Dynamics of energy fluctuations in equilibrating and driven-dissipative systems

Guy Bunin and Yariv Kafri

Department of Physics, Technion—Israel Institute of Technology, Haifa 32000, Israel
E-mail: buning@tx.technion.ac.il and kafri@physics.technion.ac.il

Received 30 October 2012, in final form 26 December 2012
Published 12 February 2013
Online at stacks.iop.org/JPhysA/46/095002

Abstract
When two isolated systems are brought into contact, they relax to equilibrium via energy exchange. In another setting, when one of the systems is driven and the other is large, the first system reaches a steady state which is not described by the Gibbs distribution. Here, we derive expressions for the size of energy fluctuations as a function of time in both settings, assuming that the process is composed of many small steps of energy exchange. In both cases, the results depend only on the average energy flows in the system and the density of states, independent of any other microscopic detail. In the steady state, we also derive an expression relating three key properties: the relaxation time of the system, the energy injection rate and the size of the fluctuations. The framework is modular and allows generalizations to other setups, such as a system coupled to two baths at different temperatures.

PACS numbers: 02.50.+r, 05.40.-a, 05.70.Ln
(Some figures may appear in colour only in the online journal)

1. Introduction
In this paper, we consider two closely related non-equilibrium problems. In the first problem, two systems which are coupled to each other but isolated otherwise are allowed to exchange energy, see figure 1(a). The systems start with arbitrary initial energies and eventually reach equilibrium. It is natural to ask: how do the initial energies evolve in time as the two systems approach equilibrium? For example, one might imagine measuring the energy of a teacup as it cools, or the equilibration of a mesoscopic system of two atomic gases, initially prepared at two different temperatures. In the second problem we consider, one of the two systems is also driven by an external protocol, see figure 1(b). This is achieved, for example, by applying a time-varying field which repeatedly returns to its initial form. When the second system is much larger than the first, it acts as a dissipative bath, and the first system eventually settles to a non-equilibrium steady state. This scenario serves as a generic model for driven-dissipative...
systems, which describe a broad range of phenomena [1–3]. Here, one can ask how the first system reaches a steady state, and what are the properties of this non-Gibbsian steady state.

As the dynamics of a system are affected by the detailed microscopic state, repeating the same experiment will lead to different outcomes. Specifically, a measurement of the energy as a function of time will yield different results, see figure 1(c). The variations between experiments might average-out in large, thermodynamic systems, or when the driving protocol applied is quasi-static. However, they can be significant when the drive is not quasi-static and in small or mesoscopic systems, which are of current experimental interest [4–6]. The dependence of these fluctuations on the dynamics makes general statements scarce, and one typically has to resort to the study of specific models.

The setup in figure 1(b) but without interaction with system 2 (so that system 1 is isolated) was recently studied in [7]. It was shown that if the drive induces many small (but still irreversible) changes in the system’s energy, then the distribution of energies at a given time depends only on the energy injection rate, and is independent of any other details of the system or drive.

Here, we quantify energy fluctuations in a much broader setting of the two canonical scenarios described above. The scenario of equilibrating systems is modeled by two isolated systems which are allowed to interact. In the scenario of a driven-dissipative system, one of the systems is driven as in [7], and the second is taken to be very large, so as to model a bath. The resulting theory is exact to order $1/N$, where $N$ is the number of degrees of freedom of the smaller system, thus being highly relevant to mesoscopic systems, where energy fluctuations may be significant. Conditions on the drive are similar to those in [7]. For the two coupled systems, we require that the energy flow between the two systems is slow, so that each of the two systems separately is close to microcanonical equilibrium at its energy (see below for the exact conditions). In addition, in the driven-dissipative case we assume that the driving and the coupling to the bath are statistically independent processes. The results are insensitive to almost all microscopic details of the systems, depending only on the average energy flows from the drive to the system and between the systems as a function of time, and on the density of states. We stress that the assumptions made do not imply that the combined system (composed of systems 1 and 2) is close to equilibrium, but only that each of the systems separately is close to equilibrium within its energy shell. Our main results are (1) equations (3) and (4), which quantify the variance of the energy fluctuations as the system approaches its steady state.
state (which is equilibrium when no drive is present), (2) equation (5), which relates three main quantities at the steady state: the variance of the energy fluctuations, the average rate of energy flow through the system and the relaxation time of energy fluctuations. In addition, we show that the approach is modular, and may be applied to other settings. As an example, we study a system connected to two external baths. The validity of the results and their use is illustrated in a system of colliding hard spheres and an interacting spin system.

Before turning to a systematic derivation, we first sketch the approach used. To understand the energy exchange behavior, note that the assumptions of weak interaction and slow energy exchange between the systems lead to processes comprised of many small independent energy exchanges. This, much like a particle in water, leads to a diffusive (Langevin) behavior; in our case the diffusion is in the energy plane (see section 3 for a systematic derivation). The state of the combined system is specified by a point in the \( (E_1, E_2) \) plane, and the diffusion takes place in this plane. However, in both scenarios, of equilibrating systems and driven-dissipative systems, one can specify the state by considering \( E_1 \) alone. For equilibrating systems, this is possible when the initial total energy \( E_{\text{total}} = E_1 + E_2 \) is fixed, so that \( E_2 \) can be considered to be a function of \( E_1 \). In the case of driven-dissipative systems, \( E_2 \) drops completely from the equations when we take system 2 to be much larger than system 1. This is because system 2 acts as a thermal bath whose properties are insensitive to the changes in \( E_2 \) (see section 3 for details).

