Fluctuation and dissipation of work in a Joule experiment.

B. Cleuren,[1] C. Van den Broeck,[1] and R. Kawai[2]

[1] *Hasselt University - B-3590 Diepenbeek, Belgium*

[2] *Department of Physics, University of Alabama at Birmingham, Birmingham, Alabama 35294, USA*

(Dated: February 8, 2007)

We elucidate the connection between various fluctuation theorems by a microcanonical version of the Crooks relation. We derive the microscopically exact expression for the work distribution in an idealized Joule experiment, namely for a convex object moving at constant speed through an ideal gas. Analytic results are compared with molecular dynamics simulations of a hard disk gas.

PACS numbers: 05.70.Ln, 05.40.-a, 05.20.-y

Microscopic time-reversibility implies, in a system at equilibrium, the basic symmetry of detailed balance, stating that any process and its time reverse occur equally frequently. In the *linear* regime outside equilibrium this property entails, as Onsager has first shown, a relation between fluctuation and dissipation, since in this regime one cannot distinguish between the average regression following an external perturbation or an equilibrium fluctuation. Over the past decade, time-reversibility of deterministic or stochastic dynamics has been shown to imply relations between fluctuation and dissipation in systems *far* from equilibrium, taking the form of a number of intriguing equalities, the fluctuation theorem [1, 2], the Jarzynski equality [3] and the Crooks relation [4]. In this letter, we want to stress the relation between these results and discuss their relevance by a microscopically exact study of a Joule experiment.

Our theoretical starting point will be the derivation of a microcanonical version of the Crooks relation. This result has the advantage that we can consider from the onset an isolated system, thereby dispensing with the need for considering a heat bath. In the limit of an infinitely large system, we recover the three above mentioned equalities. The validity and experimental observability of these relations and their interconnection can be discussed from the exact result for distribution of work when moving a convex object through an ideal gas.

Consider an isolated system at time $t = 0$ in microcanonical equilibrium at energy E. Its Hamiltonian depends on a control parameter, which is varied during the time interval $[0, t]$ following a specified protocol. An initial position x_i of the system in phase space will evolve according to Hamiltonian dynamics into a final position x_f. The corresponding initial and final values of the Hamiltonian are denoted by $H_i(x_i)$ and $H_f(x_f)$, respectively. During this process, an amount of work $W = H_f(x_f) - H_i(x_i)$ is delivered to the system. Due to the microcanonical sampling of the initial state from the energy shell $H_i = E$, the work W is a random variable with the following probability density:

$$P_E(W) = \frac{\int dx_i \delta(H_i(x_i) - E)\delta(W - H_f(x_f) + H_i(x_i))}{\Omega_i(E)}, \tag{1}$$

where $\Omega_i(E) = \int dx_i \delta(H_i(x_i) - E) = \exp\{S_i(E)/k_B\}$ is the volume of the energy shell and $S_i(E)$ is the entropy at the initial equilibrium state. Consider now the time-reversed protocol. A phase point $\tilde{x}_f$, where the tilde refers to velocity inversion, will evolve in time to the final phase space position $\tilde{x}_i$. We now average over a microcanonical sampling in the energy shell $H_f(\tilde{x}_f) = E + W$. The probability distribution $\tilde{P}_{E+W}(-W)$ for a work $-W$ in this time reversed protocol is given by:

$$\tilde{P}_{E+W}(-W) = \frac{\int d\tilde{x}_f \delta(H_f(\tilde{x}_f) - E - W)\delta(H_f(\tilde{x}_f) - H_i(\tilde{x}_i) - W)}{\Omega_f(E + W)}. \tag{2}$$

Since the Jacobian for the transformation from dx_i to $d\tilde{x}_f$ is one, the integrals in Eqs. (1) and (2) are identical, and the following microcanonical Crooks relation follows:

$$\frac{P_E(W)}{\tilde{P}_{E+W}(-W)} = \frac{\Omega_f(E + W)}{\Omega_i(E)} = e^{\frac{S_f(E+W)-S_i(E)}{k_B}}. \tag{3}$$

