\tau$ in the past. (By Liouville's theorem, the probability density is a constant of the adiabatic motion.)⁸ The probability is the exponential of the reservoir contribution to the total entropy, eqn (4.11), (the points in phase space are microstates and have zero entropy), and hence we have

$$\tilde{\mathcal{P}}_{\text{tent}}(\Gamma|\beta_0, \beta_1) = \frac{1}{\tilde{Z}_{\text{ss}}(\beta_0, \beta_1)} e^{-\beta_0 E_0(\Gamma)} e^{-\beta_1 E_1^-(\Gamma)}. \quad (4.39)$$

Here \tilde{Z} is the normalizing partition function, which is discussed further below. This probability gives the correct structure at time $t = -\tau$ for a system in thermal contact with the reservoirs, but it does not give the correct dynamics. However, the adiabatic evolution of the dynamics is much faster than the evolution of the structure, since $\dot{E}_1^0 \rightarrow \dot{E}_1^0$ in time τ_{short} , but generally $|\tau_{\text{short}} \dot{E}_1^0| \ll |\dot{E}_1^0|$. Hence allowing this probability to evolve adiabatically to the present time $t = 0$, will give the correct steady state dynamics, (provided $\tau \gtrsim \tau_{\text{short}}$), without greatly changing the structure, (provided $\tau \lesssim \tau_{\text{long}}$).

This probability distribution is somewhat inconvenient from the computational point of view. Worse, it depends sensitively upon the time interval τ . It may be corrected by noting that $E_0^-(\Gamma) = E_0(\Gamma)$, and that

$$\begin{aligned} E_1^-(\Gamma) &= \frac{1}{2}[E_1^+(\Gamma) + E_1^-(\Gamma)] - \frac{1}{2}[E_1^+(\Gamma) - E_1^-(\Gamma)] \\ &\approx E_1(\Gamma) - W_1^{\text{mir}}(\Gamma). \end{aligned} \quad (4.40)$$

For $|\tau \dot{E}_1| \ll |E_1|$, which we assume throughout, the first bracketed term on the right hand side can be approximated by $E_1(\Gamma)$. The difficulties associated with the tentative probability distribution are removed by this replacement: it becomes insensitive to the value of τ , (see section 4.5.2 below), and it gives the correct structure at the present time rather than at a time in the past (see section 4.5.4 below). For reasons that will become clearer below, we shall call the second term the mirror work. With these we give the steady state probability for a system with an imposed thermal gradient as⁷

$$\mathcal{P}_{\text{ss}}(\Gamma|\beta_0, \beta_1) = \frac{e^{-\beta_0 E_0(\Gamma)} e^{-\beta_1 E_1(\Gamma)}}{h^{3N} N! Z_{\text{ss}}(\beta_0, \beta_1)} e^{\beta_1 W_1^{\text{mir}}(\Gamma)}. \quad (4.41)$$

The tentative probability distribution, eqn (4.39), is an approximation to this steady state probability distribution. The mirror work has odd parity, $W_1(\Gamma^\dagger) = -W_1(\Gamma)$, which is reflected in its name, and so the steady state probability can be written as the product of even and odd factors, $\varphi_{\text{ss}}(\Gamma|\beta_0, \beta_1) = \varphi_e(\Gamma|\beta_0, \beta_1)\varphi_o(\Gamma|\beta_1)$.

The normalizing partition function is

$$Z_{\text{ss}}(\beta_0, \beta_1) = \int \frac{d\Gamma}{h^{3N} N!} e^{-\beta_0 E_0(\Gamma)} e^{-\beta_1 E_1(\Gamma)} e^{\beta_1 W_1^{\text{mir}}(\Gamma)}, \quad (4.42)$$

where h is Planck's constant and N is the number of atoms in the sub-system, and where the possible dependence on τ has not been exhibited. The non-equilibrium unconstrained total entropy is $S_{\text{total,ss}}(\beta_0, \beta_1) = k_B \ln Z_{\text{ss}}(\beta_0, \beta_1)$, and its derivatives generate non-equilibrium statistical mechanical averages, just as in the equilibrium case.⁸ The non-equilibrium constrained free energy is $-T_0$ times the constrained total entropy, $F = -T_0 S_{\text{total,ss,cons}}$. The latter is the sub-system-dependent part of the reservoir entropy (the exponent of the probability density), plus the entropy of the isolated sub-system constrained to be in the macrostate. Hence

$$\begin{aligned} F(E_0, E_1, E_1^\pm|\beta_0, \beta_1) &= E_0 + \beta_1[E_1 - (E_1^+ - E_1^-)/2] \\ &\quad \beta_0 - T_0 S_0^{(2)}(E_0, E_1, E_1^+, E_1^-), \end{aligned} \quad (4.43)$$

where the final term is the constrained second entropy of the isolated sub-system, (see also section 4.6.4 below). The constrained free energy is strictly greater than $-k_B T_0 \ln Z_{\text{ss}}$. But in the thermodynamic limit the minimum value of the constrained free energy (with respect to the constraints, which are written to the left of the vertical bar), is strictly greater than, but almost equal to, the partition function with negligible error, $F^-(\beta_0, \beta_1) \equiv F(\bar{E}_0, \bar{E}_1, \bar{E}_1^\pm|\beta_0, \beta_1) \gtrsim -k_B T_0 \ln Z_{\text{ss}}(\beta_0, \beta_1)$. The state giving this minimum value is by definition the most likely thermodynamic state. The derivatives of the free energy in the most likely macrostate generate various relationships between thermodynamic quantities, just as in the equilibrium case.⁸

4.5. Properties of the mirror work

4.5.1. Definition and rationalization. The mirror work may be written in several equivalent forms,

$$\begin{aligned} W_1^{\text{mir}}(\Gamma) &= \frac{1}{2}[E_1^+(\Gamma) - E_1^-(\Gamma)] \\ &= \frac{1}{2} \int_{-\tau}^{\tau} dt' \dot{E}_1^0(\Gamma_0(t'|\Gamma)) \\ &= \frac{1}{2} \int_{-\tau}^0 dt' [\dot{E}_1^0(\Gamma_0(t'|\Gamma)) - \dot{E}_1^0(\Gamma_0(t'|\Gamma^\dagger))]. \end{aligned} \quad (4.44)$$

The second equality shows that only the even part of the integrand contributes, and the past contribution to date is exactly half the total contribution. The third equality demonstrates that the first two equalities do not violate time causality.

We can give a physical rationalization of the mirror work as the 'internal' part of the change in E_1 , since only the 'external' part of the change in E_1 should contribute to the reservoirs' entropy. That is, since E_1 is not a conserved variable, the total change in energy moment arises from its adiabatic evolution,

which costs no reservoir entropy, and from perturbations by the reservoir, which do. The total change in moment is already accounted for by the even exponent, and the mirror work is essentially the internal change that must be subtracted. This particular physical interpretation of W_1^{mir} is discussed in more detail below.

The mirror work is equal to the change in reservoir entropy over an interval. This follows since in the steady state there is no net change in the total change in energy moment. Hence the internal change in moment over an interval must be cancelled by the change in the sub-system moment induced by the reservoir, $\Delta^0 E_1 = -\Delta^{s(r)} E_1$. But by energy conservation, the change in the sub-system moment induced by the reservoir must be equal and opposite to the change in the reservoir moment, $\Delta^{s(r)} E_1 = -\Delta E_{1r}$. Hence $\Delta W_1^{\text{mir}} = \Delta^0 E_1 = \Delta E_{1r}$, and we have $\beta_1 \Delta W_1^{\text{mir}} = \Delta S_r / k_B$. Accordingly, we may indeed interpret ΔW_1^{mir} as a type of thermodynamic work performed by the reservoir, which accounts for the notation used for this term.

4.5.2 Insensitivity to τ . The non-zero contributions to the integral for the mirror work come from about the origin. Eqn (3.22) shows that in the inertial regime the coarse rate of change of a thermodynamic variable is linearly proportional to the departure of that variable from its equilibrium value, and eqn (3.29) shows that the coarse velocity over the interval is equal to the instantaneous velocity in the intermediate regime. Hence we have

$$\dot{E}_1^0(\Gamma_0(t|\Gamma)) \sim \text{sign}(t) \overline{\dot{E}_1^0}(E_1(\Gamma)), \tau_{\text{short}} |t| \tau_{\text{long}}. \quad (4.45)$$

(This is valid if $|t \overline{\dot{E}_1^0}| \ll |E_1|$, which in practice is always the case and will be assumed throughout.) Also the over-line in this and similar contexts will henceforth denote the most likely flux going *forward* in time. In this case, eqn (4.14) and (4.19) give the most likely flux as

$$\overline{\dot{E}_1^0}(E_1) = \frac{-\lambda V k_B T_0^2}{\langle \dot{E}_1^0 \rangle_0} E_1 \equiv -c E_1, \text{ (future)}, \quad (4.46)$$

where the thermal conductivity is related to the transport coefficient by $\sigma_{\text{ss}} = -1/2\lambda V T_0^2$.

Since the asymptote is odd in time, it integrates to zero and the mirror work can be written

$$W_1^{\text{mir}}(\Gamma) = \frac{1}{2} \int_{-\tau}^{\tau} dt' [\dot{E}_1^0(\Gamma_0(t'|\Gamma)) - \text{sign}(t') \overline{\dot{E}_1^0}(E_1(\Gamma))] \quad (4.47)$$

This shows explicitly that the mirror work is dominated by the brief inertial part of the trajectory, and that it is independent of τ for $\tau \gtrsim \tau_{\text{short}}$.

In so far as the time integral that is W_1^{mir} is dominated by short times, then the temporally even part of $\dot{E}_1(t)$ must vanish by $|t| \lesssim \tau_{\text{short}}$, the inertial time. Hence a reasonable estimate of integral is

$$W_1^{\text{mir}}(\Gamma) \approx \frac{\tau_{\text{short}}}{2} \dot{E}_1^0(\Gamma). \quad (4.48)$$

This is consistent with an early ansatz,³ which was shown to be a reasonable approximation in certain regimes.⁵

An estimate of the inertial time was given above as eqn (3.19).⁴ For the present problem of heat flow it becomes

$$\begin{aligned} \tau_{\text{short}} &= \frac{-2\langle \dot{E}_1^0(t) E_1(0) \rangle_0}{\langle (\dot{E}_1^0)^2 \rangle_0}, \quad t \gg \tau_{\text{short}}, \\ &= \frac{2\lambda V k_B T_0^2}{\langle (\dot{E}_1^0)^2 \rangle_0}. \end{aligned} \quad (4.49)$$

The long time limit was given as eqn (3.17), and for the present case it is explicitly⁴

$$\tau_{\text{long}} = \frac{\langle E_1^2 \rangle_0}{2\lambda V k_B T_0^2}. \quad (4.50)$$

4.5.3 Adiabatic evolution. Let $\Gamma' = \Gamma + \Delta_t \dot{\Gamma}$, be the adiabatic evolution of Γ after an infinitesimal time step. The adiabatic evolution of W_1^{mir} can be obtained from

$$\begin{aligned} W_1^{\text{mir}}(\Gamma) &= \frac{1}{2} \int_{-\tau}^{\tau} dt \dot{E}_1^0(\Gamma_0(t|\Gamma)) \\ &= \frac{1}{2} \int_{-\tau-\Delta_t}^{\tau-\Delta_t} dt'' \dot{E}_1^0(\Gamma_0(t'' + \Delta_t|\Gamma)) \\ &= \frac{1}{2} \int_{-\tau-\Delta_t}^{\tau-\Delta_t} dt'' \dot{E}_1^0(\Gamma_0(t''|\Gamma')) \\ &= W_1^{\text{mir}}(\Gamma') - \frac{\Delta_t}{2} [\dot{E}_1^0(\Gamma_0(\tau|\Gamma')) \\ &\quad - \dot{E}_1^0(\Gamma_0(-\tau|\Gamma'))] \\ &= W_1^{\text{mir}}(\Gamma') - \Delta_t \overline{\dot{E}_1^0}(E_1(\Gamma')). \end{aligned} \quad (4.51)$$

Hence the adiabatic rate of change is⁷

$$\dot{W}_1^{\text{mir}}(\Gamma) = \overline{\dot{E}_1^0}(E_1(\Gamma)), \quad (4.52)$$

which has even parity, as it ought. This result is exact for τ in the intermediate regime.

This result for the adiabatic evolution of W_1^{mir} provides us with a physical interpretation of the steady state probability, eqn (4.41). The difficulty with heat flow is the lack of conserved variables; the change in energy moment ΔE_1 occurs both internally by the adiabatic evolution of the isolated sub-system, $\Delta^0 E_1$, and externally by exchange with the reservoir, $\Delta^{s(r)} E_1 = -\Delta E_{1r}$. This is in contrast to the canonical Boltzmann distribution where the total energy is conserved, so that $\Delta E_0 = \Delta^0 E_0 + \Delta^{s(r)} E_0 - \Delta E_{0r}$, which is why the Boltzmann exponent is $\Delta S_r / k_B = -\beta_0 E_0$. In the present case, the change in entropy of the reservoirs is

$$\begin{aligned} \Delta S_r / k_B &= \beta_0 \Delta E_{0r} + \beta_1 \Delta E_{1r} \\ &= -\beta_0 \Delta E_0 - \beta_1 [\Delta E_1 - \Delta^0 E_1] \\ &\approx -\beta_0 \Delta E_0 - \beta_1 [\Delta E_1 - \Delta_t \overline{\dot{E}_1^0}] \\ &\approx -\beta_0 \Delta E_0 - \beta_1 [\Delta E_1 - \Delta W_1^{\text{mir}}]. \end{aligned} \quad (4.53)$$

From the final approximate expression, we see that subtracting the change in W_1^{mir} from the total change in moment is equivalent to identifying that part of the change in energy moment that is due to the reservoir. It is only this external influence that affects the reservoir entropy and contributes to the steady state probability.

In section 4.6.2 below we show that the steady state probability distribution is approximately stationary during the adiabatic evolution of the isolated sub-system. This confirms the above interpretation that the exponent reflects the entropy of the reservoirs only, and that the contribution from internal changes of the sub-system have been correctly removed.