One can therefore consider the probability distribution \( P(E_1) \) of \( E_1 \), which will satisfy a Fokker–Planck (FP) (drift-diffusion) equation

\[
\partial_t P = -\partial_{E_1}(A_1 P) + \frac{1}{2}\partial_{E_1}^2(B_{11} P),
\]

where \( A_1(E_1), B_{11}(E_1) \) are functions which depend on system details. They do, however, obey the following relation, which is the key to the results which are described below:

\[
2A_1 - 2\beta_2/\beta_1 A_F = (\beta_1 - \beta_2)B_{11},
\]

where \( A_F \equiv A_1 + A_2 \) is the rate of energy injected into the system by the drive, \( \beta_i(E_i) = \partial_{E_i}S_i \), for \( i = 1, 2 \) are the (microcanonical) inverse temperatures of system \( i \) at energy \( E_i \) and \( S_i(E_i) \) is the entropy of system \( i \). Note that \( \beta_1 \) and \( \beta_2 \) are well-defined functions, depending only on the density of states of the system, and unrelated to the driving mechanism and the interaction between the systems. Moreover, \( \beta_1 \) can be very different from \( \beta_2 \), so the combined system may be arbitrarily far from equilibrium. When system 2 is a bath, \( A_2 \) is only a function of \( E_1 \). In the case of equilibrating systems \( A_F = 0 \), and equation (2) reduces to \( 2A_1 = (\beta_1 - \beta_2)B_{11} \).

Equation (2) is exact up to corrections of order \( 1/N \), and requires that the energy transfers be slow compared with the relaxation times of each of the systems separately.

We comment that formally slowness enters as a set of conditions on the smallness of higher order cumulant of the work during the relaxation time of the system. For processes which involve an external drive, let \( \Delta E_F \) be the size of the random jump in energy associated with this process. As discussed in the derivation section, we require that \( \beta_1^2 \langle (\Delta E_F)^3 \rangle \ll \langle \Delta E_F \rangle \). Similarly, let \( \Delta E_B \) be the amount of energy exchanged between the systems in a single jump. Then we require \( (\beta_1 - \beta_2)^2 \langle \Delta E_B^3 \rangle \ll \langle \Delta E_B \rangle \). These conditions can also be used to reason for the validity of the FP equation. By the independence of the steps, the third cumulant of the total energy is expected to grow linearly with the number of steps \( N_{\text{steps}} \), and therefore its effect on the energy distribution grows as \( N_{\text{steps}}^{1/3} \), and is therefore small. For derivations of FP equations along this line of reasoning see, for example, [12]. With equations (1) and (2), we derive all the results described below.

The paper is organized as follows. In section 2, we summarize the main results of this work. In section 3, we derive these results, which follow from equation (2) in the two scenarios
of equilibrating and driven-dissipative systems. In section 4, we apply the main results to two example systems.

2. Results

2.1. Approach to the steady state

We start by considering the approach of the combined system (composed of systems 1 and 2) to its steady state. If no driving is present (scenario 1), then this steady state is thermal equilibrium. We derive an expression for the evolution of the variance $\sigma_1^2 = \langle E_1^2 \rangle - \langle E_1 \rangle^2$ during the entire equilibration process. The result is derived in section 3 by solving for the evolution of the first two moments of equation (1), and using the key relation, equation (2).

We find for the equilibrating systems that the variance is given by

$$\sigma_{1,eq}^2 (\langle E_1 \rangle) = \sigma_{1,0}^2 \frac{A_1^2 (\langle E_1 \rangle)}{A_1^2 (\langle E_1 \rangle_0)} + 2A_1^2 (\langle E_1 \rangle) \int_{\langle E_1 \rangle_0}^{\langle E_1 \rangle} \frac{1}{A_1^2 (E') (\beta_1 - \beta_2)} dE'. \quad (3)$$

Here, $\langle E_1 \rangle_0$ and $\sigma_{1,0}^2$ are $\langle E_1 \rangle$ and $\sigma_1^2$, respectively, at the initial time. Recall that $E_{total} = E_1 + E_2$ is held constant in this expression. It is easy to extend these results when $E_{total}$ varies between experiments. It is interesting to note that when $\beta_2$ is set to zero, this expression is identical to that obtained for a single-driven isolated system [7]. This means that within this theory, driving a system is formally equivalent to attaching it to a bath with infinite temperature. It is straightforward to show, that when system 2 is a bath, so that $\beta_2$ can be taken to be a constant, the width $\sigma_1^2$ approaches the equilibrium value: $(\theta_{E_1, \beta_1})_{E_{eq}}^{-1} = k_B T^2 C$, where $E_{eq}$ is the equilibrium value of $\langle E_1 \rangle$ and $C$ is the heat capacity (see e.g., [13]). To see this, note that at equilibrium $A_1$ must vanish, and $\beta_1 = \beta_2$. Therefore, the entire expression for $\sigma_1^2$ is controlled by the final approach of $E$ to $E_{eq}$ where $\beta_1 \approx \beta_2 + (\beta_2 - \beta_1) E_{eq} / (E_{eq} - E_{eq})$. $A_1(E') = \frac{dA_1}{dT} \int_{E_{eq}}^{E_{eq}} (E - E_{eq})$ and the equilibrium expression follows. Note that away from the final equilibration regime $A_1(E)$ need not be linear.

In the case of driven-dissipative systems (when system 2 is large), we obtain for the variance

$$\sigma_{1,dr}^2 (\langle E_1 \rangle) = \sigma_{1,0}^2 \frac{A_1^2 (\langle E_1 \rangle)}{A_1^2 (\langle E_1 \rangle_0)} + 2A_1^2 (\langle E_1 \rangle) \int_{E_{10}}^{\langle E_1 \rangle} \frac{Z (E')}{A_1^2 (E')} dE', \quad (4)$$

where $Z = [1 - \beta_2 A_F / (\beta_1 A_1)] / (\beta_1 - \beta_2)$. Equations (3) and (4) are our main results for the approach to the steady state. They predict the size of fluctuations in $E_1$ around its average value. They depend only on the rates of energy injection into the system $A_F$ (which is zero for equilibrating systems) and the rate of energy transfer to the bath $A_2$. In principle, both these quantities can be measured separately: as a consequence of the independence of the drive and interaction between the systems, $A_F$ can be measured by the rate of energy absorption when system 1 is isolated, and $A_2$ in an equilibration experiment without the drive.

