International Journal of Thermophysics, Vol. 22, No. 1, 2001

Fluctuation Theorem for Heat Flow[1]

D. J. Searles[2, 3] and D. J. Evans[4]

Thermal conduction in a classical many-body system which is in contact with two isothermal reservoirs maintained at different temperatures is considered. The probability that when observed for a finite time, the heat flux of a finite system flows in the reverse direction to that required by the Second Law of Thermodynamics is calculated from first principles. Analytical expressions are given for the probability of observing Second Law violating fluctuations in this system. These expressions constitute an application of the fluctuation theorem to the problem of thermal conduction. The expressions are tested using nonequilibrium molecular dynamics simulations of heat flow between thermostated walls.

KEY WORDS: fluctuation theorem; heat flow; Second Law of Thermodynamics; simulation.

1. INTRODUCTION

In a nonequilibrium system, thermodynamic, X_i, or mechanical, F_e fields, do work on the system which prevent it from relaxing to equilibrium. This work is proportional to the product of the thermodynamic or mechanical force, the system volume, V, and the dissipative flux, J. The Second Law of Thermodynamics implies that for large systems the average work done by the external forces and fields and the associated total entropy production are positive. This is in spite of the fact that the microscopic equations of motion are reversible. Recently there has been some progress toward understanding the microscopic origin of this irreversibility. The fluctuation

[1] Paper presented at the Fourteenth Symposium on Thermophysical Properties, June 25–30, 2000, Boulder, Colorado, U.S.A.

[2] School of Science, Griffith University, Nathan, Brisbane, QLD 4111, Australia.

[3] To whom correspondence should be addressed.

[4] Research School of Chemistry, Australian National University, Canberra, ACT 0200, Australia.

0195-928X/01/0100-0123$19.50/0

theorem [1–5] (FT) gives a formula for the logarithm of the probability ratio that in a thermostated nonequilibrium system subject to a dissipative mechanical field, the time-averaged dissipative flux takes on a value, A, compared to minus the value, namely, $-A$. This formula is an analytic expression for the probability, for a finite system and for a finite time, that the dissipative flux flows in the reverse direction to that required by the Second Law of Thermodynamics. This theorem is quite general and has been shown to apply to classes of both deterministic and stochastic systems.

To date, the fluctuation theorem has been applied almost exclusively to mechanical rather than thermal nonequilibrium systems (for an exception, see Ref. 6). These mechanical systems were thermostated using the deterministic and stochastic thermostats that have been developed for nonequilibrium molecular dynamics computer simulation over the last two decades. The use of these model mathematical thermostats has been felt by some to reduce the relevance of the FT, since these thermostats do not actually occur in nature; they are mathematical devices developed to calculate transport coefficients correctly. Recently we have proposed a local version of the FT and applied it to Poiseuille flow of a fluid between thermostated walls [5]. In this system the mathematical thermostats operate only inside walls that are remote from the fluid, so the question raised by the use of artificial thermostats is thereby removed. The local FT applies to the fluid system which is not subject to any artificial dynamics or thermostating. In that paper the dissipative field, gravity, was, however, still mechanical.

In the present paper we again develop a fluctuation theorem for a system where the thermostats are remote from the actual system under consideration, but we go further. We consider the application of the theorem to a *thermal* transport process, where it is the boundary conditions which prevent the system from relaxing to equilibrium. The example we consider is thermal conduction in a cell which is in contact with thermostated reservoirs which maintain a constant temperature difference across the thermal conduction cell.

The thought experiment we have in mind is the following. At $t<0$ we have three equilibrium systems, H, 0, C, at temperature T_0. At $t=0$ a thermostat is applied to the H and C regions to bring them to temperatures T_{H} and T_{C}, where for simplicity $T_0=(T_{\mathrm{H}}+T_{\mathrm{C}})/2$. Again for simplicity we assume that each of the systems is composed of atoms with the same interatomic interaction and that the number of atoms in the T_{H}, T_{C} systems, N_{H}, N_{C}, is equal to $N_{\mathrm{T}}, \neq N_0$. At $t>0$ we expect that the 0-system will be driven away from equilibrium as heat flows from the hot reservoir H through the 0-system toward the cold reservoir. After

relaxation of initial transients which last a time, τ_{M}, we expect the 0 system to relax not to equilibrium but to a unique steady state defined by N_0, T_{H}, T_{C} and the conduction cell's geometrical dimensions. We do not consider the situation where for large temperature gradients nonsteady behavior may occur (e.g., Rayleigh–Bernard).