4.5.4 Relationship with Green–Kubo. We now show that the steady state probability density, eqn (4.41), gives the Green–Kubo expression for the thermal conductivity. Fourier’s law, eqn (4.27), gives the heat flux due to an applied temperature gradient. Since the heat flux is the rate of change of energy moment, the average of the latter in the steady state can be used to obtain the thermal conductivity. Linearizing the exponents for small temperature gradients, $\beta_1 \ll 1$, we obtain

$$\begin{aligned} \langle \dot{E}_1^0(\Gamma) \rangle_{\text{ss}} &= \frac{\int d\Gamma e^{-\beta_0 E_0(\Gamma)} e^{-\beta_1 E_1(\Gamma)} e^{\beta_1 W_1^{\text{mir}}(\Gamma)} \dot{E}_1^0(\Gamma)}{\int d\Gamma e^{-\beta_0 E_0(\Gamma)} e^{-\beta_1 E_1(\Gamma)} e^{\beta_1 W_1^{\text{mir}}(\Gamma)}} \\ &= \frac{\int d\Gamma e^{-\beta_0 E_0(\Gamma)} [1 - \beta_1 E_1(\Gamma) + \beta_1 W_1^{\text{mir}}(\Gamma)] \dot{E}_1^0(\Gamma)}{\int d\Gamma e^{-\beta_0 E_0(\Gamma)} [1 - \beta_1 E_1(\Gamma) + \beta_1 W_1^{\text{mir}}(\Gamma)]} \quad (4.54) \\ &= \frac{\beta_1 \int d\Gamma e^{-\beta_0 E_0(\Gamma)} W_1^{\text{mir}}(\Gamma) \dot{E}_1^0(\Gamma)}{\int d\Gamma e^{-\beta_0 E_0(\Gamma)}} \\ &= \beta_1 \langle W_1^{\text{mir}}(\Gamma) \dot{E}_1^0(\Gamma) \rangle_0, \end{aligned}$$

Since \dot{E}_1^0 is of odd parity, the integral of it times the two even terms from the linearization of the exponents vanishes. The integral over $-\beta_1 E_1(\Gamma)$ in the denominator vanishes because it is odd in z .

Using the definition of the mirror work and Fourier’s law, eqn (4.27), this gives the thermal conductivity as

$$\begin{aligned} \lambda(\tau) &= \frac{1}{2V k_B T_0^2} \int_{-\tau}^{\tau} dt \langle \dot{E}_1^0(\Gamma_0(t|\Gamma)) \dot{E}_1^0(\Gamma) \rangle_0 \\ &= \frac{1}{V k_B T_0^2} \int_0^{\tau} dt \langle \dot{E}_1^0(\Gamma_0(t|\Gamma)) \dot{E}_1^0(\Gamma) \rangle_0 \quad (4.55) \end{aligned}$$

The second equality follows because the time correlation function of two quantities of the same parity is an even function of time. In the intermediate regime, this may be recognised as the Green–Kubo expression for the thermal conductivity,¹¹ which in turn is equivalent to the Onsager expression for the transport coefficients.⁴

This result is a very stringent test of the present expression for the steady state probability distribution, eqn (4.41). There is one, and only one, exponent that is odd, linear in β_1 , and that satisfies the Green–Kubo relation.

We can likewise test the even part of the probability by taking the steady state average of an even function, namely E_1 . Similar arguments show that

$$\langle E_1(t) \rangle_{\text{ss}} = -\beta_1 \langle E_1(t)^2 \rangle_0 = S^{-1} T_1^{-1}, \quad (4.56)$$

which is in full agreement with the fluctuation expression for the first entropy, eqn (4.16). This confirms that the even part of the steady state probability distribution, eqn (4.41), is correct since there is one, and only one, even exponent that is linear in β_1 that will yield this result.

We now consider similar analysis for the tentative steady state probability, eqn (4.39). This will also yield the Green–Kubo formula when the average of \dot{E}_1^0 is taken, because the odd parity part of that distribution is identical to the mirror work and so it gives the same result as that obtained for the steady state probability distribution. However, the part with even parity is different and in this case one obtains

$$\begin{aligned} \langle E_1(t) \rangle_{\text{tent}} &= -\beta_1 \langle E_1(t) E_1(t-\tau) \rangle_0 \\ &= -\beta_1 \langle E_1(t)^2 \rangle_0 + \beta_1 \tau \langle E_1(t) \dot{E}_1^0(t; -\tau) \rangle_0 \\ &\approx -\beta_1 \langle E_1(t)^2 \rangle_0 - \beta_1 \tau \left\langle E_1(t) \overline{\dot{E}_1^0(E_1(t))} \right\rangle_0 \\ &= -\beta_1 (1 - \tau c) \langle E_1(t)^2 \rangle_0, \end{aligned} \quad (4.57)$$

where eqn (4.46) for the most likely velocity has been used. This differs from the correct result by the factor $1 - \tau c$ compared to 1. This confirms that the tentative probability distribution yields an average moment that, during the adiabatic evolution to the present time, has decreased from what it should be. Of course the term τc equals $-\tau \dot{E}_1^0/E_1$, which we have been assuming to be small for τ in the intermediate regime. Nevertheless the analysis does show the sensitivity of the tentative probability distribution, eqn (4.39), to the value chosen for τ , and it does show the improvement offered by the steady state probability distribution, eqn (4.41).

4.6. Microstate transitions

4.6.1 Transition probability. We now explore the nature of the microscopic transitions in the steady state system. For these it is essential to account for the stochastic perturbations from the reservoirs. We consider the transition between the microstates $\Gamma \rightarrow \Gamma''$ in an infinitesimal time step Δ_t . This comprises a deterministic transition $\Gamma \rightarrow \Gamma'$ due to the internal forces of the sub-system, followed by a stochastic transition $\Gamma' \rightarrow \Gamma''$ due to the perturbations by the reservoir.

The deterministic transition is just the adiabatic evolution of the isolated sub-system, $\Gamma' = \Gamma + \Delta_t \dot{\Gamma}$. In terms of the conditional transition probability this is

$$A_d(\Gamma'|\Gamma) = \delta(\Gamma' - \Gamma - \Delta_t \dot{\Gamma}) \quad (4.58)$$

The stochastic transition probability can be written as the product of an even and an odd function,

$$A_s(\Gamma''|\Gamma') = A_e(\Gamma''|\Gamma') A_o(\Gamma''|\Gamma'). \quad (4.59)$$

The parity of a transition refers to the reversibility or irreversibility of the transition. That is, even transition probabilities satisfy

$$A_e(\Gamma''|\Gamma') = A_e(\Gamma'^{\dagger}|\Gamma''^{\dagger}), \quad (4.60)$$

and odd transition probabilities satisfy

$$A_o(\Gamma''|\Gamma') = 1/A_o(\Gamma'^{\dagger}|\Gamma''^{\dagger}). \quad (4.61)$$

We can use entropic arguments to obtain an expression for the odd transition probability. The change in entropy of the reservoirs during the stochastic transition $\Gamma \rightarrow \Gamma''$ is

$$\begin{aligned} \Delta S/k_B &= -\beta_0 \Delta E_0 - \beta_1 [\Delta E_1 - \Delta^0 E_1] \\ &= -\beta_0 (E_0'' - E_0) - \beta_1 (E_1'' - E_1 - \Delta_t \dot{E}_1^0) \\ &= -\beta_0 (E_0'' - E_0') - \beta_1 (E_1'' - E_1'). \end{aligned} \quad (4.62)$$

Here $\Delta E_1 \equiv E_1'' - E_1$ is the total change in moment, and $\Delta^0 E_1 \equiv E_1' - E_1$ is the deterministic change in moment. It is only the stochastic part of the change in moment that is due to perturbations by the reservoirs and that contributes to the change in entropy of the reservoirs. Accordingly, we may take the odd stochastic transition probability to be the exponential of half the change in entropy,⁶

$$\Lambda_o(\Gamma''|\Gamma') = e^{-\beta_0(E_0'' - E_0')/2} e^{-\beta_1(E_1'' - E_1')/2} \quad (4.63)$$

The reason for taking half the change in entropy is that when multiplied by the even part of the probability density, this gives an unconditional transition probability that obeys dynamic reversibility, $\Lambda_o(\Gamma''|\Gamma') e^{-\beta_0 E_0'} e^{-\beta_1 E_1'} = e^{-\beta_0(E_0'' + E_0')/2} e^{-\beta_1(E_1'' + E_1')/2}$. Such stochastic transition probabilities were originally used in equilibrium contexts.^{17,18}

The even stochastic transition probability may be taken as

$$\Lambda_e(\Gamma''|\Gamma') = \Theta_\Delta(|\Gamma'' - \Gamma'|). \quad (4.64)$$

Here Θ_Δ is a short ranged, even function, (such as a Heaviside step function, or a Gaussian), with range Δ that represents the strength of the stochastic perturbations from the reservoir. It is normalized, and does not depend on the temperatures to leading order.

It is possible to multiply the even stochastic transition probability by a factor of $e^{\beta_1(W_1^{\text{mir}''} - W_1^{\text{mir}'})/2}$. This does not affect the normalization of Θ_Δ to leading order. All of the results given below would be unchanged if this factor were to be added. This includes the microscopic transition theorem established in ref. 6, which is independent of the even transition probability.

With these expressions, the stochastic conditional transition probability is explicitly

$$\Lambda_s(\Gamma''|\Gamma') = \Theta_\Delta(|\Gamma'' - \Gamma'|) e^{-\beta_0(E_0'' - E_0')/2} e^{-\beta_1(E_1'' - E_1')/2}. \quad (4.65)$$

The exponent is half the change in reservoir entropy during the transition.

4.6.2 Stationary steady state probability. For steady heat flow the probability density does not depend explicitly on time, and so it must be stationary under the transition probability given above. Under the adiabatic evolution for a time-step Δ_t , the probability density becomes

$$\begin{aligned} \mathcal{P}(\Gamma'|A_t) &= \mathcal{P}_{ss}(\Gamma'|\beta_0, \beta_1) \\ &= \mathcal{P}_{ss}(\Gamma'|\beta_0, \beta_1) e^{\beta_1 \Delta_t [E_1^0(\Gamma') - \dot{W}_1^{\text{mir}}(\Gamma')]} \end{aligned} \quad (4.66)$$

Hence the steady state probability distribution is preserved to the extent that the exponent shown explicitly is zero. (To leading order we can use Γ or Γ' in this term.)

We have already seen that $\dot{W}_1^{\text{mir}}(\Gamma) = \overline{\dot{E}_1^0}(\Gamma)$. The overbar signifies the most likely state, and it is in such states that the steady state probability is significantly greater than zero. Conversely, states with $\dot{E}_1^0 \neq \overline{\dot{E}_1^0}$ occur rarely, with the probability getting smaller as the difference increases. (Fluctuations away from equilibrium or the steady state are relatively negligible.⁸) Hence in the most likely state the final term vanishes, and for nearby states it is small. Accordingly the present steady state probability density is almost stationary during its adiabatic evolution.

During the stochastic step, the transition integral becomes

$$\begin{aligned} &\int d\Gamma' \Lambda_s(\Gamma''|\Gamma') \mathcal{P}_{ss}(\Gamma'|\beta_0, \beta_1) e^{\beta_1 \Delta_t [E_1^0(\Gamma') - \dot{W}_1^{\text{mir}}(\Gamma')]} \\ &= \int d\Gamma' \Theta_\Delta(|\Gamma'' - \Gamma'|) e^{-\beta_0(E_0'' - E_0')/2} e^{-\beta_1(E_1'' - E_1')/2} \\ &\quad \times \frac{e^{-\beta_0 E_0'} e^{-\beta_1 E_1'} e^{\beta_1 W_1^{\text{mir}'}}}{Z_{ss}} e^{\beta_1 \Delta_t [E_1^0(\Gamma') - \dot{W}_1^{\text{mir}}(\Gamma')]} \\ &= \frac{e^{-\beta_0 E_0''} e^{-\beta_1 E_1''} e^{\beta_1 W_1^{\text{mir}''}}}{Z_{ss}} \int d\Gamma' \Theta_\Delta(|\Gamma'' - \Gamma'|) \\ &\quad \times e^{\beta_1 (W_1^{\text{mir}' - W_1^{\text{mir}''})} e^{-\beta_0(E_0' - E_0'')/2} e^{-\beta_1(E_1' - E_1'')/2} \\ &\quad \times e^{\beta_1 \Delta_t [E_1^0(\Gamma') - \dot{W}_1^{\text{mir}}(\Gamma')]} \\ &= \mathcal{P}_{ss}(\Gamma''|\beta_0, \beta_1) \int d\Gamma' \Lambda_s(\Gamma'|\Gamma'') \\ &\quad \times e^{\beta_1 (W_1^{\text{mir}'' - W_1^{\text{mir}'})} e^{\beta_1 \Delta_t [E_1^0(\Gamma') - \dot{W}_1^{\text{mir}}(\Gamma')]} \\ &= \mathcal{P}_{ss}(\Gamma''|\beta_0, \beta_1). \end{aligned} \quad (4.67)$$

This is the required result which shows the stationarity of the steady state probability under the present transition probability.

The final equality holds to order Δ^2 . The normalization of the conditional transition probability, $\int d\Gamma' \Lambda_s(\Gamma'|\Gamma'') = 1$, is unchanged to order Δ^2 by the factor of $e^{\beta_1(W_1^{\text{mir}''} - W_1^{\text{mir}'})}$. We have also neglected the term $\beta_1 \Delta_t [\dot{E}_1^0(\Gamma') - \dot{W}_1^{\text{mir}}(\Gamma')]$ on the grounds that the fluctuations away from the most likely state are relatively negligible in the thermodynamic limit.

4.6.3 Forward and reverse transitions. In the present case of steady heat flow, the unconditional microscopic transition probability is

$$\begin{aligned} \mathcal{P}(\Gamma'' \leftarrow \Gamma|\Delta_t) &= \Lambda_s(\Gamma''|\Gamma') \mathcal{P}_{ss}(\Gamma|\beta_0, \beta_1) \\ &= \Theta_\Delta(|\Gamma'' - \Gamma'|) e^{-\beta_0(E_0'' + E_0)/2} \\ &\quad \times e^{-\beta_1(E_1'' + E_1 - \Delta_t \dot{E}_1)/2} e^{\beta_1 W_1^{\text{mir}}}/Z_{ss}(\beta_0, \beta_1) \\ &= \Theta_\Delta(|\Gamma'' - \Gamma'|) \sqrt{\mathcal{P}_{ss}(\Gamma''|\beta_0, \beta_1) \mathcal{P}_{ss}(\Gamma|\beta_0, \beta_1)} \\ &\quad \times e^{-\beta_1(W_1^{\text{mir}''} - W_1^{\text{mir}})/2} e^{\Delta_t \beta_1 E_1^0/2}, \end{aligned} \quad (4.68)$$

for an infinitesimal Δ_t . The \dot{E}_1^0 in the final term can be replaced by $(\dot{E}_1^0 + \dot{E}_1^0)/2$ to this order.