2.2. Steady-state fluctuations

The framework described above can also be used to study fluctuations in the steady state of driven-dissipative systems, specifically fluctuations of $E_1$ around $\langle E_1 \rangle$. Typical fluctuations around the steady state are expected to decay exponentially (see section 3), $(e_1(t_1) e_1(t_2)) = \frac{\beta_1}{2} e^{-|t_2 - t_1|/\tau}$ where $e_1 \equiv E_1 - \langle E_1 \rangle$. $\tau$ is the relaxation time to the steady state. When $A_F \neq 0$, namely for a driven-dissipative system, we find that

$$\tau A_F = \frac{\beta_1}{\beta_2} (\beta_2 - \beta_1) \sigma_1^2. \quad (5)$$
This is our main result for the steady state of driven-dissipative systems. \( A_F \) is the rate of energy injected into the system from the drive. In the steady state, this energy is then dissipated into the bath. This expression therefore relates three central quantities characterizing the steady state: the size of the energy fluctuations \( \sigma_1^2 \), the rate of energy dissipation \( A_F \) and \( \tau \) which is the relaxation time in the steady state.

2.3. Modularity

The above results can be easily generalized to analyzing other settings. For example, consider a setup where a system is attached to two baths, at temperatures \( \beta_2 \) and \( \beta_3 \), see figure 2. The system approaches a steady state. Here one can show that the size of the variance during the approach is given by equation (4) now with \( Z(E_1) \) given by

\[
Z(E_1) = \frac{A_2}{\beta_1 - \beta_2} + \frac{A_3}{\beta_3 - \beta_1},
\]

where \( A_2(E_1) \) and \( A_3(E_1) \) are the average energy transfer rates and \( \beta_1(E_1) \) is the microcanonical temperature of the system at \( E_1 \). For the steady state, one obtains

\[
\tau A_F = \frac{(\beta_1 - \beta_2)(\beta_3 - \beta_1)}{\beta_3 - \beta_2} \sigma_1^2,
\]

where \( A_F \) is the average rate of energy passing from one bath, through the system, to the other bath. Other more elaborate setups may be constructed similarly.

3. Derivation

To derive the results presented in section 2, we consider the evolution of the energies in the \((E_1, E_2)\) plane, where \( E_1 \) and \( E_2 \) are the energies of systems 1 and 2, respectively. Consider a series of changes in \( E_1 \) and \( E_2 \), each taking place over a time interval \( \Delta t \). The changes in \( E_1 \) and \( E_2 \) result from the interaction between the systems and from the drive. The drive is implemented by varying a parameter \( \lambda \) in the Hamiltonian which repeatedly returns to some baseline value after a time interval \( \Delta t \). As in [7] we assume that the energy changes due to the drive in the time interval \( \Delta t \) are small, in a sense specified below. In addition, we assume that \( \tau_R \ll \Delta t \), where \( \tau_R \) is the relaxation time of each of the isolated systems separately, and without any external drive. Thus, after time \( \Delta t \) many typical relaxation times \( \tau_R \) have passed, and the system has forgotten all information about its state except its total energy (which is assumed to be the only conserved quantity), and is hence close to microcanonical equilibrium. The energy changes \( \Delta E_1, \Delta E_2 \) during non-overlapping time segments of duration \( \Delta t \) are then independent, and depend only on the total energy in that time.

The changes in \( \Delta E_1, \Delta E_2 \) must obey certain relations. These relations follow, ultimately, from Liouville’s equation, or the unitarity of the dynamics in quantum cases. Consider first scenario (b) in figure 1, but with no interaction between systems 1 and 2. The drive consists of many small perturbations to system 1, where the Hamiltonian is varied and repeatedly
returns to its original form. In [7] it was shown, that in this case, the change in energy of
the system $\Delta E_F$ when the drive returns to its original state must obey $\langle \exp (-\beta_1 \Delta E_F) \rangle = 1$. Here, $\beta_i (E_i) = \partial E_i / \partial S_i$, for $i = 1, 2$ are the (microcanonical) inverse temperatures of system $i$ at
energy $E_i$. $S_i$ is the entropy of system $i$ and the angular brackets denote averaging over the
initial microscopic states with energy $E_i$. This relation is essentially a Jarzynski relation for
isolated systems. Assuming that the changes are small [7] (for a different approach see [10]),
we take the logarithm $\ln \langle \exp (-\beta_1 \Delta E_F) \rangle = 0$ and expand to second order to find

$$2 \langle \Delta E_F \rangle = \beta_1 \left( \langle \Delta E_F^2 \rangle - \langle \Delta E_F \rangle^2 \right) + O (1/N).$$

(8)

where $N$ is the number of degrees of freedom. This holds when $\beta_1^2 \langle \Delta E_F^3 \rangle \ll \langle \Delta E_F \rangle$ and
$\beta_1 \langle \Delta E_F \rangle \ll C_i$, where $C_i = \partial E_i / \partial S_i$ is the extensive specific heat. The first condition
ensures that there are no corrections from the third cumulant in the expansion, the second
relation ensures that $\beta_1$ does not change significantly in one jump and is usually a much
weaker condition, and ensures the validity of the Jarzynski relation. These conditions and
others cited below are derived in appendix A.

We expand on this result by using a variant of the exchange fluctuation relation [11]
which describes the energy flow between two systems. For a weak interaction between
the two systems, typically corresponding to slow exchange of energy between them, the
entropy of the combined system is additive $S(E_1, E_2) = S_1(E_1) + S_2(E_2)$. In this case
$\langle \exp (- (\beta_2 - \beta_1) \Delta E_B) \rangle = 1$, where $\beta_1 (E_1), \beta_2 (E_2)$ are the microcanonical temperatures of the
(otherwise isolated) systems, and $\Delta E_B$ is the amount of energy transferred from system
1 to system 2. Under small work assumptions, we expand $\ln \langle \exp (- (\beta_2 - \beta_1) \Delta E_B) \rangle = 0$ to obtain

$$2 \langle \Delta E_B \rangle = (\beta_2 - \beta_1) \left( \langle \Delta E_B^2 \rangle - \langle \Delta E_B \rangle^2 \right) + O (1/N).$$

(9)

This holds when $(\beta_2 - \beta_1)^2 \langle \Delta E_B^3 \rangle \ll \langle \Delta E_B \rangle$, and $\langle \Delta E_B^2 \rangle \ll \min \left( \beta_2^2 C_i, \beta_1^2 C_i \right)$. This is the key relation which will be used below to describe the evolution of fluctuations in
equilibrating systems (scenario (a) in figure 1).