For this system we derive expressions for the logarithm of the probability that the total time-averaged entropy production $\bar{\Sigma}_t \equiv (1/t) \int_0^t ds\, \Sigma(s)$, in the conduction cell takes on a value, A, compared to minus that value. If the instantaneous irreversible entropy production is calculated as $\Sigma \equiv \sigma V = \sum J_i V X_i$, where V is the system volume, σ is the so-called entropy source strength, and the sum is over the product of all conjugate thermodynamic fluxes, J_i, and forces, X_i, then

$$\ln\left[\frac{p(\bar{\Sigma}_t = A)}{p(\bar{\Sigma}_t = -A)}\right] = \frac{At}{k_{\mathrm{B}}} \tag{1}$$

We give a generalized expression for the entropy production so that Eq. (1) is correct for systems where the imposed temperature gradient may be arbitrarily large. In the weak gradient limit, this expression reduces to the usual expression from linear irreversible thermodynamics.

From Eq. (1) it is trivial to derive an expression for the probability that for a finite time, the Second Law of Thermodynamics is violated $\bar{\Sigma}_t < 0$. If $\langle \cdots \rangle_{\bar{\Sigma}_t > 0}$ denotes an average over all fluctuations in which the time-integrated entropy production is positive, then

$$\left[\frac{p(\bar{\Sigma}_t > 0)}{p(\bar{\Sigma}_t < 0)}\right] = \langle e^{-\bar{\Sigma}_t t/k_{\mathrm{B}}} \rangle_{\bar{\Sigma}_t > 0}^{-1} = \langle e^{-\bar{\Sigma}_t t/k_{\mathrm{B}}} \rangle_{\bar{\Sigma}_t < 0} \tag{2}$$

and the probability of Second Law violations becomes exponentially small with increased time of violation, t, and with the number of particles (since Σ is extensive).

2. MICROSCOPIC DESCRIPTION OF THERMAL CONDUCTION

Experimentally there are a number of ways in which the thermal walls can be thermostated at their initial temperatures. If the walls are made of high-thermal conductivity material, a coolant may be circulated through channels in the reservoirs. Alternatively, if the heat capacity of the reservoirs is huge compared to that of the thermal conduction cell, then the temperature variation in the two reservoirs over relevant observation times may be regarded as insignificant. For theoretical analysis both of these mechanisms are too complex. Instead, we employ the so-called Nosé–

Hoover thermostat in the reservoir regions to maintain these regions at a fixed temperature. Although this thermostat does not exist in nature, its impact on the system of interest, namely, the thermal conduction cell, is only indirect. One could argue that the properties of the thermal conduction cell should be independent of whether the reservoirs are maintained at a fixed temperature by virtue of the circulation of a coolant, the use of large heat-capacity reservoirs, or the use of a fictitious thermostat such as the Nosé–Hoover thermostat.

The aim is to derive fluctuation formulae for the transient response. We consider the system initially at equilibrium (because then the phase-space distribution function is known). At this stage the whole system is at the same temperature (equal to the mean temperature of the steady-state system). The temperature gradient is then applied, and a heat flux develops.

The equations of motion for all the particles in the combined systems, H, 0, C are

$$\begin{aligned} \dot{\mathbf{q}}_i &= \mathbf{p}_i \\ \dot{\mathbf{p}}_i &= \mathbf{F}_i - \alpha_{\mathrm{H}} \mathbf{p}_i S_i - \alpha_{\mathrm{C}} \mathbf{p}_i T_i \end{aligned} \tag{3}$$

and

$$\frac{d\alpha_{\mathrm{H/C}}}{dt} = \frac{1}{Q}\left(\sum_{i \in \mathrm{H/C}} \frac{\mathbf{p}_i^2}{m} - (g+1)\, k_{\mathrm{B}} T_{\mathrm{H/C}} \right) \tag{4}$$