In an appropriate thermodynamic limit, entailing $E \to \infty$, the work distributions converge to functions $P(W)$ and $\tilde{P}(-W)$, independent of the energy of the system, while the temperature T, $\partial S/\partial E = 1/T$, is a well defined constant (i.e., same for the initial and final microcanonical distribution). Since $\Delta F = \Delta E - T\Delta S$ and $\Delta E = W$ the internal energy difference, one recovers the canonical Crooks relation[4]:

$$\frac{P(W)}{\tilde{P}(-W)} = \exp\left\{\frac{\Delta S}{k_B}\right\} = e^{\beta(W - \Delta F)}, \tag{4}$$

from which the Jarzynski relation $\langle \exp(-\beta W)\rangle = \exp(-\beta \Delta F)$ follows by integration. Note that in the aforementioned thermodynamic limit, entailing $W/E \to 0$, $\Delta F = W - T(S_f(E + W) - S_i(E)) \to -T(S_f(E) - S_i(E))$ is the free energy difference between final and initial state at same energy E, independent of W. Note also that for a protocol of asymptotically long duration $t \to \infty$, that leads the system into a nonequilibrium steady state, one can write $\Delta S = t\sigma$, where σ is the entropy production per time unit, while $W = \Delta F + T\Delta S \approx Tt\sigma$. Eq. (4) then can be rewritten $P(\sigma)/\tilde{P}(-\sigma) \sim \exp(t\sigma/k_B)$, which in the particular case of a time-symmetric schedule with $P = \tilde{P}$ reduces to the Evans-Cohen-Gallavotti fluctuation theorem [1, 2]. As an application of the above result, we turn to an exactly solvable microscopic model of a Joule experiment: an ideal gas at equilibrium in an infinitely large container receives an amount of mechanical energy W by moving a closed convex body through it during a time duration t at a fixed speed V along a fixed horizontal axis x (see also Fig. 1). We will calculate explicitly the probability distribution $P(W)$ for this work. For simplicity and for comparison with molecular dynamics, we restrict ourselves to a two-dimensional system, with vertical axis y. The shape of the object is completely specified by the circumference S and the form factor $F(\theta)$, with $F(\theta)d\theta$ defined as the fraction of the circumference with polar angle between θ and $\theta + d\theta$, the angle being measured counterclockwise from the x-axis (see Ref.[5]). When a gas particle hits the object, the amount of work ΔW supplied by the external force is equal to the increase in kinetic energy of the particle. The post-collisional speed is found in terms of the (precollisional) speed $\vec{v} = (v_x, v_y)$ from the conservation of total energy and total momentum in the x-direction. The resulting work contribution is $\Delta W = -2mV \sin^2\theta(v_x - V - v_y \cot\theta)$. Note that this quantity is a random variable, through its dependence on the speed of the incoming particle and of the inclination θ of the impact point. In the case of an ideal gas at equilibrium (or in the limit of a extremely dilute gas, with the mean free path of the object much larger that its linear dimension, the so-called large Knudsen number regime), the subsequent collisions are independent random events. Hence the total work $W(t)$ after a time t, being the sum of uncorrelated identically distributed ΔW's, is a stochastic process with independent increments. The time evolution of the work distribution $P(W,t)$ is described by the following Master equation:

$$\partial_t P(W,t) = \int_{-\infty}^{+\infty} T(\Delta W)(e^{-\Delta W \partial_W} - 1)P(W,t)d\Delta W. \quad (5)$$

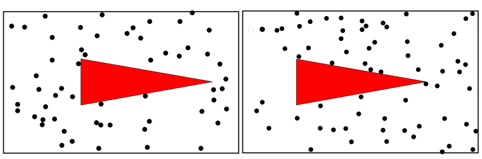

FIG. 1: Snapshots for a triangular object moving to the left (lhs panel) and to right (rhs panel). Note the void behind the object when it moves in the direction of its arrow (rhs panel).

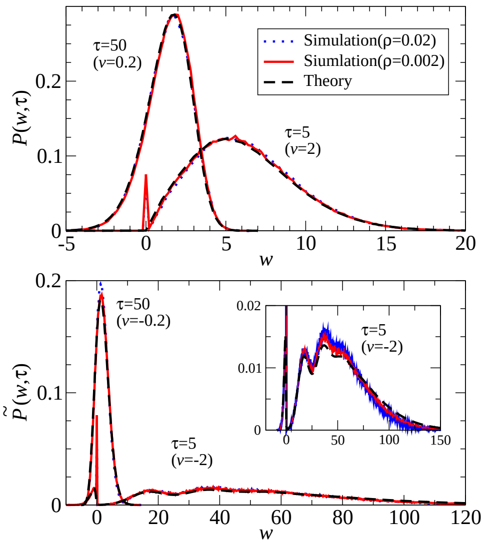

FIG. 2: Work distributions for a triangular object moving to the $+x$ (upper panel) and $-x$ direction (lower panel). Inset: detail of the multiple peak structure for $\tau = 5$.