In the scenario of a driven-dissipative system, we combine these two results. We take
system 2 to be large, so as to model a bath, and assume further that the driving mechanism and
the interaction with the bath produce independent changes in $E_1$, e.g., when the drive and
interaction processes act on different modes, on different parts of the systems, etc. Then the
two relations (8) and (9) may be combined. The change in $E_1$ is $\Delta E_1 = \Delta E_F - \Delta E_B$ and the
change in $E_2$ is given by $\Delta E_2 = \Delta E_B$. We define for $i, j = 1, 2$,

$$A_i = \langle \Delta E_i \rangle / \Delta t, \quad B_{ij} = \langle \Delta E_i \Delta E_j \rangle - \langle \Delta E_i \rangle \langle \Delta E_j \rangle / \Delta t,$$

where $A_i$ is simply the average energy change per unit time in system $i$. Note that $A_i, B_{ij}$ are functions of $(E_1, E_2)$.

The assumption of independence of $\Delta E_F, \Delta E_B$ gives $\langle \Delta E_F^2 \rangle = \langle \Delta E_B^2 \rangle + \langle \Delta E_B^2 \rangle$, which combined with equations (8) and (9) gives equation (2), where $A_F \equiv A_1 + A_2$ is the rate of energy injected into the system by the drive. As are equations (8) and (9), this result is exact up to corrections of order $1/N$, and requires $\beta_1^2 \langle \Delta E_F \rangle \langle \Delta E_B \rangle \ll C_i$. In the case of equilibrating systems, $A_F = 0$ and equation (2) reduces to $2A_1 = (\beta_1 - \beta_2) B_{11}$.

Under conditions of many small independent jumps, and much like when a Brownian
particle is displaced by many small kicks, the jumps in the $(E_1, E_2)$ result in an FP equation.
The FP equation is a drift-diffusion for the probability distribution $P_{12}(E_1, E_2)$, which reads

$$\partial_t P_{12} = -\sum_{i=1}^2 \partial E_i (A_i P_{12}) + \frac{1}{2} \sum_{i,j=1}^2 \partial E_i \partial E_j (B_{ij} P_{12}),$$

(10)
where \( A_i, B_j \) for \( i, j = 1, 2 \) are precisely the functions of \( \langle E_1, E_2 \rangle \) described above. Equation (10) is valid under the same above conditions on the third cumulants of \( \Delta E_1, \Delta E_2 \) [12]. It is then more convenient to work with the marginal probability distribution of \( E_i \): 

\[
P_i(E_i) \equiv \int dE_2 P_{12}(E_1, E_2).
\]

As in the different scenarios only \( E_1 \) is needed, we can integrate equation (10) over \( E_2 \) to obtain equation (1).

We derive an expression for the evolution of the variance \( \sigma^2_i = \langle E_i^2 \rangle - \langle E_i \rangle^2 \) during the entire equilibration process. Proceeding similarly to [7], we take the first two moments with respect to \( E_1 \) of equation (1):

\[
\begin{align*}
\partial_t \langle E_i \rangle &= \langle A_i \rangle, \\
\partial_t \sigma^2_i &= \langle B_{11} \rangle + 2 \langle \langle A_1 E_1 \rangle - \langle A_1 \rangle \langle E_1 \rangle \rangle.
\end{align*}
\]

The assumptions on the smallness of \( \Delta E_i \) prescribed above imply that \( \Delta E_i \) can scale at most as \( O(N) \); therefore, \( A_1 = \langle \Delta E_i \rangle / \Delta t \) grows as \( O(N) \) at most, and using equation (2), we see that \( \sigma^2_i \) grows as \( O(N) \) at most; hence the distributions are narrow, their widths’ scaling subextensively, at most as \( N^{1/2} \). When \( N \) scales as \( N \) the width scales as at equilibrium.

Therefore, up to corrections of order \( 1/N \), \( \langle A_i \rangle \) can be taken to depend on \( \langle E_i \rangle \) alone, and the change of \( \langle E_i \rangle \) in time will be monotonic (since when \( A \) changes sign, \( \partial_t \langle E_i \rangle = 0 \) and the change in \( \langle E_i \rangle \) stops). Combining the two equalities in equation (11) and linearizing \( A_1 \) within the width of the probability distribution, we find

\[
\frac{\partial \sigma^2_i}{\partial \langle E_i \rangle} = 2Z_1(\langle E_i \rangle) + 2 \frac{\partial_\sigma_1 A_1(\langle E_i \rangle)}{\langle A_1 \rangle} \sigma^2_i(\langle E_i \rangle),
\]

where \( Z_1(\langle E_i \rangle) \equiv B_{11}/(2A_1) \). Solving the ordinary differential equation (12) and using equation (2), we find for the equilibrating systems that the variance is given by equation (3).