We consider the forward transition, $\Gamma \rightarrow \Gamma' \rightarrow \Gamma''$, and its reverse $\Gamma'' \rightarrow \Gamma' \rightarrow \Gamma$. Note that $\Gamma'' \neq \Gamma'$, but that $|\Gamma'' - \Gamma'| = |\Gamma' - \Gamma''|$. The ratio of the forward to the reverse

transition probabilities is

$$\frac{\mathcal{P}(\Gamma'' \leftarrow \Gamma|\Delta_t)}{\mathcal{P}(\Gamma^\dagger \leftarrow \Gamma''|\Delta_t)} = e^{\beta_1(W_1^{\text{mir}}(\Gamma'') + W_1^{\text{mir}}(\Gamma))} e^{\Delta_t \beta_1(E_1^0(\Gamma'') + \dot{E}_1^0(\Gamma))/2}, \quad (4.69)$$

the even terms cancelling. The exponent of the final term is part of the change in reservoir entropy (the thermodynamic work done) on the trajectory, eqn (4.62).

We now consider a trajectory $[\Gamma] = \{\Gamma_0, \Gamma_1, \dots, \Gamma_f\}$, over a time interval $t_f = f\Delta_t$, and its reverse, $[\Gamma^\dagger] = \{\Gamma_f^\dagger, \Gamma_{f-1}^\dagger, \dots, \Gamma_0^\dagger\}$. The adiabatic change in E_1 over the trajectory is

$$\Delta^0 E_1[\Gamma] = \frac{\Delta_t}{2} [\dot{E}_1^0(\Gamma_0) + \dot{E}_1^0(\Gamma_f)] + \Delta_t \sum_{i=1}^{f-1} \dot{E}_1^0(\Gamma_i). \quad (4.70)$$

Clearly $\Delta^0 E_1[\Gamma^\dagger] = -\Delta^0 E_1[\Gamma]$. In view of eqn (4.62), for large $f\Delta_t$ this gives the change in the reservoirs' entropy over the trajectory, $\Delta S_r/k_B = -\beta_0 \Delta E_0 - \beta_1 \Delta E_1 + \beta_1 \Delta^0 E_1 \approx \beta_1 \Delta^0 E_1$, since the surviving term is the only one that grows with the length of the trajectory.

The unconditional probability of the trajectory is

$$\begin{aligned} \mathcal{P}[\Gamma] &= \prod_{i=1}^f [\Theta_{\Delta}(|\Gamma_{i+1} - \Gamma_i'|) e^{-\beta_0(E_{0,i} - E_{0,i-1}')/2} e^{-\beta_1(E_{1,i} - E_{1,i-1}')/2}] \\ &\quad \times \mathcal{P}_{\text{ss}}(\Gamma_0|\beta_0, \beta_1) \\ &= \prod_{i=1}^f [\Theta_{\Delta}(|\Gamma_{i+1} - \Gamma_i'|) e^{\beta_1 \Delta^0 E_1/2} e^{-\beta_1(W_{1,f}^{\text{mir}} - W_{1,0}^{\text{mir}})/2}] \\ &\quad \times \sqrt{\mathcal{P}_{\text{ss}}(\Gamma_f|\beta_0, \beta_1) \mathcal{P}_{\text{ss}}(\Gamma_0|\beta_0, \beta_1)}, \end{aligned} \quad (4.71)$$

since $E'_{1,i-1} = E_{1,i-1} + \Delta_t \dot{E}_{1,i-1}^0$. The exponent $\Delta^0 \dot{E}_1$ should really be replaced by $\Delta^0 E_1 + \Delta_t [E_1^0(\Gamma_0) - \dot{E}_1^0(\Gamma_f)]/2$, but this correction is negligible for large f , and in any case it is fixed up by the ratio that is now taken.

The ratio of the probability of the forward and reverse trajectory in the steady state is

$$\begin{aligned} \frac{\mathcal{P}[\Gamma]}{\mathcal{P}[\Gamma^\dagger]} &= e^{\beta_1 W_1^{\text{mir}}(\Gamma_f)} e^{\beta_1 W_1^{\text{mir}}(\Gamma_0)} e^{\beta_1 \Delta^0 E_1[\Gamma]} \\ &\approx e^{\beta_1 \Delta^0 E_1[\Gamma]}, \end{aligned} \quad (4.72)$$

the even terms cancelling. The final approximation is valid for large t_f , since the retained term scales with $f\Delta_t$. The final exponent is the change in the entropy of the reservoirs, eqn (4.62) in this limit. The result for the ratio of the probabilities of the forward and reverse trajectories may be called the *reverse transition theorem*.⁶

It is possible to consider trajectories that evolve to the steady state rather than begin in the steady state. This provides slightly neater forms for certain ratios in the short time limit, even though arguably the results are less physically applicable. We let the starting point of the trajectory be drawn from a probability density with even parity. The most appropriate density applies to the so-called static state,³

$$\mathcal{P}_{\text{st}}(\Gamma|\beta_0, \beta_1) = \frac{e^{-\beta_0 E_0(\Gamma)} e^{-\beta_1 E_1(\Gamma)}}{Z_{\text{st}}(\beta_0, \beta_1)}. \quad (4.73)$$

In this case the trajectory probability during the evolution to the steady state is

$$\begin{aligned} \mathcal{P}_{\text{st}}[\Gamma] &= \prod_{i=1}^f [\Theta_{\Delta}(|\Gamma_{i+1} - \Gamma_i'|)] e^{\beta_1 \Delta^0 E_1/2} \\ &\quad \times \sqrt{\mathcal{P}_{\text{st}}(\Gamma_f|\beta_0, \beta_1) \mathcal{P}_{\text{st}}(\Gamma_0|\beta_0, \beta_1)}, \end{aligned} \quad (4.74)$$

and the ratio of the forward and reverse trajectories is

$$\frac{\mathcal{P}_{\text{st}}[\Gamma]}{\mathcal{P}_{\text{st}}[\Gamma^\dagger]} = e^{\beta_1 \Delta^0 E_1[\Gamma]}. \quad (4.75)$$

The probability of observing the entropy of the reservoirs change by ΔS over a period t is related to the probability of observing the opposite change by

$$\begin{aligned} \mathcal{P}(\Delta S|\beta_0, \beta_1, t_f) &= \int d[\Gamma] \delta(\Delta S - \Delta S[\Gamma]) \mathcal{P}[\Gamma] \\ &\approx k_B^{-1} \int d[\Gamma^\dagger] \delta(\Delta S/k_B - \beta_1 \Delta^0 E_1[\Gamma]) \\ &\quad \times \mathcal{P}[\Gamma^\dagger] e^{\beta_1 \Delta^0 E_1[\Gamma]} \\ &= e^{\Delta S/k_B} \mathcal{P}(-\Delta S|\beta_0, \beta_1, t_f). \end{aligned} \quad (4.76)$$

(Using instead trajectories that begin and end with the static probability distribution, we have the exact result, $\wp_{\text{st}}(\Delta^0 E_1|\beta_0, \beta_1, t_f) = e^{\beta_1 \Delta^0 E_1} \wp_{\text{st}}(-\Delta^0 E_1|\beta_0, \beta_1, t_f)$.) This result says in essence that the probability of a positive increase in entropy is exponentially greater than the probability of a decrease in entropy during heat flow. In essence this is the temperature gradient version of the fluctuation theorem that was originally derived by Evans *et al.* for mechanical work (see section 6.2 below).^{19,20} The present derivation for heat flow is based upon the microscopic transition probability given by Attard.⁶

Closely related to the fluctuation theorem is the work theorem due to Jarzynski.²¹ In this context we are motivated to consider the average of the exponential of the negative of the heat flux,

$$\begin{aligned} \langle e^{-\beta_1 \Delta^0 E_1} \rangle_{\text{ss}, t_f} &= \int d[\Gamma] e^{-\beta_1 \Delta^0 E_1[\Gamma]} \mathcal{P}([\Gamma]|\beta_0, \beta_1, t_f) \\ &= \int d[\Gamma^\dagger] \mathcal{P}([\Gamma^\dagger]|\beta_0, \beta_1, t_f) \\ &\quad \times e^{\beta_1(W_1^{\text{mir}}(\Gamma_f) + W_1^{\text{mir}}(\Gamma_0))} \\ &\approx \int d\Gamma_f^\dagger \mathcal{P}(\Gamma_f^\dagger|\beta_0, \beta_1) e^{-\beta_1 W_1^{\text{mir}}(\Gamma_f^\dagger)} \\ &\quad \times \int d\Gamma_0^\dagger \mathcal{P}(\Gamma_0^\dagger|\beta_0, \beta_1) e^{-\beta_1 W_1^{\text{mir}}(\Gamma_0^\dagger)} \\ &= \left(\frac{Z_{\text{st}}(\beta_0, \beta_1)}{Z_{\text{ss}}(\beta_0, \beta_1)} \right)^2. \end{aligned} \quad (4.77)$$

Here it has been assumed that the trajectory is long enough that the ends are uncorrelated. This result shows that this particular average is not extensive in time, (*i.e.* it does not scale with t_f). In essence the right hand side is the exponential of twice the difference in free energies of the static and steady state systems. If the trajectories were begun from the static distribution, the right hand side would be unity.

4.6.4 Second entropy. We now derive the relationship between the steady state probability density and the second entropy for heat flow. Three relevant quantities appear in the steady state probability density, eqn (4.41): E_1 , E_1^+ , and E_1^- , since the mirror work term is $W_1^{\text{mir}} = (E_1^+ - E_1^-)/2$. Accordingly, the steady state probability density really describes the sequential transition, $E_1^- \rightarrow E_1 \rightarrow E_1^+$, with each interval of duration τ . The exponent represents the reservoir contribution to the total entropy, since a sub-system microstate, a point in its phase space, has zero entropy. Hence the total second entropy for this sequential transition is just the logarithm of the sum over the macrostate of the weights of the microstates, the latter being the unnormalised steady state probability,

$$\begin{aligned} S_{\text{total}}^{(2)}(E_1^+, E_1, E_1^-|\tau)/k_B &= \ln \int d\Gamma \exp[-\beta_0 E_0(\Gamma) - \beta_1 E_1(\Gamma)] \\ &\quad + \beta_1 \{E_1^+(\Gamma) - E_1^-(\Gamma)\} / 2 \delta(E_1 - E_1(\Gamma)) \\ &\quad \times \delta(E_1^+ - E_1^+(\Gamma)) \delta(E_1^- - E_1^-(\Gamma)) \\ &= S_0^{(2)}(E_1^+, E_1, E_1^-|\tau)/k_B \\ &\quad - \beta_1 E_1 + \beta_1 (E_1^+ - E_1^-)/2. \end{aligned} \quad (4.78)$$

The second entropy of the isolated system during the adiabatic sequential transition is

$$\begin{aligned} S_0^{(2)}(E_1^+, E_1, E_1^-|\tau) &= \frac{\sigma_{\text{ss}}}{2\tau} (E_1^+ - E_1)^2 + \frac{\sigma_{\text{ss}}}{2\tau} (E_1 - E_1^-)^2 \\ &\quad + \frac{S^{(1)}}{2} E_1^2, \tau > \tau_{\text{short}}. \end{aligned} \quad (4.79)$$

This is eqn (3.11) applied to the present problem. The final term can be replaced by the first or ordinary entropy, $S^{(1)}(E_1)$, which in turn to leading order can be replaced by $[S^{(1)}(E_1^-) + S^{(1)}(E_1^+)]/2$ without changing the following results for the coarse velocity over the *entire* interval.

The derivatives of the total second entropy are

$$\frac{\partial S_{\text{total}}^{(2)}}{\partial E_1^+} = \sigma_{\text{ss}} \frac{E_1^+ - E_1}{\tau} + \frac{\beta_1}{2}, \quad (4.80)$$

$$\frac{\partial S_{\text{total}}^{(2)}}{\partial E_1^-} = -\sigma_{\text{ss}} \frac{E_1 - E_1^-}{\tau} - \frac{\beta_1}{2}, \quad (4.81)$$

and

$$\frac{\partial S_{\text{total}}^{(2)}}{\partial E_1} = -\sigma_{\text{ss}} \frac{E_1^+ - E_1}{\tau} + \sigma_{\text{ss}} \frac{E_1 - E_1^-}{\tau} - \beta_1 + S^{(1)} E_1. \quad (4.82)$$

Maximizing the total second entropy by setting these derivatives to zero, we see that the most likely moment is

$$\bar{E}_1 = (S^{(1)})^{-1} \beta_1. \quad (4.83)$$

and that the most likely flux is

$$\bar{E}_1^+ = \bar{E}_1^- = -\frac{\sigma_{\text{ss}}^{-1}}{2} \beta_1, \quad (4.84)$$

These agree with eqn (4.16) and (4.24). As mentioned above, the second entropy form of the transport coefficient is related to the thermal conductivity by $\sigma_{\text{ss}} = -1/2\lambda VT_0^2$.

We may conclude two things from this analysis. First, the present steady state probability distribution is entirely consistent with the second entropy analysis of heat transport. This demonstrates again the necessity of the mirror work term W_1^{mir} . Second, the steady state probability distribution really describes transitions rather than phase space microstates *per se*, since the exponent is so closely related to the second entropy. This surprising result contrasts with equilibrium systems where it is the first entropy or state weight that is relevant to the probability density.

5. Monte Carlo simulations of heat flow

The availability of a phase space probability distribution for the steady state means that it is possible to develop Monte Carlo algorithms for the computer simulation of non-equilibrium systems. We give two algorithms here, one based on the probability distribution and the Metropolis algorithm,⁷ and the other based on the transition probability and the von Neumann algorithm.⁶ Some aspects of the system geometry and potential are common to both methods.

5.1. System details

A Lennard-Jones fluid was simulated, with the pair potential cut and shifted at $R_{\text{cut}} = 2.5$. No tail correction was used. The shift to make the potential zero at the cut-off is necessary for consistency between the Monte Carlo and molecular dynamics aspects of the computations.

All quantities were made dimensionless using the well-depth ε_{LJ} , the diameter σ_{LJ} , and the time constant $\tau_{\text{LJ}} = \sqrt{(m_{\text{LJ}}\sigma_{\text{LJ}}^2/\varepsilon_{\text{LJ}})}$, where m_{LJ} is the mass. In addition, Boltzmann's constant was set equal to unity.