The probability per unit time $T(\Delta W)$ for a change in W by an amount ΔW can be calculated following the basic methods of kinetic theory, and is given by:

$$T(\Delta W) = \int_0^{2\pi} SF(\theta)d\theta \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} dv_x dv_y \\ \times \rho\phi(v_x,v_y)H[(\vec{V}-\vec{v}).\vec{e}_\perp(\theta)]|(\vec{V}-\vec{v}).\vec{e}_\perp(\theta)| \\ \times \delta[\Delta W + 2mV\sin^2\theta(v_x - V - v_y\cot\theta)], \quad (6)$$

where H denotes the Heaviside function, $\vec{V} = (V, 0)$, $\vec{e}_\perp(\theta)$ is the vector orthogonal to the circumference at the orientation θ, while the delta functions picks out the incoming speeds that give rise to the requested work contribution ΔW. The gas is characterized by the (uniform) density ρ and the Maxwellian velocity distribution $\phi(v_x, v_y)$ at temperature T. The solution of the Master equation is found by Fourier transform, and is most easily expressed in terms of the following dimensionless variables:

$$w = \beta W, \quad v = V\left(\frac{\beta m}{2}\right)^{1/2}, \quad \tau = \frac{S\rho t}{(2\beta m)^{1/2}}, \quad (7)$$

being the work and speed measured in terms of the thermal energy and speed of the gas particles, and the time in terms of the average time between collisions. The cumulant generating function $G(q) = \log\langle e^{-iqw}\rangle$ reads:

$$G(q) = \log\int_{-\infty}^{+\infty} e^{-iqw}P(w,\tau)dw = \sum_{n=1}^{\infty}\frac{(-iq)^n}{n!}\kappa_n \\ = \tau\int_0^{2\pi} d\theta F(\theta)v\sin\theta\Big\{\Big(\mathrm{erf}[(1-2iq)v\sin\theta]+1\Big) \\ \times(1-2iq)e^{-4q(i+q)v^2\sin^2\theta} - 1 - \mathrm{erf}(v\sin\theta)\Big\}. \quad (8)$$

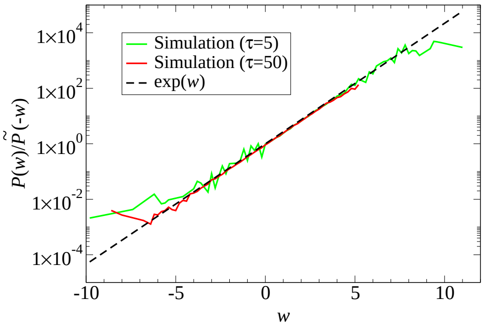

FIG. 3: Crooks relation (4) for a triangular object with velocity $v = 0.2(\tau = 50)$ and $v = 2(\tau = 5)$.

Note that, as expected, the work distribution is invariant under velocity inversion $v \to -v$ for symmetric objects $F(\theta) = F(2\pi - \theta)$.

Having obtained the explicit form of the work distribution, we turn to the verification of the fluctuation-dissipation relations. It will suffice to check the Crooks relation which, in the present problem with $\Delta F = 0$, reads $P(w) = \exp(w)\tilde{P}(-w)$. The tilde here corresponds to velocity inversion. By Fourier transformation, the Crooks relation is equivalent with the symmetry condition $\tilde{G}(-q-i) = G(q)$, which is indeed verified as, $\tilde{G}(-q-i) - G(q) = -2\tau v \int_0^{2\pi} F(\theta)\sin\theta d\theta = 0$. The last equality follows from the fact that the object is closed. The explicit result (8) allows, by expansion of $G(q)$, to evaluate the cumulants κ_n of the random variable w:

$$\kappa_1 = \langle w\rangle = \tau v\int_0^{2\pi} d\theta F(\theta)\left\{\frac{4v\sin^2\theta e^{-v^2\sin^2\theta}}{\sqrt{\pi}} + 2\sin\theta(1+2v^2\sin^2\theta)(1+\mathrm{erf}[v\sin\theta])\right\}, \tag{9a}$$

$$\kappa_2 = \langle \delta w^2\rangle = 8\tau v^2\int_0^{2\pi} d\theta F(\theta)\sin^2\theta\left\{\frac{2e^{-v^2\sin^2\theta}(1+v^2\sin^2\theta)}{\sqrt{\pi}} + v\sin\theta(3+2v^2\sin^2\theta)(1+\mathrm{erf}[v\sin\theta])\right\}, \tag{9b}$$