where g is the degrees of freedom of the system, S_i and T_i are switches equal to 1 or 0: S_i is only nonzero for particles in region H, and T_i is only nonzero for particles in region C. For simplicity, assume that the walls are sufficiently dense that the particles from region 0 do not penetrate either of the reservoir regions. The details of the interatomic forces implicit in the $\{\mathbf{F}_i\}$ are described in Section 5. It is important to note that in the 0-region and the H, 0 and C, 0 interfaces, the equations of motion can be made arbitrarily realistic by improved modeling of the interatomic forces. In the 0-region there are no unphysical forces.

In the thermal reservoirs where either S_i or $T_i = 1$, the thermostating terms in the equations of motion are unphysical in the sense that the additional terms do not exist in nature (as discussed above). The additional so-called Nosé–Hoover thermostat ensures that the reservoir regions are maintained at constant kinetic temperatures, T_{H}, T_{C}. In the long-time limit

$$\lim_{t \to \infty} \frac{d\bar{\alpha}_{\mathrm{H/C},\,t}}{dt} = 0 \Rightarrow T_{\mathrm{H/C}} \equiv \frac{1}{(3N_{\mathrm{T}} + 1)\, k_{\mathrm{B}}} \sum_{i \in \mathrm{H/C}} \overline{\frac{\mathbf{p}_{i_t}^2}{m}} \tag{5}$$

where we use the notation $\bar{B}_t = (1/t)\int_0^t ds\, B(s)$ for the time-average of any phase variable, B. The extensive constant Q controls the time scale for fluctuations in the kinetic temperatures, T_{H}, T_{C}.

3. TRANSIENT FLUCTUATION THEOREM FOR HEAT CONDUCTION

We assume that the composite system is at equilibrium at $t=0$ and that the initial phase space distribution, $f(\mathbf{\Gamma}, t)$, is canonical:

$$f(\mathbf{\Gamma}, 0) = \frac{\exp[-\beta_0[H_0(\mathbf{\Gamma}) + Q(\alpha_{\mathrm{H}}^2 + \alpha_{\mathrm{C}}^2)/2]]}{\int d\mathbf{\Gamma}\, d\alpha_{\mathrm{H}}\, d\alpha_{\mathrm{C}} \exp[-\beta_0[H_0(\mathbf{\Gamma}) + Q(\alpha_{\mathrm{H}}^2 + \alpha_{\mathrm{C}}^2)/2]]} \tag{6}$$

where $\beta_0 = 1/kT_0$ and $H_0 = \sum p_i^2/2m + \Phi(q)$ is the internal energy. We note that in an ergodic equilibrium system, Nosé–Hoover dynamics is expected to generate phases, $\mathbf{\Gamma}$, which are distributed canonically.

The phase-space compression factor, $\Lambda(\mathbf{\Gamma})$, defined from the Liouville equation,

$$\frac{df(\mathbf{\Gamma}, t)}{dt} \equiv -f(\mathbf{\Gamma}, t)\, \Lambda(\mathbf{\Gamma}) \tag{7}$$

is

$$\Lambda \equiv \frac{\partial}{\partial \mathbf{\Gamma}} \cdot \dot{\mathbf{\Gamma}} + \frac{\partial}{\partial \alpha_{\mathrm{H}}} \cdot \dot{\alpha}_{\mathrm{H}} + \frac{\partial}{\partial \alpha_{\mathrm{C}}} \cdot \dot{\alpha}_{\mathrm{C}} = -dN_{\mathrm{H}}\alpha_{\mathrm{H}} - dN_{\mathrm{C}}\alpha_{\mathrm{C}} \tag{8}$$

where d is the Cartesian dimension. Thus,

$$\begin{aligned} f(\mathbf{\Gamma}(t), t) &= f(\mathbf{\Gamma}(0), 0) \exp\left[-\int_0^t ds\, \Lambda(s)\right] \\ &= f(\mathbf{\Gamma}(0), 0) \exp\left[\int_0^t ds\, dN_{\mathrm{H}}\alpha_{\mathrm{H}}(s) + dN_{\mathrm{C}}\alpha_{\mathrm{C}}(s)\right] \end{aligned} \tag{9}$$

From the equations of motion we see that the rate of change of the internal energy is

$$\dot{H}_0 = \sum \mathbf{p}_i \cdot \dot{\mathbf{p}}_i/m - \mathbf{F}_i \cdot \dot{\mathbf{q}}_i = -2K_{\mathrm{H}}\alpha_{\mathrm{H}} - 2K_{\mathrm{C}}\alpha_{\mathrm{C}} \tag{10}$$

where K_{a} is the instantaneous kinetic energy of region a.

