Thermal contact through a diathermal wall: a solvable toy model

# Thermal contact through a diathermal wall: a solvable toy model

## F Cornu[1,3] and M Bauer[2]

[1] Laboratoire de Physique Théorique, UMR 8627 du CNRS,
Université Paris-Sud, Bâtiment 210, F-91405 Orsay, France
[2] Institut de Physique Théorique de Saclay (CEA/DSM/IPhT, Unité de
Recherche Associée au CNRS), CEA Saclay, F-91191 Gif-sur-Yvette Cedex,
France
E-mail: Francoise.Cornu@u-psud.fr and Michel.Bauer@cea.fr

**Abstract.** A diathermal wall between two heat baths at different temperatures can be mimicked by a layer of independent spin pairs with some internal energy and where each spin $\sigma_a$ is flipped by thermostat $a$ $(a = 1, 2)$. The transition rates are determined from the modified detailed balance. Generalized heat capacities, excess heats, the housekeeping entropy flow and the thermal conductivity in the steady state are calculated. The joint probability distribution of the heat cumulated exchanges at any time is computed explicitly. We obtain the large deviation function of heat transfer via a variety of approaches. In particular, by a saddle-point method performed accurately, we obtain the explicit expressions not only of the large deviation function, but also of the amplitude prefactor, in the long-time probability density for the heat current. The following physical properties are discussed: the effects of typical time scales of the mesoscopic dynamics which do not appear in equilibrium statistical averages and the limit of strict energy dissipation towards a thermostat when its temperature goes to zero.

**Keywords:** exact results, transport processes/heat transfer (theory), current fluctuations, large deviations in non-equilibrium systems

---

[3] Author to whom any correspondence should be addressed.

## Contents

# 1. Introduction

## 1.1. Issues at stake

Though a rather specialized topic within non-equilibrium physics, the problem of heat transfer is in itself a vast subject, and one of foremost theoretical and practical interest. Heat exchanges between two bodies at different inverse temperatures $\beta_1$ and $\beta_2$ can be settled either by a piece of material with macroscopic width (or length in the case of electric wires) or by a thin interface with a microscopic width. In the first case the heat flow from one body to the other is described by the phenomenological Fourier law, and a microscopic description from an out-of-equilibrium statistical mechanics approach is a wide research field beyond the scope of the present paper. Here we rather address the thermal contact situation, where heat flows through a diathermal interface between two

bodies, which may be either an immaterial interface between two solids or a diathermal wall between two fluids.

A theoretical, microscopic understanding of non-equilibrium phenomena is still lacking today. However, when compared with the most general non-equilibrium physics world, certain specific settings have some advantages and this is the case for thermal contact: when only quantities conserved by microscopic dynamics (energy for a thermal contact) are exchanged between two bodies, then the corresponding cumulative transfers obey some universal symmetry relations. There are a number of relevant viewpoints on these issues. A few of which are the following.

- In the framework of Hamiltonian dynamics of finite systems but within an approximation where interfacial interaction energies can be neglected [1], it can be shown that, if a body 1 at initial inverse temperature $\beta_1$ is put in contact with a body 2 at initial inverse temperature $\beta_2$ during some finite time $t$, then the (fluctuating) heat amount $\mathcal{Q}_2(t)$ coming out of body 2 and going into body 1 ($\mathcal{Q}_2$ being positive if heat indeed goes out of body 2) obeys a detailed fluctuation relation at finite time (transient regime): the probability that the latter heat amount measured during $t$ is equal to $\mathcal{Q}_2$, $P(\mathcal{Q}_2; t)$, is related to the probability of observing the opposite value by $P(\mathcal{Q}_2; t)/P(-\mathcal{Q}_2; t) = \exp[(\beta_1 - \beta_2)\mathcal{Q}_2]$.

- In the case where bodies 1 and 2 are thermal baths (so that their temperatures are well defined and time independent), and at the mesoscopic level where the energy $\mathcal{E}$ of the interface is taken into account, the fluctuating cumulative current $\mathcal{Q}_2(t)/t$ associated with the heat amount $\mathcal{Q}_2(t)$ received by the interface from thermal bath 2 during some time interval $t$ still obeys a fluctuation relation in the long-time limit (stationary regime), as shown in the context of a Markovian stochastic evolution where the transition rates obey the modified detailed balance recalled in (2.3) [2, 3]. Indeed, in the case of an interface with a non-negligible internal energy $\mathcal{E}$, by virtue of energy conservation $\mathcal{Q}_1(t) + \mathcal{Q}_2(t) = \mathcal{E}(t) - \mathcal{E}(t = 0)$ does not vanish, but if the values taken by $\mathcal{E}$ are bounded, $\lim_{t \to +\infty} \mathcal{Q}_2(t)/t = -\lim_{t \to +\infty} \mathcal{Q}_1(t)/t$, and then the current $\mathcal{J}(t) \equiv \mathcal{Q}_2(t)/t$ has a large deviation function $f_{\mathcal{Q}_2}(\mathcal{J})$, whose definition is recalled in section 6, and which obeys the fluctuation relation

$$f_{\mathcal{Q}_2}(\mathcal{J}) = f_{\mathcal{Q}_2}(-\mathcal{J}) + (\beta_1 - \beta_2)\mathcal{J}. \tag{1.1}$$

The latter equation is an example of the various fluctuation relations that can be derived for Markov jump processes in continuous time (see review in [4]) in the framework of the so-called 'stochastic thermodynamics' of small systems (see for instance an extended review in [5] and a recent brief introduction in [6]).

Though relation (1.1) is universal in the considered class of models, the expression of $f_{\mathcal{Q}_2}$ depends on the characteristics of the model for the thermal contact, namely the internal energy of the interface and the choice of the transition rates of its stochastic evolution under heat exchanges.

In the absence of any theoretical framework which would play the role of Gibbs' statistical ensemble theory for equilibrium states and allow one to predict the influence of the dynamics, one has searched for solvable models which could give some hints about properties of fluctuating cumulative currents finer than the symmetries described by the fluctuation relations and which might nevertheless be rather generic. Models where

the contact system between two thermal baths has a stochastic dynamics have been introduced: the system may be an Ising spin lattice (see for instance [7]–[10]) or a particle with Langevin stochastic dynamics [11, 12]. In the latter case the heat exchanges are described as the work performed by a force including a friction as well as a random noise component. The latter interpretation of heat has been proposed and investigated by Sekimoto [13] for ratchet models, and has been used again in the interpretation of the Hatano–Sasa identity [14] as well as in the investigation of heat fluctuations in Brownian transducers [15].

Generic properties are to be obeyed by those among the thermal contact models that have the following features: the system has a finite number of possible configurations, the heat exchanges are described as changes in the populations of energy levels, and the configurations evolve under a stochastic master equation with transition rates bound to obey the modified detailed balance (2.3). This is the case for the long-time fluctuation relation (1.1) and for the following finite-time detailed fluctuation relation, valid for a protocol where a piece of material with non-negligible internal energy, initially at inverse temperature $\beta_0$, is put as a diathermal interface between two thermal baths at inverse temperatures $\beta_1$ and $\beta_2$ at time $t = 0$: the joint probability for measuring the heat amounts $\mathcal{Q}_1$ and $\mathcal{Q}_2$ during a finite time $t$ satisfies the relation $P(\mathcal{Q}_1, \mathcal{Q}_2; t)/P(-\mathcal{Q}_1, -\mathcal{Q}_2; t) = \exp[(\beta_0 - \beta_1)\mathcal{Q}_1 + (\beta_0 - \beta_2)\mathcal{Q}_2]$ [16].

In the present paper, within the latter class of models, we consider a very simple one:

- The interface consists of two layers.
- Each layer consists in a number of independent identical microscopic systems, which we call spins, because we assume that they have only two states. Without loss of generality these states can be labeled $\pm 1$, so that we are dealing with (classical Ising) spins.

Of course, in a more realistic thermal contact, some interactions between spins in the same layer, reflecting the (two-dimensional) geometry of the interface, would be present. But there is no obvious reason to believe that these interactions would qualitatively change the physics of heat transfer.

- Each spin in a layer is coupled to a single spin in the other layer, building what we call a spin pair, and this coupling makes it possible to transfer energy through the interface, hence the name 'diathermal wall'.
- Locality makes it physically natural to assume that heat bath 1 (resp. 2) can only flip a dynamical variable (spin) in layer 1 (resp. 2).

Within our simplistic model, the law of large numbers allows one to quantify how fluctuations of the heat flow are suppressed when the size of the interface goes from microscopic to mesoscopic and macroscopic. We shall not embark on this study in the present paper, and we shall use independence to concentrate on the heat exchanges at the level of a single spin pair, i.e. a pointlike thermal contact: we deal with a single spin pair $(\sigma_1, \sigma_2)$, with spin $\sigma_1$ in layer 1 coupled to spin $\sigma_2$ in layer 2.

- The energy $\mathcal{E}(\sigma_1, \sigma_2)$ changes when a contact dynamical variable is flipped and, assuming an energy-conserving dynamics, this means that some energy comes from, or is given to, the heat bath responsible for the flip.

In general, the energy $\mathcal{E}(\sigma_1, \sigma_2)$ for the contact dynamical variables could take four distinct values, but we just concentrate on the case $\mathcal{E}(\sigma_1, \sigma_2) = \frac{1}{2}(1 - \sigma_1\sigma_2)\Delta e$, where $\Delta e > 0$ is the energy gap. In the language of spins, this means the absence of external magnetic fields. More abstractly, it implies a twofold symmetry. As the toy model has only four states, solving it can be reduced in some sense to the diagonalization of a $4 \times 4$ matrix, and the twofold symmetry of the energy functional allows to reduce this task to the diagonalization of a pair of $2 \times 2$ matrices.

An interest of our specific solvable model is that it plays the role of a pedagogical example where the general statements are made very explicit. As already mentioned, though the fluctuation relations entail a constraint upon large deviation functions, they do not allow one to determine them. The analytical calculation of the large deviation functions may provide a deeper understanding in the information that they contain. We shall see that, within the model, the computation of large deviation functions for the energy variations in the baths can be remarkably simple or tricky, depending on the kind of techniques one uses.

Moreover, in the absence of principles for out-of-equilibrium statistical mechanics, the formulae obtained for the solvable model can give a flavor of some physical effects. For instance, the time scales of the microscopic dynamics, which do not appear in equilibrium averages, play a role in the properties of non-equilibrium stationary states. Moreover, the model can be considered in the limit where the temperature of the cold thermostat vanishes; then the strict dissipation of energy towards the zero-temperature bath gives rise to specific phenomena.

## 1.2. Contents of the paper

The results of the explicit analytical calculations for the solvable model where the system is reduced to two spins are the following.

In the case where the spin system involves only two spins, the transition rates are determined by the modified detailed balance (2.3) up to the typical inverse times $\nu_a$ of spin flips by each thermal bath $a$, characterized by its temperature $T_a$. For an Ising interaction between the two spins, the transition rates for the energy exchanges with one bath take a form similar to that introduced by Glauber [17] in his investigation of the time-dependent statistics of the Ising chain in contact with a single thermal bath. Most of the time, our results will hold whatever the values of $T_1$ and $T_2$ are. However, it is sometimes convenient to know in which direction heat flows from one reservoir to the other on the average, and then we shall always assume that $T_1 \leq T_2$. By symmetry, this induces no loss of generality anyway: the results for $T_1 \geq T_2$ can be retrieved by permuting $T_1$ with $T_2$ and $\nu_1$ with $\nu_2$.

The non-equilibrium stationary state (NESS) of the model happens to have a very specific property (section 3.1): since the transition rates are invariant under the simultaneous flips of both spins, the configuration probability distribution in the NESS coincides with an equilibrium canonical distribution at some inverse temperature $\beta_\star$.