A spatial neighbor table was used with cubic cells of side-length ≈ 0.6 .³ At the beginning of the simulation a list of neighbor cells within the cut-off of each cell was calculated and stored. The neighborhood volume composed of such small neighbor cells can be made to closely approximate the cut-off sphere. This reduces the enveloping neighborhood volume from 27 large cubes, (each of size R_{cut} , neighborhood volume $27R_{\text{cut}}^3$), to approximately 667 small cubes giving a neighborhood volume of the order of $(4\pi/3)(R_{\text{cut}} + 0.6)^3$. The number of neighbors required for a force or potential calculation with these small cells is almost a factor of three smaller than for the conventional cells of length R_{cut} .

Both a uniform bulk fluid and an inhomogeneous fluid were simulated. The latter was in the form of a slit pore, terminated in the z -direction by uniform Lennard-Jones walls. The distance between the walls for a given number of atoms was chosen so that the uniform density in the center of the cell was equal to the nominal bulk density. The effective width of the slit pore used to calculate the volume of the sub-system was taken as the region where the density was non-zero. For the bulk fluid in all directions, and for the slit pore in the lateral directions, periodic boundary conditions and the minimum image convention were used.

The energy per atom consists of kinetic energy, singlet, and pair potential terms,

$$\varepsilon_i = \frac{1}{2m} \mathbf{p}_i \cdot \mathbf{p}_i + w(q_{iz}) + \frac{1}{2} \sum_{j=1}^N (j \neq i) u(q_{ij}), \quad (5.1)$$

where $w(z)$ is the wall-potential, (if present). In terms of this the zeroth moment is just the total energy,

$$E_0(\Gamma) = \sum_{i=1}^N \varepsilon_i, \quad (5.2)$$

and the first moment is

$$E_1(\Gamma) = \sum_{i=1}^N q_{iz} \varepsilon_i. \quad (5.3)$$

The adiabatic rate of change of the first energy moment is

$$\dot{E}_1^0(\Gamma) = \sum_{i=1}^N \dot{q}_{iz} \varepsilon_i + q_{iz} \dot{\varepsilon}_i, \quad (5.4)$$

where the velocity is $\dot{q}_{iz} = p_{iz}/m_{LJ}$. Using Hamilton's equations, it is readily shown that

$$\dot{\varepsilon}_i = \frac{-1}{2} \sum_{j=1}^N (j \neq i) u'(q_{ij}) \frac{\mathbf{q}_{ij} \cdot [\mathbf{p}_i + \mathbf{p}_j]}{mq_{ij}}. \quad (5.5)$$

This holds whether or not the singlet potential is present. In the case of periodic boundary conditions, it is quite important to use the minimum image convention for all the separations that appear in this expression. This may be rewritten in the convenient form⁶

$$\dot{E}_1^0 = \sum_{iz} \kappa_{iz} p_{iz} / m, \quad (5.6)$$

where

$$\kappa_i = \frac{\varepsilon_i}{m} \hat{\mathbf{z}} - \sum_{j=1}^N (j \neq i) u'(q_{ij}) \frac{[q_{iz} - q_{jz}]}{2mq_{ij}} \mathbf{q}_{ij}. \quad (5.7)$$

The Monte Carlo algorithms require ΔE_0 , and ΔE_1 , and $\Delta \dot{E}_1^0$. In attempting to move atom n in phase space, we identify the n -dependent contribution to these formulae and only calculate the change in this for each attempted move.

In calculating $E_1(t)$ on a trajectory, it is essential to integrate $\dot{E}_1^0(t)$ over the trajectory rather than use the expression for $E_1(\Gamma(t))$ given above. This is because \dot{E}_1^0 is insensitive to the periodic boundary conditions, whereas E_1 depends on whether the coordinates of the atom are confined to the central cell, or whether the itinerant coordinate is used, and problems arise in both cases when the atom leaves the central cell on a trajectory.

Because the starting position of each trajectory was taken from a Boltzmann-weighted distribution in $6N$ -dimensional phase space, the center of mass velocity of the system (the total linear momentum) was generally non-zero. Prior to commencing each molecular dynamics trajectory, the z -component of the center of mass velocity was zeroed at constant kinetic energy by shifting and rescaling the z -component of the momenta. (Only the z -component of the first energy moment was used.) It was found that a non-zero center of mass velocity made a non-negligible contribution to the conductivity. Conventional molecular dynamics simulations are performed with zero center of mass velocity, which is of course the most appropriate model of reality. For the bulk case the total z -momentum was conserved at zero along the molecular

dynamics trajectory. For the inhomogeneous simulations, the momentum during the adiabatic evolution was not conserved due to collisions with the walls. In this case an additional external force was applied to each atom that was equal and opposite to the net wall force per atom, which had the effect of conserving the z -component of the total linear momentum at zero along the molecular dynamics trajectory.

5.2. Metropolis algorithm

We performed umbrella sampling Monte Carlo simulations in $6N$ -dimensional phase space, where $N = 120$ – 500 atoms.⁷ We used the Metropolis algorithm with umbrella sampling. The weight density used was

$$\omega(\Gamma) = e^{-\beta_0 E_0(\Gamma)} e^{-\beta_1 E_1(\Gamma)} e^{\alpha \beta_1 \dot{E}_1^0(\Gamma)}. \quad (5.8)$$

We corrected for the umbrella weight used to generate the configurations by using the exact steady state probability density, eqn (4.41), to calculate the averages, (see below). The final term obviously approximates $\beta_1 W_1^{\text{mir}}$, but is about a factor of 400 faster to evaluate. We fixed α at 0.08, although it would be possible to optimize this choice, or to determine α on the fly.⁵

A trial move of an atom consisted of a small displacement in its position and momentum simultaneously. Step lengths of 0.9 in velocity and 0.09 in position gave an acceptance rate of about 50%. A cycle consisted of one trial move of each atom.

We collected averages after every 50 cycles. For this we required the mirror work which was obtained from the adiabatic Hamiltonian trajectory generated forward and backward in time, starting at the current configuration. We used a second order integrator,

$$q_{nz}(t + \Delta_t) = q_{nz}(t) + \Delta_t \dot{q}_{nz}(t) + \frac{\Delta_t^2}{2m} F_{nz}(t), \quad (5.9)$$

$$p_{nz}(t + \Delta_t) = p_{nz}(t) + \frac{\Delta_t}{2} [F_{nz}(t) + F_{nz}(t + \Delta_t)],$$

where n labels the atom, $\alpha = x, y$, or z labels the component, and $F_{nz}(t) \equiv F_{nz}(\mathbf{q}^N(t))$ is the force, which does not depend upon the momenta. Obviously one evaluates $F_{nz}(t + \Delta_t)$ after evaluating the new positions and before evaluating the new momenta. Typically, the time step was $\Delta_t = 10^{-3}$. The zeroth energy moment in general increased by less than 1% over the trajectory.

Labelling the current configuration by i , the trajectory is $\Gamma_0(t | \Gamma_i)$. We calculated the running integral for $W_1^{\text{mir}}(\Gamma_i; t)$ along the trajectory using both the trapezoidal rule and Simpson's rule, with indistinguishable results. We calculated the average flux as a function of the time interval,

$$\langle \dot{E}_1^0 \rangle_\tau = \frac{\sum_i \dot{E}_1^0(\Gamma_i) e^{-\alpha \beta_1 \dot{E}_1^0(\Gamma_i)} e^{\beta_1 W_1^{\text{mir}}(\Gamma_i; \tau)}}{\sum_i e^{-\alpha \beta_1 \dot{E}_1^0(\Gamma_i)} e^{\beta_1 W_1^{\text{mir}}(\Gamma_i; \tau)}}. \quad (5.10)$$

Notice how the umbrella weight used in the Metropolis scheme is cancelled here. We obtained the thermal conductivity as a function of the time interval, $\lambda(\tau) = \langle \dot{E}_1^0 \rangle_\tau / \beta_1 V k_B T_0^2$. Compared to implementing the steady state probability directly in the Metropolis algorithm, not only is the umbrella method orders of magnitude faster in generating configurations, but it also allows results as a function of τ to be

collected, and it reduces the correlation between consecutive, costly trajectories, by inserting many cheap, umbrella steps. Of the order of 50 000 trajectories were generated for each case studied.

5.3. von Neumann algorithm

The transition probability described in section 4.6 provides a basis for an alternative Monte Carlo algorithm that may be called stochastic molecular dynamics.^{6,17,18} It avoids the necessity of full knowledge of the steady state probability density, and it avoids having to calculate the mirror work, which can be costly. Basically one follows a trajectory in time in which deterministic Hamiltonian steps are interspersed with stochastic Monte Carlo steps. That is, from the current configuration $\Gamma = \Gamma(t)$, one generates Γ' using Hamilton's equations for a time step Δ_t , and then one generates Γ'' stochastically according to A_s and takes $\Gamma(t + \Delta_t) = \Gamma''$. von Neumann's algorithm for sampling a probability distribution may be used for the stochastic step.^{6,11} Un-weighted averages over the trajectory are taken.

We have only tested this stochastic molecular dynamics algorithm for heat flow in a rather limited sense.⁶ Instead of using the adiabatic trajectory for the deterministic step, (*i.e.* Hamilton's equations for an isolated system), modified equations of motion were used that constrained the zeroth and first energy moment to be constant. Although physically plausible, this procedure is somewhat *ad hoc* and is not consistent with the formal analysis of section 4.6. These deterministic equations were used in conjunction with the stochastic transition probability given above in section 4.6.2. It was found that for large stochastic step lengths the direction of the heat flux was randomized, and consequently the thermal conductivity was reduced. Several corrections for this effect were explored.⁶

5.4. Non-equilibrium molecular dynamics

Perhaps the most common computer simulation method for non-equilibrium systems is the non-equilibrium molecular dynamics (NEMD) method.^{22,23} This typically consists of Hamilton's equations of motion augmented with an artificial force designed to mimic particular non-equilibrium fluxes, and a constraint force or thermostat designed to keep the kinetic energy or temperature constant. Here we give a brief derivation and critique of the main thrust of that method.

Following ref. 11, we consider a perturbation to the Hamiltonian, $\mathcal{F}(t)\mathcal{A}(\Gamma)$, switched on at a certain time,

$$\mathcal{F}(t) = \begin{cases} \mathcal{F}, & t > 0, \\ 0, & t < 0. \end{cases} \quad (5.11)$$

Let the Hamiltonian for the non-equilibrium state be represented by

$$\mathcal{H}_{\text{ne}}(\Gamma) = \mathcal{H}_0(\Gamma) + \mathcal{F}(t)\mathcal{A}(\Gamma), \quad (5.12)$$

in which case the non-equilibrium equations of motion are

$$\dot{q}_{ix} = p_{ix}/m + \mathcal{F}(t)\mathcal{A}_{pix} \quad (5.13)$$

and

$$\dot{p}_{ix} = f_{ix} - F(t)\mathcal{A}_{qix}, \quad (5.14)$$

with the subscripts on the non-equilibrium potential denoting a derivative. (A thermostat is generally also added.) The function \mathcal{A} has even parity so that the equations of motion are time reversible. It is straightforward to show that at a given point in phase space, the rate of change of the non-equilibrium potential is the same in the non-equilibrium and the natural system,

$$\dot{\mathcal{A}}(\Gamma) = \dot{\mathcal{A}}^0(\Gamma). \quad (5.15)$$

The average of an odd function of phase space at a time $t > 0$ after switching on the non-equilibrium perturbation, assuming that the system is initially Boltzmann distributed, is

$$\begin{aligned} \langle \mathcal{B}(t) \rangle_{\text{ne}} &= \frac{\int d\Gamma_1 e^{-\beta\mathcal{H}_0(\Gamma_1)} \mathcal{B}(\Gamma_{\mathcal{F}}(t|\Gamma_1))}{\int d\Gamma_1 e^{-\beta\mathcal{H}_0(\Gamma_1)}} \\ &= \frac{\int d\Gamma_{\mathcal{F}} e^{-\beta\mathcal{H}_0(\Gamma_{\mathcal{F}})} e^{-\beta\mathcal{F}[\mathcal{A}(\Gamma_{\mathcal{F}}) - \mathcal{A}(\Gamma_1)]} \mathcal{B}(\Gamma_{\mathcal{F}})}{\int d\Gamma_{\mathcal{F}} e^{-\beta\mathcal{H}_{\mathcal{F}}(\Gamma_{\mathcal{F}})} e^{\beta\mathcal{F}[\mathcal{A}(\Gamma_{\mathcal{F}}) - \mathcal{A}(\Gamma_1)]}} \\ &= \frac{-\beta\mathcal{F} \int d\Gamma_{\mathcal{F}} e^{-\beta\mathcal{H}_0(\Gamma_{\mathcal{F}})} [\mathcal{A}(\Gamma_{\mathcal{F}}) - \mathcal{A}(\Gamma_1)] \mathcal{B}(\Gamma_{\mathcal{F}})}{\int d\Gamma_{\mathcal{F}} e^{-\beta\mathcal{H}_{\mathcal{F}}(\Gamma_{\mathcal{F}})}} \\ &= -\beta\mathcal{F} \langle [\mathcal{A}(\Gamma_{\mathcal{F}}) - \mathcal{A}(\Gamma_1)] \mathcal{B}(\Gamma_{\mathcal{F}}) \rangle_0 \\ &\approx -\beta\mathcal{F} \int_0^t dt' \langle \dot{\mathcal{A}}(t') \mathcal{B}(0) \rangle_0, \end{aligned} \quad (5.16)$$

where $\Gamma_{\mathcal{F}} \equiv \Gamma_{\mathcal{F}}(t|\Gamma_1)$. This result uses the fact that the non-equilibrium Hamiltonian is a constant on the non-equilibrium trajectory, and that the Jacobean of the transformation along a non-equilibrium trajectory is unity. It also linearizes everything with respect to \mathcal{F} , and neglects terms with total odd parity.

The parity of the time correlation functions discussed in section 2.4 has been used here. That discussion was predicated on an even probability distribution and reversible equations of motion. Hence those properties of the time correlation function also hold for the present case of the Boltzmann distribution and modified equations of motion.

There is an approximation implicit in the final line. The subscript zero implies an average for an isolated system (*i.e.* on an adiabatic or bare Hamiltonian trajectory), whereas the actual trajectory used to obtain this result is the modified one, $\Gamma_{\mathcal{F}}(t|\Gamma_1) \neq \Gamma_0(t|\Gamma_1)$. In so far as these are the same to leading order, this difference may be neglected.