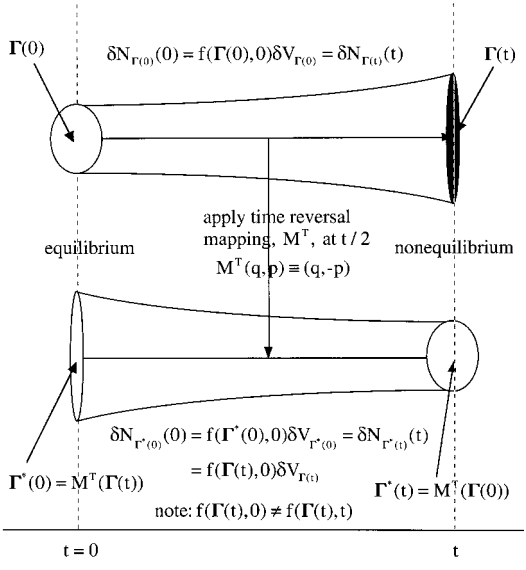

Fig. 1. Schematic diagram showing the phase space contraction as a function of time for a nonequilibrium system.

The probability ratio of observing trajectories and antitrajectories originating from phase regions $\delta\mathbf{\Gamma}(0)$, $\delta\mathbf{\Gamma}^*(0)$, respectively, is given by the probability density at the initial phase points multiplied by the initial phase volume. The phase volume at the initial point of the antitrajectory is equal to that about the final point of the original trajectory (see Fig. 1). The ratio of these phase volumes at the beginning and end is just the phase-space contraction.

In general,

$$\begin{aligned}\frac{\Pr(\delta\mathbf{\Gamma}(0))}{\Pr(\delta\mathbf{\Gamma}^*(0))} &= \frac{f(\mathbf{\Gamma}(0),0)}{f(\mathbf{\Gamma}^*(0),0)}\frac{\delta V(\mathbf{\Gamma}(0),0)}{\delta V(\mathbf{\Gamma}^*(0),0)} \\ &= \frac{f(\mathbf{\Gamma}(0),0)}{f(\mathbf{\Gamma}(t),0)}\frac{\delta V(\mathbf{\Gamma}(0),0)}{\delta V(\mathbf{\Gamma}(t),t)} \\ &= \frac{f(\mathbf{\Gamma}(0),0)}{f(\mathbf{\Gamma}(t),0)}e^{-\bar{\Lambda}_t t} \end{aligned} \tag{11}$$

Now consider $f(\mathbf{\Gamma}(0),0)/f(\mathbf{\Gamma}(t),0)$ for this system. For an NH extended canonical distribution,

$$\frac{f(\mathbf{\Gamma}(0),0)}{f(\mathbf{\Gamma}(t),0)} = \frac{\exp[-\beta_0\{H_0(\mathbf{\Gamma}(0)) + \tfrac{1}{2}Q(\alpha_{\mathrm{H}}(0)^2 + \alpha_{\mathrm{C}}(0)^2)\}]}{\exp[-\beta_0\{H_0(\mathbf{\Gamma}(t)) + \tfrac{1}{2}Q(\alpha_{\mathrm{H}}(t)^2 + \alpha_{\mathrm{C}}(t)^2)\}]}$$

$$= \exp\left[\beta_0 \int_0^t ds[\dot{H}_0(s) + Q(\alpha_{\mathrm{H}}(s)\,\dot{\alpha}_{\mathrm{H}}(s) + \alpha_{\mathrm{C}}(s)\,\dot{\alpha}_{\mathrm{C}}(s))]\right]$$

$$= \exp[\beta_0(-dN_{\mathrm{H}}\bar{\alpha}_{\mathrm{H},t}t/\beta_{\mathrm{H}} - 2N_{\mathrm{C}}\bar{\alpha}_{\mathrm{C},t}t/\beta_{\mathrm{C}}) + O(1)] \tag{12}$$

The $O(1)$ corrections will be dependent on any constraints imposed on the wall particles (see Section 5).