The linear and nonlinear static responses are explicitly calculated (sections 3.2 and 3.3). The expressions for the generalized heat capacities involve not only the temperatures of the energy reservoirs $a = 1, 2$ but also the typical inverse time scales $\nu_a$ of the heat exchange dynamics with each reservoir. The $\nu_a$ are also called kinetic parameters in the following. In the vicinity of equilibrium the mean heat current is proportional to the

difference $T_2 - T_1$ between the bath temperatures; then a linear thermal conductivity can be defined. When the system is far from equilibrium the mean heat current is a bounded function of the thermostat temperatures (saturation phenomenon); one can introduce a nonlinear thermal conductivity which vanishes in the limit where the relative temperature difference goes to infinity. The expression of the housekeeping entropy flow is given, and the excess mean heats, which are defined in terms of the measurable averages of the cumulative heats [18], are explicitly calculated, for the static response protocol, from the average heat amounts received by the system from each bath during a finite time $t$, and which are determined in section 4.3.

The joint probability distribution for received cumulative heats $\mathcal{Q}_1$ and $\mathcal{Q}_2$ is determined at any finite time and for any initial distribution probability through a generating function method (see section 4). Other distribution probabilities are then derived from its expression (4.40), and the explicit results are summarized in section 4.2. The results are given in terms of two integrals in the complex plane. The system obeys the finite-time symmetry (4.68) enforced by the modified detailed balance for the ratio of the probabilities to measure some given heat amounts $\mathcal{Q}_1$ and $\mathcal{Q}_2$ or their opposite values when the system is initially prepared in an equilibrium state. But it also satisfies another finite-time symmetry specific to the model for the ratio of the same probabilities when the system has any initial distribution probability. The latter fluctuation relation (4.72) is more subtle, as it involves the initial probability distribution for the product of the spins (or equivalently for the energy of the spin pair).

The cumulants for the cumulative heat $\mathcal{Q}_2$ are studied in section 5 from the characteristic function of the probability density for $\mathcal{Q}_2$. The relation between the characteristic function of a probability density $\Pi(\mathcal{Q}; t)$ with the generating function for the probability function $P(\mathcal{Q}; t)$ when the variable $\mathcal{Q}$ can take only discrete values is recalled in section 5.1.1. The explicit formulæ for the first four cumulants per unit time in the infinite-time limit are given in (5.15). Even at equilibrium the cumulants are not those of a Gaussian.

The large deviation function for the cumulative heat current $\mathcal{Q}_2/t$ is calculated by three different methods (section 6): from the Gärtner–Ellis theorem (section 6.1), from a saddle-point method (section 6.2) and from Laplace's method on a discrete sum (section 6.3). The second and third methods rely explicitly on the discrete nature of heat exchanges in the model and on an ad hoc definition of large deviation functions discussed in [16]; but they allow one to compute subdominant contributions as well. The first method is straightforward, one just has to check that the general applicability hypotheses (recalled in detail below) are fulfilled, which is easy in our case. The third method is also simple because it deals with a sum of non-negative terms, so no compensation is possible. The saddle-point method, however, is remarkably tricky in our case, for reasons that we shall detail below. The expressions in terms of various parameter sets are given in (6.10) and (6.16). In order to readily obtain the large deviation function in the case where the temperature of the colder bath vanishes, its expressions for positive and negative currents are explicitly distinguished in (6.13) and (6.14).

The limit where the kinetic parameter of one thermostat becomes infinitely large with respect to the kinetic parameter of the other thermostat is studied in section 7. In this limit the stationary distribution of the spins is the equilibrium canonical probability at the temperature of the 'fast' heat bath, while the typical inverse time scale in the mean

instantaneous heat current is the kinetic parameter $\nu_{\text{slow}}$ of the 'slow' heat bath. The probability distribution for the heat amount received from the slow thermostat, $\mathcal{Q}_{\text{slow}}$, at any finite time $t$ is that of an asymmetric random walk with the inverse time scale $\nu_{\text{slow}}$. As a consequence the probability distribution for $\mathcal{Q}_{\text{slow}}$ obeys a fluctuation relation at any finite time (see (7.20)). The probability distributions of $\sigma_1\sigma_2$ and $\mathcal{Q}_{\text{slow}}$ are independent of each other, and this mean-field property is interpreted as a kinetic effect in the considered limit. The very simple forms of the infinite-time cumulants per unit time are given. The long-time distribution of the cumulative heat current is exhibited: it vanishes exponentially fast over a time scale given by the inverse of the large deviation function (7.32), with an amplitude which is explicitly calculated.

In the limit where the temperature of the colder thermostat vanishes (section 8) the microreversibility is broken, but the system still reaches a stationary state where all configurations have a non-vanishing weight, because the Markov matrix is still irreducible. The large deviation function is expressed in (8.9). In the limit where the kinetic parameter of one thermostat becomes infinitely large with respect to the kinetic parameter of the other thermostat, the probability distribution for the heat amount $\pm\mathcal{Q}_{\text{slow}}$, with sign $-$ $(+)$ if the slow thermostat is the cold (hot) one, becomes a Poisson process at any finite time $t$, because the zero-temperature thermostat can only absorb energy (strict dissipation towards the zero-temperature bath). Again, the very simple forms of the infinite-time cumulants per unit time are given, as well as the large deviation function (8.21).

## 2. Model

The physical system we deal with in this paper is a toy model of thermal contact, consisting of two heat baths, generically at different temperatures, put indirectly in contact via a small subsystem made of two interacting Ising spins $\sigma_1$ and $\sigma_2$. Each spin $\sigma_a$, $a = 1, 2$ is in contact with a single bath denoted by $a$. We aim at a statistical description, where the details of what happens in the heat baths is not observed, but only the evolution of the two spins, i.e. of the configuration $\mathcal{C} \equiv (\sigma_1, \sigma_2)$. We assume that this evolution is described by a Markov process (in continuous time) with transition rate $(\mathcal{C}'|\mathbb{W}|\mathcal{C})$ from configuration $\mathcal{C}$ to configuration $\mathcal{C}'$.

As the system is out-of-equilibrium, the form of $\mathbb{W}$ is not a direct consequence of known physical laws, and it is unclear whether a nature-given preferred choice exists. So we start with a purely technical and down-to-earth description of our choice for the transition rates, which we shall use for all later explicit computations. The general principles and steps that guided us to the modified detailed balance that the transition rates must obey have been given in [16]. The main ideas are the following.

### 2.1. Constraints upon transition rates arising from microscopic discrete ergodic energy-conserving dynamics

As usual, we view a heat bath as an ideal limit of some large but finite system. So the system we describe is obtained via a limiting procedure from a large system made of two large parts and a small part, which is reduced to the two Ising spins $\sigma_1$ and $\sigma_2$, each one directly in contact with one of the large parts.

We expect that in this limit many details become irrelevant, so we assume for the sake of the argument that the degrees of freedom in the large parts are discrete.

As in classical statistical mechanics, we take the viewpoint that the statistical description of $\sigma_1, \sigma_2$ is an effective mesoscopic description arising from a deterministic, energy-conserving dynamics for the whole system. With discrete variables, there is no general definition of time-reversal invariance, but we impose that the dynamics is ergodic.

We also want the dynamics to reflect the fact that the two large parts interact only indirectly: there is an interaction energy $\mathcal{E}(\sigma_1, \sigma_2)$ between the two spins and the spin $\sigma_a$ is flipped thanks to energy exchanges with the large part $a$ ($a = 1, 2$). Defining the operator $\mathbb{F}_a$ as the operator flipping the spin $\sigma_a$ while leaving the other spin unchanged (e.g $\mathbb{F}_1(\sigma_1, \sigma_2) = (-\sigma_1, \sigma_2)$), in this process the energy of the large part $a$ is changed from $E_a$ to $E'_a$ according to the energy conservation law

$$E'_a - E_a = \begin{cases} -\left[\mathcal{E}(\mathcal{C}') - \mathcal{E}(\mathcal{C})\right] & \text{if } \mathcal{C}' = \mathbb{F}_a\mathcal{C} \\ 0 & \text{otherwise,} \end{cases} \tag{2.1}$$

while the energy of the other large part is unchanged.

As shown in [16], when the large parts are described at a statistical level and in a transient regime where the large parts are described in the thermodynamic limit, the transition rate $(\mathcal{C}'|\mathbb{W}|\mathcal{C})$ from configuration $\mathcal{C}$ to configuration $\mathcal{C}'$ obeys three constraints: first the graph associated with the transition rates is connected; second there is microscopic reversibility for any couple of configurations $(\mathcal{C}, \mathcal{C}')$,

$$(\mathcal{C}'|\mathbb{W}|\mathcal{C}) \neq 0 \qquad \Leftrightarrow \qquad (\mathcal{C}|\mathbb{W}|\mathcal{C}') \neq 0; \tag{2.2}$$

third the ratio of transition rates obeys the so-called modified detailed balance (MDB),

$$\text{for } \mathcal{C}' = \mathbb{F}_a\mathcal{C} \quad \frac{(\mathcal{C}'|\mathbb{W}|\mathcal{C})}{(\mathcal{C}|\mathbb{W}|\mathcal{C}')} = e^{-\beta_a[\mathcal{E}(\mathcal{C}') - \mathcal{E}(\mathcal{C})]}. \tag{2.3}$$

We remind the reader that the latter relation is also referred to in the literature as the 'generalized detailed balance'.

### 2.2. Determination of transition rates

The transition rates are non-zero only if the initial and final configurations $\mathcal{C}$ and $\mathcal{C}'$ differ only by the flip of one spin: $(\mathcal{C}'|\mathbb{W}|\mathcal{C}) = 0$ unless either $\mathcal{C}' = \mathbb{F}_1\mathcal{C}$ or $\mathcal{C}' = \mathbb{F}_2\mathcal{C}$. Since $\sigma_a$ can take only the two values $+1$ and $-1$, the transition rate where $\sigma_a$ is flipped takes the generic form

$$(\mathbb{F}_a\mathcal{C}|\mathbb{W}|\mathcal{C}) = \frac{\nu_a(\sigma_b)}{2}\left[1 - \sigma_a\Gamma_a(\sigma_b)\right]. \tag{2.4}$$

The four parameters $\nu_a, \Gamma_a$, $a = 1, 2$, are *a priori* arbitrary, except that the $\nu$ are $> 0$ and the $\Gamma$ are of absolute value $\leq 1$.

Taking for simplicity an interaction energy between the spins

$$\mathcal{E}(\sigma_1, \sigma_2) = \frac{1 - \sigma_1\sigma_2}{2}\Delta e, \tag{2.5}$$

where $\Delta e > 0$ is the energy gap between the two energy levels, one gets from the modified detailed balance in the form (2.3) that

$$\text{for } \mathcal{C}' = \mathbb{F}_a\mathcal{C} \quad \frac{(\mathcal{C}'|\mathbb{W}|\mathcal{C})}{(\mathcal{C}|\mathbb{W}|\mathcal{C}')} = e^{-\sigma_1\sigma_2\beta_a\Delta e}, \tag{2.6}$$

a condition similar to the one obtained by Glauber [17] in the equilibrium case. As $e^{2x} = (1 + \tanh x)/(1 - \tanh x)$, the generic form (2.4) of $(\mathbb{F}_a\mathcal{C}|\mathbb{W}|\mathcal{C})$ has to satisfy

$$(\mathbb{F}_a\mathcal{C}|\mathbb{W}|\mathcal{C}) = \frac{\nu_a(\sigma_b)}{2}\left[1 - \sigma_1\sigma_2\gamma_a\right] \tag{2.7}$$

with

$$\gamma_a \equiv \tanh\left(\beta_a\frac{\Delta e}{2}\right). \tag{2.8}$$

If $\beta_1$ and $\beta_2$ are finite, $0 \leq \gamma_1 < 1$ and $0 \leq \gamma_2 < 1$, and the microscopic reversibility condition (2.2) is also satisfied. Without loss of generality we could, and will sometimes, assume that $T_1 \leq T_2$. Then $\gamma_1 \geq \gamma_2$.