$$\kappa_3 = 16\tau v^3\int_0^{2\pi} d\theta F(\theta)\sin^3\theta\left\{\frac{2e^{-v^2\sin^2\theta}v\sin\theta(5+2v^2\sin^2\theta)}{\sqrt{\pi}} + (3+12v^2\sin^2\theta+4v^4\sin^4\theta)(1+\mathrm{erf}[v\sin\theta])\right\}, \tag{9c}$$

$$\kappa_4 = 64\tau v^4\int_0^{2\pi} d\theta F(\theta)\sin^4\theta\left\{\frac{2e^{-v^2\sin^2\theta}(4+9v^2\sin^2\theta+2v^4\sin^4\theta)}{\sqrt{\pi}} + v\sin\theta(15+20v^2\sin^2\theta+4v^4\sin^4\theta)(1+\mathrm{erf}[v\sin\theta])\right\}. \tag{9d}$$

We next turn to two limits of particular interest. In the quasi-static limit $v \to 0$, expansion of $G(q)$ in v leads to:

$$G(q) \approx -\frac{8q(i+q)}{\sqrt{\pi}}v^2\int_0^{2\pi} d\theta F(\theta)\sin^2\theta + O(v^3). \tag{10}$$

Hence $P(w)$ converges to $\delta(W)$ in the strict quasi-static limit ($G(q) = 0$). Keeping the leading term in Eq. (10), one concludes that $P(w)$ is Gaussian. This level of perturbation corresponds to the linear Gaussian regime around equilibrium. The average work reduces to the familiar Joule heating, which in original variables, reads: $\langle W\rangle \approx \gamma V^2 t$. Here γ is the friction coefficient, the proportionality factor in friction force versus speed: $F_{friction} = \gamma V$. The expression for the friction coefficient γ agrees with a direct calculation of this quantity [5], namely $\gamma = 4S\rho\sqrt{k_B T m/2\pi}\int_0^{2\pi} F(\theta)\sin^2\theta d\theta$. The second moment reproduces an equilibrium fluctuation-dissipation relation. In original variables: $\beta\langle\delta W^2\rangle = 2\langle W\rangle$. The Jarzynski equality implies more generally for a Gaussian shape of W that the so-called fluctuation-dissipation ratio $R = \beta\langle\delta W^2\rangle/2(\langle W\rangle - W_{rev})$ be equal to one [6]. In the present case, the reversible work, $W_{rev} = \Delta F$, is equal to zero. The quasi-static limit has to be compared but also contrasted with the long time limit $\tau \to \infty$. We first note that all the cumulants are proportional to τ, a property characteristic for processes with independent increments. As a result, one finds in the limit $\tau \to \infty$ that W converges to $\langle W\rangle$ (the law of large numbers), and more precisely that the random variable $(W - \langle W\rangle)/\sqrt{\langle\delta W^2\rangle}$ converges to a normal random variable (central limit theorem). We stress however that, while in this case the dominant part of the probability mass is indeed rendered correctly by a Gaussian ansatz centered around $\langle W\rangle$, the validity of the fluctuation theorems rests on the contribution of the so-called extreme non-Gaussian deviations, as all higher order cumulants contribute equally (all proportional to τ) to the Crooks relation. In particular, the average work is not related in any obvious way to the free energy difference nor does the fluctuation-dissipation ratio verify the near equilibrium result $R = 1$. We finally turn to a comparison of the above analytic results with those from hard disk molecular dynamics for a dilute gas with $N = 2000$ disks of diameter $d = 1$ and mass $m = 1$ (see Refs. [7] for the detailed simulation methods.) The initial positions and velocities of the disks are sampled from a microcanoni-

TABLE I: Comparison between molecular dynamics simulation and theory ($v\tau = 10$).

	v	κ_1		κ_2		κ_3		κ_4		R		$\langle e^{-w} \rangle$
		Simulation	Theory	Simulation	Theory	Simulation	Theory	Simulation	Theory	Simulation	Theory	Simulation
Disk	0.01	0.190	0.226	0.394	0.451	0.004	0.002	−0.009	0.005	1.04	0.998	1.01
	0.50	13.0	12.0	38.3	35.7	0.353	0.385	0.191	0.229	1.48	1.49	0.31
	1.00	30.1	27.8	169.	160.	0.466	0.509	0.234	0.300	2.80	2.88	0.033
	2.00	86.3	80.2	1310.	1280.	0.458	0.507	0.225	0.282	7.57	7.96	0.0043
Triangle	0.01	0.078	0.078	0.156	0.153	−0.076	−0.111	0.012	0.018	1.004	0.981	1.00
	−0.01	0.078	0.079	0.159	0.160	0.002	0.111	0.024	0.017	1.01	1.02	1.00
	0.50	2.73	2.76	2.97	2.99	−0.566	−0.530	1.04	1.03	0.544	0.542	0.77
	−0.50	5.43	5.63	21.48	23.2	0.585	0.629	0.371	0.464	1.98	2.06	1.01
	1.00	4.32	4.14	5.05	4.14	0.395	0.189	0.965	0.737	0.583	0.500	0.53
	−1.00	14.9	15.6	106.	120.	0.623	0.679	0.383	0.511	3.57	3.85	1.00
	2.00	6.09	6.08	10.9	11.0	0.614	0.627	0.417	0.446	0.891	0.903	0.11
	−2.00	48.5	52.0	855.	1040.	0.568	0.646	0.292	0.443	8.83	10.0	1.10

