Non-equilibrium thermodynamics of stochastic systems with odd and even variables

Richard E. Spinney and Ian J. Ford
Department of Physics and Astronomy, UCL, Gower Street, London WC1E 6BT, UK and London Centre for Nanotechnology, 17-19 Gordon Street London WC1H 0AH, UK
(Dated: December 20, 2011)

The total entropy production of stochastic systems can be divided into three quantities. The first corresponds to the excess heat, whilst the second two comprise the house-keeping heat. We denote these two components the transient and generalised house-keeping heat and we obtain an integral fluctuation theorem for the latter, valid for all Markovian stochastic dynamics. A previously reported formalism is obtained when the stationary probability distribution is symmetric for all variables that are odd under time reversal which restricts consideration of directional variables such as velocity.

PACS numbers: 05.70.Ln,05.40.-a

arXiv:1201.0904v2 [cond-mat.stat-mech] 10 Feb 2012

For over 100 years the statement of the second law of thermodynamics stood simply as the Clausius inequality. However in recent years advances in technology have encouraged the thermodynamic consideration of small systems which has led to the generalisation of the concept of entropy production: it may be associated with individual dynamical realisations revealing a wealth of relations valid out of equilibrium. Such extensions had their origins in the dissipation function of Evans et al. for thermostatted systems that led to the Fluctuation Theorem [1–4] with similar, but asymptotic relations for chaotic systems [5] which were extended to Langevin dynamics [6] followed by general Markovian stochastic systems [7]. Crooks and Jarzynski [8–10] then derived work relations for a variety of dynamics which held for finite times. These were followed by similar generalised relations for the entropy production associated with transitions between stationary states [11], the total entropy production [12] and the heat dissipation required to maintain a stationary state [13]. More recently the relationship between the latter quantities has been explored [14–17] resulting in a formalism involving a division of the total entropy change into two distinct terms, the adiabatic and non-adiabatic entropy productions [18–20], each of which obeys appropriate fluctuation relations and which map onto the house-keeping and excess heats, respectively, of Oono and Paniconi [21]. We seek to take such a formalism and generalise its scope by the explicit inclusion of both even (e.g. spatial) and odd (e.g. momentum) variables that transform differently under time reversal. In doing so we define a new quantity which obeys an integral fluctuation theorem for all time.

Specifically, we consider the dynamics of a general set of variables $\boldsymbol{x} = (x^1, x^2, \dots x^n)$ that behave differently under time reversal such that $\boldsymbol{\varepsilon x} = (\varepsilon^1 x^1, \varepsilon^2 x^2, \dots \varepsilon^n x^n)$ where $\varepsilon^i = \pm 1$ for even and odd variables x^i respectively. Odd variables arise in the discussion of directional quantities and consequently such a consideration is essential when discussing velocities, from the most simple lattice Boltzmann model to considerations of full phase space. The entropy production of a path of duration τ depends on two probabilities. The first is the path probability, $P^{\mathrm{F}}[\vec{\boldsymbol{x}}]$, defined as the probability of the forward trajectory, $\vec{\boldsymbol{x}} = \boldsymbol{x}(t)$ for $0 \le t \le \tau$, with a distribution of starting configurations, $P^{\mathrm{F}}(\boldsymbol{x}(0), 0)$, that acts as an initial condition for the general master equation (relevant examples arise, for example, in the context of full phase space [22, 23] and in lattice Boltzmann models):

$$\frac{\partial P^{\mathrm{F}}(\boldsymbol{x}, t)}{\partial t} = \sum_{\boldsymbol{x}'} T(\boldsymbol{x}|\boldsymbol{x}', \lambda^{\mathrm{F}}(t)) P^{\mathrm{F}}(\boldsymbol{x}', t) \tag{1}$$

where $T(\boldsymbol{x}|\boldsymbol{x}', \lambda^{\mathrm{F}}(t))$ is a matrix of transition rates between configurations $\boldsymbol{x}'$ and $\boldsymbol{x}$, defining the normal dynamics, parameterised by the forward protocol λ^{F} at time t. We use notation $T(\boldsymbol{x}|\boldsymbol{x}) = -\sum_{\boldsymbol{x}' \neq \boldsymbol{x}} T(\boldsymbol{x}'|\boldsymbol{x})$ which describes the mean escape rate. The path probability of some sequence of N transitions to configurations $\boldsymbol{x}_i$ from $\boldsymbol{x}_{i-1}$ at times t_i, such that $t_0 = 0$ and $t_{N+1} = \tau$, can then be computed as a function of transition rates and exponential waiting times

$$P^{\mathrm{F}}[\vec{\boldsymbol{x}}] = P^{\mathrm{F}}(\boldsymbol{x}_0, 0) e^{\int_{t_0}^{t_1} dt' T(\boldsymbol{x}_0|\boldsymbol{x}_0, \lambda^{\mathrm{F}}(t'))} \times \prod_{i=1}^{N} T(\boldsymbol{x}_i|\boldsymbol{x}_{i-1}, \lambda^{\mathrm{F}}(t_i)) dt_i e^{\int_{t_i}^{t_{i+1}} dt' T(\boldsymbol{x}_i|\boldsymbol{x}_i, \lambda^{\mathrm{F}}(t'))}. \tag{2}$$