For the sake of simplicity, in the following we assume that $\nu_a$ depends only on the properties of the thermostat and not on the value of $\sigma_b$. (This choice enforces the equality between the transition rate from $(\sigma_1, \sigma_2)$ and that from $(-\sigma_1, -\sigma_2)$, which are two configurations with the same energy.) Apart from simplicity, we have no convincing argument that this should be THE nature-given preferred choice. Anyway, we write

$$(\mathbb{F}_a\mathcal{C}|\mathbb{W}|\mathcal{C}) = \frac{\nu_a}{2}\left[1 - \sigma_1\sigma_2\gamma_a\right]. \tag{2.9}$$

This ends the argument explaining our choice of transition rates and gives a physical interpretation of the parameters: $\gamma_a$ is formed with the energy scale in the two-spin system and the temperature of bath $a$, while $\nu_a$ describes a rate at which bath $a$ attempts to flip spin $\sigma_a$.

We notice that, though the transition rate expressions have been derived from hypotheses implying the microscopic reversibility (2.2), these expressions still make sense if $\beta_2 < \beta_1 = +\infty$. (The limit $\beta_1 \to +\infty$, where microscopic reversibility is broken, is discussed in section 8.)

Moreover, even if $\beta_1 = +\infty$, the Markov matrix $\mathbb{M}$ defined by

$$(\mathcal{C}'|\mathbb{M}|\mathcal{C}) = \begin{cases} (\mathcal{C}'|\mathbb{W}|\mathcal{C}) & \text{if } \mathcal{C}' \neq \mathcal{C} \\ -\sum_{\mathcal{C}''}(\mathcal{C}''|\mathbb{W}|\mathcal{C}) & \text{if } \mathcal{C}' = \mathcal{C} \end{cases} \tag{2.10}$$

is irreducible; namely any configuration $\mathcal{C}'$ can be reached by a succession of jumps with non-zero transition rates from any configuration $\mathcal{C}$.

## 3. Non-equilibrium stationary state (NESS) as a canonical distribution with an effective temperature

### 3.1. Stationary state distribution

The master equation which rules the evolution of the probability $P(\mathcal{C}; t)$ can be written in terms of the Markov matrix $\mathbb{M}$ defined in (2.10) as

$$\frac{\mathrm{d}P(\mathcal{C}; t)}{\mathrm{d}t} = \sum_{\mathcal{C}'}(\mathcal{C}|\mathbb{M}|\mathcal{C}')P(\mathcal{C}'; t). \tag{3.1}$$

In the basis where the probability $P(\sigma_1, \sigma_2; t)$ is represented by the column vector

$$|P(t)) = \begin{pmatrix} P(++; t) \\ P(--; t) \\ P(+-; t) \\ P(-+; t) \end{pmatrix} \tag{3.2}$$

the matrix $\mathbb{M}$ takes the form

$$\mathbb{M} = \frac{\nu_1 + \nu_2}{2} \begin{pmatrix} -1 + \gamma_\star & 0 & \overline{\nu}_2(1 + \gamma_2) & \overline{\nu}_1(1 + \gamma_1) \\ 0 & -1 + \gamma_\star & \overline{\nu}_1(1 + \gamma_1) & \overline{\nu}_2(1 + \gamma_2) \\ \overline{\nu}_2(1 - \gamma_2) & \overline{\nu}_1(1 - \gamma_1) & -1 - \gamma_\star & 0 \\ \overline{\nu}_1(1 - \gamma_1) & \overline{\nu}_2(1 - \gamma_2) & 0 & -1 - \gamma_\star \end{pmatrix}. \tag{3.3}$$

In the latter equation we have introduced the dimensionless inverse time scales

$$\overline{\nu}_a = \frac{\nu_a}{\nu_1 + \nu_2} \qquad \text{for } a = \{1, 2\}, \tag{3.4}$$

and we have set

$$\gamma_\star = \overline{\nu}_1 \gamma_1 + \overline{\nu}_2 \gamma_2. \tag{3.5}$$

The Markov matrix $\mathbb{M}$ is irreducible (even if $\gamma_1 = 1$, namely $T_1 \to 0$): for any pair of configurations $\mathcal{C}$ and $\mathcal{C}'$, there exists a succession of spin flips, with non-zero transition rates, which allows one to make the system evolve from $\mathcal{C}$ to $\mathcal{C}'$. Henceforth, according to the Perron–Frobenius theorem there exists a single stationary state distribution $P_{\text{st}}(\mathcal{C})$, and it is non-zero for every configuration $\mathcal{C}$.

Moreover, since the system is made of two discrete variables which can take only the values $\pm 1$ and since the transition rates are invariant under the simultaneous flips $\sigma_1 \to -\sigma_1$ and $\sigma_2 \to -\sigma_2$, the stationary distribution $P_{\text{st}}(\sigma_1, \sigma_2)$ takes the form $P_{\text{st}}(\sigma_1, \sigma_2) = a + d\sigma_1\sigma_2$. Indeed, the generic form of $P(\sigma_1, \sigma_2)$ reads $P(\sigma_1, \sigma_2) = a + b\sigma_1 + c\sigma_2 + d\sigma_1\sigma_2$. On the other hand, the invariance of the transition rates under the simultaneous flips $\sigma_1 \to -\sigma_1$ and $\sigma_2 \to -\sigma_2$ entails that if $a + b\sigma_1 + c\sigma_2 + d\sigma_1\sigma_2$ is a stationary solution, $a - b\sigma_1 - c\sigma_2 + d\sigma_1\sigma_2$ is also a stationary solution. But, since $\mathbb{M}$ is irreducible, there is only one stationary solution, so that $b = c = 0$. By solving explicitly the master equation (3.1) and using the normalization of a probability distribution, the stationary solution proves to be

$$P_{\text{st}}(\sigma_1, \sigma_2) = \tfrac{1}{4}[1 + \gamma_\star \sigma_1 \sigma_2] \tag{3.6}$$

where $\gamma_\star$ is defined in (3.5).

The stationary distribution of the model has the following remarkable property: it coincides with some equilibrium distribution. More precisely, the stationary state distribution is equal to the canonical state distribution at the effective inverse temperature $\beta_\star$

$$P_{\text{st}}(\sigma_1, \sigma_2) = P_{\text{can}}^{\beta_\star}(\sigma_1, \sigma_2), \tag{3.7}$$

where $\beta_\star$ is determined by the relation

$$\gamma_\star = \tanh\left(\beta_\star \frac{\Delta e}{2}\right), \tag{3.8}$$

and

$$P_{\text{can}}^{\beta}(\mathcal{C}) = \frac{\mathrm{e}^{-\beta\mathcal{E}(\mathcal{C})}}{Z(\beta)}, \tag{3.9}$$

where $Z(\beta)$ is the canonical partition function at the inverse temperature $\beta$, $Z(\beta) = \sum_{\mathcal{C}} \mathrm{e}^{-\beta\mathcal{E}(\mathcal{C})}$. We notice that the canonical form for the state distribution implies that $\beta_{\star}$ obeys the canonical ensemble relation, which is equivalent to the definition of the inverse temperature in the microcanonical ensemble, namely

$$\beta_{\star} = \frac{\partial S^{\text{SG}}[P_{\text{st}}]}{\partial \langle \mathcal{E} \rangle_{\text{st}}}, \tag{3.10}$$

where $\langle \mathcal{E} \rangle_{\text{st}} \equiv \sum_{\mathcal{C}} \mathcal{E}(\mathcal{C}) P_{\text{st}}(\mathcal{C})$ is the stationary mean value of the energy and $S^{\text{SG}}[P_{\text{st}}]$ is the value of the dimensionless Shannon–Gibbs entropy in the stationary state. The dimensionless Shannon–Gibbs entropy (where the Boltzmann constant is set equal to 1) is defined from the configuration probability distribution $P(\mathcal{C};t)$ as

$$S^{\text{SG}}[P(t)] \equiv -\sum_{\mathcal{C}} P(\mathcal{C};t) \ln P(\mathcal{C};t). \tag{3.11}$$

Its evolution has been recalled in [16].

### 3.2. Linear static response to a variation of some external parameter

In the present section we consider the static linear response of some observable $\mathcal{O}$ to a change of some external parameter, namely the inverse temperature $\beta_a$ or the typical inverse time scale $\nu_a$ of bath $a$, with $a = 1, 2$.

In the protocols for the study of static linear response, the system is prepared in some stationary state at time $t_0 = 0^-$ and the external parameters are instantaneously changed by infinitesimal amounts at time $t = 0$. Then, in the infinite-time limit, the system reaches another stationary state corresponding to the new values of the external parameters.

*3.2.1. Relation with static correlations for a 'canonical' NESS.* Since the non-equilibrium stationary distribution given by (3.6) involves only one parameter, namely $\beta_{\star}$, the linear response coefficient $\partial \langle \mathcal{O} \rangle_{\text{st}} / \partial g_{\text{ext}}$ for the mean value of an observable $\mathcal{O}$ in the stationary distribution when some external parameter $g_{\text{ext}}$ is varied is proportional to $\partial \langle \mathcal{O} \rangle_{\text{st}} / \partial \beta_{\star}$, namely $\partial \langle \mathcal{O} \rangle_{\text{st}} / \partial g_{\text{ext}} = (\partial \beta_{\star} / \partial g_{\text{ext}}) \times (\partial \langle \mathcal{O} \rangle_{\text{st}} / \partial \beta_{\star})$. Moreover, by virtue of (3.7), the stationary distribution is the canonical distribution at the inverse temperature $\beta_{\star}$. Henceforth the coefficient $\partial \langle \mathcal{O} \rangle_{\text{st}} / \partial \beta_{\star}$ is merely opposite to the correlation between $\mathcal{O}$ and the energy $\mathcal{E}$ according to the canonical equilibrium identity

$$\frac{\partial \langle \mathcal{O} \rangle_{\text{can}}^{\beta_{\star}}}{\partial \beta_{\star}} = -\left[ \langle \mathcal{O}\mathcal{E} \rangle_{\text{can}}^{\beta_{\star}} - \langle \mathcal{O} \rangle_{\text{can}}^{\beta_{\star}} \langle \mathcal{E} \rangle_{\text{can}}^{\beta_{\star}} \right], \tag{3.12}$$

where $\langle \mathcal{O} \rangle_{\text{can}}^{\beta_{\star}}$ denotes an average with respect to the canonical distribution $P_{\text{can}}^{\beta_{\star}}$. As a result, the relation valid for responses to the variation of any external parameter in the non-equilibrium stationary state reads

$$\frac{\partial \langle \mathcal{O} \rangle_{\text{st}}}{\partial g_{\text{ext}}} = -\frac{\partial \beta_{\star}}{\partial g_{\text{ext}}} \left[ \langle \mathcal{O}\mathcal{E} \rangle_{\text{st}} - \langle \mathcal{O} \rangle_{\text{st}} \langle \mathcal{E} \rangle_{\text{st}} \right]. \tag{3.13}$$

*3.2.2. Dependence of the mean energy upon the time scales of the microscopic dynamics.*
The main difference between the response of the mean energy in non-equilibrium and
equilibrium states arises for the response to a variation of the time scales of the microscopic
dynamics which rules the heat exchanges with the baths. When $\beta_1 = \beta_2$ the equilibrium
mean energy $\langle \mathcal{E} \rangle_{\mathrm{eq}} = \langle \mathcal{E} \rangle_{\mathrm{can}}^{\beta_1}$ depends only on the thermodynamic temperature common
to both baths. In contrast, in the non-equilibrium case the stationary mean energy
$\langle \mathcal{E} \rangle_{\mathrm{st}}$ also depends on both inverse time scales $\nu_1$ and $\nu_2$. Indeed, since the stationary
probability corresponds to the effective canonical distribution (3.7), the stationary mean
energy reads

$$\langle \mathcal{E} \rangle_{\mathrm{st}} = (1 - \gamma_\star) \frac{\Delta e}{2} = (1 - \overline{\nu}_1 \gamma_1 - \overline{\nu}_2 \gamma_2) \frac{\Delta e}{2}. \tag{3.14}$$