The choice of \mathcal{A} , \mathcal{B} and \mathcal{F} is dictated by the Green-Kubo relation for the particular flow of interest. For heat flow we identify $\mathcal{B}(\Gamma) = \dot{E}_1^0(\Gamma)$, $\dot{\mathcal{A}}(\Gamma) = -\dot{E}_1^0(\Gamma)$, and $\mathcal{F} = T_0/T_1 = \beta_1/\beta_0$.

Depending on your point of view, it is either a strength or a weakness of the NEMD method that it gives a uniform structure for the non-equilibrium system, (*e.g.* for heat flow the sub-system does not acquire the applied temperature gradient, nor does it have gradients in energy or density). On the one hand such imposed uniformity makes the simulations compatible with periodic boundary conditions, and it does not affect the dynamics in the linear regime. On the other hand the incorrect structure precludes reliable results for the dynamics in the non-linear regime when the two are coupled.

It is possible to develop NEMD equations that do correctly account for the structure by analyzing the linear response of functions of opposite parity to that used above, as was done above at the end of section 4.5.4.

In the practical implementation of the NEMD method, it is usual to set the momentum derivative of the non-equilibrium potential to zero, $\mathcal{A}_{pix} = 0$.^{22,24} Presumably the reason for imposing this condition is that it preserves the classical relationship between velocity and momentum, $\dot{q}_{ix} = p_{ix}/m$. In view of this condition, the rate of change of the non-equilibrium potential reduces to

$$\dot{\mathcal{A}}(\Gamma) = \sum_{ix} A_{qix} p_{ix}/m, \quad (\text{NEMD}). \quad (5.17)$$

The Green–Kubo result demands that this be equated to the negative of the natural rate of change of the first energy moment, eqn (5.6), which means that

$$\mathcal{A}_{qix} = -\kappa_{ix}, \quad (\text{NEMD}). \quad (5.18)$$

However, this leads to the contradiction that $\dot{\mathcal{A}}^0(t) = -\dot{E}_1^0(t)$, but $A^0(t) \neq -\dot{E}_1^0(t)$.

The problem arises because one does not have the freedom to make the momentum derivative zero. One can see this from the usual condition on second derivatives,

$$\frac{\partial A}{\partial q_{ix} \partial p_{jy}} = \frac{\partial A}{\partial p_{jy} \partial q_{ix}}. \quad (5.19)$$

From the fact that

$$\frac{\partial \kappa_{ix}}{\partial p_{jy}} = \delta_{ij} \delta_{xz} p_{jy}/m, \quad (5.20)$$

one concludes that the momentum derivative of the non-equilibrium potential must be non-zero. It is in fact equal to

$$\mathcal{A}_{pix} = -\frac{\partial E_1}{\partial p_{ix}} = -q_{ix} p_{ix}/m. \quad (5.21)$$

This means that

$$\mathcal{A}_{qix} = -\kappa_{ix} + q_{ix} f_{ix}, \quad (5.22)$$

where f_i is the force on atom i . With these, $\dot{\mathcal{A}}^0(\Gamma) = -\dot{E}_1^0(\Gamma)$, and $\mathcal{A}^0(\Gamma) = -E_1^0(\Gamma)$. Using these forces, the non-equilibrium trajectory is properly derived from the non-equilibrium Hamiltonian, and the adiabatic incompressibility of phase space is assured, $\nabla \cdot \dot{\Gamma}_{\mathcal{F}} = 0$.

5.5. Monte Carlo results

5.5.1 Structure. In Fig. 1 we show the profiles induced in a bulk system by an applied temperature gradient. These Monte Carlo results³ were obtained using the static probability distribution, eqn (4.73). Clearly the induced temperature is equal to the applied temperature. Also, the slopes of the induced density and energy profiles can be obtained from the susceptibility, as one might expect since in the linear regime there is a direct correspondence between the slopes and moments.³

We give the energy susceptibility in Fig. 2. This was again obtained using the static probability distribution. In this case the susceptibility was obtained directly from the ratio of the induced energy moment to the applied temperature gradient, eqn (4.16), and from the fluctuations, eqn (4.14), with indis-

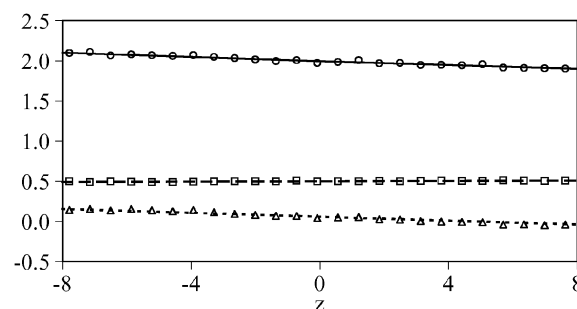


Fig. 1 Induced temperature (top), number density (middle), and energy density (bottom) profiles for $\beta_1 = 0.0031$ and $T_0 = 2$, $\rho = 0.5$. The symbols are Monte Carlo results using the static probability distribution, eqn (4.73), and the lines are either the applied temperature, or else the profiles predicted from the simulated susceptibility.³

tinguishable results. (In the latter formula E_1 was replaced by its departure from equilibrium, $\delta E_1 \equiv E_1 - \langle E_1 \rangle_{st}$.) The line passing through the points was obtained from bulk properties,³ which shows that the non-equilibrium structure is directly related to that of an equilibrium system.

5.5.2 Dynamics. In Fig. 3 we show the decay of the first energy moment following a fluctuation.⁴ We induced the fluctuation by sampling the static probability distribution, eqn (4.73), which has no preferred direction in time, and we used the configurations as starting points for independent trajectories. We calculated the trajectories adiabatically forward and backward in time. The point of the figure is that on short time scales, $t \lesssim \tau_{\text{short}}$, the moment displays a quadratic dependence on time, $E_1(t) \sim t^2$, whereas on long time scales, $t \gtrsim \tau_{\text{short}}$, it decays linearly in time, $E_1(t) \sim |t|$. This is consistent with conclusions drawn from the second entropy analysis of fluctuations, sections 3.2 and 3.3.

We test in Fig. 4 the dependence of the thermal conductivity on the time interval used to calculate $W_1^{\text{mir}}(\Gamma; \tau)$.⁷ These are Monte Carlo simulations using the Metropolis algorithm, umbrella sampling, and the steady state probability distribution, eqn (4.41). It can be seen that the thermal conductivity is independent of the integration limit for W_1^{mir} for $\tau \gtrsim 1$. This asymptotic or plateau value is ‘the’ thermal conductivity. The value of τ required to reach the respective plateaus here appear comparable to straight Green–Kubo equilibrium calculations,⁵

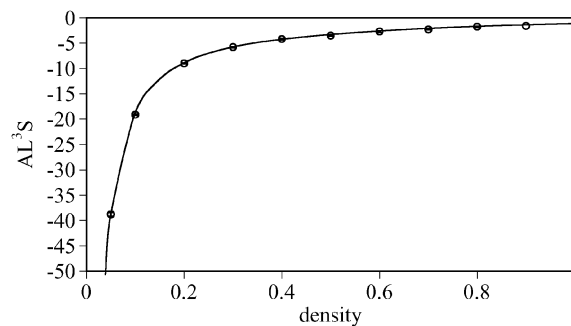


Fig. 2 Susceptibility of the energy moment at $T_0 = 2$. The symbols are static Monte Carlo results³ and the curve is obtained from a local thermodynamic approximation³ using the bulk susceptibilities from a Lennard-Jones equation of state.³⁹

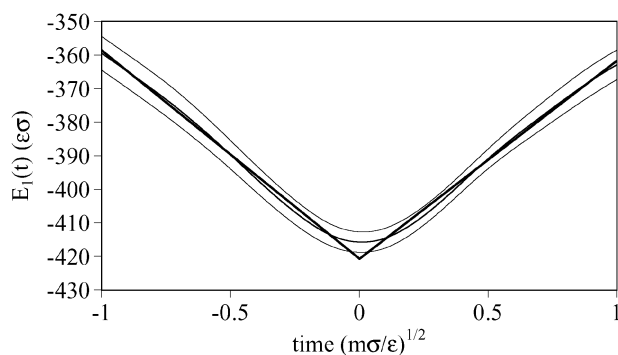


Fig. 3 Molecular dynamics simulation of the decay forward and backward in time of the fluctuation of the first energy moment of a Lennard-Jones fluid (the central curve is the average moment, the enveloping curves are estimated standard error, and the lines are best fits). The starting positions of the adiabatic trajectories are obtained from Monte Carlo sampling of the static probability distribution, eqn (4.73). The density is 0.80, the temperature is $T_0 = 2$, the initial imposed thermal gradient is $\beta_1 = 0.02$. Figure taken from ref. 4 (Copyright, American Institute of Physics, (2005)).

but the present steady state simulations used about one third the number of trajectories for comparable statistical error.

In the text we assumed that the change in moment over the relevant time scales was negligible, $\tau|\dot{E}_1| \ll |E_1|$. In the case of $\rho = 0.8$ at the largest value of τ in Fig. 2, $\langle E_1 \rangle_{ss} = -432$, and $\langle \dot{E}_1 \rangle_{ss} = 161$, and so this assumption is valid in this case. Indeed, the reason we made this assumption was that on long time scales the moment must return to zero and the rate of change of moment must begin to decrease. There is no evidence of this occurring in any of the cases over the full interval shown in Fig. 4.

In Table 1 we show the values of the relaxation time calculated using eqn (4.49) and (4.50). Both the inertial time and long time decrease with increasing density. This is in agreement with the trend of the curves in Fig. 4. Indeed, the actual estimates of the relaxation times in Table 1 are in semi-quantitative agreement with the respective boundaries of the plateaux in Fig. 4. The estimate of τ_{long} , the upper limit on τ that may be used in the present theory, is perhaps a little conservative.

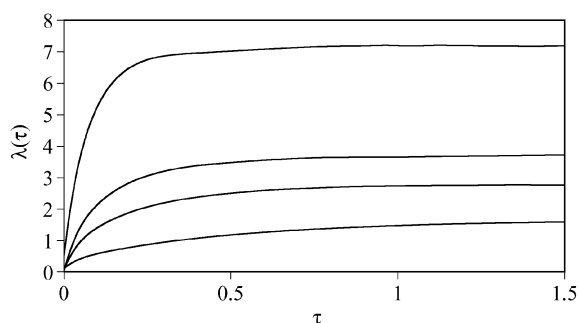


Fig. 4 The dependence of the thermal conductivity on the time interval for the mirror work, $W_1^{\text{mir}}(T, \tau)$. The curves are $\lambda(\tau) = \langle \dot{E}_1^0(0) \rangle_{\tau} / V k_B T_0^2 \beta_1$ for densities of, from bottom to top, 0.3, 0.5, 0.6, and 0.8, and $T_0 = 2$. Figure taken from ref. 7 (Copyright, American Institute of Physics, (2006)).

Table 1 Thermal conductivity and relaxation times for various densities at $T_0 = 2$. The standard error of the last few digits is in parentheses. Data taken from ref. 7

ρ	λ	τ_{short} eqn (4.49)	τ_{long} eqn (4.50)
0.3	1.63(8)	0.404(19)	3.22(16)
0.5	2.78(13)	0.233(11)	5.31(34)
0.6	3.76(16)	0.197(9)	3.41(18)
0.8	7.34(18)	0.167(4)	1.36(3)

In Fig. 5 we compare the thermal conductivity obtained from non-equilibrium Monte Carlo simulations⁷ with previous NEMD results.^{25,24} The good agreement between the two approaches validates the present phase space probability distribution. Of course, since we showed analytically that the present steady state probability gave the Green-Kubo formula, the results in Fig. 3 test the simulation algorithm rather than the probability distribution *per se*. The number of time steps that we required for an error of about 0.1 was about 3×10^7 , (typically 2×10^5 independent trajectories, each of about 75 time steps forward and backward to get into the intermediate regime). This obviously depends upon the size of the applied thermal gradient (the statistical error decreases with increasing gradient), but appears comparable to that required by NEMD simulations.²⁴ No attempt was made to optimize the present algorithm in terms of the number of Monte Carlo cycles between trajectory evaluations or the value of the umbrella parameter.

Fig. 5 also shows results for the thermal conductivity obtained for the slit pore, where the simulation cell was terminated by uniform Lennard-Jones walls. The results are consistent with those obtained for a bulk system using periodic boundary conditions. This indicates that the density inhomogeneity induced by the walls has little effect on the thermal conductivity.

We give in Fig. 6 results for the thermal conductivity obtained with the stochastic molecular dynamics algorithm.⁶ The three algorithms mentioned in the caption all used constrained equations of motion, $\dot{E}_0 = \dot{E}_1 = 0$, $\dot{E}_1^0 \neq 0$, and the stochastic transition probability of section 4.6. Algorithm A made no further corrections, algorithm B made the stochastic step at constant \dot{E}_1^0 by an appropriate rotation of the new

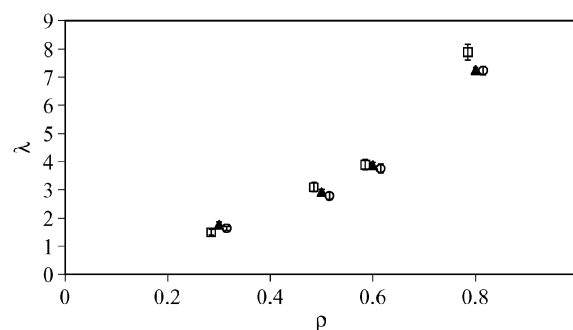


Fig. 5 Non-equilibrium Monte Carlo results for the thermal conductivity ($T_0 = 2$). The circles and squares are the present steady state results for bulk and inhomogeneous systems, respectively (horizontally offset by ± 0.015 for clarity), and the triangles are NEMD results.^{24,25} Figure taken from ref. 7 (Copyright, American Institute of Physics, (2006)).