### 3.1. Steady-state fluctuations

At the steady state the probability distribution \( P_s(\langle E_i \rangle) \) is independent of time. As \( \partial_t \langle E_i \rangle = \langle A_1 \rangle = A \langle \langle E_i \rangle \rangle \), at the steady state \( \langle \langle E_i \rangle \rangle \) must vanish. Using the narrowness of the distribution (recall that the width scales at most as \( N^{1/2} \)), we expand \( A_1 \) and \( B_{11} \) to lowest order in \( \epsilon_i = E_1 - E_i^0 \):

\[
A_1 = -\frac{1}{\tau} \epsilon_1, \quad B_{11} = B_s,
\]

where \( B_s \) and \( \tau \) are constants. Equivalently, in this regime the FP equation describes the Brownian motion of the energy in a harmonic potential \( \epsilon_i = -\epsilon_1 / \tau + \sqrt{\beta_1} \eta \), where the white noise \( \eta(t) \) satisfies \( \langle \eta(t) \eta(t') \rangle = \delta(t - t') \). \( \tau \) is then interpreted as the relaxation time, as can be seen from the two time correlation functions \( \langle \epsilon_1(t_1) \epsilon_1(t_2) \rangle = \frac{\beta_1}{\tau} e^{-|t_2-t_1|/\tau} \). The variance of the energy fluctuations is given by \( \sigma^2_i = \langle \epsilon_1(t)^2 \rangle = B_s \tau / 2 \).

When \( A_F = 0 \), equations (2) and (13) imply that \( \epsilon_i = -\frac{\epsilon_1}{\tau} (\beta_1 - \beta_2) \). Then expanding \( \beta_1 \) around \( \beta_2 \) as done above, we find that \( \sigma^2_i = B_s \tau / 2 = -\langle \partial_\sigma \beta_1 \rangle \) which again reproduces the canonical distribution width. The present derivation gives a dynamic interpretation to this formula.

When \( A_F \neq 0 \) namely for a driven-dissipative system we find, using \( A_1 \langle E_i^0 \rangle = 0 \) in equation (2), and \( \sigma_i^2 = B_s \tau / 2 \), we obtain equation (5).

### 4. Applications

We now present two examples of systems where the above framework applies. The first is a gas of hard spheres, with two different particle types which form the two
systems, or the system and the bath, see figure 3(a). The second system is a chain of interacting planar spins (the XY-model), attached to two baths at different temperatures, see figure 3(b).

4.1. Hard sphere system—molecular dynamics simulations

We now illustrate our main results on a gas of hard-sphere particles in a box, simulated by an event-driven molecular dynamics simulation [8]. The simulations demonstrate the validity of the theory for fully deterministic dynamics (apart from bath particles, where a bath is present). In addition, the simulations involve relatively small systems, with just few tens of particles, thus demonstrating the applicability of the predictions to small systems. The systems are completely mixed, making this setup more challenging.

The hard-sphere gas is composed of $N_1$ particles of mass $m_1$ and $N_2$ particles of mass $m_2$, all of equal size, corresponding to systems 1 and 2 respectively, see figure 3(a). Although the entropy of the two systems between collisions indeed factorizes, the collision process involves a strong interaction, which changes the velocities of the particles by a significant amount. A collision calculation shows that if the two masses are very different, then the energy transfer in each collision is small. In this case, energy transfer occurs over many collisions, fulfilling the assumption of timescale separation (see above). In what follows, we take $m_1 = 10^{-4}, m_2 = 1$. (Throughout we use arbitrary units.) The large ratio of masses allows us to demonstrate the validity of the results on a single set of system parameters, for a wide range of driving strengths. The results still hold for a smaller mass ratios, albeit in a narrower range of validity. The box is a unit cube with reflecting boundary conditions, and the particles are taken to occupy a volume fraction of 0.05.
We first consider the approach to equilibrium of two systems in contact, figure 1(a), to be compared with the predictions of equation (3). We take $N_1 = 30$ for the first system and $N_2 = 20$ for the second system. $N_1$ and $N_2$ are chosen to be relatively small in order to test the theory on a mesoscopic system. The initial velocities are sampled from a Maxwell–Boltzmann distribution with $\beta_1 = 60$ and $\beta_2 = 3$, corresponding to average energies per particle of $\langle E_1 \rangle / N_1 = 0.025$ and $\langle E_2 \rangle / N_2 = 0.5$. We start all runs from a fixed total energy $E_{\text{total}} = \langle E_1 \rangle + \langle E_2 \rangle$, by performing a (small) rescaling of the $m_2$-particles’ velocities. Gathering statistics over many runs, we calculate at each time the average energy $\langle E_1 \rangle (t)$ and the variance $\sigma_1^2 (t)$. The function $A_1 (\langle E_1 \rangle)$ is obtained by plotting $A_1 (t) = d \langle E_1 \rangle / dt$ as a function of $\langle E_1 \rangle (t)$. Given $A_1 (\langle E_1 \rangle)^1$, we use equation (3) to predict $\sigma_1^2 (\langle E_1 \rangle)$, and find a good fit with the simulation results, see figures 4(a) and (b).

We now turn to the driven-dissipative scenario and test equation (5) for the steady state. We run simulations on a system with $N_1 = 10$ and $N_2 = 50$ particles. System 1 is driven by applying short impulses to the $m_1$-particles, changing their velocity by $\Delta v = F \delta t / m_1$, where $F$ is a constant force and $\delta t$ is the impulse duration. This impulse is applied at a constant rate. In order to mimic the behavior of a very large system 2, the velocity of the $m_2$-particles is changed upon reflection from the wall [9], so as to maintain a constant $\langle E_2 \rangle$. The quantities $\tau, A_F, \beta_1$ and $\sigma_1^2$ are computed from the numerics. The energy of system 2 is maintained at $\langle E_2 \rangle / N_2 = 1/2$, or $\beta_M = 3$. Figure 4(c) shows the results obtained for the two sides of equation (5) as a function of $\beta_1 / \beta_2$ for different strengths of the drive. Good agreement is found over a wide range of drive strengths and temperature differences. In this range, $A_F$ increases by a factor of 1000, $\sigma_1^2$ by a factor of 17 and the relaxation time $\tau$ decreases by a factor of 3.