Changing $\nu_a$ means changing the physical connection between thermal bath $a$ and the
spin system. The linear response of the stationary energy associated with a variation of
the inverse time scale $\nu_a$ is determined by the coefficient

$$\frac{\partial \langle \mathcal{E} \rangle_{\mathrm{st}}}{\partial \nu_a} = -\frac{\nu_b}{(\nu_1 + \nu_2)^2} \gamma_a \frac{\Delta e}{2} \qquad \text{for } \{a, b\} = \{1, 2\}. \tag{3.15}$$

*3.2.3. Stationary mean energy and generalized heat capacities.* The heat capacity $C_{\mathrm{eq}}$ is
a measurable quantity defined as the ratio

$$C_{\mathrm{eq}}(T) = \frac{\langle \delta \mathcal{Q} \rangle}{\mathrm{d}T}, \tag{3.16}$$

where $\langle \delta \mathcal{Q} \rangle$ is the mean heat amount received by the system in transformations which
involve only heat transfers and make the system go from an equilibrium state at
temperature $T$ to another equilibrium state at temperature $T + \mathrm{d}T$, while all other
thermodynamic parameters which determine the equilibrium state are kept constant.
($\langle \delta \mathcal{Q} \rangle = \lim_{t \to +\infty} \langle \mathcal{Q} \rangle_t$ in the protocol mentioned at the beginning of section 3.2.)
According to the energy conservation, $\langle \delta \mathcal{Q} \rangle = \langle \mathcal{E} \rangle_{\mathrm{eq}}^{T+\mathrm{d}T} - \langle \mathcal{E} \rangle_{\mathrm{eq}}^{T}$ and the heat capacity
is related to a partial derivative of the equilibrium mean energy,

$$C_{\mathrm{eq}}(T) = \frac{\partial \langle \mathcal{E} \rangle_{\mathrm{eq}}}{\partial T} = -\beta^2 \frac{\partial \langle \mathcal{E} \rangle_{\mathrm{eq}}}{\partial \beta}. \tag{3.17}$$

When the system is in a stationary non-equilibrium state induced by thermal
contact with two heat reservoirs at respective temperatures $T_1$ and $T_2$, we can introduce
measurable heat capacities by similar definitions. When the temperature $T_1$ of thermal
bath 1 is changed by $\mathrm{d}T_1$, while the temperature $T_2$ of thermal bath 2 is kept fixed, and
when the system evolves from a stationary state to another one only by heat transfers,
then the generalized heat capacity $C_{\mathrm{st}}^{[1]}$ is defined as

$$C_{\mathrm{st}}^{[1]}(T_1, T_2) = \frac{\langle \delta (\mathcal{Q}_1 + \mathcal{Q}_2) \rangle}{\mathrm{d}T_1}. \tag{3.18}$$

According to conservation energy, $\delta\langle\mathcal{Q}_1+\mathcal{Q}_2\rangle = \langle\mathcal{E}\rangle_{\text{st}}^{T_1+\text{d}T_1,T_2} - \langle\mathcal{E}\rangle_{\text{st}}^{T_1,T_2}$ and the heat capacity is related to a partial derivative of the stationary mean energy

$$C_{\text{st}}^{[1]}(T_1,T_2) = \left.\frac{\partial\langle\mathcal{E}\rangle_{\text{st}}^{T_1,T_2}}{\partial T_1}\right|_{T_2}. \tag{3.19}$$

In the present model the expression (3.14) of the stationary mean energy takes the very specific form

$$\langle\mathcal{E}\rangle_{\text{st}} = \overline{\nu}_1\langle\mathcal{E}\rangle_{\text{eq}}^{T_1} + \overline{\nu}_2\langle\mathcal{E}\rangle_{\text{eq}}^{T_2}. \tag{3.20}$$

Indeed the relation $\overline{\nu}_1 + \overline{\nu}_2 = 1$ and the expression of the equilibrium mean energy at the inverse temperature $\beta$,

$$\langle\mathcal{E}\rangle_{\text{eq}}^T = (1-\gamma)\frac{\Delta e}{2} \qquad \text{when } \beta_1 = \beta_2 = \beta, \tag{3.21}$$

allow one to rewrite the mean energy expression (3.14) in the non-equilibrium stationary state in the form (3.20). By virtue of the specific decomposition (3.20) of the mean energy, the heat capacities $C_{\text{st}}^{[a]}(T_1,T_2)$ read

$$C_{\text{st}}^{[a]}(T_1,T_2) = \overline{\nu}_a C_{\text{eq}}(T_a) \qquad \text{with } a = \{1,2\}, \tag{3.22}$$

where, according to the relation (3.17) and the expression (3.21) of $\langle\mathcal{E}\rangle_{\text{eq}}^{\text{T}}$,

$$C_{\text{eq}}(T_a) = \left[1 - \tanh^2\left(\frac{\beta_a\Delta e}{2}\right)\right]\left(\frac{\beta_a\Delta e}{2}\right)^2. \tag{3.23}$$

More generally, when the temperatures $T_1$ and $T_2$ of both thermostats are varied independently

$$\langle\mathcal{E}\rangle_{\text{st}}^{T_1+\text{d}T_1,T_2+\text{d}T_2} - \langle\mathcal{E}\rangle_{\text{st}}^{T_1,T_2} = C_{\text{st}}^{[1]}(T_1,T_2)\text{d}T_1 + C_{\text{st}}^{[2]}(T_1,T_2)\,\text{d}T_2. \tag{3.24}$$

If $T_1$ and $T_2$ are increased by the same infinitesimal quantity $\text{d}T$ the corresponding heat capacity, defined as $C_{\text{st}}(T_1,T_2) \equiv \delta\langle\mathcal{Q}_1 + \mathcal{Q}_2\rangle/\text{d}T$, is equal to the sum $C_{\text{st}}^{[1]}(T_1,T_2) + C_{\text{st}}^{[2]}(T_1,T_2)$. For the present model $C_{\text{st}}(T_1,T_2) = \overline{\nu}_1 C_{\text{eq}}(T_1) + \overline{\nu}_2 C_{\text{eq}}(T_2)$. In the limit where $T_1 = T_2 = T$, by virtue of the relation $\overline{\nu}_1 + \overline{\nu}_2 = 1$, we retrieve the equilibrium heat capacity $C_{\text{eq}}(T)$, as it should be.

*3.2.4. Stationary heat current and linear thermal conductivity.* The instantaneous heat current $j_a(\mathcal{C})$ received from heat bath $a$ when the system jumps out of the configuration $\mathcal{C}$ has been defined in [16] as

$$j_a(\mathcal{C}) \equiv j_{\delta q_a}(\mathcal{C}) \equiv \sum_{\mathcal{C}'}(\mathcal{C}'|\mathbb{W}|\mathcal{C})\,\delta q_a(\mathcal{C}' \leftarrow \mathcal{C}), \tag{3.25}$$

where $\delta q_a(\mathcal{C}' \leftarrow \mathcal{C})$ is the heat received from thermal bath $a$ when the system evolves from configuration $\mathcal{C}$ to configuration $\mathcal{C}' = \mathbb{F}_a\mathcal{C}$, where $\mathbb{F}_a$ is the flip caused by thermal bath $a$, namely

$$\begin{cases} \delta q_a(\mathcal{C}' \leftarrow \mathcal{C}) = [\mathcal{E}(\mathcal{C}') - \mathcal{E}(\mathcal{C})] & \text{if } \mathcal{C}' = \mathbb{F}_a\mathcal{C} \\ \delta q_a(\mathcal{C}' \leftarrow \mathcal{C}) = 0 & \text{otherwise.} \end{cases} \tag{3.26}$$

In the present model $j_2(\sigma_1, \sigma_2) = [\mathcal{E}(\sigma_1, -\sigma_2) - \mathcal{E}(\sigma_1, \sigma_2)] (\sigma_1, -\sigma_2|\mathbb{W}|\sigma_1, \sigma_2) = \nu_2 [\sigma_1\sigma_2 - \gamma_2] (\Delta e/2)$.

In the stationary state the mean energy is constant, so that the mean currents received from both baths cancel, $\langle j_1 \rangle_{st} + \langle j_2 \rangle_{st} = 0$. For the stationary state probability distribution (3.6), one has $\langle \sigma_1\sigma_2 \rangle_{st} = \gamma_\star = \overline{\nu}_1\gamma_1 + \overline{\nu}_2\gamma_2$ and

$$\langle j_2 \rangle_{st} = \overline{\nu}_1\overline{\nu}_2 (\gamma_1 - \gamma_2) \frac{(\nu_1 + \nu_2)\Delta e}{2}, \tag{3.27}$$

where $\gamma_1 - \gamma_2$ may be rewritten as

$$\gamma_1 - \gamma_2 = \tanh\left(\frac{(\beta_1 - \beta_2)\Delta e}{2}\right) \left[1 - \tanh\left(\frac{\beta_1\Delta e}{2}\right)\tanh\left(\frac{\beta_2\Delta e}{2}\right)\right]. \tag{3.28}$$

When $T_1 \leq T_2$, $\langle j_2 \rangle_{st} \geq 0$, as it should: the mean heat current flows from the hot bath to the cold bath. Note that $\langle j_2 \rangle_{st}$ is a bounded function of $T_1$ and $T_2$. Thus, in the generic case $\langle j_2 \rangle_{st}$ is not proportional to the bath temperatures difference $T_2 - T_1$. As for any system, the linear dependence upon $T_2 - T_1$ (or $\beta_1 - \beta_2$) appears in the limit where $(\beta_1 - \beta_2)\Delta e \ll 1$. In the high-temperature regime, where both $\beta_1\Delta e \ll 1$ and $\beta_2\Delta e \ll 1$, the condition $(\beta_1 - \beta_2)\Delta e \ll 1$ is satisfied and $\langle j_2 \rangle_{st}$ is proportional to $T_2 - T_1$.

When $\beta_1 = \beta_2$ the system is at equilibrium and $\langle j_2 \rangle_{eq} = 0$. Moreover, as shown in [16], the partial derivatives of the current obey the generic symmetry

$$\left.\frac{\partial \langle j_2 \rangle_{st}}{\partial \beta_1}\right|_{\beta_2} (\beta_1 = \beta, \beta_2 = \beta) = -\left.\frac{\partial \langle j_2 \rangle_{st}}{\partial \beta_2}\right|_{\beta_1} (\beta_1 = \beta, \beta_2 = \beta). \tag{3.29}$$

This property can also be checked from the expression (3.27) of $\langle j_2 \rangle_{st}$. It entails that

$$\langle j_2 \rangle_{st} \underset{(T_1, T_2) \to (T, T)}{\sim} (T_2 - T_1)\left.\frac{\partial \langle j_2 \rangle_{st}}{\partial T_2}\right|_{T_1} (T, T). \tag{3.30}$$

In other words, when $T_1$ and $T_2$ independently tend to the same value $T$, at first order in the independent variables $T_1 - T$ and $T_2 - T$ the ratio $\langle j_2 \rangle_{st}/(T_2 - T_1)$ depends on $T$ but is independent of the ways $T_1 - T$ and $T_2 - T$ vanish.