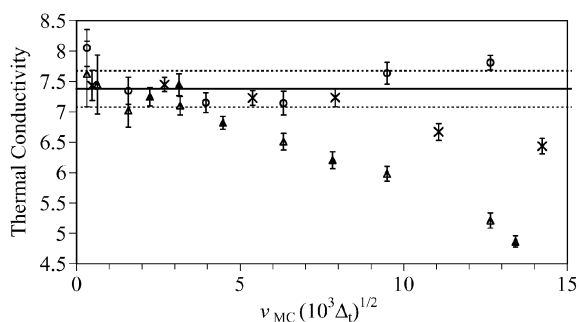


Fig. 6 Stochastic molecular dynamics results for the thermal conductivity ($\rho = 0.8$, $T_0 = 2$). The abscissa is the Monte Carlo time step, $\Delta_{MC} \propto v_{MC}\Delta_t$, divided by the square root of the molecular dynamics time step, $\sqrt{\Delta_t}$. All data are for a time step of $\Delta_t = 10^{-2}$, except for the filled triangles, which have $\Delta_t = 5 \times 10^{-3}$. The triangles are algorithm A, the circles are algorithm B, and the crosses are algorithm C, (see text). The total length of the error bar is twice the standard error on the mean. The horizontal lines are an equilibrium molecular dynamics result and error, 7.38 ± 0.3 .²⁴ Figure taken from ref. 6 (Copyright, American Institute of Physics, (2006)).

momenta, and algorithm C used an umbrella weight to stochastically cancel the randomization of \dot{E}_1^0 that occurs in the bare algorithm; see ref. 6 for full details. We conclude from the data in the figure that the stochastic molecular dynamics algorithm works well for small Monte Carlo steps, but the bare algorithm A underestimates the thermal conductivity for large steps. Algorithms B and C both correct this deficiency.

In the light of these results, it would be worthwhile to implement and test the stochastic molecular dynamics method using the actual adiabatic equations of motion and the stochastic transition probability of section 4.6. In this algorithm one avoids the costly computation of the mirror work, and so there is the possibility that it will prove more efficient computationally than the direct Monte Carlo approach.

6. Non-equilibrium mechanical work

An important class of non-equilibrium systems are those in which work, either steady or varying, is performed on the sub-system while it is in contact with a heat reservoir. Such work is represented by a time-dependent Hamiltonian, $\mathcal{H}_\mu(\Gamma, t)$, where $\mu(t)$ is the work parameter. (For example, this could represent the position of a piston, or the strength of an electric field, *etc.*)

6.1. Phase space probability

We begin with the probability distribution proposed by Yamada and Kawasaki.^{26,27} If the system were isolated from the thermal reservoir during its evolution, and if the system were Boltzmann distributed at $t - \tau$, then the probability distribution at time t would be

$$\begin{aligned} \tilde{\mathcal{P}}_\mu(\Gamma|\beta, t) &= \tilde{Z}^{-1} e^{-\beta \mathcal{H}_\mu(\Gamma, t-\tau)} \\ &= \tilde{Z}^{-1} e^{-\beta \mathcal{H}_\mu^{(\Gamma, t)}} e^{\beta \tilde{W}_\mu(\Gamma, t)}, \end{aligned} \quad (6.1)$$

where $\beta = 1/k_B T$ is the inverse temperature of the reservoir. Here $\Gamma_- = \Gamma_\mu(t - \tau|\Gamma, t)$ is the starting point of the adiabatic

trajectory, and the work done is $\tilde{W}_\mu(\Gamma, t) = \mathcal{H}_\mu(\Gamma, t) - \mathcal{H}_\mu(\Gamma_-, t - \tau)$. This result invokes Liouville's theorem, namely that the probability density is conserved on an adiabatic trajectory.⁸ This expression, which is intended to represent the probability of the phase point Γ at t , is just the analogue of the tentative probability distribution discussed for steady heat flow, eqn (4.39). Hence it suffers from the same deficiencies identified in that case: since it does not take into account the influence of the heat reservoir while the work is being performed, the structure of the sub-system evolves adiabatically away from the true non-equilibrium structure even while the dynamics evolve to their correct non-equilibrium state. A modified thermostatted form of the Yamada–Kawasaki distribution has been given, but it is said to be computationally intractable.^{22,28,29} It is unclear whether artifacts are introduced by the artificial thermostat. Based on calculations performed with the Yamada–Kawasaki distribution, some have even come to doubt the very existence of a non-equilibrium probability distribution.^{30–32}

In view of these difficulties with the Yamada–Kawasaki distribution and its modifications, one seeks a non-equilibrium probability analogous to that described above for the case of heat flow. Hence one requires ‘mirror work’ that has odd parity. To obtain this one extends the work path into the future by making it even about t ,

$$\mu_t^{\text{mir}}(t') \equiv \begin{cases} \mu(t'), & t' \leq t, \\ \mu(2t - t'), & t' > t. \end{cases} \quad (6.2)$$

We denote the corresponding adiabatic trajectory that is at Γ at time t by

$$\Gamma_{\mu_t^{\text{mir}}}(t'|\Gamma, t) = \begin{cases} \Gamma_\mu(t'|\Gamma, t), & t' \leq t, \\ (\Gamma_\mu(2t - t'|\Gamma^\dagger, t))^\dagger, & t' > t. \end{cases} \quad (6.3)$$

One defines $\Gamma_\pm(\Gamma) \equiv \Gamma_{\mu_t^{\text{mir}}}(t \pm \tau|\Gamma, t)$, which have the property $\Gamma_\pm(\Gamma^\dagger) = (\Gamma_\mp(\Gamma))^\dagger$.

With these the mirror work is⁷

$$\begin{aligned} W_\mu^{\text{mir}}(\Gamma, t) &= \frac{1}{2} \left[\mathcal{H}_{\mu_t^{\text{mir}}}(\Gamma_+, t + \tau) - \mathcal{H}_{\mu_t^{\text{mir}}}(\Gamma_-, t - \tau) \right] \\ &= \frac{1}{2} \int_{t-\tau}^{t+\tau} dt' \dot{\mathcal{H}}_{\mu_t^{\text{mir}}}(\Gamma_{\mu_t^{\text{mir}}}(t'|\Gamma, t), t') \\ &= \frac{1}{2} \int_{t-\tau}^t dt' [\dot{\mathcal{H}}_\mu(\Gamma_\mu(t'|\Gamma, t), t') \\ &\quad - \dot{\mathcal{H}}_\mu(\Gamma_\mu(t'|\Gamma^\dagger, t), t')]. \end{aligned} \quad (6.4)$$

(In this section, $\dot{\mathcal{H}}_\mu$ means the adiabatic rate of change of the Hamiltonian.) By construction, the mirror work has odd phase space parity, $W_\mu^{\text{mir}}(\Gamma, t) = -W_\mu^{\text{mir}}(\Gamma^\dagger, t)$. With it the non-equilibrium probability distribution for mechanical work is⁷

$$\mathcal{P}_\mu(\Gamma|\beta, t) = \frac{e^{-\beta \mathcal{H}_\mu(\Gamma, t)} e^{\beta W_\mu^{\text{mir}}(\Gamma, t)}}{h^{3N} N! Z_\mu(\beta, t)}. \quad (6.5)$$

The probability distribution is normalized by $Z_\mu(\beta, t)$, which is a time-dependent partition function whose logarithm gives the non-equilibrium total entropy, which may be used as a generating function.

We expect that on long time scales,

$$\dot{\mathcal{H}}_{\mu}^{\text{mir}}(\Gamma_{\mu}^{\text{mir}}(t'|\Gamma, t), t') \sim \text{sign}(t') \overline{\mathcal{H}}_{\mu}(t), |t'| \gtrsim \tau_{\text{short}}, \quad (6.6)$$

where the most likely rate of doing work at time t appears. This assumes that the change in energy is negligible on the relevant time scales, $\tau|\dot{\mathcal{H}}_{\mu}| \ll |\mathcal{H}_{\mu}|$. Since this asymptote is odd in time, one concludes that the mirror work is independent of τ for τ in the intermediate regime, and that W_{μ}^{mir} is dominated by the region $t' \approx t$.

In view of this and the fact that $d\Gamma_{\mu}(t'|\Gamma, t)/dt = 0$, we find that the rate of change of the mirror work along a Hamiltonian trajectory is

$$\dot{W}_{\mu}^{\text{mir}}(\Gamma, t) = \overline{\mathcal{H}}_{\mu}(t) \quad (6.7)$$

We conclude that the non-equilibrium probability distribution is stationary during adiabatic evolution on the most likely points of phase space.

6.2. Transition probability

We now analyze the time evolution of the probability distribution, including the stochastic perturbations from the thermal reservoir. As above, a single prime is used to denote the adiabatic development in time Δ_t , $\Gamma \rightarrow \Gamma'$, and a double prime to denote the final stochastic position due to the influence of the reservoir, $\Gamma' \rightarrow \Gamma''$. The conditional transition probability may be taken to be

$$A_{\mu}(\Gamma''|\Gamma, t) = \Theta_{\Delta}(|\Gamma'' - \Gamma'|) e^{-\beta(\mathcal{H}_{\mu}'' - \mathcal{H}_{\mu}')/2} \quad (6.8)$$

This transition probability preserves the non-equilibrium phase space probability density, eqn (6.5), during its time evolution,

$$\wp_{\mu}(\Gamma''|\beta, t + \Delta_t) = \int d\Gamma A_{\mu}(\Gamma''|\Gamma, t) \wp_{\mu}(\Gamma|\beta, t). \quad (6.9)$$

This result may be readily confirmed using the fact that $\mathcal{H}_{\mu}' - W_{\mu}^{\text{mir}} = \mathcal{H}_{\mu} - W_{\mu}^{\text{mir}}$, (at least for those phase points most likely to occur), together with the usual normalization requirements on the transition probability.

The final, odd term in the transition probability is identical to that given previously,⁶ and hence the microscopic transition theorem is obeyed, (see eqn (9) and (11) of ref. 6), and it yields the fluctuation theorem^{19,20} and the work theorem,²¹ (see Section I.C of ref. 6), as we now show. The unconditional microscopic transition probability is

$$\begin{aligned} \mathcal{P}_{\mu}(\Gamma'' \leftarrow \Gamma|t, \Delta_t) &= A_s(\Gamma''|\Gamma') \mathcal{P}_{\mu}(\Gamma|\beta, t) \\ &= \Theta_{\Delta}(|\Gamma'' - \Gamma'|) e^{-\beta(\mathcal{H}_{\mu}'' + \mathcal{H}_{\mu}')/2} \\ &\times e^{\beta\Delta_t \dot{\mathcal{H}}_{\mu}(\Gamma, t)/2} e^{\beta W_{\mu}^{\text{mir}}(\Gamma, t)} / Z_{\mu}(\beta, t) \\ &= \Theta_{\Delta}(|\Gamma'' - \Gamma'|) \sqrt{\mathcal{P}(\Gamma''|\beta, t + \Delta_t, [\mu]) \mathcal{P}(\Gamma|\beta, t, [\mu])} \\ &\times e^{-\beta[W_{\mu}^{\text{mir}}(\Gamma'', t + \Delta_t) - W_{\mu}^{\text{mir}}(\Gamma, t)]/2} e^{\beta\Delta_t \dot{\mathcal{H}}_{\mu}(\Gamma, t)/2} \\ &\times \sqrt{Z_{\mu}(\beta, t + \Delta_t) / Z_{\mu}(\beta, t)}, \end{aligned} \quad (6.10)$$

for an infinitesimal Δ_t . The $\dot{\mathcal{H}}_{\mu}$ can be replaced by $(\dot{\mathcal{H}}_{\mu}'' + \dot{\mathcal{H}}_{\mu}')/2$ to this order.

We consider the forward transition, $\Gamma \rightarrow \Gamma' \rightarrow \Gamma''$, and its reverse $\Gamma'' \rightarrow \Gamma' \rightarrow \Gamma$ on the mirror path. (It's necessary to use the mirror path centered at $t + \Delta_t$ because time goes from $t + \Delta_t \rightarrow t + 2\Delta_t$, and $\mu(t + \Delta_t) \rightarrow \mu(t)$). The ratio of the forward to the reverse transition probabilities is

$$\begin{aligned} \frac{\mathcal{P}_{\mu}(\Gamma'' \leftarrow \Gamma|t, \Delta_t)}{\mathcal{P}_{\mu}^{\text{mir}}(\Gamma \leftarrow \Gamma''|t + \Delta_t, \Delta_t)} &= e^{\beta(W_{\mu}^{\text{mir}}(\Gamma'') + W_{\mu}^{\text{mir}}(\Gamma))} \\ &\times e^{\beta\Delta_t[\dot{H}_{\mu}(\Gamma'', t + \Delta_t) + \dot{H}_{\mu}(\Gamma, t)]/2} \frac{Z_{\mu}(\beta, t + \Delta_t)}{Z_{\mu}(\beta, t)}, \end{aligned} \quad (6.11)$$

the even terms cancelling. The ratio of partition functions can be written as the exponential of the change in free energy.

We now consider a trajectory $[\Gamma] = \{\Gamma_0, \Gamma_1, \dots, \Gamma_f\}$, at times $0, \Delta_t, \dots, f\Delta_t$, and its reverse, $[\Gamma^{\dagger}] = \{\Gamma_f^{\dagger}, \Gamma_{f-1}^{\dagger}, \dots, \Gamma_0^{\dagger}\}$. The work done on the sub-system on the trajectory is just the total adiabatic change in the Hamiltonian,

$$\begin{aligned} \Delta^0 \mathcal{H}_{\mu}[\Gamma] &= \frac{\Delta_t}{2} \left[\dot{\mathcal{H}}_{\mu}^0(\Gamma_0) + \dot{\mathcal{H}}_{\mu}^0(\Gamma_f) \right] \\ &+ \Delta_t \sum_{i=1}^{f-1} \dot{\mathcal{H}}_{\mu}^0(\Gamma_i). \end{aligned} \quad (6.12)$$

Clearly $\Delta^0 \mathcal{H}_{\mu}[\Gamma^{\dagger}] = -\Delta^0 \mathcal{H}_{\mu}[\Gamma]$. In view of eqn (4.62), for large $f\Delta_t$ this gives the change in the reservoirs' entropy over the trajectory, $\Delta S_r/k_B = -\beta\Delta \mathcal{H}_{\mu} + \beta\Delta^0 \mathcal{H}_{\mu} \approx \beta\Delta^0 \mathcal{H}_{\mu}$.

The unconditional probability of the trajectory is

$$\begin{aligned} \mathcal{P}_{\mu}[\Gamma] &= \prod_{i=1}^f \left[\Theta_{\Delta}(|\Gamma_{i+1} - \Gamma_i|) e^{-\beta(H_{\mu,i} - H'_{\mu,i-1})/2} \right] \\ &\times \mathcal{P}_{\mu}(\Gamma_0|\beta, 0) \\ &= \prod_{i=1}^f \left[\Theta_{\Delta}(|\Gamma_{i+1} - \Gamma_i|) \right] \\ &\times e^{\beta_1 \Delta^0 H_{\mu}[\Gamma]/2} e^{-\beta(W_{\mu,f}^{\text{mir}} - W_{\mu,0}^{\text{mir}})/2} \\ &\times \sqrt{\mathcal{P}_{\mu}(\Gamma_f|\beta, t_f) \mathcal{P}_{\mu}(\Gamma_0|\beta, 0)} \\ &\times \sqrt{Z_{\mu}(\beta, t_f) / Z_{\mu}(\beta, 0)}, \end{aligned} \quad (6.13)$$

since $\mathcal{H}'_{\mu,i-1} = \mathcal{H}_{\mu,i-1} + \Delta_t \dot{\mathcal{H}}_{\mu,i-1}$. The boundary terms \mathcal{H}'_{μ} for $\Delta^0 \mathcal{H}_{\mu}$ has been neglected, which is valid for large f .