Combining Eqs. (11) and (12) gives

$$\frac{\Pr(\delta\mathbf{\Gamma}(0))}{\Pr(\delta\mathbf{\Gamma}^*(0))} = \exp\left[dN_{\mathrm{T}}\left(\frac{T_{\mathrm{C}} - T_{\mathrm{H}}}{T_{\mathrm{C}} + T_{\mathrm{H}}}\right)\int_0^t ds[\alpha_{\mathrm{H}}(s) - \alpha_{\mathrm{C}}(s)]\right] \tag{13}$$

Clearly, the probability ratio of observing conjugate values for the time-averaged difference in the thermostat multipliers is

$$\frac{\Pr(\bar{\alpha}_{\mathrm{C},t} - \bar{\alpha}_{\mathrm{H},t} = A)}{\Pr(\bar{\alpha}_{\mathrm{C},t} - \bar{\alpha}_{\mathrm{H},t} = -A)} = \exp\left[dN_{\mathrm{T}}\frac{T_{\mathrm{H}} - T_{\mathrm{C}}}{T_{\mathrm{C}} + T_{\mathrm{H}}}At\right] \tag{14}$$

In deriving Eq. (14), it is not necessary to assume that all transient trajectory segments that have the specified value of $\bar{\alpha}_{\mathrm{C},t} - \bar{\alpha}_{\mathrm{H},t}$ originate in the same small contiguous subregion of phase space, $\delta\mathbf{\Gamma}(0)$ or $\delta\mathbf{\Gamma}^*(0)$, and, hence, Eq. (14) is valid even when there are multiple islands of phase space which generate the specified conjugate values of $\bar{\alpha}_{\mathrm{C},t} - \bar{\alpha}_{\mathrm{H},t}$.

Equation (14) is a statement of the transient fluctuation theorem for heat flow between Nosé–Hoover thermostated walls. If the steady state exists and is unique, then a steady-state fluctuation theorem is valid asymptotically [9].

$$\lim_{t\to\infty} \ln\left[\frac{\Pr(\bar{\alpha}_{\mathrm{C},t} - \bar{\alpha}_{\mathrm{H},t} = A)}{\Pr(\bar{\alpha}_{\mathrm{C},t} - \bar{\alpha}_{\mathrm{H},t} = -A)}\right] \Big/ \left[dN_{\mathrm{T}}\frac{T_{\mathrm{H}} - T_{\mathrm{C}}}{T_{\mathrm{C}} + T_{\mathrm{H}}}At\right] = 1 \tag{15}$$

These two equations are valid outside the linear regime. The only caveat is that the steady-state formula requires the existence of a unique steady state, regardless of the initial $t = 0$ equilibrium phase, $\mathbf{\Gamma}(0)$. Equations (14) and (15) are clearly consistent with the Second Law of Thermodynamics in that it is exponentially more probable for heat to flow from hot to cold, in which case, $\bar{\alpha}_{\mathrm{C},t} > 0$, $\bar{\alpha}_{\mathrm{H},t} < 0$.

4. NONLINEAR RESPONSE THEORY FOR HEAT CONDUCTION

In order to understand this system better, we will calculate the time-dependent response of an arbitrary phase function, $B(\mathbf{\Gamma})$. Following Yamada and Kawasaki [7, 8], the distribution function for the system considered in this work, at time t after the application of a temperature gradient, is given by