1 Incidentally, $\langle E_1 \rangle (t)$ is found to fit very well to a function of the form $E = E_{\text{int}} \tanh^2 (a + bt)$, or $A_1 \propto \langle E \rangle^{1/2} (\langle E \rangle - E_{\text{int}})$, allowing us to write a closed analytical expression for $\sigma_1^2 (\langle E \rangle)$. This can be understood by calculating the rate of energy transfer, taking the collision rate to depend only on the velocity of the lighter, and hence faster, $m_1$ particles.
4.2. XY spin chain

In the second example, we consider a one-dimensional XY model. In this model, a two-dimensional vector is associated with each site. The interaction between neighboring sites \( i \) and \( j \) is given by \( H_{ij} = 1 - \cos(\theta_i - \theta_j) \), where \( \theta_i \) is the angle of the vector at site \( i \). The total Hamiltonian is \( H = \sum_{i,j} H_{ij} \), running over all neighboring pairs; in the case of a linear spin chain with \( N \) spins, this becomes \( H = \sum_{i=1}^{N-1} \cos z_i \), where \( z_i = \theta_{i+1} - \theta_i \).

We attach the system to baths at both ends of the system, with inverse temperatures \( \beta_2 \) and \( \beta_3 \), and demonstrate the derivation of equation (7) in this case. A bath is modeled by adding an interaction to the leftmost (and similarly for the rightmost) spin \( \theta_1 \), which applies random changes to the angle of that spin. The changes in the angle \( \theta_1 \) are taken in the range \(-\alpha \leq \delta \theta_1 \leq \alpha \), with \( \alpha \ll 1 \), and at rates

\[
    w(\delta \theta_1) = w_0 e^{-\beta_2 \Delta(\delta \theta_1)/2} e^{-\beta_2 \Delta(-\delta \theta_1)/2},
\]

where \( \Delta(\delta \theta) = -[\cos (z_1 + \delta \theta) - \cos z_1] \) is the change in energy due to the change in \( \theta_1 \) and \( w_0 \) is a constant setting the overall timescale. This interaction functions as a bath with temperature \( \beta_2 \). The last spin, \( \theta_N \), is attached to a bath at temperature \( \beta_3 \), by applying a similar protocol with \( \beta_2 \) replaced by \( \beta_3 \).

We now calculate the first two moments of the energy transfer from the bath to the system. Note that the steady-state energy and fluctuations will be given by the combined effect of the two baths. We now use the assumptions described above, namely that the energy exchanges with the bath are slow compared with the relaxation time of the system. Here this enters by assuming that \( \alpha \ll 1 \) at a fixed \( w_0 \). As a consequence, the system at a given time is approximately in a microcanonical equilibrium. This means that up to \( 1/N \) corrections \( z_1 \) is distributed with a Boltzmann weight: \( P(z_1) \propto \exp(\beta_1 \cos z_1) \), where \( \beta_1 = \partial_\Omega S_1 \) is the microcanonical temperature of system 1. The average change in the energy is given by

\[
    A_{12}(E_1) = \int_{-\alpha}^{\alpha} \frac{d(\delta \theta_1)}{2\pi} \int_0^{2\pi} d\theta_2 \; P(z_1) \; w(\delta \theta_1) \; \delta e(\delta \theta_1)
    = \frac{\alpha^2}{3} I_1(\beta_1) \frac{\beta_1 - \beta_2}{\beta_1} + O(\alpha^5)
\]

\[
    B_{12}(E_1) = \int_{-\alpha}^{\alpha} \frac{d(\delta \theta_1)}{2\pi} \int_0^{2\pi} d\theta_2 \; P(z_1) \; w(\delta \theta_1) [\delta e(\delta \theta_1)]^2
    = \frac{A_{12}}{\beta_1 - \beta_2} + O(\alpha^5),
\]

where \( I_n(\beta) \) is the modified Bessel of the first kind. The dependence on energy enters through the microcanonical energy–temperature relation: \( E_1 = -N \partial_\epsilon F(\beta_1) \), \( A_{12}, B_{12} \) satisfy \( 2A_{12} = (\beta_1 - \beta_2) B_{12} \), which is precisely equation (2) for this system (as \( A_F = 0 \) for two coupled systems without any drive). Similarly, if the spin \( \theta_N \) is attached to a bath at temperature \( \beta_3 \), then \( 2A_{13} = (\beta_1 - \beta_3) B_{13} \). The last two relations are valid more generally, as discussed in previous sections, and encode all the information on the system that is required to derive equations (6) and (7). For example, to obtain equation (7), we attach the system to two baths at temperatures \( \beta_2, \beta_3 \) acting independently at both sides of the chain. Then (using the conventions of figure 2), \( A_1 = A_{12} - A_{13} \) and \( B_1 = B_{12} + B_{13} \). At the steady state one has (see section 3) \( \sigma_1^2 = B_1 \tau/2 \), which together with the relations \( 2A_{12} = (\beta_1 - \beta_2) B_{12} \), and \( 2A_{13} = (\beta_1 - \beta_3) B_{13} \) gives equation (7).
5. Conclusions

In this paper, we considered several canonical non-equilibrium scenarios, in which energy is transferred between two or more systems, or as work from a drive. We find a specific regime where universal statements on the energy fluctuations during relaxation to the steady state and at the steady state can be made. The results are applicable to systems in various fields. The framework can be extended to other more elaborate settings, such as ones involving a larger number of systems.

Acknowledgments

We are grateful to Luca D’Alessio, Dov Levine, Daniel Podolsky, Anatoli Polkovnikov and Yair Shokef for many useful discussions and comments. The work was supported by a BSF grant. YK thanks G M Schutz.

Appendix

A.I. Derivation of equation (3)

In what follows, we first derive equation (3) of the main text, and then proceed to carefully analyze the conditions for its validity. To derive equation (3) we look at two times \( t, t + \Delta t \) at which the driving protocol has returned to its original state, i.e. where \( H(t) = H(t + \Delta t) \), where \( H \) is the Hamiltonian of the combined system. As discussed in section 1 ‘framework and assumptions’, we assume that both subsystems are relaxed in their respective energy shells, with energies \( E_1 \) and \( E_2 \), at times \( t \) and \( t + \Delta t \). We denote the changes in \( E_1, E_2 \) during the time interval \( \Delta t \) by \( \Delta E_1, \Delta E_2 \) respectively, and define \( \Delta E \) and \( \Delta F \) via

\[
\Delta E_1 = \Delta E_F - \Delta E_B, \quad \Delta E_2 = -\Delta E_B.
\]

\( \Delta E_B \) is the energy transferred from system 1 to system 2, and \( \Delta E_F \) is the work done on system 1 by the external drive.