The ratio of the probability of the forward and reverse trajectory is

$$\begin{aligned} \frac{\mathcal{P}_{\mu}[\Gamma]}{\mathcal{P}_{\mu}^{\text{mir}}[\Gamma^{\dagger}]} &= e^{\beta W_{\mu}^{\text{mir}}(\Gamma_f)} e^{\beta W_{\mu}^{\text{mir}}(\Gamma_0)} \\ &\times e^{\beta\Delta^0 \mathcal{H}_{\mu}[\Gamma]} \frac{Z_{\mu}(\beta, t_f)}{Z_{\mu}(\beta, 0)}. \end{aligned} \quad (6.14)$$

This is the mechanical version of the *reverse transition theorem*.⁶

It is possible to simplify these results for the case when it is valid to draw the start and end point of the trajectories from a probability density with even parity such as the Boltzmann distribution,

$$\mathcal{P}_{\text{B}}(\Gamma|\beta, t) = \frac{e^{-\beta\mathcal{H}_{\mu}(\Gamma, t)}}{Z_{\text{B}}(\beta, \mu(t))}. \quad (6.15)$$

In this case the trajectory probability is

$$\begin{aligned} \mathcal{P}_B[\Gamma] &= \prod_{i=1}^f [\Theta_{\Delta}(|\Gamma_{i+1} - \Gamma'_i|)] e^{\beta \Delta^0 \mathcal{H}_{\mu}/2} \\ &\times \sqrt{\mathcal{P}_B(\Gamma_f|\beta, t_f) \mathcal{P}_B(\Gamma_0|\beta, 0)} \\ &\times \sqrt{Z_B(\beta, t_f)/Z_B(\beta, 0)}, \end{aligned} \quad (6.16)$$

and the ratio of the forward and reverse trajectories is

$$\frac{\mathcal{P}_B[\Gamma]}{\mathcal{P}_B[\Gamma^{\dagger}]} = e^{\beta \Delta^0 \mathcal{H}_{\mu}[\Gamma]} \frac{Z_B(\beta, t_f)}{Z_B(\beta, 0)}. \quad (6.17)$$

This modification of the termini of the trajectory is identical to that assumed in the fluctuation^{19,20} and work²¹ theorems discussed next. The type of work paths for which this modification is valid is discussed at the end of this section.

The probability of observing the entropy of the reservoir change by ΔS over a period t_f is related to the probability of observing the opposite change by

$$\begin{aligned} \mathcal{P}_{\mu}(\Delta S|\beta, t_f) &= \int d[\Gamma] \delta(\Delta S - \Delta S[\Gamma]) \mathcal{P}[\Gamma] \\ &\approx k_B^{-1} \int d[\Gamma^{\dagger}] \delta(\Delta S/k_B - \beta \Delta^0 \mathcal{H}_{\mu}[\Gamma]) \\ &\times \mathcal{P}[\Gamma^{\dagger}] e^{\beta \Delta^0 \mathcal{H}_{\mu}[\Gamma]} \frac{Z_{\mu}(\beta, t_f)}{Z_{\mu}(\beta, 0)} \\ &= e^{\Delta S/k_B} \frac{Z_{\mu}(\beta, t_f)}{Z_{\mu}(\beta, 0)} \mathcal{P}_{\mu}(-\Delta S|\beta, t_f). \end{aligned} \quad (6.18)$$

(Using instead trajectories that begin and end with the Boltzmann distribution, we have the exact result, $\wp_B(\Delta^0 \mathcal{H}_{\mu}|\beta, t_f) = e^{\beta \Delta^0 \mathcal{H}_{\mu}} \wp_B(-\Delta^0 \mathcal{H}_{\mu}|\beta, t_f) Z_B(\beta, t_f)/Z_B(\beta, 0)$. This result says in essence that the probability of a positive increase in entropy is exponentially greater than the probability of a decrease in entropy during mechanical work. This is in essence the fluctuation theorem that was originally derived by Evans *et al.*^{19,20} A derivation has also been given by Crooks,^{33,34} and the theorem has been verified experimentally.³⁵ The present derivation is based upon the microscopic transition probability given by Attard.⁶

For the case of the work theorem²¹ we consider the average of the exponential of the negative of the heat flow,

$$\begin{aligned} \langle e^{-\beta \Delta^0 \mathcal{H}_{\mu}} \rangle_{|\mu|, t_f} &= \int d[\Gamma] e^{-\beta \Delta^0 \mathcal{H}_{\mu}[\Gamma]} \mathcal{P}_{\mu}([\Gamma]|\beta, t_f) \\ &= \int d[\Gamma^{\dagger}] \mathcal{P}_{\mu}([\Gamma^{\dagger}]|\beta, t_f) e^{\beta(W_{\mu}^{\text{mir}}(\Gamma_f) + W_{\mu}^{\text{mir}}(\Gamma_0))} \\ &\times Z_{\mu}(\beta, t_f)/Z_{\mu}(\beta, 0) \\ &\approx \int d\Gamma_f^{\dagger} \mathcal{P}_{\mu}(\Gamma_f^{\dagger}|\beta) e^{-\beta W_{\mu}^{\text{mir}}(\Gamma_f^{\dagger})} Z_{\mu}(\beta, t_f) \\ &\times \int d\Gamma_0^{\dagger} \mathcal{P}_{\mu}(\Gamma_0^{\dagger}|\beta) e^{-\beta W_{\mu}^{\text{mir}}(\Gamma_0^{\dagger})} / Z_{\mu}(\beta, 0) \\ &= \frac{Z_B(\beta, \mu(t_f)) Z_B(\beta, \mu(0))}{Z_{\mu}(\beta, 0)^2}. \end{aligned} \quad (6.19)$$

Here it has been assumed that the trajectory is long enough that the ends are uncorrelated. This result shows that this

particular average is not extensive in time, (*i. e.* it does not scale with t_f). If the termini of the trajectories are drawn from a Boltzmann distribution, the result becomes

$$\langle e^{-\beta \Delta^0 \mathcal{H}_{\mu}} \rangle_{B, |\mu|, t_f} = \frac{Z_B(\beta, \mu(t_f))}{Z_B(\beta, \mu(0))} = e^{-\beta \Delta F(\beta)}, \quad (6.20)$$

where the exponent is the difference between the Helmholtz free energies of the system at the final and initial values of the work parameter. This is Jarzynski's work theorem in its original form,²¹ which has been rederived in different fashions^{20,36,37} and verified experimentally.³⁸

When are the simplified results valid? If the work path has buffer regions at its beginning and end during which the work parameter is fixed for a time $\gtrsim \tau_{\text{short}}$, then the sub-system will have equilibrated at the initial and final values of μ in each case. Hence the mirror work, which has odd parity, is on average zero and can be neglected, and the probability distribution reduces to Boltzmann's.

7. Non-equilibrium quantum statistical mechanics

We consider a quantum system with time-dependent, Hermitian, Hamiltonian operator $\hat{\mathcal{H}}(t)$. We can develop a theory for non-equilibrium quantum statistical mechanics by analogy with the classical case. In the quantum case it is the wave functions ψ that are the microstates analogous to a point in phase space. The complex conjugate ψ^* plays the role of the cf7.6onjugate point in phase space, since, according to Schrodinger, it has equal and opposite time derivative to ψ .

We define the mirror work operator,

$$\hat{W}^{\text{mir}}(t) = [\hat{E}_+(t) - \hat{E}_-(t)]/2 \quad (7.1)$$

where the past and future energy operators are

$$\hat{E}_{\pm}(t) = \hat{\Theta}(\mp \tau; t) \hat{\mathcal{H}}_t^{\text{mir}}(t \pm \tau) \hat{\Theta}(\pm \tau; t) \quad (7.2)$$

and where the time-shift operator is

$$\hat{\Theta}(\tau; t) = \exp\left[\frac{-i}{\hbar} \int_t^{t+\tau} dt' \hat{\mathcal{H}}_t^{\text{mir}}(t')\right]. \quad (7.3)$$

The mirror Hamiltonian operator has been continued into the future,

$$\hat{\mathcal{H}}_t^{\text{mir}}(t') \equiv \begin{cases} \mathcal{H}(t'), & t' \leq t, \\ \mathcal{H}(2t - t'), & t' > t, \end{cases} \quad (7.4)$$

and the manipulation of the operators derived from it is facilitated by the symmetry about t , $\hat{\mathcal{H}}_t^{\text{mir}}(t') = \hat{\mathcal{H}}_t^{\text{mir}}(2t - t')$.

With these definitions, the non-equilibrium density operator for a sub-system of a thermal reservoir of inverse temperature β is⁷

$$\hat{\rho}(t) = \frac{1}{Z(t)} \exp -\beta[\hat{\mathcal{H}}(t) - \hat{W}^{\text{mir}}(t)], \quad (7.5)$$

where the normalization factor is $Z(t) = \text{TR}\{\exp -\beta[\hat{\mathcal{H}}(t) - \hat{w}^{\text{mir}}(t)]\}$. Accordingly, the average of an observable at time t is

$$\langle \hat{O} \rangle_t = \text{TR}\{\hat{\rho}(t) \hat{O}(t)\}, \quad (7.6)$$

and the present density operator can be said to provide a basis for non-equilibrium quantum statistical mechanics.

8. Selected applications

The theory for non-equilibrium statistical mechanics reviewed here can be applied in a large number of practical situations, including to the flow of heat that was treated in the text in depth. However, the variety and range of time-dependent phenomena are too numerous to cover in any detail. Prior to concluding this review, however, it does seem appropriate to survey two applications of the theory that may be of broad interest, namely to non-equilibrium thermodynamics, and to the origin and evolution of life.

8.1. Non-equilibrium thermodynamics

The establishment of thermodynamics as a discipline was the great triumph of 19th century science. The power of thermodynamics lies in the generality of the laws that form its basis. From these few principles can be set the rules that bind all physical phenomena.

Thermodynamics, however, has in a sense been superseded by statistical mechanics. The thermodynamic laws are a consequence of the statistical behavior of molecules rather than the other way around. For this reasons violations of the thermodynamic laws can occur at the molecular level, although they become increasingly unlikely as the size of the system is increased. Fluctuation theory itself is a violation of the second law of thermodynamics, since obviously the transition away from equilibrium represents the spontaneous movement of the system to a lower entropy state (see for example section 4.6.3). The probability of such a violation is inversely proportional to the exponential of the size of the system, which shows why the laws of thermodynamics hold at the macroscopic level.

Because of the historical primacy of the laws of thermodynamics, and because their very generality lies beyond the details of specific molecular interactions, and because they are valid on large-scales, it is of interest to enquire how they should be modified or added to in the light of the present theory for non-equilibrium statistical mechanics. It is emphasized however that the full non-equilibrium theory is the mathematical formulation given above, and that the following discussion is meant more to summarize and encapsulate those equations than to provide a precise epistemological basis for non-equilibrium thermodynamics.

Conventionally, entropy is introduced by the second law of thermodynamics, which is about the direction in which changes occur:

during spontaneous changes of an isolated system the entropy increases.

Equilibrium may be defined as the relaxed or unconstrained state, where no macroscopic changes in structure occur. The second law implies that the equilibrium state is the state of maximum entropy, and it provides a variational principle that determines the physical state of the system. We can draw many conclusions from the second law, including the concavity of the entropy, the stability of the equilibrium state, and the high probability of the equilibrium state. Similarly implicit is the fact that the entropy is well-defined in the non-equilibrium state (*i.e.* during the approach to equilibrium); such a con-

strained entropy and the trivially related constrained thermodynamic potential may be calculated explicitly using statistical mechanics.⁸

In the text we drew a distinction between the first and second entropy. Hence in the light of the foregoing mathematical analysis, we are inclined to restate the second law:

during spontaneous changes in the static structure of an isolated system the first entropy increases.

Static structure means the spatial distribution of conserved variables (energy, number, charge, volume, *etc.*) and includes the moments of an isolated system, and also the related case of exchange between a sub-system and a reservoir that comprise an isolated total system. We would also stress that this second law refers to the equilibrium state, and to the approach to the equilibrium state, but it does not refer to the rate of such an approach.

A non-equilibrium or dynamic corollary of the second law can be formulated:

during spontaneous changes in the dynamic structure of an isolated system the second entropy increases.

Dynamic structure here means fluxes in the conserved variables, such as the rate of change of their spatial distribution. The steady state may be defined as the relaxed or unconstrained state where no macroscopic changes in flux occur. The dynamic second law implies that the steady state is the state of maximum second entropy. In so far as the second entropy is an extensive function of the time interval over which it is defined, then this is equivalent to the steady state being the state that produces second entropy most rapidly.

This dynamic corollary to the second law then tells why a flux develops and its direction, and it provides a variational principle that determines the physical value of the flux when it reaches the steady state. These fluxes are driven by gradients in the entropy potential, which are the second entropy derivatives with respect to the fluxes, and which may correspond to spatially separated reservoirs with different values of the conjugate thermodynamic field variables. Implicit in this dynamic corollary of the second law of thermodynamics is the concavity of the second entropy, the stability of the steady state, the high probability of the steady state, and the well-defined nature of the second entropy and related thermodynamic potentials in the transition to the steady state, and for states with artificially constrained fluxes.

The dynamic second law is applicable in the transient regime during the approach to the steady state following the switching on of an external force or the development of a fluctuation. The second entropy is perfectly well-defined in this regime, just as the first entropy is well-defined during the approach to equilibrium; in both cases the constraint formalism used here and in ref. 8 may be used. This is also true for time varying external thermodynamic or mechanical forces, where the unconstrained system naturally varies over time, which is to say that it is in a non-steady, non-equilibrium state. This physical state also obeys a variational principle, namely the maximization of the full second entropy, $\sigma_{\text{total}}(\Delta^0 \mathbf{x}, \Delta \mathbf{x}_r, \mathbf{x} | \mathbf{X}, t, \tau)$, without assuming extensivity in the time interval or a

quadratic form in the fluxes. Even more generally, one would invoke a functional maximization over the path $x(t)$ of $\sigma_{\text{total}}([\dot{x}^0], [\dot{x}^r], x(0)|[X], \tau,)$ to find the most likely state of the transient system at any given time.