$$\begin{aligned} f(\mathbf{\Gamma}, t) &= \exp\left[-\int_0^t ds\, \Lambda(-s) \right] \exp[-\beta(H(-t) + \tfrac{1}{2} Q(\alpha_{\mathrm{H}}^2(-t) + \alpha_{\mathrm{C}}^2(-t)))] \\ &= f(\mathbf{\Gamma}, 0) \exp\left\{ -\beta \int_0^{-t} ds[\, \dot{H}(s) + Q(\alpha_{\mathrm{H}}(s)\, \dot{\alpha}_{\mathrm{H}}(s) + \alpha_{\mathrm{C}}(s)\, \dot{\alpha}_{\mathrm{C}}(s)))] \right\} \\ &\quad \times \exp \int_0^{-t} ds\, \Lambda(s) \\ &= f(\mathbf{\Gamma}, 0) \exp\left\{ -\frac{dN_{\mathrm{T}}(T_{\mathrm{H}} - T_{\mathrm{C}})}{T_{\mathrm{H}} + T_{\mathrm{C}}} \int_0^{-t} ds[\, \alpha_{\mathrm{H}}(s) - \alpha_{\mathrm{C}}(s)] \right\} \end{aligned} \tag{16}$$

From this distribution function, the transient time correlation function (TTCF) expression for the ensemble average of a phase variable, B, is given by

$$\langle B(t) \rangle = \langle B(0) \rangle - \frac{dN_{\mathrm{T}}(T_{\mathrm{H}} - T_{\mathrm{C}})}{T_{\mathrm{H}} + T_{\mathrm{C}}} \int_0^t ds \langle B(s)[\, \alpha_{\mathrm{H}}(0) - \alpha_{\mathrm{C}}(0)] \rangle \tag{17}$$

By comparing with the Kawasaki distribution function for a system driven by an external mechanical force (e.g., see Eq. (7.25) of Ref. 8), we see that although the system is a thermal nonequilibrium system where boundary conditions rather than external mechanical forces drive the system away from equilibrium, there is a formal resemblance of the nonlinear response to that obtained if we applied a mechanical field,

$$F_{\mathrm{e}} = \frac{k_{\mathrm{B}}(T_{\mathrm{H}} - T_{\mathrm{C}})}{2} \tag{18}$$

to the system. In this case the intensive dissipative flux J can be identified as

$$J(\mathbf{\Gamma}) = dn_{\mathrm{T}}[\, \alpha_{\mathrm{H}}(\mathbf{\Gamma}) - \alpha_{\mathrm{C}}(\mathbf{\Gamma})] \tag{19}$$

Thus, the transient fluctuation theorem of Eq. (14) then takes on the standard form,

$$\frac{\Pr(\bar{J}_t = A)}{\Pr(\bar{J}_t = -A)} = \exp[-\beta A V F_e t] \tag{20}$$

Further, the integrated form of the fluctuation formula can be written as

$$p(\bar{J}_t < 0) = \frac{\langle e^{\bar{J}_t \beta V F_e t} \rangle_{\bar{J}_t > 0}}{1 + \langle e^{\bar{J}_t \beta V F_e t} \rangle_{\bar{J}_t > 0}} \tag{21}$$

Equation (20) shows that if A is negative, then as the system size or time interval grows, the probability of observing this negative flux relative to that of observing the corresponding positive flux increases exponentially. In the limit of infinite t or infinite system, where any fluctuations in the phase variables disappear, Eqs. (20) and (21) predict a negative value of the dissipative flux. Since in this limit, $\dot{H}_0 = 0$ and hence $2K_H\alpha_H + 2K_C\alpha_C = 0$ [see Eq. (10)], it is straightforward to show that in this limit, $\beta J V F_e = -dN_T(\alpha_H + \alpha_C)$, which is equal to the phase-space contraction and proportional to the total spontaneous entropy production [$I_\sigma = k_B(dN\alpha_H + dN\alpha_C)$]. In these limits, the heat flux per unit area at the top and bottom walls must be equal and if region C is above region H using $J_q A = dQ/dt$, we obtain $J_q A = dN_T(\alpha_H + \alpha_C)[k_B T_H T_C/(T_H - T_C)]$. The heat flux is therefore positive in these limits and with this geometry and will flow from the hot to the cold wall. Therefore, the fluctuation theorem given in the forms of Eqs. (14), (20), and (21) predicts that in the limit of infinite time interval or the thermodynamic limit, the spontaneous entropy production must be positive, the phase space must contract, and heat must flow from the hot to the cold wall.