We compare this probability to that of another trajectory $\vec{\boldsymbol{x}}^*$, protocol λ^*, initial condition $P^*(\boldsymbol{x}^*(0), 0)$ and chosen dynamics, denoted P^*, and write

$$A[\vec{\boldsymbol{x}}] = \ln \left[P^{\mathrm{F}}[\vec{\boldsymbol{x}}] / P^*[\vec{\boldsymbol{x}}^*] \right]. \tag{3}$$

Such a quantity may obey an integral fluctuation theorem (IFT) which may be derived by explicit summation over all possible paths, $\vec{\boldsymbol{x}}$, for which $P^{\mathrm{F}}[\vec{\boldsymbol{x}}] \neq 0$ as follows

$$\langle \exp[-A[\vec{\boldsymbol{x}}]] \rangle^{\mathrm{F}} = \sum_{\vec{\boldsymbol{x}}} P^{\mathrm{F}}[\vec{\boldsymbol{x}}] \exp[-A[\vec{\boldsymbol{x}}]] = \sum_{\vec{\boldsymbol{x}}} P^{\mathrm{F}}[\vec{\boldsymbol{x}}] \frac{P^*[\vec{\boldsymbol{x}}^*]}{P^{\mathrm{F}}[\vec{\boldsymbol{x}}]} = \sum_{\vec{\boldsymbol{x}}^*} P^*[\vec{\boldsymbol{x}}^*] = 1. \tag{4}$$

We assume a one to one mapping between $\vec{\boldsymbol{x}}$ and $\vec{\boldsymbol{x}}^*$ (a condition equivalent to a Jacobian of unity in the transformation) so that we may consider the summation over

$\vec{\boldsymbol{x}}^*$ to be equivalent to that over $\vec{\boldsymbol{x}}$. We also require that $P^*[\vec{\boldsymbol{x}}^*] = 0$ for all $P^{\rm F}[\vec{\boldsymbol{x}}] = 0$ such that the final summation contains all possible paths $\vec{\boldsymbol{x}}^*$, meaning the required normalisation of $P^*[\vec{\boldsymbol{x}}^*]$ then yields the result of unity. A key result is the implication $\langle A[\vec{\boldsymbol{x}}]\rangle^{\rm F} \geq 0$ by Jensen's inequality.

A common choice for P^*, and that used to construct the total entropy production, is that of the normal dynamics under the reversed protocol, denoted $P^* = P^{\rm R}$. Given the specification of the normal dynamics we point out that all further specifications, including the choice of protocol, can be systematically derived from the appropriate path transformation $\vec{\boldsymbol{x}}^*$ which we must choose carefully in conjunction with the dynamics so as to obey the above conditions. At this point we must be clear that given a transition $\boldsymbol{x} \to \boldsymbol{x}'$ under the normal dynamics, the transition $\boldsymbol{x}' \to \boldsymbol{x}$ is not, in general, possible under those same dynamics. Explicitly, we can construct models such that $T(\boldsymbol{x}'|\boldsymbol{x}) \neq 0$ whilst $T(\boldsymbol{x}|\boldsymbol{x}') = 0$ (as an intuitive example: Hamiltonian dynamics cannot produce a negative positional step whilst the velocity is positive). The correct path, $\vec{\boldsymbol{x}}^*$, to consider is the time reversed trajectory proper which includes a reversal of sign for all odd variables. This is the choice $\boldsymbol{x}^*(t) = \boldsymbol{x}^\dagger(t) = \boldsymbol{\varepsilon x}(\tau - t)$ and it satisfies the condition $P^*[\vec{\boldsymbol{x}}^*] = P^{\rm R}[\vec{\boldsymbol{x}}^\dagger] = 0$ for all $P^{\rm F}[\vec{\boldsymbol{x}}] = 0$ required for an IFT. The reversed protocol $\lambda^* = \lambda^{\rm R}$ may be similarly obtained from the forward protocol, which may be treated as an even dynamical variable, meaning it transforms to yield $\lambda^*(t) = \varepsilon\lambda^{\rm F}(\tau - t) = \lambda^{\rm F}(\tau - t) = \lambda^{\rm R}(t)$. And finally we require the choice of initial condition for the reverse path. This may be informed physically: we seek to characterise the irreversibility of the forward path and so initiate the reverse behaviour by time reversing the coordinates, $\boldsymbol{x}(\tau)$, and distribution, $P^{\rm F}(\boldsymbol{x}(\tau), \tau)$, at the end of the forward process and evolve forward in time from there. The distribution can also be found by applying the transformation rules used to obtain the trajectory $\vec{\boldsymbol{x}}^\dagger$ from $\vec{\boldsymbol{x}}$ such that $P^*(\boldsymbol{x}^*(0), 0) = P^R(\boldsymbol{x}^\dagger(0), 0) = \hat{\boldsymbol{\varepsilon}} P^{\rm F}(\boldsymbol{\varepsilon x}(\tau), \tau) = P^{\rm F}(\boldsymbol{\varepsilon\varepsilon x}(\tau), \tau) = P^{\rm F}(\boldsymbol{x}(\tau), \tau)$ where $\hat{\boldsymbol{\varepsilon}}$ denotes the time reversal operation on the distribution. In this instance the path probability is therefore