Under these conditions, Liouville’s theorem or the unitarity of the dynamics, together with micro-reversibility of the dynamics, imply a Crooks relation for the combined isolated system [14–16, 7]:

\[
P_{E_1, E_2} (\Delta E_1, \Delta E_2) e^{S_1(E_1) + S_2(E_2)} = \bar{P}_{E_1 + \Delta E_1, E_2 + \Delta E_2} (\Delta E_1, -\Delta E_2) e^{S_1(E_1 + \Delta E_1) + S_2(E_2 + \Delta E_2)},
\]

where \( P_{E_1, E_2} (\Delta E_1, \Delta E_2) \) is the probability of a transition \( (E_1, E_2) \to (E_1 + \Delta E_1, E_2 + \Delta E_2) \), and \( \bar{P} \) is defined similarly, only with respect to the reversed protocol, defined by the dynamics generated by the time-reversed Hamiltonian \( \bar{H}(t') = H(t + \Delta t - t') \). Here, we have used the assumption of the additivity on the entropy, valid for weak interactions, \( S(E_1, E_2) = S_1(E_1) + S_2(E_2) \), see the main text.

We approximate \( \bar{P}_{E_1 + \Delta E_1, E_2 + \Delta E_2} \approx P_{E_1, E_2} \); this introduces errors which are of order \( 1/N \), as shown below. By the assumption of independence of the driving and interaction mechanisms, \( \Delta E_B \) and \( \Delta E_F \) are statistically independent, and we can write

\[
P_{E_1, E_2} (\Delta E_1, \Delta E_2) = P^F_{E_1, E_2} (\Delta E_F, 0) P^B_{E_1, E_2} (\Delta E_B, \Delta E_F).
\]

The Crooks relation then reads

\[
P^F_{E_1, E_2} (\Delta E_F, 0) P^B_{E_1, E_2} (\Delta E_B, \Delta E_F) e^{\beta_1 \Delta E_F - (\beta_2 - \beta_1) \Delta E_B} = \bar{P}^F_{E_1, E_2} (\Delta E_F, 0) \bar{P}^B_{E_1, E_2} (\Delta E_B, \Delta E_F).
\]

In the quantum mechanical setting, we assume that the initial and final density matrices are diagonal, as in [7], and that there is no significant entanglement between the two systems.
Integrating over $\Delta E_F$, $\Delta E_B$ gives a Jarzynski relation
\[ \langle e^{-\beta_i \Delta E_F} \rangle \langle e^{-(\beta_i-\beta_o)\Delta E_B} \rangle = 1. \quad (A.1) \]
Rearranging the equation, noting that $P^B (-\Delta E_B, \Delta E_B) = P^B (\Delta E_B, -\Delta E_B)$ and integrating gives
\[ \langle e^{-\beta_i \Delta E_F} \rangle = \langle e^{-(\beta_i-\beta_o)\Delta E_B} \rangle. \quad (A.2) \]
Together the two relations yield
\[ \langle e^{-\beta_i \Delta E_F} \rangle = \langle e^{-(\beta_i-\beta_o)\Delta E_B} \rangle = 1. \]
Taking the logarithm of these relations and expanding to second order gives relations (1) and (2) of the main text.

A2. Conditions for validity of equation (3)

We now discuss the conditions for the validity of the above derivation. In it we have made the following approximations. First, we have assumed that $P_{E_i+\Delta E_i,E_i+\Delta E_i} \simeq P_{E_i,E_i}$. To find the next correction, we take $P_{E_i+\Delta E_i,E_i+\Delta E_i} \simeq P_{E_i,E_i} - \Delta E_i \partial E_i P_{E_i,E_i} = -\Delta E_i \partial E_i P_{E_i,E_i}$. Following the above derivation, one obtains equation (1) together with the first correction
\[ 2\langle \Delta E_F \rangle = (\beta_1 + \partial E_i) \langle \Delta E_F^2 \rangle, \]
where $\langle \Delta E_F^2 \rangle = \langle \Delta E_F^2 \rangle - \langle \Delta E_F \rangle^2$ is the second cumulant of $\Delta E_F$. As $\beta_1$ is an intensive and $E_i$ extensive, the term $\partial E_i \langle \Delta E_F^2 \rangle$ introduces corrections of order $1/N$ which we neglect. Similarly, equation (2) becomes
\[ 2\langle \Delta E_B \rangle = (\beta_2 - \beta_1 + \partial E_i - \partial E_i) \langle \Delta E_B^2 \rangle, \]
and the corrections to equation (2) are of order $1/N$ for the same reason.