The ultimate explanation for the irreversibility inherent in these equations is the assumption of causality in calculating the probabilities. We calculated the probability of observing fluctuations from the probabilities of observation of the *initial* equilibrium phases from which these fluctuations were generated. Had we made the corresponding anticausal assumption, then an anti-Second Law would have been derived.

5. SIMULATIONS

In order to test the fluctuation formula given by Eq. (14), simulations of a two-dimensional fluid between walls were carried out. The system consisted of three sections: a fluid region of 64 particles between two walls each containing 32 particles. The complete system was initially in a square

box with periodic boundary conditions in the direction parallel to the walls. The equations of motion for all the particles are given by Eqs. (3) and (4).

For the particles in the fluid region (labeled the 0 region), the switches S_i and T_i were set to zero at all times, and therefore, these particles obeyed Newtonian mechanics. The forces on these particles were due solely to their interactions with other particles via the WCA pair potential [10]. The particle density of the fluid region was initially set to $n = 0.4$. All units are reduced Lennard–Jones units.

The wall particles were thermostated using the Nosé–Hoover thermostat, and forces were applied so that their density was maintained at a higher value of $n = 1.2$. One wall was designated the hot wall, H, and the other the cold wall, C. In the hot wall, the switches were set to $S_i = 1$ and $T_i = 0$, whereas in the cold wall they were set to $S_i = 0$ and $T_i = 1$. These particles interacted with other particles via a WCA pair potential. In addition, a spring potential was applied to prevent the walls from diffusing [$U(r_{ij}) = \frac{1}{2}k(r_{ij} - r_{eq})^2$] and each layer of particles in the wall was subject to a layer force, using the method of Todd et al. [11]. These forces are nonphysical and designed to ensure that the wall remained intact throughout the simulation. During an equilibration period, the temperature in the Nosé–Hoover thermostat was set to $T = 1.0$ for both the hot and the cold walls. After this period, the Nosé–Hoover thermostat was set to $T_H = 1.1$ and $T_C = 0.9$ to create a temperature gradient across the cell.

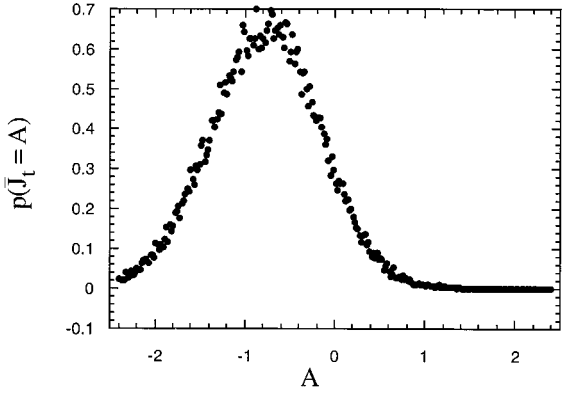

Fig. 2. Histogram of the values of $\bar{J}_t$ obtained from simulations of a fluid between two walls to which a temperature gradient is applied at time 0. The density of the fluid is $n = 0.4$, and the walls are thermostated at $T_H = 1.1$ and $T_C = 0.9$. The fluid consists of 64 particles, and each wall consists of 32 particles.

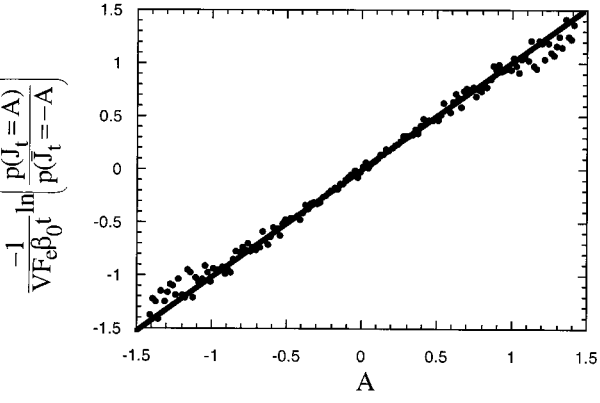

Fig. 3. Plot of $-1/(VF_e\beta_0 t)\ln(p(\bar{J}_t = A)/p(\bar{J}_t = -A))$ versus A carried out in order to test Eq. (20) for a system consisting of a fluid between two walls to which a temperature gradient is applied at time 0. The behavior predicted by Eq. (20) is shown by the line. The density of the fluid is $n = 0.4$, and the walls are thermostated at $T_H = 1.1$ and $T_C = 0.9$. The fluid consists of 64 particles, and each wall consists of 32 particles.