$$P^{\rm R}[\vec{\boldsymbol{x}}^\dagger] = P^{\rm R}(\boldsymbol{x}_0^\dagger, 0) e^{\int_{t_0}^{t_1} dt' T(\boldsymbol{x}_0^\dagger|\boldsymbol{x}_0^\dagger, \lambda^{\rm R}(t'))} \times \prod_{i=1}^{N} T(\boldsymbol{x}_i^\dagger|\boldsymbol{x}_{i-1}^\dagger, \lambda^{\rm R}(t_i)) dt_i e^{\int_{t_i}^{t_{i+1}} dt' T(\boldsymbol{x}_i^\dagger|\boldsymbol{x}_i^\dagger, \lambda^{\rm R}(t'))}. \quad (5)$$

We have $\boldsymbol{x}_i^\dagger = \boldsymbol{\varepsilon x}_{N-i}$ so we may rearrange to give

$$P^{\rm R}[\vec{\boldsymbol{x}}^\dagger] = P^{\rm F}(\boldsymbol{x}_N, \tau) e^{\int_{t_N}^{t_{N+1}} dt' T(\boldsymbol{\varepsilon x}_0|\boldsymbol{\varepsilon x}_0, \lambda^{\rm R}(t'))} \times \prod_{i=1}^{N} e^{\int_{t_{N-i}}^{t_{N-i+1}} dt' T(\boldsymbol{\varepsilon x}_i|\boldsymbol{\varepsilon x}_i, \lambda^{\rm R}(t'))} T(\boldsymbol{\varepsilon x}_{i-1}|\boldsymbol{\varepsilon x}_i, \lambda^{\rm R}(t_{N-i+1})) dt_i. \quad (6)$$

We then perform a change of variable $t' \to \tau - t'$ and use $\lambda^{\rm R}(t_i) = \lambda^{\rm F}(t_{N-i+1})$ such that

$$P^{\rm R}[\vec{\boldsymbol{x}}^\dagger] = P^{\rm F}(\boldsymbol{x}_N, \tau) e^{-\int_{t_1}^{t_0} dt' T(\boldsymbol{\varepsilon x}_0|\boldsymbol{\varepsilon x}_0, \lambda^{\rm F}(t'))} \times \prod_{i=1}^{N} e^{-\int_{t_{i+1}}^{t_i} dt' T(\boldsymbol{\varepsilon x}_i|\boldsymbol{\varepsilon x}_i, \lambda^{\rm F}(t'))} T(\boldsymbol{\varepsilon x}_{i-1}|\boldsymbol{\varepsilon x}_i, \lambda^{\rm F}(t_i)) dt_i. \quad (7)$$

A comparison of $P^{\rm F}[\boldsymbol{x}]$ and $P^{\rm R}[\boldsymbol{x}^\dagger]$ characterises the irreversibility of the forward path and defines the total entropy production (using units $k_B = 1$)

$$\begin{aligned}\Delta S_{\rm tot} &= \ln P^{\rm F}[\vec{\boldsymbol{x}}] - \ln P^{\rm R}[\vec{\boldsymbol{x}}^\dagger] \\ &= \ln \frac{P^{\rm F}(\boldsymbol{x}_0, 0)}{P^{\rm F}(\boldsymbol{x}_N, \tau)} + \sum_{i=0}^{N} \ln \frac{e^{\int_{t_i}^{t_{i+1}} dt' T(\boldsymbol{x}_i|\boldsymbol{x}_i, \lambda^{\rm F}(t'))}}{e^{\int_{t_i}^{t_{i+1}} dt' T(\boldsymbol{\varepsilon x}_i|\boldsymbol{\varepsilon x}_i, \lambda^{\rm F}(t'))}} \\ &\quad + \sum_{i=1}^{N} \ln \frac{T(\boldsymbol{x}_i|\boldsymbol{x}_{i-1}, \lambda^{\rm F}(t_i))}{T(\boldsymbol{\varepsilon x}_{i-1}|\boldsymbol{\varepsilon x}_i, \lambda^{\rm F}(t_i))}\end{aligned} \quad (8)$$

which by its definition and Eq. (4) obeys [12]

$$\langle \exp\left[-\Delta S_{\rm tot}\right]\rangle^{\rm F} = 1. \quad (9)$$

We find that this form of $\Delta S_{\rm tot}$ is more complicated than previous descriptions [18, 24] unless $\boldsymbol{\varepsilon x} = \boldsymbol{x}$. Note that if detailed balance holds, such that $P^{\rm eq}(\boldsymbol{x}) T(\boldsymbol{x}'|\boldsymbol{x}) = P^{\rm eq}(\boldsymbol{\varepsilon x}') T(\boldsymbol{\varepsilon x}|\boldsymbol{\varepsilon x}')$, we expect $P^{\rm eq}$, the equilibrium state for a given $\lambda^{\rm F}(t)$, to satisfy $P^{\rm eq}(\boldsymbol{x}) = P^{\rm eq}(\boldsymbol{\varepsilon x})$ due to time-reversal invariance, along with $T(\boldsymbol{x}|\boldsymbol{x}) = T(\boldsymbol{\varepsilon x}|\boldsymbol{\varepsilon x})$. For a system in equilibrium, we therefore conclude that $\Delta S_{\rm tot} = 0$ for all paths.