As a consequence, for a non-equilibrium stationary state near equilibrium, namely when the temperature difference between the thermostats is such that $(\beta_1 - \beta_2)\Delta e \ll 1$, one can define the thermal conductivity as

$$\kappa_{th} \equiv \lim_{(T_1, T_2) \to (T, T)} \frac{\langle j_2 \rangle_{st}}{T_2 - T_1} = \left.\frac{\partial \langle j_2 \rangle_{st}}{\partial T_2}\right|_{T_1} (T, T). \tag{3.31}$$

From (3.27) we get the expression for the thermal conductivity,

$$\kappa_{th} = \frac{\nu_1\nu_2}{\nu_1 + \nu_2} \left[1 - \tanh^2\left(\frac{\beta\Delta e}{2}\right)\right] \left(\frac{\beta\Delta e}{2}\right)^2. \tag{3.32}$$

We remind the reader that the thermal conductivity, which is a positive transport coefficient, is related to the kinetic coefficient (also called Onsager coefficient) introduced

in phenomenological irreversible thermodynamics as

$$L \equiv \lim_{\mathcal{F} \to 0} \frac{\langle j_2 \rangle_{\mathrm{st}}}{\mathcal{F}}, \tag{3.33}$$

where, as recalled in [16], the thermodynamic force $\mathcal{F}$ can be defined from the stationary entropy production rate, which is opposite to the exchange entropy flow, $\mathrm{d}_{\mathrm{int}}S^{\mathrm{SG}}/\mathrm{d}t|_{\mathrm{st}} = -\mathrm{d}_{\mathrm{exch}}S/\mathrm{d}t|_{\mathrm{st}}$, through the relation

$$\frac{\mathrm{d}_{\mathrm{int}}S^{\mathrm{SG}}}{\mathrm{d}t}\bigg|_{\mathrm{st}} = -\frac{\mathrm{d}_{\mathrm{exch}}S}{\mathrm{d}t}\bigg|_{\mathrm{st}} = \mathcal{F}\langle j_2 \rangle_{\mathrm{st}} \tag{3.34}$$

when there is only one independent mean instantaneous current. In the case of the thermal contact $\mathcal{F} = \beta_1 - \beta_2$. Therefore the relation between the kinetic coefficient and the thermal conductivity defined in (3.31) reads

$$L = \frac{\kappa_{\mathrm{th}}}{\beta^2}. \tag{3.35}$$

Now we compare the results about the linear static response in non-equilibrium stationary states which are either in the vicinity of equilibrium or far away from equilibrium. When the system is far from equilibrium, namely when $(\beta_1 - \beta_2)\Delta e \gg 1$, (3.27) leads to

$$\frac{\partial \langle j_2 \rangle_{\mathrm{st}}}{\partial \beta_2}\bigg|_{\beta_1} = -\frac{\nu_1 \nu_2}{\nu_1 + \nu_2}\left[1 - \tanh\left(\frac{\beta_2 \Delta e}{2}\right)^2\right]\left(\frac{\Delta e}{2}\right)^2 \tag{3.36}$$

$$\frac{\partial \langle j_2 \rangle_{\mathrm{st}}}{\partial \beta_1}\bigg|_{\beta_2} = \frac{\nu_1 \nu_2}{\nu_1 + \nu_2}\left[1 - \tanh\left(\frac{\beta_1 \Delta e}{2}\right)^2\right]\left(\frac{\Delta e}{2}\right)^2. \tag{3.37}$$

The linear response coefficients $\partial \langle j_2 \rangle_{\mathrm{st}}/\partial T_2|_{T_1}$ and $\partial \langle j_2 \rangle_{\mathrm{st}}/\partial T_1|_{T_2}$ are no longer opposite to each other. As a consequence, when $T_1$ and $T_2$ are varied independently, the corresponding variation of the stationary mean instantaneous current $\langle j_2 \rangle_{\mathrm{st}}^{[\beta_1, \beta_2]}$ at first order reads

$$\langle j_2 \rangle_{\mathrm{st}}^{[\beta_1', \beta_2']} - \langle j_2 \rangle_{\mathrm{st}}^{[\beta_1, \beta_2]} \underset{(T_1', T_2') \to (T_1, T_2)}{\sim} (T_1' - T_1)\frac{\partial \langle j_2 \rangle_{\mathrm{st}}^{[\beta_1, \beta_2]}}{\partial T_1}\bigg|_{T_2} + (T_2' - T_2)\frac{\partial \langle j_2 \rangle_{\mathrm{st}}^{[\beta_1, \beta_2]}}{\partial T_2}\bigg|_{T_1}. \tag{3.38}$$

The latter variation depends not only on $T_1$, $T_2$ and the variation of the temperature difference $(T_1' - T_2') - (T_1 - T_2)$ but also on the way in which $T_1'$ and $T_2'$ are varied around the given values $T_1$ and $T_2$.

### 3.3. Nonlinear static response in the NESS

*3.3.1. Nonlinear thermal conductivity.* When the system is far from equilibrium, instead of introducing the linear response $\langle j_2 \rangle_{\mathrm{st}}^{[\beta_1', \beta_2']} - \langle j_2 \rangle_{\mathrm{st}}^{[\beta_1, \beta_2]}$ with $(\beta_1' - \beta_1)\Delta e \ll 1$ and $(\beta_2' - \beta_2)\Delta e \ll 1$ (and the associated linear response coefficients $\partial \langle j_2 \rangle_{\mathrm{st}}^{[\beta_1, \beta_2]}/\partial \beta_a$), one may rather consider a nonlinear thermal conductivity, defined as

$$\kappa_{\mathrm{th}}^{\mathrm{nlin}} = \frac{\langle j_2 \rangle_{\mathrm{st}}^{[\beta_1, \beta_2]}}{T_2 - T_1}. \tag{3.39}$$

From (3.27)

$$\kappa_{\text{th}}^{\text{nlin}} = \frac{1}{T_1 T_2} \frac{\nu_1 \nu_2}{\nu_1 + \nu_2} \frac{\tanh\left((\beta_1 - \beta_2)\Delta e/2\right)}{(\beta_1 - \beta_2)\Delta e/2}$$

$$\times \left[1 - \tanh\left(\frac{\beta_1 \Delta e}{2}\right) \tanh\left(\frac{\beta_2 \Delta e}{2}\right)\right] \left(\frac{\Delta e}{2}\right)^2. \tag{3.40}$$

According to the expression (3.27), $\langle j_2 \rangle_{\text{st}}$ is a bounded function of $T_1$ and $T_2$, so that $\kappa_{\text{th}}^{\text{nlin}}$ vanishes when $T_2 - T_1$ becomes very large with respect to either $T_1$ or $T_2$.

We also notice that when both thermostats are at very high temperature, namely when $\beta_1 \Delta e \ll 1$ and $\beta_2 \Delta e \ll 1$, $\langle j_2 \rangle_{\text{st}}$ is proportional to $\beta_1 - \beta_2$, with a coefficient independent of the temperatures. As a consequence, the partial derivatives $\partial \langle j_2 \rangle_{\text{st}}^{[\beta_1,\beta_2]}/\partial \beta_1$ and $\partial \langle j_2 \rangle_{\text{st}}^{[\beta_1,\beta_2]}/\partial \beta_2$ are opposite to each other, as in the symmetry property (3.29) in the very vicinity of the equilibrium limit. Then the difference (3.38) is proportional to the difference $(\beta_1' - \beta_1) - (\beta_2' - \beta_2)$,

$$\langle j_2 \rangle_{\text{st}}^{[\beta_1',\beta_2']} - \langle j_2 \rangle_{\text{st}}^{[\beta_1,\beta_2]} \underset{\substack{(\beta_1',\beta_2') \to (\beta_1,\beta_2) \\ \beta_1 \Delta e \to 0, \, \beta_2 \Delta e \to 0}}{\sim} \left[(\beta_1' - \beta_1) - (\beta_2' - \beta_2)\right] \left.\frac{\partial \langle j_2 \rangle_{\text{st}}}{\partial \beta_1}\right|_{\beta_2}. \tag{3.41}$$

Besides, the thermal conductivity (3.40) behaves as

$$\kappa_{\text{th}}^{\text{nlin}} \underset{\substack{\beta_1 \Delta e \to 0 \\ \beta_2 \Delta e \to 0}}{\sim} \frac{1}{T_1 T_2} \frac{\nu_1 \nu_2}{\nu_1 + \nu_2} \left(\frac{\Delta e}{2}\right)^2. \tag{3.42}$$

*3.3.2. Housekeeping entropy flow and mean excess heats.* In the long-time limit, whatever the initial configuration probability $P_0$ may be, the system reaches a stationary state where the Markovian stochastic dynamics enforces that the cumulated heats received from each thermostat, namely the random variables $\mathcal{Q}_1(t)$ and $\mathcal{Q}_2(t)$, have averages $\langle \mathcal{Q}_1(t) \rangle_{P_0}$ and $\langle \mathcal{Q}_2(t) \rangle_{P_0}$ which both grow linearly in time with opposite coefficients, $-\langle \mathcal{Q}_1(t) \rangle_{P_0} \underset{t \to +\infty}{\sim} \langle \mathcal{Q}_2(t) \rangle_{P_0} \underset{t \to +\infty}{\sim} t \langle j_2 \rangle_{\text{st}}^{[\beta_1,\beta_2]}$. Then

$$\lim_{t \to +\infty} \frac{\beta_1 \langle \mathcal{Q}_1(t) \rangle_{P_0} + \beta_2 \langle \mathcal{Q}_2(t) \rangle_{P_0}}{t} = -(\beta_1 - \beta_2)\langle j_2 \rangle_{\text{st}}^{[\beta_1,\beta_2]} = \left.\frac{\mathrm{d}_{\text{exch}} S}{\mathrm{d}t}\right|_{\text{st}}, \tag{3.43}$$

where the stationary exchange entropy flow appears by virtue of (3.34). Meanwhile the sum $\mathcal{Q}_1(t) + \mathcal{Q}_2(t)$ remains bounded at any time and its average tends to the heat amount corresponding to the mean energy difference between the final and initial stationary states,

$$\lim_{t \to +\infty} \langle \mathcal{Q}_1(t) + \mathcal{Q}_2(t) \rangle_{P_0} = \langle \mathcal{E} \rangle_{\text{st}}^{[\beta_1,\beta_2]} - \langle \mathcal{E} \rangle_{P_0}. \tag{3.44}$$

In the phenomenological framework of steady state thermodynamics [19], when work is supplied to the system, the total heat *given* to the system is usually expressed as the sum of an 'excess' heat $\mathcal{Q}_{\text{exc}}$ associated with the energy exchange during transitions between two different steady states and a 'housekeeping' heat $\mathcal{Q}_{\text{hk}}$ associated with the energy supplied to maintain the system in the NESS reached in the long-time limit. These two heat amounts have been discussed for a system in contact with only one thermal bath and submitted to a time-dependent external force which is described by Langevin dynamics [20, 14, 21].