In order to test the fluctuation formula, approximately 8×10^7 trajectories were simulated. For each trajectory, this involved sampling a starting point from the equilibrium distribution, applying the temperature gradient, and measuring the value of $\bar{J}_t$ for a trajectory of length $t = 1.6$. The 8×10^7 values of $\bar{J}_t$ obtained were then used to construct a frequency histogram (see Fig. 2) from which the probabilities required for testing Eq. (20) could be obtained. The histogram for this simulation is shown in Fig. 2. Equation (20) was tested by plotting $-1/(VF_e\beta_0 t)\ln(p(\bar{J}_t = A)/p(\bar{J}_t = -A))$ versus A, as shown in Fig. 3. According to the fluctuation theorem for this system [see Eq. (20)], the slope of this plot should be 1. Clearly the numerical data are consistent with the theoretical prediction.

6. CONCLUSION

We have derived a fluctuation expression for a system to which a temperature gradient is applied. The expression is consistent with the Second Law of Thermodynamics and predicts that the heat flow will occur from a hot region to a cold region. The expression was tested using numerical simulation of a two-dimensional fluid containing particles undergoing Newtonian dynamics and interacting via a WCA pair potential.

ACKNOWLEDGMENTS

We would like to thank the Australian Research Council for their support of this project. Helpful discussions and comments from Professor E. G. D. Cohen are also gratefully acknowledged.

REFERENCES

1. D. J. Evans, E. G. D. Cohen, and G. P. Morriss, *Phys. Rev. Lett.* **71**:2401 (1993).
2. D. J. Evans and D. J. Searles, *Phys. Rev. E* **50**:1645 (1994); D. J. Evans and D. J. Searles, *Phys. Rev. E* **52**:5839 (1995); D. J. Evans and D. J. Searles, *Phys. Rev. E* **53**:5808 (1996); D. J. Searles and D. J. Evans, *J. Chem. Phys.* **112**:9727 (2000); D. J. Searles and D. J. Evans, *J. Chem. Phys.* **113**:3503 (2000).
3. G. Gallavotti and E. G. D. Cohen, *J. Stat. Phys.* **80**:931 (1995); G. Gallavotti and E. G. D. Cohen, *Phys. Rev. Lett.* **74**:2694 (1995).
4. J. L. Lebowitz and H. Spohn, *J. Stat. Phys.* **95**:333 (1999); C. Maes, *J. Stat. Phys.* **95**: 367 (1999); J. Kurchan, *J. Phys. A* **31**:3719 (1998); F. Bonetto, G. Gallavotti, and P. L. Garrido, *Physica D* **105**:226 (1997); F. Bonetto, N. I. Chernov, and J. L. Lebowitz, *Chaos* **8**:823 (1998).
5. G. Ayton, D. J. Evans, and D. J. Searles, submitted for publication.
6. S. Lepri, R. Livi, and A. Politi, *Physica D* **119**:140 (1998).
7. T. Yamada and K. Kawasaki, *Prog. Theor. Phys.* **38**:1031 (1967).
8. D. J. Evans and G. P. Morriss, *Statistical Mechanics of Nonequilibrium Fluids* (Academic Press, London, 1990).
9. G. Ayton and D. J. Evans, *J. Stat. Phys.* **97**:811 (1999); E. G. D. Cohen and G. Gallavotti, *J. Stat. Phys.* **96**:1343 (1999).
10. J. D. Weeks, D. Chandler, and H. C. Anderson, *J. Chem. Phys.* **54**:5237 (1985).
11. B. D. Todd, D. J. Evans, and P. J. Daivis, *Phys. Rev. E* **52**:1627 (1995).