Next we consider alternative specifications of P^*. We consider the adjoint dynamics which lead to the same stationary state, $P^{\rm st}(\boldsymbol{x}, \lambda^{\rm F}(t))$, as the normal dynamics, but generate flux of the opposite sign in that stationary state. It can be shown [14, 18, 24] that this requires an adjoint transition rate matrix $T^{\rm ad}$ described by

$$T^{\rm ad}(\boldsymbol{x}|\boldsymbol{x}', \lambda^{\rm F}(t)) = T(\boldsymbol{x}'|\boldsymbol{x}, \lambda^{\rm F}(t)) \frac{P^{\rm st}(\boldsymbol{x}, \lambda^{\rm F}(t))}{P^{\rm st}(\boldsymbol{x}', \lambda^{\rm F}(t))}. \quad (10)$$

However, in the same way that the normal dynamics may not, in general, permit transitions $\boldsymbol{x}' \to \boldsymbol{x}$ or $\boldsymbol{\varepsilon x} \to \boldsymbol{\varepsilon x}'$, similarly the adjoint dynamics may not, in general, permit transitions $\boldsymbol{x} \to \boldsymbol{x}'$ or $\boldsymbol{\varepsilon x}' \to \boldsymbol{\varepsilon x}$. Thus we must consider the representation of the adjoint dynamics as either Eq. (10) or

$$T^{\rm ad}(\boldsymbol{\varepsilon x}'|\boldsymbol{\varepsilon x}, \lambda^{\rm F}(t)) = T(\boldsymbol{\varepsilon x}|\boldsymbol{\varepsilon x}', \lambda^{\rm F}(t)) \frac{P^{\rm st}(\boldsymbol{\varepsilon x}', \lambda^{\rm F}(t))}{P^{\rm st}(\boldsymbol{\varepsilon x}, \lambda^{\rm F}(t))} \quad (11)$$

depending on the specific transition being considered. Explicitly, when choosing $P^*[\vec{\boldsymbol{x}}^*]$, we should not consider $P^{\rm ad}[\vec{\boldsymbol{x}}]$ or $P^{\rm ad}[\vec{\boldsymbol{x}}^\dagger]$ since these might violate the required condition $P^*[\vec{\boldsymbol{x}}^*] = 0$ for all $P^{\rm F}[\vec{\boldsymbol{x}}] = 0$, required for an IFT.

Under the adjoint dynamics, however, an appropriate transformation of $\vec{\boldsymbol{x}}$ is $\boldsymbol{x}^*(t) = \boldsymbol{x}^{\rm R}(t) = \boldsymbol{x}(\tau - t)$. Applying the transformation rules used to obtain $\vec{\boldsymbol{x}}^{\mathrm{R}}$ yields the reverse protocol as before $\lambda^*(t)=\lambda^{\mathrm{F}}(\tau-t)=\lambda^{\mathrm{R}}(t)$ and the initial distribution $P^*(\boldsymbol{x}^*(0),0)=P^{\mathrm{ad,R}}(\boldsymbol{x}^{\mathrm{R}}(0),0)=P^{\mathrm{F}}(\boldsymbol{x}(\tau),\tau)$. The path probability is then

$$
\begin{aligned}
&P^{\mathrm{ad,R}}[\vec{\boldsymbol{x}}^{\mathrm{R}}]=P^{\mathrm{ad,R}}(\boldsymbol{x}_0^{\mathrm{R}},0)e^{\int_{t_0}^{t_1}dt' T^{\mathrm{ad}}(\boldsymbol{x}_0^{\mathrm{R}}|\boldsymbol{x}_0^{\mathrm{R}},\lambda^{\mathrm{R}}(t'))}\\
&\times\prod_{i=1}^{N}T^{\mathrm{ad}}(\boldsymbol{x}_i^{\mathrm{R}}|\boldsymbol{x}_{i-1}^{\mathrm{R}},\lambda^{\mathrm{R}}(t_i))dt_i e^{\int_{t_i}^{t_{i+1}}dt' T^{\mathrm{ad}}(\boldsymbol{x}_i^{\mathrm{R}}|\boldsymbol{x}_i^{\mathrm{R}},\lambda^{\mathrm{R}}(t'))}\\
&=P^{\mathrm{F}}(\boldsymbol{x}_N,\tau)e^{-\int_{t_1}^{t_0}dt' T^{\mathrm{ad}}(\boldsymbol{x}_0|\boldsymbol{x}_0,\lambda^{\mathrm{F}}(t'))} \qquad (12)\\
&\times\prod_{i=1}^{N}e^{-\int_{t_{i+1}}^{t_i}dt' T^{\mathrm{ad}}(\boldsymbol{x}_i|\boldsymbol{x}_i,\lambda^{\mathrm{F}}(t'))}T^{\mathrm{ad}}(\boldsymbol{x}_{i-1}|\boldsymbol{x}_i,\lambda^{\mathrm{F}}(t_i))dt_i.
\end{aligned}
$$