This point of the second entropy being a functional of the path underscores the essential difference between the first and second entropy. Describing the second entropy as the ‘dynamic’ entropy is a little misleading since the more precise notion is that it describes the transition over time between pairs of states (as treated in detail in the text), or more generally a path of such transitions. It is true that for the steady state the coarse velocity and the rate of second entropy production are the most important quantities to emerge. But even though these are dynamic quantities, it should be understood that in the limit $\tau \rightarrow 0$, the coarse velocity becomes the instantaneous velocity. In this limit we only require the first entropy, with ‘static structure’ redefined to include the instantaneous velocity of the constraints. The crucial difference then is that the first entropy may refer to position and instantaneous velocity, but the second entropy refers to transitions between pairs of states or a path of such transitions. This is the sense in which ‘dynamic structure’ in the corollary to the second law should be understood.

The word ‘structure’ in the restated second law hints at order, and this is particularly pertinent in the case of fluxes. So important is this particular consequence of the second entropy that it arguably merits its own corollary:

an entropy potential gradient necessarily induces order in an intervening conductor.

This order, which has both static and dynamic aspects, is the most probable state of a sub-system with an imposed thermodynamic gradient. In the case of heat flow treated in the text, the static order in the sub-system was the induced energy moment that matched the sub-system temperature gradient to that imposed by the reservoir. The dynamic order was the concerted motion of the atoms of the sub-system that conducted heat between the reservoirs in a definite direction in space and time. As we shall now see, this corollary to the dynamic second law is of relevance to the origin of life and to the direction of evolution.

8.2. Origin and evolution of life

The puzzle with life, and ecological and sociological systems more generally, is of course that they appear to be most unlikely from the point of view of random chance. They have a high degree of order, and, moreover, they evolve into ever more complex and ordered forms in apparent contradiction to the conventional second law of thermodynamics. A number of previous workers, including Prigogine, Kay, Swenson, and co-workers,^{40–42} have noted that non-equilibrium systems organize themselves to increase the entropy of the universe, and that this may provide the basis for life itself. Prigogine explored dissipative structures that are formed and maintained during energy exchange and which thereby dissipate heat and create entropy. Swenson articulated the principle that “the system will select the path or assembly of paths out of otherwise available paths that minimizes the potential or

maximizes the entropy at the fastest rate given the constraints”. This is similar to Kay who said “systems . . . away from equilibrium . . . utilize all avenues available to counter the applied gradients”, and “the emergence of life hastens the progression to equilibrium”.

The present theory applied to the problem of life and ecology reaches broadly similar conclusions from arguably a more detailed and precise mathematical basis. The main contribution of the present non-equilibrium theory is the discovery of the second entropy and the proof that ordered fluxes develop to maximize its rate of production. In contrast earlier work focused on the conventional or first entropy and asserted that the rate of its production was either minimal^{40,43} or maximal.^{41,42,44,45}

The heat flow that was examined in detail here provides the archetype of an entropy potential gradient, namely spatially separated reservoirs in which the exchange of a conserved quantity increases the entropy of the universe. The mediation of a sub-system in this exchange, the intervening conductor, necessarily leads to an ordering of the sub-system that increases the rate of total second entropy production. The ordering is both static, as the applied gradient is mirrored in the sub-system, and dynamic, due to the ordered flow. The entropy destruction due to the dynamic order increases linearly with time (in the steady state), but this is outweighed by the linear increase in entropy of the reservoirs due to the flux (see section 4.3.3 above). The optimum structural and dynamic order of the sub-system maximizes the second entropy of the universe. As such, this is a high probability state. Far from life being improbable, we see that it is actually a most likely state of the universe, and the evolution of life is in the direction of increasing the rate of production of (second) entropy. Moreover, life, and ecosystems more generally, are actually stable, provided that the entropy gradients that maintain them remain undiminished.

In summary, life mediates entropy production. There exist on earth, and elsewhere in our universe, gradients in entropy potential, such as mineral deposits, water pools, energy sources, *etc.* Such segregated regions of differential composition arise from the usual thermodynamic fluctuations and are not of themselves ordered. However, order necessarily is induced by these according to the corollary of the dynamic second law formulated above. The driving force for life is that entropy is able to be produced, and the driving force for more complex life is that entropy is able to be produced at a greater rate.

As examples, two pieces of evidence can be offered. The average surface temperature of the moon during its day is 107 °C,⁴⁶ which is substantially higher than that of the earth. Some of the decrease in the day-time temperature of the earth may be attributed to our atmosphere and weather, which are themselves ecosystems that have evolved to increase the rate of entropy production. But it may be that biological activity has also increased the rate of heat flow from the earth to space, thereby lowering the temperature compared to the moon. This is entirely consistent with the notion that an ordered sub-system which increases the rate of entropy production of the total system is actually a more probable state of the total system than a disordered sub-system.

The rôle of ecosystems in reducing surface temperature was stressed by Schneider and Kay.⁴¹ They cited data of Luvall and Holbo,⁴⁷ who measured ecosystem radiation and found that quarried and clear cut land was about 20 °C warmer than plantation and immature natural forest, and about 25 °C warmer than old growth forest. This again is consistent with the idea that biological activity serves to increase the rate of entropy production, in this case by facilitating the transfer of energy that falls on earth from the sun out into space.

Schneider and Kay⁴¹ focus on the temperature gradient between the earth and the sun as being the driving force for plant life. It is worth noting that there is also a gradient in the chemical potential of water between the surface of the earth, where water exists as a liquid in the oceans and rivers, and the atmosphere where it exists as an under-saturated vapor. Plants facilitate the transport of water from the surface to the atmosphere *via* respiration and transpiration, and this transport reduces the gradient in chemical potential than would otherwise be the case. This is especially true for higher plants such as trees.

9. Conclusion

In this article we have presented a general theory for non-equilibrium statistical mechanics. By ‘theory’ we mean an explicit formula for the phase space probability distribution for non-equilibrium systems. In the case of heat flow, the probability density turned out to be surprisingly simple, with a structural part that is the direct analogue of the Boltzmann distribution. The dynamic part, the so-called mirror work, depends on the direction of time. Such a time-dependence is essential for a non-equilibrium system, and it provides the necessary modification to the Boltzmann distribution. This mirror work was generalized to give the probability distribution for the case of mechanical, time-dependent work, and for non-equilibrium quantum systems.

Although the probability distribution has been given explicitly, the non-equilibrium theory is by no means complete. Ascending from the base the foundations need to be secured and the regime of validity identified. One open question is the practical limitation of the present theory in terms of the time-dependence of the non-equilibrium phenomenon. Is it restricted to the steady state, the quasi-steady state (slow changes in gradients or rates of doing work), or is it completely general? Is it restricted to the linear regime, and to small gradients? (Certainly formal expressions applicable in the non-linear regime, and to short-wavelength variations, have been given.)^{3,7} In terms of the stationarity of the probability density, how important are fluctuations away from the most likely velocity for the adiabatic evolution, and is it necessary to modify the stochastic transition probability to account for them? What is the precise relationship between the present second entropy maximization approach and the least entropy dissipation approach of Onsager and others for the transition probability?^{2,48–50}

A quick survey from the summit reveals a vista of non-equilibrium applications with many enticing journeys proceeding to different destinations. Further development of the Monte Carlo algorithm and the stochastic molecular dynamics

algorithm is indicated, as well as a quantitative comparison of their efficiency. Applications of the algorithms to specific thermodynamic fluxes and types of mechanical work, and to particular systems, also await. Beyond computer simulation, the availability of a non-equilibrium probability density and the associated entropy and free energy provides the opportunity to develop non-equilibrium analytic and numeric techniques such as already exist for equilibrium systems, (*e.g.* diagrammatic expansions, integral equations, density functionals, asymptotic analysis, *etc.*). Practical application of the theory to technological and industrial processes, and to laboratory measurements is called for, and experimental tests along the lines of time-dependent nanoscopic work, either mechanical or thermodynamic, should provide convincing evidence for the theory. And for the bigger picture the dynamic second law and its corollary should prove valuable when applied to the origin and evolution of ecological and biological systems.

Over time we expect to see further progress on the foundations and applications of non-equilibrium statistical mechanics. The characteristic property of a non-equilibrium system is its dependence on the direction of time: the future beckons time’s arrow. The flow of time, or the flow of material in time, increases the disorder of the universe, and it represents a state of order that is developed and maintained in the mediator of the flow. Arguably the rate of production of entropy in the universe is accelerated by human society and culture. Hopefully the present theory of non-equilibrium statistical mechanics is a contribution to the evolution of science toward greater order, rather than a manifestation of the inevitable chaos that is our universal fate.

Acknowledgements

The Australian Research Council is thanked for financial support. Dr Marianne Granfelt of the University of Linköping is thanked for hospitality during the writing of sections of this review.

References

- 1 R. P. Feynman, *Statistical Mechanics, A Set of Lectures*, Benjamin, Reading MA, 1972.
- 2 (a) L. Onsager, *Phys. Rev.*, 1931, **37**, 405; (b) L. Onsager, *Phys. Rev.*, 1931, **38**, 2265.
- 3 P. Attard, *J. Chem. Phys.*, 2004, **121**, 7076.
- 4 P. Attard, *J. Chem. Phys.*, 2005, **122**, 154101.
- 5 P. Attard, *J. Chem. Phys.*, 2005, **122**, 244105.
- 6 P. Attard, *J. Chem. Phys.*, 2006, **124**, 024109.
- 7 P. Attard, *J. Chem. Phys.*, 2006, **124**, 224103.
- 8 P. Attard, *Thermodynamics and Statistical Mechanics, Equilibrium by Entropy Maximisation*, Academic Press, London, 2002.
- 9 J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, Academic Press, London, 1986.
- 10 R. Zwanzig, *Non-equilibrium Statistical Mechanics*, Oxford University Press, Oxford, 2001.
- 11 M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Oxford University Press, Oxford, 1987.
- 12 M. S. Green, *J. Chem. Phys.*, 1954, **22**, 398.
- 13 R. Kubo, *Rep. Prog. Phys.*, 1966, **29**, 255.
- 14 H. Risken, *The Fokker-Planck Equation*, Springer-Verlag, Berlin, 1984.
- 15 I. Prigogine, *Introduction to Thermodynamics of Irreversible Processes*, Interscience, New York, 1967.

- 16 S. R. de Groot and P. Mazur, *Non-equilibrium Thermodynamics*, North-Holland, Amsterdam, 1969.
- 17 P. Attard, *J. Chem. Phys.*, 2002, **116**, 9616.
- 18 S. Boinepalli and P. Attard, *J. Chem. Phys.*, 2003, **119**, 12769.
- 19 D. J. Evans, E. G. D. Cohen and G. P. Morriss, *Phys. Rev. Lett.*, 1993, **71**, 2401.
- 20 D. J. Evans, *Mol. Phys.*, 2003, **101**, 1551.
- 21 C. Jarzynski, *Phys. Rev. Lett.*, 1997, **78**, 2690.
- 22 D. J. Evans and G. P. Morriss, *Statistical Mechanics of Nonequilibrium Liquids*, Academic Press, London, 1990.
- 23 D. J. Evans and G. P. Morriss, *Comput. Phys. Rep.*, 1984, **1**, 297.
- 24 P. J. Davis and D. J. Evans, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 1993, **48**, 1058.
- 25 D. J. Evans, *Phys. Rev. A*, 1986, **34**, 1449.
- 26 T. Yamada and K. Kawasaki, *Prog. Theor. Phys.*, 1967, **38**, 1031.
- 27 T. Yamada and K. Kawasaki, *Prog. Theor. Phys.*, 1975, **53**, 111.
- 28 G. P. Morriss and D. J. Evans, *Mol. Phys.*, 1985, **54**, 629.
- 29 G. P. Morriss and D. J. Evans, *Phys. Rev. A*, 1988, **37**, 3605.
- 30 B. L. Holian, G. Ciccotti, W. G. Hoover, B. Moran and H. A. Posch, *Phys. Rev. A*, 1989, **39**, 5414.
- 31 G. P. Morriss, *Phys. Rev. A*, 1989, **39**, 4811.
- 32 D. J. Evans and D. J. Searles, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 1995, **52**, 5839.
- 33 G. E. Crooks, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 1999, **60**, 2721.
- 34 G. E. Crooks, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 2000, **61**, 2361.
- 35 G. M. Wang, E. M. Sevick, E. Mittag, D. J. Searles and D. J. Evans, *Phys. Rev. Lett.*, 2002, **89**, 050601.
- 36 C. Jarzynski, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 1997, **56**, 5018.
- 37 G. E. Crooks, *J. Stat. Phys.*, 1998, **90**, 1481.
- 38 J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco, Jr and C. Bustamante, *Science*, 2002, **296**, 1832.
- 39 J. K. Johnson, J. A. Zollweg and K. E. Gubbins, *Mol. Phys.*, 1993, **78**, 591.
- 40 I. Prigogine, *Thermodynamics of Irreversible Processes*, Wiley, New York, 1955.
- 41 E. D. Schneider and J. J. Kay, *Math. Comput. Mod.*, 1994, **19**, 25.
- 42 R. Swenson and M. T. Turvey, *Ecol. Psychol.*, 1991, **3**, 317.
- 43 G. W. Paltridge, *Q. J. R. Meteorol. Soc.*, 1975, **101**, 475.
- 44 G. W. Paltridge, *Nature*, 1979, **279**, 630.
- 45 *Non-equilibrium Thermodynamics and the Production of Entropy*, ed. A. Kleidon and R. D. Lorenz, Springer, Berlin, 2005.
- 46 G. Heiken, D. Vaniman and B. M. French, *Lunar Sourcebook: A User's Guide to the Moon*, University of Cambridge Press, Cambridge, 1991.
- 47 J. C. Luvall and H. R. Holbo, *Remote Sens. Environ.*, 1989, **27**, 1.
- 48 L. Onsager and S. Machlup, *Phys. Rev.*, 1953, **91**, 1505.
- 49 (a) N. Hashitsume, *Prog. Theor. Phys.*, 1952, **8**, 461; (b) N. Hashitsume, *Prog. Theor. Phys.*, 1956, **15**, 369.
- 50 R. F. Fox and G. E. Uhlenbeck, *Phys. Fluids*, 1970, **13**, 1893.