By analogy, with the standard sign convention, we may introduce a 'housekeeping' entropy flow supplied to the system, which can be measured as the asymptotic behavior

$$\sigma_{\text{hk}}[P_{\text{st}}] \equiv - \lim_{t \to +\infty} \frac{\beta_1 \langle \mathcal{Q}_1(t) \rangle_{P_0} + \beta_2 \langle \mathcal{Q}_2(t) \rangle_{P_0}}{t}, \tag{3.45}$$

and which, by virtue of (3.43), coincides with the opposite of the stationary exchange entropy flow, namely with the stationary entropy production rate (see (3.34))

$$\sigma_{\text{hk}}[P_{\text{st}}] = - \frac{\mathrm{d}_{\text{exch}} S}{\mathrm{d}t} \bigg|_{\text{st}} = \frac{\mathrm{d}_{\text{int}} S}{\mathrm{d}t} \bigg|_{\text{st}}. \tag{3.46}$$

From the explicit expression (3.27) of the mean instantaneous heat current we obtain the expression for the housekeeping entropy flow (3.45)

$$\sigma_{\text{hk}}[P_{\text{st}}] = \frac{\nu_1 \nu_2}{\nu_1 + \nu_2} (\gamma_1 - \gamma_2) (\beta_1 - \beta_2) \frac{\Delta e}{2}. \tag{3.47}$$

When the system is prepared in a stationary state by thermal contact with heat reservoirs at the inverse temperatures $\beta_1^0$ and $\beta_2^0$ respectively, then $\langle \mathcal{E} \rangle_{P_0} = \langle \mathcal{E} \rangle_{\text{st}}^{[\beta_1^0, \beta_2^0]}$ and the difference in (3.44) becomes equal to $\langle \mathcal{E} \rangle_{\text{st}}^{[\beta_1, \beta_2]} - \langle \mathcal{E} \rangle_{\text{st}}^{[\beta_1^0, \beta_2^0]}$. With the standard convention, the 'excess' heats *given* to the system $Q_{\text{exc},a}\big|_{[\beta_1^0, \beta_2^0]}^{[\beta_1, \beta_2]}$ with $a = 1, 2$ can be measured as

$$Q_{\text{exc},a}\big|_{[\beta_1^0, \beta_2^0]}^{[\beta_1, \beta_2]} \equiv - \lim_{t \to +\infty} \left[ \langle \mathcal{Q}_a(t) \rangle_{P_{\text{st}}^{[\beta_1^0, \beta_2^0]}} - t \langle j_a \rangle_{\text{st}}^{[\beta_1, \beta_2]} \right]. \tag{3.48}$$

Then, by virtue of the stationary condition $\langle j_1 \rangle_{\text{st}} = - \langle j_2 \rangle_{\text{st}}$, the equality (3.44) becomes

$$-Q_{\text{exc},1}\big|_{[\beta_1^0, \beta_2^0]}^{[\beta_1, \beta_2]} - Q_{\text{exc},2}\big|_{[\beta_1^0, \beta_2^0]}^{[\beta_1, \beta_2]} = \langle \mathcal{E} \rangle_{\text{st}}^{[\beta_1, \beta_2]} - \langle \mathcal{E} \rangle_{\text{st}}^{[\beta_1^0, \beta_2^0]}. \tag{3.49}$$

The excess heats $Q_{\text{exc},a}\big|_{[\beta_1^0, \beta_2^0]}^{[\beta_1, \beta_2]}$ defined in (3.48) are explicitly calculated in subsection (4.3) from the expressions of the average heat amounts $\langle \mathcal{Q}_a(t) \rangle_{P_0}$ at any finite time $t$ (for any initial distribution $P_0$ of the two-spin configuration) with the results given in (4.52).

In the linear response regime, where the relative differences $(T_1 - T_1^0)/T_1^0$ and $(T_2 - T_2^0)/T_2^0$ are infinitesimal, by virtue of the definition (3.19) of the generalized heat capacities $C_{\text{st}}^{[a]}(T_1, T_2)$, with $a = 1, 2$,

$$-Q_{\text{exc},1}\big|_{[T_1, T_2]}^{[T_1 + \mathrm{d}T_1, T_2 + \mathrm{d}T_2]} - Q_{\text{exc},2}\big|_{[T_1, T_2]}^{[T_1 + \mathrm{d}T_1, T_2 + \mathrm{d}T_2]} \to C_{\text{st}}^{[1]}(T_1, T_2) \, \mathrm{d}T_1 + C_{\text{st}}^{[2]}(T_1, T_2) \, \mathrm{d}T_2. \tag{3.50}$$

We notice that the notion of heat capacity has been studied in the case of non-equilibrium steady states where the system is submitted to a non-conservative force and is in contact with a single thermostat [22].

## 4. Joint probability distribution for heat cumulated exchanges at finite time in the model

Instead of studying the evolution of the probability distribution $P(\mathcal{C}; t)$ of the spin configuration $\mathcal{C} = (\sigma_1, \sigma_2)$, we address directly the evolution of the joint probability

distribution $P(\mathcal{C}'|\mathcal{Q}_1, \mathcal{Q}_2, t|\mathcal{C})$ for the cumulated heats $\mathcal{Q}_1$ and $\mathcal{Q}_2$ received from the thermal baths 1 and 2 during a time $t$ when the system is in configuration $\mathcal{C} = (\sigma_1, \sigma_2)$ at time $t_0 = 0$ and in configuration $\mathcal{C}' = (\sigma'_1, \sigma'_2)$ at time $t$. In order to obtain results which hold as generally as possible, the initial probability distribution for configurations is not assumed to have the same symmetry under simultaneous spin flips as the stationary distribution.

Since the two-spin system has only two energy levels separated by the energy gap $\Delta e$, the cumulated heats $\mathcal{Q}_a$ are integer multiples of $\Delta e$ and we set

$$\mathcal{Q}_1 = -n_1 \Delta e \qquad \text{and} \qquad \mathcal{Q}_2 = n_2 \Delta e. \tag{4.1}$$

The minus sign in the definition of $\mathcal{Q}_1$ is introduced for the sake of conveniency, because the mean instantaneous heat currents $\langle j_1 \rangle_{\text{st}}$ and $\langle j_2 \rangle_{\text{st}}$ in the stationary state are opposite to each other. In other words, $n_1 \Delta e$ is the amount of heat dissipated towards heat bath 1, while $n_2 \Delta e$ is the amount of heat received from heat bath 2. With these notations $P(\mathcal{C}_{\text{f}}|\mathcal{Q}_1, \mathcal{Q}_2, t|\mathcal{C}_0)$ can be written as a matrix element of some evolution operator $\mathbb{U}(n_1, n_2; t)$ as

$$P(\mathcal{C}'|\mathcal{Q}_1, \mathcal{Q}_2, t|\mathcal{C}) = (\sigma'_1, \sigma'_2|\mathbb{U}(n_1, n_2; t)|\sigma_1, \sigma_2). \tag{4.2}$$

## 4.1. Explicit calculations

*4.1.1. Constraint from energy conservation.* According to the expression (2.5) for the interaction energy between the two spins, the energy difference between the final and the initial configurations reads

$$\mathcal{E}(\sigma'_1, \sigma'_2) - \mathcal{E}(\sigma_1, \sigma_2) = \frac{\sigma_1 \sigma_2 - \sigma'_1 \sigma'_2}{2} \Delta e, \tag{4.3}$$

and it can take only three values $0$, $+\Delta e$ and $-\Delta e$. On the other hand, according to (4.1), $\mathcal{Q}_1 + \mathcal{Q}_2 = (n_2 - n_1)\Delta e$, namely

$$\mathcal{Q}_1 + \mathcal{Q}_2 = \Delta n \times \Delta e \qquad \text{where } \Delta n \equiv n_2 - n_1. \tag{4.4}$$

Energy conservation entails that the energy variation of the two-spin system is equal to the sum of the heat amounts received from the thermostats: $\mathcal{E}(\sigma'_1, \sigma'_2) - \mathcal{E}(\sigma_1, \sigma_2) = \mathcal{Q}_1 + \mathcal{Q}_2$. As a consequence, the correspondence between the total amount of received heat and the couple of initial and final states reads

$$\begin{aligned} \Delta n = 0 \qquad &\Leftrightarrow \qquad \sigma'_1 \sigma'_2 = \sigma_1 \sigma_2 \\ (\Delta n)^2 = 1 \qquad &\Leftrightarrow \qquad \sigma_1 \sigma_2 = \Delta n \qquad \text{and} \qquad \sigma'_1 \sigma'_2 = -\Delta n. \end{aligned} \tag{4.5}$$

Therefore it is convenient to introduce the decomposition

$$\mathbb{U}(n_1, n_2; t) = \sum_{\Delta n = 0, +1, -1} \delta_{n_2, n_1 + \Delta n} \mathbb{U}(n_1, n_1 + \Delta n; t). \tag{4.6}$$

In the basis $\{(+, +), (-, -), (+, -), (-+)\}$, already used in (3.2), the correspondence (4.5) enforces that $\mathbb{U}(n_1, n_2; t)$ can be decomposed into three $4 \times 4$ matrices

$$\mathbb{U} = \mathbb{U}_{[\Delta n = 0]} + \mathbb{U}_{[\Delta n = +1]} + \mathbb{U}_{[\Delta n = -1]} \tag{4.7}$$

with

$$\mathbb{U}_{[\Delta n=0]} = \begin{pmatrix} \mathbb{A} & \mathbb{O} \\ \mathbb{O} & \mathbb{D} \end{pmatrix} \qquad \mathbb{U}_{[\Delta n=1]} = \begin{pmatrix} \mathbb{O} & \mathbb{O} \\ \mathbb{C} & \mathbb{O} \end{pmatrix} \qquad \mathbb{U}_{[\Delta n=-1]} = \begin{pmatrix} \mathbb{O} & \mathbb{B} \\ \mathbb{O} & \mathbb{O} \end{pmatrix}. \tag{4.8}$$

The subscript involving $\Delta n$ indicates the unique value of $\Delta n$ which is involved in a history where the initial and final states are $(\sigma_1, \sigma_2)$ and $(\sigma_1', \sigma_2')$ respectively. $\mathbb{O} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ and $\mathbb{A}$, $\mathbb{B}$, $\mathbb{C}$ and $\mathbb{D}$ are $2 \times 2$ matrices.

*4.1.2. Generating function method.* The evolution equation for $\mathbb{U}(n_1, n_2; t)$ is easily derived by considering the probability

$$P(\sigma_1, \sigma_2, n_1, n_2; t) = \sum_{\sigma_1', \sigma_2'} (\sigma_1, \sigma_2 | \mathbb{U}(n_1, n_2; t) | \sigma_1', \sigma_2') \, P(\sigma_1', \sigma_2'; t=0). \tag{4.9}$$

$P(\sigma_1, \sigma_2, n_1, n_2; t)$ is the probability that the system is in configuration $(\sigma_1, \sigma_2)$ at time $t$ and has received the heat amounts $\mathcal{Q}_1 = -n_1 \Delta e$ and $\mathcal{Q}_2 = n_2 \Delta e$ during the time interval $[0, t]$ when the initial probability distribution for the spins is $P(\sigma_1', \sigma_2'; t=0)$. The evolution equation for $P(\sigma_1, \sigma_2, n_1, n_2; t)$ is a generalization of the master equation (3.1) which governs the evolution of $P(\sigma_1, \sigma_2; t)$. By taking into account the explicit expression (2.9) for the transition rates we get

$$\begin{aligned} \frac{2}{\nu_1 + \nu_2} \frac{\mathrm{d}P(\sigma_1, \sigma_2, n_1, n_2; t)}{\mathrm{d}t} = & - \left[1 - \sigma_1\sigma_2(\overline{\nu}_1\gamma_1 + \overline{\nu}_2\gamma_2)\right] P(\sigma_1, \sigma_2, n_1, n_2; t) \\ & + \overline{\nu}_1 \left[1 + \sigma_1\sigma_2\gamma_1\right] P(-\sigma_1, \sigma_2, n_1 - \sigma_1\sigma_2, n_2; t) \\ & + \overline{\nu}_2 \left[1 + \sigma_1\sigma_2\gamma_2\right] P(\sigma_1, -\sigma_2, n_1, n_2 + \sigma_1\sigma_2; t) \end{aligned} \tag{4.10}$$

where the dimensionless inverse time scales $\overline{\nu}_a$ are defined in (3.4).