We then construct a quantity of the form given in Eq. (3), utilise Eq. (10) and the property $T^{\mathrm{ad}}(\boldsymbol{x}|\boldsymbol{x})=T(\boldsymbol{x}|\boldsymbol{x})$, valid by means of balance, to obtain

$$
\begin{aligned}
\Delta S_1&=\ln P^{\mathrm{F}}[\vec{\boldsymbol{x}}]-\ln P^{\mathrm{ad,R}}[\vec{\boldsymbol{x}}^{\mathrm{R}}]\\
&=\ln\frac{P^{\mathrm{F}}(\boldsymbol{x}_0,0)}{P^{\mathrm{F}}(\boldsymbol{x}_N,\tau)}+\sum_{i=1}^{N}\ln\frac{P^{\mathrm{st}}(\boldsymbol{x}_i,\lambda^{\mathrm{F}}(t_i))}{P^{\mathrm{st}}(\boldsymbol{x}_{i-1},\lambda^{\mathrm{F}}(t_i))} \qquad (13)
\end{aligned}
$$

which through its definition and Eq. (4) obeys

$$
\langle\exp[-\Delta S_1]\rangle^{\mathrm{F}}=1 \qquad (14)
$$

which exists in the literature as the Hatano-Sasa relation [11] or IFT for the non-adiabatic entropy production [18–20]. Let us now consider, once again under the adjoint dynamics, the path transformation choice $\boldsymbol{x}^*(t)=\boldsymbol{x}^{\mathrm{T}}(t)=\boldsymbol{\varepsilon}\boldsymbol{x}(t)$. Applying the transformation rules we obtain the protocol $\lambda^*(t)=\varepsilon\lambda^{\mathrm{F}}(t)=\lambda^{\mathrm{F}}(t)$ and initial distribution $P^*(\boldsymbol{x}^*(0),0)=P^{\mathrm{ad,F}}(\boldsymbol{x}^{\mathrm{T}}(0),0)=\hat{\boldsymbol{\varepsilon}}P^{\mathrm{F}}(\boldsymbol{\varepsilon}\boldsymbol{x}(0),0)=P^{\mathrm{F}}(\boldsymbol{x}(0),0)$. The path probability for this case is therefore

$$
\begin{aligned}
&P^{\mathrm{ad,F}}[\vec{\boldsymbol{x}}^{\mathrm{T}}]=P^{\mathrm{ad,F}}(\boldsymbol{x}_0^{\mathrm{T}},0)e^{\int_{t_0}^{t_1}dt' T^{\mathrm{ad}}(\boldsymbol{x}_0^{\mathrm{T}}|\boldsymbol{x}_0^{\mathrm{T}},\lambda^{\mathrm{F}}(t'))}\\
&\times\prod_{i=1}^{N}T^{\mathrm{ad}}(\boldsymbol{x}_i^{\mathrm{T}}|\boldsymbol{x}_{i-1}^{\mathrm{T}},\lambda^{\mathrm{F}}(t_i))dt_i e^{\int_{t_i}^{t_{i+1}}dt' T^{\mathrm{ad}}(\boldsymbol{x}_i^{\mathrm{T}}|\boldsymbol{x}_i^{\mathrm{T}},\lambda^{\mathrm{F}}(t'))}\\
&=P^{\mathrm{F}}(\boldsymbol{x}_0,0)e^{\int_{t_0}^{t_1}dt' T^{\mathrm{ad}}(\boldsymbol{\varepsilon}\boldsymbol{x}_0|\boldsymbol{\varepsilon}\boldsymbol{x}_0,\lambda^{\mathrm{F}}(t'))} \qquad (15)\\
&\times\prod_{i=1}^{N}T^{\mathrm{ad}}(\boldsymbol{\varepsilon}\boldsymbol{x}_i|\boldsymbol{\varepsilon}\boldsymbol{x}_{i-1},\lambda^{\mathrm{F}}(t_i))dt_i e^{\int_{t_i}^{t_{i+1}}dt' T^{\mathrm{ad}}(\boldsymbol{\varepsilon}\boldsymbol{x}_i|\boldsymbol{\varepsilon}\boldsymbol{x}_i,\lambda^{\mathrm{F}}(t'))}.
\end{aligned}
$$

By Eq. (3) this then allows us to define

$$
\begin{aligned}
\Delta S_2&=\ln P^{\mathrm{F}}[\vec{\boldsymbol{x}}]-\ln P^{\mathrm{ad,F}}[\vec{\boldsymbol{x}}^{\mathrm{T}}]\\
&=\sum_{i=0}^{N}\ln\frac{e^{\int_{t_i}^{t_{i+1}}dt' T(\boldsymbol{x}_i|\boldsymbol{x}_i,\lambda^{\mathrm{F}}(t'))}}{e^{\int_{t_i}^{t_{i+1}}dt' T(\boldsymbol{\varepsilon}\boldsymbol{x}_i|\boldsymbol{\varepsilon}\boldsymbol{x}_i,\lambda^{\mathrm{F}}(t'))}}\\
&+\sum_{i=1}^{N}\ln\frac{P^{\mathrm{st}}(\boldsymbol{\varepsilon}\boldsymbol{x}_{i-1},\lambda^{\mathrm{F}}(t_i))}{P^{\mathrm{st}}(\boldsymbol{\varepsilon}\boldsymbol{x}_i,\lambda^{\mathrm{F}}(t_i))}\frac{T(\boldsymbol{x}_i|\boldsymbol{x}_{i-1},\lambda^{\mathrm{F}}(t_i))}{T(\boldsymbol{\varepsilon}\boldsymbol{x}_{i-1}|\boldsymbol{\varepsilon}\boldsymbol{x}_i,\lambda^{\mathrm{F}}(t_i))} \qquad (16)
\end{aligned}
$$