The second approximation involves taking $S_i (E_i + \Delta E_i) - S_i (E_i) \simeq \beta_1 \Delta E_i$. Again, we incorporate the next correction, $S_i (E_i + \Delta E_i) - S_i (E_i) \simeq \beta_1 \Delta E_i + \frac{1}{2} S''_i (E_i) \langle \Delta E_i^2 \rangle$. To this order equations (A.1) and (A.2) read
\[ \langle e^{-\beta_i \Delta E_F + \frac{1}{2} S_i' \Delta E_i^2} \rangle \langle e^{-(\beta_i-\beta_o)\Delta E_B + \frac{1}{2} (S_i' + S_i) \Delta E_B^2} \rangle = 1 + S''_i \langle \Delta E_F \rangle \langle \Delta E_B \rangle, \]
\[ \langle e^{-\beta_i \Delta E_F + \frac{1}{2} S_i' \Delta E_i^2} \rangle \langle 1 + S''_i \langle \Delta E_F \rangle \langle \Delta E_B \rangle \rangle = \langle e^{-(\beta_i-\beta_o)\Delta E_B} \rangle, \]
respectively, which reduce to the above relations if $S''_i \langle \Delta E_F \rangle \langle \Delta E_B \rangle \ll 1$. Since $S''_i (E_i) = -\beta_i^2 (E_i) / C_v (E_i)$ (here the Boltzmann constant $k_B = 1$, making $C_v$ unitless) this condition reads
\[ \beta_i^2 \langle \Delta E_F \rangle \langle \Delta E_B \rangle \ll C_v. \]
Now we can expand $\ln \langle e^{-\beta_i \Delta E_F + \frac{1}{2} S_i' \Delta E_i^2} \rangle = \ln \langle e^{-(\beta_i-\beta_o)\Delta E_F + \frac{1}{2} (S_i' + S_i) \Delta E_B^2} \rangle = 0$ to next order and find
\[ -\beta_1 \langle \Delta E_F \rangle + \frac{1}{2} S''_i \langle \Delta E_F \rangle^2 + \frac{1}{2} \beta_i^2 \langle \Delta E_F \rangle^2 + \frac{1}{3!} \beta_i^3 \langle \Delta E_F \rangle^3 = 0. \]
As $\langle \Delta E_F^2 \rangle = \langle \Delta E_F \rangle^2 + \langle \Delta E_F \rangle^2$, this becomes
\[ -\beta_1 \langle \Delta E_F \rangle + \frac{1}{2} S''_i \langle \Delta E_F \rangle^2 + \frac{1}{2} \beta_i^2 \langle \Delta E_F \rangle^2 + \frac{1}{3!} \beta_i^3 \langle \Delta E_F \rangle^3 = 0. \]
The term $\beta_i^2 + S''_i = \beta_i^2 [1 + O (1/N)]$, since $C_v = O (N)$. Comparing the first and last terms gives the condition $\frac{1}{2} \beta_i^2 / C_v \langle \Delta E_F \rangle^2 \ll \beta_1 \langle \Delta E_F \rangle$ or
\[ \beta_1 \langle \Delta E_F \rangle \ll C_v. \]
In addition, the third cumulant must be small, $\beta_1^3 \langle \Delta E_F^3 \rangle_c \ll \beta_1 \langle \Delta E_F \rangle$ or
$$\beta_1^2 \langle \Delta E_F^3 \rangle_c \ll \langle \Delta E_F \rangle.$$  
A similar argument for $\ln \left[e^{-\left(\beta_2 - \beta_1\right)\Delta E + \frac{1}{2}(S_1' + S_2')\Delta E_F^2}\right] = 0$ yields the requirement
$$\langle \Delta E_B \rangle \ll \frac{|\beta_2 - \beta_1|}{\frac{C_{v_1}}{\beta_1^2} + \frac{C_{v_2}}{\beta_2^2}}.$$  
For $\beta_2 - \beta_1$ of order $\beta_1$ and $\beta_2$, this is satisfied if $\beta_i \langle \Delta E_B \rangle \ll C_{v_j}$ for $i, j = 1, 2$. Close to $\beta_1 = \beta_2$, both $\langle \Delta E_B \rangle$ and $|\beta_2 - \beta_1|$ go to zero, and this condition may be hard to test. However, $\langle \Delta E_B^2 \rangle_c$ remains regular (see discussion in section 2.2 in the main text), and using
$$2 \langle \Delta E_B \rangle = (\beta_2 - \beta_1) \langle \Delta E_B^2 \rangle_c,$$  
we obtain an equivalent relation
$$\langle \Delta E_B^2 \rangle_c \ll \frac{C_{v_1} C_{v_2}}{\beta_2^2 C_{v_2} + \beta_2^2 C_{v_1}};$$  
as
$$\frac{C_{v_1} C_{v_2}}{\beta_2^2 C_{v_2} + \beta_2^2 C_{v_1}} < \min\left(\beta_2^{-2} C_{v_1}, \beta_1^{-2} C_{v_2}\right),$$  
we obtain the condition
$$\langle \Delta E_B^2 \rangle_c \ll \min\left(\beta_2^{-2} C_{v_1}, \beta_1^{-2} C_{v_2}\right)$$  
cited in the main text.

References

[1] Pöschel T and Brilliantov N (ed) 2003 Granular Gas Dynamics (Berlin: Springer)
[2] Srebro Y and Levine D 2004 Phys. Rev. Lett. 93 240601
[3] Dalla Torre E G, Demler E, Giamarchi T and Altman E 2010 Nature Phys. 6 806–10
[4] Bloch I, Dalibard J and Zwerger W 2008 Rev. Mod. Phys. 80 885–964
[5] Blatt R and Wineland D J 2008 Nature 453 1008–15
[6] Petta J R, Taylor J M, Johnson A C, Yacoby A, Lukin M D, Marcus C M, Hanson M P and Gossard A C 2008 Phys. Rev. Lett. 100 067601
[7] Bunin G, D’Alessio L, Kafri Y and Polkovnikov A 2011 Nature Phys. 7 913–7
[8] Allen M P and Tildesley D J 1989 Computer Simulation of Liquids (Oxford: Oxford University Press)
[9] Lebowitz J L and Spohn H 1978 J. Stat. Phys 19 6
[10] Jarzynski C 1992 Phys. Rev. A 46 7498–509
[11] Ott E 1979 Phys. Rev. Lett. 42 1628–31
[12] Cohen D 2000 Ann. Phys., NY 283 175–231
[13] Gardiner C W 1994 Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences (Berlin: Springer)
[14] Kardar M 2007 Statistical Physics of Particles (Cambridge: Cambridge University Press)
[15] Crooks G E 1998 J. Stat. Phys. 90 1481–7
[16] Campisi M, Hanggi P and Talkner P 2011 Rev. Mod. Phys. 83 771–91