The operator on the right-hand side of the evolution equation (4.10) is partially diagonalized by considering the generating function $P(\sigma_1, \sigma_2, z_1, z_2; t) = \sum_{n_1=-\infty}^{+\infty} \sum_{n_2=-\infty}^{+\infty} z_1^{n_1} z_2^{n_2} P(\sigma_1, \sigma_2, n_1, n_2; t)$, which is absolutely convergent for $z_1$ and $z_2$ of modulus 1. Considering the latter generating function is equivalent to introducing

$$\widehat{\mathbb{U}}(z_1, z_2; t) \equiv \sum_{n_1=-\infty}^{+\infty} \sum_{n_2=-\infty}^{+\infty} z_1^{n_1} z_2^{n_2} \mathbb{U}(n_1, n_2; t). \tag{4.11}$$

Since $P(\sigma_1, \sigma_2, n_1, n_2; t=0) = \delta_{n_1,0} \delta_{n_2,0} P(\sigma_1, \sigma_2; t=0)$, we infer that $\widehat{\mathbb{U}}(z_1, z_2; t=0) = \mathbb{I}_4$, where $\mathbb{I}_4$ denotes the identity $4 \times 4$ matrix. The inversion formula which allows one to retrieve $\mathbb{U}(n_1, n_2; t)$ is

$$\mathbb{U}(n_1, n_2; t) = \oint_{|z_1|=1} \frac{\mathrm{d}z_1}{2\pi\mathrm{i}} \oint_{|z_2|=1} \frac{\mathrm{d}z_2}{2\pi\mathrm{i}} \frac{\widehat{\mathbb{U}}(z_1, z_2; t)}{z_1^{n_1+1} z_2^{n_2+1}}. \tag{4.12}$$

The decomposition (4.6) of $\mathbb{U}(n_1, n_2; t)$ leads to a similar decomposition for $\widehat{\mathbb{U}}(z_1, z_2; t)$

$$\widehat{\mathbb{U}}(z_1, z_2; t) = \sum_{\Delta n=0,+1,-1} \widehat{\mathbb{U}}_{[\Delta n]}(z_1, z_2; t). \tag{4.13}$$

The decomposition (4.7) of $\mathbb{U}(n_1, n_2; t)$ into three $4 \times 4$ matrices (enforced by the constraint (4.5) due to energy conservation) is also valid for $\widehat{\mathbb{U}}(z_1, z_2; t)$. Moreover $\widehat{\mathbb{U}}_{[\Delta n]}(z_1, z_2; t)$ has

necessarily the following dependence upon $z_2$ and $z_1 z_2$: $\widehat{\mathbb{U}}_{[\Delta n]}(z_1, z_2; t) = z_2^{\Delta n} \widehat{\mathbb{V}}_{[\Delta n]}(z_1 z_2; t)$.
Therefore, by using the change of variable $z_1 \longrightarrow z = z_1 z_2$ in (4.12) one gets

$$\mathbb{U}(n_1, n_1 + \Delta n; t) = \oint_{|z|=1} \frac{\mathrm{d}z}{2\pi \mathrm{i}} \frac{\widehat{\mathbb{U}}_{[\Delta n]}(z_1 = z, z_2 = 1; t)}{z^{n_1+1}}. \tag{4.14}$$

*4.1.3. Diagonalization.* The evolution of $\widehat{\mathbb{U}}(z_1, z_2; \tau)$ with the dimensionless time variable

$$\tau = \frac{\nu_1 + \nu_2}{2} t \tag{4.15}$$

reads

$$\frac{\mathrm{d}\widehat{\mathbb{U}}(z_1, z_2; \tau)}{\mathrm{d}\tau} = \mathbb{A}(z_1, z_2)\widehat{\mathbb{U}}(z_1, z_2; \tau), \tag{4.16}$$

where, from the evolution equation (4.10),

$$\mathbb{A}(z_1, z_2) = -\mathbb{I}_4 + \begin{pmatrix} \gamma_\star & 0 & b_2 & b_1 \\ 0 & \gamma_\star & b_1 & b_2 \\ c_2 & c_1 & -\gamma_\star & 0 \\ c_1 & c_2 & 0 & -\gamma_\star \end{pmatrix} \tag{4.17}$$

with the following notations: $b_1 = \bar{\nu}_1(1+\gamma_1)z_1$, $b_2 = \bar{\nu}_2(1+\gamma_2)(1/z_2)$, $c_1 = \bar{\nu}_1(1-\gamma_1)(1/z_1)$ and $c_2 = \bar{\nu}_2(1-\gamma_2)z_2$.

Since the transition rates are invariant under the simultaneous changes of both spin signs, it is convenient to consider the transformed matrix

$$\mathbb{A}'(z_1, z_2) = \mathbb{P}^{-1}\mathbb{A}(z_1, z_2)\mathbb{P} \tag{4.18}$$

with

$$\mathbb{P}^{-1} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}. \tag{4.19}$$

The matrix $\mathbb{A}'(z_1, z_2)$ corresponds to two sets of decoupled equations,

$$\mathbb{A}'(z_1, z_2) = -\mathbb{I}_4 + \begin{pmatrix} \mathbb{B}_+ & \mathbb{O} \\ \mathbb{O} & \mathbb{B}_- \end{pmatrix} \tag{4.20}$$

where $\mathbb{B}_\epsilon = \begin{pmatrix} \gamma_\star & \epsilon b_1 + b_2 \\ \epsilon c_1 + c_2 & -\gamma_\star \end{pmatrix}$ for $\epsilon = \pm$. As $\mathbb{B}_\epsilon$ is traceless, $\mathbb{B}_\epsilon^2(z_1, z_2)$ is proportional to $\mathbb{I}_2$.
Explicitly

$$\mathbb{B}_\epsilon^2(z_1, z_2) = \Delta_\epsilon(z_1 z_2)\mathbb{I}_2 \tag{4.21}$$

with

$$\Delta_\epsilon(z) = 1 - 2A + \epsilon\left[(A+B)z + (A-B)\frac{1}{z}\right] \tag{4.22}$$

where

$$A = \bar{\nu}_1 \bar{\nu}_2 \left(1 - \gamma_1 \gamma_2\right) \qquad \text{and} \qquad B = \bar{\nu}_1 \bar{\nu}_2 \left(\gamma_1 - \gamma_2\right). \tag{4.23}$$

We notice that $A \geq |B|$. As a consequence,

$$e^{\tau \mathbb{A}'(z_1, z_2)} = e^{-\tau} \times \begin{pmatrix} e^{\tau \mathbb{B}_+(z_1, z_2)} & \mathbb{O} \\ \mathbb{O} & e^{\tau \mathbb{B}_-(z_1, z_2)} \end{pmatrix} \tag{4.24}$$

where

$$e^{\tau \mathbb{B}_\epsilon(z_1, z_2)} = \cosh(\tau \sqrt{\Delta_\epsilon(z_1 z_2)}) \mathbb{I}_2 + \frac{\sinh(\tau \sqrt{\Delta_\epsilon(z_1 z_2)})}{\sqrt{\Delta_\epsilon(z_1 z_2)}} \mathbb{B}_\epsilon(z_1, z_2). \tag{4.25}$$

Moreover the eigenvalues of the matrix $\frac{1}{2}(\nu_1 + \nu_2)\mathbb{A}(z_1, z_2)$ are, with the notations $\epsilon = \pm$ and $\eta = \pm$,

$$\mu^{(\epsilon, \eta)}(z) = \frac{\nu_1 + \nu_2}{2} \left[ -1 + \eta \sqrt{\Delta_\epsilon(z)} \right]. \tag{4.26}$$

*4.1.4. Results for the generating function.* From (4.18) we can calculate $\widehat{\mathbb{U}}(z_1, z_2; \tau) = e^{\tau \mathbb{A}(z_1, z_2)} = \mathbb{P} e^{\tau \mathbb{A}'(z_1, z_2)} \mathbb{P}^{-1}$. From the explicit expressions (4.24) and (4.25) we get the matrices $\widehat{\mathbb{U}}_{[\Delta n]}(z_1, z_2; \tau)$ defined in (4.13). The 16 matrix elements can be written in the compact form

$$\begin{aligned} (\sigma_1, \sigma_2 | \widehat{\mathbb{U}}_{[\Delta n = 0]}(z_1, z_2; \tau) | \sigma_1, \sigma_2) &= C^+(z_1 z_2; \tau) + \sigma_1 \sigma_2 \gamma_\star S^+(z_1 z_2; \tau) \\ (-\sigma_1, -\sigma_2 | \widehat{\mathbb{U}}_{[\Delta n = 0]}(z_1, z_2; \tau) | \sigma_1, \sigma_2) &= C^-(z_1 z_2; \tau) + \sigma_1 \sigma_2 \gamma_\star S^-(z_1 z_2; \tau) \\ (-\sigma_1, \sigma_2 | \widehat{\mathbb{U}}_{[\Delta n]}(z_1, z_2; \tau) | \sigma_1, \sigma_2) &\underset{\Delta n = \pm 1}{=} \delta_{\sigma_1 \sigma_2, \Delta n} z_2^{\Delta n} F_{\Delta n}^+(z_1 z_2; \tau) \\ (\sigma_1, -\sigma_2 | \widehat{\mathbb{U}}_{[\Delta n]}(z_1, z_2; \tau) | \sigma_1, \sigma_2) &\underset{\Delta n = \pm 1}{=} \delta_{\sigma_1 \sigma_2, \Delta n} z_2^{\Delta n} F_{\Delta n}^-(z_1 z_2; \tau) \end{aligned} \tag{4.27}$$

with

$$C^\pm(z; \tau) = e^{-\tau} \frac{1}{2} \left[ \cosh(\tau \sqrt{\Delta_+(z)}) \pm \cosh(\tau \sqrt{\Delta_-(z)}) \right] \tag{4.28}$$

$$S^\pm(z; \tau) = e^{-\tau} \frac{1}{2} \left[ \frac{\sinh(\tau \sqrt{\Delta_+(z)})}{\sqrt{\Delta_+(z)}} \pm \frac{\sinh(\tau \sqrt{\Delta_-(z)})}{\sqrt{\Delta_-(z)}} \right] \tag{4.29}$$

and

$$F_{\Delta n}^\pm(z; \tau) = \frac{1}{z^{\Delta n}} \bar{\nu}_1 \left(1 - \Delta n\, \gamma_1\right) S^\pm(z; \tau) + \bar{\nu}_2 \left(1 - \Delta n\, \gamma_2\right) S^\mp(z; \tau). \tag{4.30}$$

*4.1.5. Results for the joint probability.* $\mathbb{U}(n_1, n_1 + \Delta n; t)$ is derived from $\widehat{\mathbb{U}}_{[\Delta n]}(z_1, z_2; t)$ through the unit circle integral in (4.14). In fact, because of the parity property of the cosh and sinh functions, the functions $C^\pm(z; \tau)$ and $S^\pm(z; \tau)$ are functions not of $\sqrt{\Delta_\epsilon(z)}$ but only of $\Delta_\epsilon(z)$. According to the definition (4.22) of $\Delta_\epsilon(z)$, the only singular points of $\Delta_\epsilon(z)$ are $z = 0$ and $+\infty$, and the same is true for the integrands in $\oint_{|z|=1} (\mathrm{d}z/2\pi\mathrm{i}) z^{-(m+1)} C^\pm(z; \tau)$

and $\oint_{|z|=1}(\mathrm{d}z/2\pi\mathrm{i})z^{-(m+1)}S^{\pm}(z;\tau)$. Changing $z$ into $-z$ we get that

$$\oint_{|z|=1}\frac{\mathrm{d}z}{2\pi\mathrm{i}}\frac{1}{z^{m+1}}\cosh\left(\tau\sqrt{\Delta_-(z)}\right) = (-1)^m\oint_{|z|=1}\frac{\mathrm{d}z}{2\pi\mathrm{i}}\frac{1}{z^{m+1}}\cosh\left(\tau\sqrt{\Delta_+(z)}\right). \qquad (4.31)$$

As a consequence

$$\oint_{|z|=1}\frac{\mathrm{d}z}{2\pi\mathrm{i}}\frac{1}{z^{m+1}}C^{\pm}(z;\tau) = \frac{1\pm(-1)^m}{2}\mathrm{e}^{-\tau}c_m(\tau) \qquad (4.32)$$

and

$$\oint_{|z|=1}\frac{\mathrm{d}z}{2\pi\mathrm{i}}\frac{1}{z^{m+1}}S^{\pm}(z;\tau) = \frac{1\pm(-1)^m}{2}\mathrm{e}^{-\tau}s_m(\tau), \qquad (4.33)$$

where

$$c_m(\tau) \equiv \oint_{|z|=1}\frac{\mathrm{d}z}{2\pi\mathrm{i}}\frac{1}{z^{m+1}}\cosh\left(\tau\sqrt{\Delta_+(z)}\right) \qquad (4.34)$$

and

$$s_m(\tau) \equiv \oint_{|z|=1}\frac{\mathrm{d}z}{2\pi\mathrm{i}}\frac{1}{z^{m+1}}\frac{\sinh(\tau\sqrt{\Delta_+(z)})}{\sqrt{\Delta_+(z)}}. \qquad (4.35)$$