which similarly must obey

$$
\langle\exp[-\Delta S_2]\rangle^{\mathrm{F}}=1. \qquad (17)
$$

Unlike ΔS_1, the quantity ΔS_2 is new in the literature. We must immediately recognise that $\Delta S_{\mathrm{tot}}\neq\Delta S_1+\Delta S_2$ differing by a quantity

$$
\Delta S_3=\sum_{i=1}^{N}\ln\frac{P^{\mathrm{st}}(\boldsymbol{x}_{i-1},\lambda^{\mathrm{F}}(t_i))P^{\mathrm{st}}(\boldsymbol{\varepsilon}\boldsymbol{x}_i,\lambda^{\mathrm{F}}(t_i))}{P^{\mathrm{st}}(\boldsymbol{x}_i,\lambda^{\mathrm{F}}(t_i))P^{\mathrm{st}}(\boldsymbol{\varepsilon}\boldsymbol{x}_{i-1},\lambda^{\mathrm{F}}(t_i))} \qquad (18)
$$

such that $\Delta S_{\mathrm{tot}}=\Delta S_1+\Delta S_2+\Delta S_3$. If $\boldsymbol{\varepsilon}\boldsymbol{x}=\boldsymbol{x}$ then $\Delta S_3=0$ and ΔS_2 reduces to the adiabatic entropy production appearing in [18–20]. More importantly we must recognise that $\Delta S_{\mathrm{tot}}-\Delta S_1=\Delta S_2+\Delta S_3=\ln P^{\mathrm{ad,R}}[\vec{\boldsymbol{x}}^{\mathrm{R}}]-\ln P^{\mathrm{R}}[\vec{\boldsymbol{x}}^{\dagger}]$ or $\Delta S_{\mathrm{tot}}-\Delta S_2=\Delta S_1+\Delta S_3=\ln P^{\mathrm{ad,F}}[\vec{\boldsymbol{x}}^{\mathrm{T}}]-\ln P^{\mathrm{R}}[\vec{\boldsymbol{x}}^{\dagger}]$ cannot be written in the form required for Eq. (4) and so do not obey an IFT and do not necessarily have any bounds on the sign of their mean. We proceed by following the formalism of Seifert [12, 25] and write

$$
\Delta S_{\mathrm{tot}}=\ln\frac{P^{\mathrm{F}}(\boldsymbol{x}(0),0)}{P^{\mathrm{F}}(\boldsymbol{x}(\tau),\tau)}+\frac{\Delta Q}{T_{\mathrm{env}}}=\Delta S_{\mathrm{sys}}+\frac{\Delta Q}{T_{\mathrm{env}}}, \qquad (19)
$$

where T_{env} is the temperature of the environment, and that of Oono and Paniconi, such that total heat transfer to the environment, ΔQ, is the sum of the excess heat and house-keeping heat $\Delta Q=\Delta Q_{\mathrm{ex}}+\Delta Q_{\mathrm{hk}}$ [21]. The house-keeping heat is associated with the entropy production in stationary states and arises from a non-equilibrium constraint that breaks detailed balance. The sum $\Delta S_2+\Delta S_3$ is manifestly the entropy production in the stationary state and since we are considering Markov systems, both ΔS_2 and ΔS_3 are only non-zero when detailed balance is broken. Hence it is sensible to associate $\Delta S_2+\Delta S_3$ with the house-keeping heat such that

$$
\Delta Q_{\mathrm{hk}}=(\Delta S_2+\Delta S_3)T_{\mathrm{env}}. \qquad (20)
$$

ΔS_1 is zero for all trajectories in the stationary state consolidating the definition of the excess heat as the heat transfer associated with an entropy flow that exactly cancels the change in system entropy in the stationary state such that

$$
\Delta Q_{\mathrm{ex}}=(\Delta S_1-\Delta S_{\mathrm{sys}})T_{\mathrm{env}}. \qquad (21)
$$