cal ensemble (initial "temperature" $T = 1$) in a square box of L=1000 (i.e., initial gas density $\rho = 0.002$) with periodic boundary conditions in both x and y directions. Averages were taken over 400,000 realizations. Two simple shapes are considered for the object: a disk with a diameter $A = 10$ ($F(\theta) = 1/2\pi$) and an isosceles triangle with a base length $A = 10$ and an apex angle $\phi = 20°$ [$F(\theta)$ is the sum of three delta contributions at the angles $\theta = \pi - \phi/2$, $\theta = 3\pi/2$ and $\theta = \phi/2$], cf. Fig. 1. For the disk, we have by symmetry that $P = \tilde{P}$. The latter distributions can however be very different for the triangle. An illustrative comparison between analytical and simulation results is shown in Figs. 2-3 and Table I for a wide range of values of v and τ. Overall, qualitative agreement is observed. In particular, the progressive change in the general shape of the probability distribution $P(w)$ from its quasi-static Gaussian shape to a very complicated multi-peaked distribution for the triangle is well reproduced by the present theory. The comparison of moments given in Table I confirms this agreement. In the linear response regime ($R \sim 1$), the molecular dynamics simulation reproduce quite well the Jarzynski equality, meaning that such simulations could be used as an accurate estimate of the free energy difference (which happens to be zero here). However, as the velocity increases, the Jarzynski equality is not reproduced because its validity rests on the contribution of extremely rare events. Note however the following exception: the triangular object moving to the left satisfies the Jarzynski equality surprisingly well even when its velocity is greater than the mean velocity of the gas particles. The intuitive explanation is that negative work, corresponding to particles hitting the triangle while it moves away from them, can be more easily realized by the collision on the elongated side of the triangle (see the left panel of Fig. 1). This observation is also supported at the level of the Crooks relation, cf. Fig 3, and is in agreement with a general argument (C. Jarzynski, private communication) that the Jarzynksi equality is verified more easily when operating at higher dissipation, cf. Table I.

We conclude with a few relevant comments. Only in the quasi-static limit (including the linear regime around equilibrium) is the convergence of $P(W)$ to a delta function and to a Gaussian strong enough to be allowed to interchange this limit with the Jarzynski average. One concludes that $\langle W \rangle$ converges to ΔF and the fluctuation-dissipation ratio R converges to 1. In the limit $\tau \to \infty$, W, being the sum of independent increments, both law of large numbers and central limit theorem apply, but the Jarzynski average is dominated by the extreme non-Gaussian fluctuations. The present study also illustrates that one has to go far out of equilibrium, in case the speed of the object comparable to the speed of the particles, to see significant deviations from linear response theory. Furthermore, one will, in the latter case (R significantly different from 1), typically observe deviations from the Jarzynski equality because it becomes practically impossible to sample extremely rare events. We however found an interesting exception to this rule: for the motion of a highly asymmetric object, the negative work contributions of the not-so-rare collisions on its elongated tail allow to verify the Jarzynski equality outside the regime of linear response. Concomitantly, the Jarzynski equality can in such cases be used to estimate free energy differences from far from equilibrium measurements.

[1] D. Evans, E. G. D. Cohen, and G. P. Morris, Phys. Rev. Lett. **71**, 2401 (1993).

[2] G. Gallavotti and E. G. D. Cohen, Phys. Rev. Lett. **74**, 2694 (1995).

[3] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).

[4] G. E. Crooks, Phys. Rev. E **60**, 2721 (1999).

[5] P. Meurs, C. Van den Broeck, and A. Garcia, Phys. Rev. E **70**, 051109 (2004).

[6] F. Ritort, Semin. Poincaré **2**, 195 (2003).

[7] C. Van den Broeck, P. Meurs, and R. Kawai, New J. Phys. **7**, 10 (2005).