However, the prevailing definition of the house-keeping heat does not make clear its properties when the system is not in a stationary state. A reported formalism suggests that it is associated with the adiabatic entropy production which serves as a general measure of the breakage of detailed balance [18–20]. When considering cases where $\boldsymbol{\varepsilon}\boldsymbol{x}=\boldsymbol{x}$, this is a consistent approach and the mean house-keeping heat obeys strict positivity requirements suggesting the entropy additively increases due to non-equilibrium constraints and a lack of detailed balance on top of that arising from relaxation. However, with the inclusion of odd variables this simple picture no longer holds, with an ambiguity illustrated by the fact that any of ΔS_2, ΔS_3 or $\Delta S_2+\Delta S_3$ could be argued to be a measure of the departure from detailed balance. In the light of Eq. (17) we propose that it is sensible to divide the house-keeping heat into two quantities which map onto ΔS_2 and ΔS_3. It is important to observe that, on average, the rate of change of ΔS_3 vanishes in the stationary state by means of balance: the path integral over an increment in ΔS_3 explicitly vanishes. Consequently we define the 'transient house-keeping heat' and the 'generalised house-keeping heat'

$$\Delta Q_{\mathrm{hk,T}} = \Delta S_3 T_{\mathrm{env}} \qquad \Delta Q_{\mathrm{hk,G}} = \Delta S_2 T_{\mathrm{env}} \tag{22}$$

such that $\Delta Q_{\mathrm{hk}} = \Delta Q_{\mathrm{hk,T}} + \Delta Q_{\mathrm{hk,G}}$. Since $\langle d\Delta S_3/d\tau\rangle^{\mathrm{F,st}} = 0$, the generalised house-keeping heat, when averaged, has the mean properties previously attributed to the house-keeping heat: it describes the heat flow required to maintain a non-equilibrium stationary state and is rigorously non-negative. Our central result therefore is

$$\langle \exp\left[-\Delta Q_{\mathrm{hk,G}}/T_{\mathrm{env}}\right]\rangle^{\mathrm{F}} = 1 \tag{23}$$

so $\langle \Delta Q_{\mathrm{hk,G}}\rangle^{\mathrm{F}} \geq 0$ for all times, protocols and initial conditions. As a corollary we also state that in general

$$\langle \exp\left[-\Delta Q_{\mathrm{hk}}/T_{\mathrm{env}}\right]\rangle^{\mathrm{F}} \neq 1 \tag{24}$$

providing no bounds on $\langle \Delta Q_{\mathrm{hk}}\rangle^{\mathrm{F}}$ except in the stationary state when $\Delta S_1 = 0$ and $\Delta Q_{\mathrm{hk}}/T_{\mathrm{env}} = \Delta S_{\mathrm{tot}}$ or generally when $P^{\mathrm{st}}(\boldsymbol{\varepsilon x}, \lambda^{\mathrm{F}}(t)) = P^{\mathrm{st}}(\boldsymbol{x}, \lambda^{\mathrm{F}}(t))$. As such the view that the mean rate of entropy production is the sum of two specific non-negative contributions as in [18–20], is incomplete. The contribution associated with a non-equilibrium constraint requires further unravelling, particularly when out of stationarity.

To explore the nature of the house-keeping heat we consider its behaviour in the approach to the stationary state of a simple model of particle dynamics on a ring. The phase space consists of L identical spatial positions $X_1, X_2 \ldots X_L$ and two velocities labelled $+$ and $-$ as shown in Fig. 1 with the time reversal properties $\boldsymbol{\varepsilon} X_i\pm = X_i\mp$ necessitated by the one-way nature of many of the transitions. The stationary state probabilities that arise from these dynamics are $P^{\mathrm{st}}(X_i+) = A/(L(A+B))$ and $P^{\mathrm{st}}(X_i-) = B/(L(A+B))$. Any difference between the velocity reversal rates A and B gives rise to a non-equilibrium stationary state by providing a stationary particle current, which for $A > B$ runs from left to right. Such dynamics amount to a very simple lattice Boltzmann model. Contributions ΔS_2 and ΔS_3 associated with particle behaviour consisting of instantaneous transitions and waiting periods are indicated. We consider particle behaviour over a small time interval dt, and compute the mean entropy production rates to leading order in dt. Examining the path probability in Eq. (2) we need only consider $N = 0$ or $N = 1$ transitions. Identifying

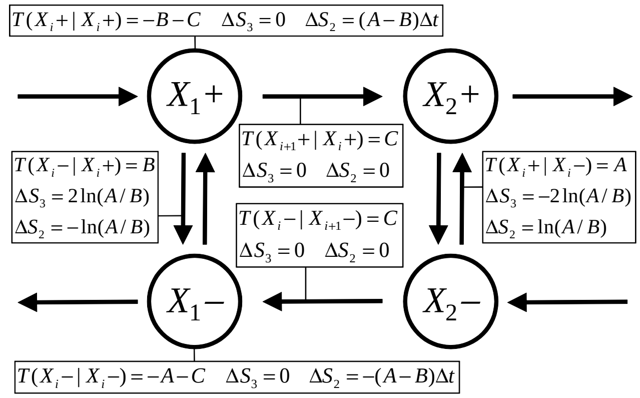

FIG. 1. Allowed moves between positions X_i and $\pm$ velocity states are shown by arrows, with associated rates T. Periodic boundaries allow jumps from X_L+ to X_1+ and X_1- to X_L-. A given path contributes to the transient and generalised house-keeping heats, $T_{\mathrm{env}}\Delta S_3$ and $T_{\mathrm{env}}\Delta S_2$, respectively, due to transitions between, and residence times Δt at, each phase space point, as indicated. These correspond to individual terms in the summations in Eqs. (16) and (18).

leading order terms in the products of P, T, exponentiated waiting times and ΔS_3 that make up the average of the form given in Eq. (4) yields

$$\frac{d\langle \Delta S_3\rangle^{\mathrm{F}}}{dt} = \sum_{i=1}^{L} 2P(X_i+)B\ln\frac{A}{B} + 2P(X_i-)A\ln\frac{B}{A}. \tag{25}$$

For non-stationary P its sign is unbounded: for example if all the probability were uniformly distributed initially amongst the $+$ velocity states it would equal $2B\ln(A/B)$, whilst if it were distributed over the $-$ states it would be $-2A\ln(A/B)$ instead. Such non-zero contributions to ΔS_3 require an asymmetric stationary state in odd variables which thus explains their absence when the stationary velocity distribution is assumed to be symmetric, such as in overdamped Langevin descriptions (see [13] and examples in [20]). However, in the stationary state with $P = P^{\mathrm{st}}$, $d\langle \Delta S_3\rangle^{\mathrm{F}}/dt$ is demonstrably equal to zero as claimed. By similar means

$$\frac{d\langle \Delta S_2\rangle^{\mathrm{F}}}{dt} = \sum_{i=1}^{L} P(X_i+)\left[A - B - B\ln\frac{A}{B}\right] + P(X_i-)\left[B - A - A\ln\frac{B}{A}\right] \tag{26}$$

which is positive for all positive A and B and reduces to $d\langle \Delta S_2\rangle^{\mathrm{F,st}}/dt = (A-B)^2/(A+B)$ in the stationary state. We note that the sum of Eqs. (25) and (26) has no bound on its sign and relates to the inequality in Eq. (24). Further, $d\langle \exp\left[-\Delta S_2\right]\rangle^{\mathrm{F}}/dt = 0$ and $\langle \exp\left[-\Delta S_2(t=0)\right]\rangle^{\mathrm{F}} = 1$ which explicitly demonstrates the expected IFT for any normalised $P(X_i\pm)$. Finally, we note that for $A = B$, all contributions vanish in detail as this corresponds to equilibrium where there is no entropy production.

We have extended the formalism found in [11, 13, 18–20] and split the total entropy production into two rigorously positive contributions and a third contribution which has no bounds on its sign. We have argued that this final quantity is, in the mean, a transient contribution to the house-keeping heat and it is the mean generalised house-keeping heat that is rigorously positive for all times. It is not straightforward to consolidate this with the two causes of time reversal asymmetry namely relaxation to the stationary state and imposed non-equilibrium constraints: ΔS_3 exists only in the presence the latter, but is, in the mean, its own measure of relaxation to the stationary state. It could be argued that the non-adiabatic entropy production and Hatano-Sasa relation do not fully capture the entropy production due to transitions between stationary states, but associating ΔS_3 with one or other form of entropy production is not entirely satisfactory as it occurs when the line between them is blurred. Nevertheless, either interpretation elucidates a new layer of complexity in the theory of entropy production in stochastic systems. Further exploration in the context of continuous stochastic processes is to be reported elsewhere [26]. The authors acknowledge financial support from EPSRC.

[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 **52**, 5839 (1995).

[3] D. J. Evans and D. J. Searles, Adv. Phys. **51**, 1529 (2002).

[4] D. Carberry, J. C. Reid, G.M.Wang, E. Sevick, D. J. Searles, and D. J. Evans, Phys. Rev. Lett. **92**, 140601 (2004).

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

[6] J. Kurchan, J. Phys. A: Math. Gen. **31**, 3719 (1998).

[7] J. L. Lebowitz and H. Spohn, J. Stat. Phys. **95**, 333 (1999).

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

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

[10] G. E. Crooks, J. Stat. Phys. **90**, 1481 (1998).

[11] T. Hatano and S. Sasa, Phys. Rev. Lett. **86**, 3463 (2001).

[12] U. Seifert, Phys. Rev. Lett. **95**, 040602 (2005).

[13] T. Speck and U. Seifert, J. Phys. A: Math. Gen. **38**, L581 (2005).

[14] V. Y. Chernyak, M. Chertkov, and C. Jarzynski, J. Stat. Mech. P08001 (2006).

[15] M. Esposito, U. Harbola, and S. Mukamel, Phys. Rev. E **76**, 031132 (2007).

[16] H. Ge, Phys. Rev. E **80**, 021137 (2009).

[17] H. Ge and H. Qian, Phys. Rev. E **81**, 051133 (2010).

[18] M. Esposito and C. Van den Broeck, Phys. Rev. Lett. **104**, 090601 (2010).

[19] M. Esposito and C. Van den Broeck, Phys. Rev. E **82**, 011143 (2010).

[20] C. Van den Broeck and M. Esposito, Phys. Rev. E **82**, 011144 (2010).

[21] Y. Oono and M. Paniconi, Prog. Theor. Phys. Suppl. **130**, 29 (1998).

[22] L. Brenig and C. Van den Broeck, Phys. Rev. A **21**, 1039 (1980).

[23] N. G. van Kampen, AIP Conference Proceedings **27**, 153 (1976).

[24] R. J. Harris and G. M. Schütz, J. Stat. Mech. P07020 (2007).

[25] Seifert, U., Eur. Phys. J. B **64**, 423 (2008).

[26] R. E. Spinney and I. J. Ford, In preparation.