Eventually, the matrix elements of $\mathbb{U}(n_1, n_1 + \Delta n; t)$ are derived from the expressions (4.27) for the matrix elements of $\widehat{\mathbb{U}}_{[\Delta n]}(z_1, z_2; t)$ with the result

$$
\begin{aligned}
(\sigma_1, \sigma_2|\mathbb{U}(n,n;t)|\sigma_1,\sigma_2) &= \delta_{\mathrm{even}}(n)\times U^{(0)}(n,\sigma_1\sigma_2;\tau)\\
(-\sigma_1,-\sigma_2|\mathbb{U}(n,n;t)|\sigma_1,\sigma_2) &= \delta_{\mathrm{odd}}(n)\times U^{(0)}(n,\sigma_1\sigma_2;\tau)\\
(-\sigma_1,\sigma_2|\mathbb{U}(n,n+\Delta n;t)|\sigma_1,\sigma_2) &\underset{\Delta n=\pm1}{=} \delta_{\sigma_1\sigma_2,\Delta n}\times\delta_{\mathrm{odd}}(n)\times U^{(1)}(n,\Delta n;\tau)\\
(\sigma_1,-\sigma_2|\mathbb{U}(n,n+\Delta n;t)|\sigma_1,\sigma_2) &\underset{\Delta n=\pm1}{=} \delta_{\sigma_1\sigma_2,\Delta n}\times\delta_{\mathrm{even}}(n)\times U^{(1)}(n,\Delta n;\tau),
\end{aligned}
\qquad (4.36)
$$

where $\delta_{\mathrm{even}}(n) = \frac{1}{2}\left[1+(-1)^n\right]$ and $\delta_{\mathrm{odd}}(n) = \frac{1}{2}\left[1-(-1)^n\right]$ while

$$
\begin{aligned}
U^{(0)}(n,\sigma_1\sigma_2;\tau) &= \mathrm{e}^{-\tau}\left[c_n(\tau)+\gamma_\star\sigma_1\sigma_2 s_n(\tau)\right]\\
U^{(1)}(n,\Delta n;\tau) &= \mathrm{e}^{-\tau}\left[\overline{\nu}_2\left(1-\gamma_2\Delta n\right)s_n(\tau)+\overline{\nu}_1\left(1-\gamma_1\Delta n\right)s_{n+\Delta n}(\tau)\right].
\end{aligned}
\qquad (4.37)
$$

We notice that the parity condition factors $\frac{1}{2}\left[1\pm(-1)^n\right]$ have a simple interpretation. During a history such that spin $\sigma_1$ is in the same state (in flipped states) in the initial and final configurations, thermal bath 1 has flipped spin $\sigma$ an even (odd) number of times, so that the corresponding sum $n\Delta e$ of the successive amounts $\pm\Delta e$ dissipated towards thermal bath 1 is necessarily an even (odd) multiple of $\Delta e$.

## 4.2. Various explicit probabilities

The probability that the system is in configuration $(\sigma_1,\sigma_2)$ at time $t$ when the initial configuration is distributed according to the law $P_0$ can be calculated, by virtue of the definition (4.11), as

$$P_{P_0}(\sigma_1,\sigma_2;t) = \sum_{\Delta n=-1,0,+1}\sum_{\sigma_1',\sigma_2'}\left(\sigma_1,\sigma_2|\widehat{\mathbb{U}}_{[\Delta n]}(z_1=1,z_2=1;t)|\sigma_1',\sigma_2'\right)P_0(\sigma_1',\sigma_2'). \qquad (4.38)$$

The matrix elements of $\widehat{\mathbb{U}}_{[\Delta n]}(z_1 = 1, z_2 = 1; t)$ are derived from (4.27) where, according to (4.22), $\Delta_+(z = 1) = 1$ while $\Delta_-(z = 1) = 1 - 4A$ with $A = \bar{\nu}_1 \bar{\nu}_2 (1 - \gamma_1 \gamma_2)$. Using the identities $P_0(\sigma_1, \sigma_2) + P_0(-\sigma_1, -\sigma_2) = \frac{1}{2}[1 + \sigma_1 \sigma_2 \gamma_0]$, where $\gamma_0 = \langle \sigma_1 \sigma_2 \rangle_{P_0}$, and $P_0(\sigma_1, \sigma_2) - P_0(-\sigma_1, -\sigma_2) = \frac{1}{2}[\sigma_1 \langle \sigma_1' \rangle_{P_0} + \sigma_2 \langle \sigma_2' \rangle_{P_0}]$, where $\langle \sigma_1' \rangle_{P_0}$ (resp. $\langle \sigma_2' \rangle_{P_0}$) is the average value of the first (resp. second) spin at time 0, a straightforward calculation leads to

$$
\begin{aligned}
P_{P_0}(\sigma_1, \sigma_2; t) = {} & \tfrac{1}{4}[1 + \sigma_1 \sigma_2 \gamma_\star] + \tfrac{1}{4}\sigma_1 \sigma_2 [\gamma_0 - \gamma_\star] e^{-2\tau} \\
& + \tfrac{1}{8}\left[\sigma_1 \langle \sigma_1' \rangle_{P_0} + \sigma_2 \langle \sigma_2' \rangle_{P_0}\right]\left[\left(1 + \sigma_1 \sigma_2 \frac{\gamma_\star}{\alpha}\right) e^{-(1-\alpha)\tau}\right. \\
& + \left.\left(1 - \sigma_1 \sigma_2 \frac{\gamma_\star}{\alpha}\right) e^{-(1+\alpha)\tau}\right] + \tfrac{1}{8}\left[-\sigma_1 \langle \sigma_1' \rangle_{P_0} + \sigma_2 \langle \sigma_2' \rangle_{P_0}\right] \\
& \times \frac{1}{\alpha}\left[\bar{\nu}_1 - \bar{\nu}_2 + \sigma_1 \sigma_2 (\bar{\nu}_1 \gamma_1 - \bar{\nu}_2 \gamma_2)\right]\left[e^{-(1-\alpha)\tau} - e^{-(1+\alpha)\tau}\right],
\end{aligned}
\tag{4.39}
$$

where $\alpha \equiv \sqrt{1 - 4A}$. When $P_0$ is invariant under the simultaneous reversal of the spins $\sigma_1$ and $\sigma_2$, $P_0(\sigma_1, \sigma_2) = (1/4)\left[1 + \sigma_1 \sigma_2 \gamma_0\right]$ and only the terms in the first line of (4.39) contribute. Then the evolution of $P_{P_0}(\sigma_1, \sigma_2; t)$ towards the stationary distribution $P_{\mathrm{st}}(\sigma_1, \sigma_2) = \frac{1}{4}\left[1 + \sigma_1 \sigma_2 \gamma_\star\right]$ involves only one time scale, namely $1/(\nu_1 + \nu_2)$ (recall that $\tau = \frac{1}{2}(\nu_1 + \nu_2)t$).

For any initial probability distribution $P_0$ of the spins, the probability $P_{P_0}(\sigma_1, \sigma_2, n_1, n_2; t)$ that at time $t$ the system is in configuration $(\sigma_1, \sigma_2)$ and has received a heat amount $\mathcal{Q}_1 = -n_1 \Delta e$ from bath 1 and $\mathcal{Q}_2 = n_2 \Delta e$ from bath 2 is calculated from (4.36) with the result

$$
\begin{aligned}
P_{P_0}(\sigma_1, \sigma_2, n, n; t) = {} & U^{(0)}(n, \sigma_1 \sigma_2; \tau)\left[\delta_{\mathrm{even}}(n) P_0(\sigma_1, \sigma_2) + \delta_{\mathrm{odd}}(n) P_0(-\sigma_1, -\sigma_2)\right] \\
P_{P_0}(\sigma_1, \sigma_2, n, n + \Delta n; t) \underset{\Delta n = \pm 1}{=} {} & \delta_{\sigma_1 \sigma_2, -\Delta n} U^{(1)}(n, \Delta n; \tau)\left[\delta_{\mathrm{even}}(n) P_0(\sigma_1, -\sigma_2)\right. \\
& \left. + \delta_{\mathrm{odd}}(n) P_0(-\sigma_1, \sigma_2)\right].
\end{aligned}
\tag{4.40}
$$

Various joint probabilities can be derived from these expressions.

The joint probability $P_{P_0}(n_1, n_2; t)$ that at time $t$ the system has received a heat amount $\mathcal{Q}_1 = -n_1 \Delta e$ from bath 1 and a heat amount $\mathcal{Q}_2 = n_2 \Delta e$ from bath 2 is

$$
\begin{aligned}
P_{P_0}(n, n; \tau) = {} & e^{-\tau}\left[c_n(\tau) + \gamma_\star \gamma_0 s_n(\tau)\right] \\
P_{P_0}(n, n + \Delta n; \tau) \underset{\Delta n = \pm 1}{=} {} & e^{-\tau} \tfrac{1}{2}[1 + \Delta n\, \gamma_0] \\
& \times \left[\bar{\nu}_2 \left(1 - \Delta n\, \gamma_2\right) s_n(\tau) + \bar{\nu}_1 \left(1 - \Delta n\, \gamma_1\right) s_{n+\Delta n}(\tau)\right],
\end{aligned}
\tag{4.41}
$$

where $\gamma_0$ has been defined before (4.39).

The joint probability $P_{P_0}(\sigma_1 \sigma_2 = \pm 1, n_1; t)$ that at time $t$ the system is in a configuration where $\sigma_1 \sigma_2$ is equal to $\pm 1$ and has received a heat amount $\mathcal{Q}_1 = -n_1 \Delta e$ from bath 1 is seen via (4.40) to have the value

$$
\begin{aligned}
P_{P_0}(\sigma_1 \sigma_2 = +1, n_1; \tau) = {} & \tfrac{1}{2}e^{-\tau}\left\{(1 + \gamma_0)\, c_{n_1}(\tau) + \left[2\gamma_\star + (\bar{\nu}_2 - \bar{\nu}_1 \gamma_1)(1 - \gamma_0)\right] s_{n_1}(\tau)\right. \\
& \left. + \bar{\nu}_1 (1 + \gamma_1)(1 - \gamma_0)\, s_{n_1 - 1}(\tau)\right\}
\end{aligned}
\tag{4.42}
$$

and

$$
\begin{aligned}
P_{P_0}(\sigma_1 \sigma_2 = -1, n_1; \tau) = {} & \tfrac{1}{2}e^{-\tau}\left\{(1 - \gamma_0)\, c_{n_1}(\tau) + \left[-2\gamma_\star + (\bar{\nu}_2 + \bar{\nu}_1 \gamma_1)(1 + \gamma_0)\right] s_{n_1}(\tau)\right. \\
& \left. + \bar{\nu}_1 (1 - \gamma_1)(1 + \gamma_0)\, s_{n_1 + 1}(\tau)\right\}.
\end{aligned}
\tag{4.43}
$$

The expressions for $P_{P_0}(\sigma_1\sigma_2 = \pm1, n_2; \tau)$ are obtained from the latter equations by making the exchanges $\overline{\nu}_1 \leftrightarrow \overline{\nu}_2$ and $\gamma_1 \leftrightarrow \gamma_2$ and the replacements $c_{n_1} \to c_{n_2}$, $s_{n_1} \to s_{n_2}$, and $s_{n_1-1} \to s_{n_2+1}$ for $\sigma_1\sigma_2 = 1$, resp. $s_{n_1+1} \to s_{n_2-1}$ for $\sigma_1\sigma_2 = -1$.

From these expressions we get the probability distribution for only one heat amount $\mathcal{Q}_1$ or $\mathcal{Q}_2$

$$P_{P_0}(n_1; \tau) = \mathrm{e}^{-\tau}\Big\{ c_{n_1}(\tau) + [\overline{\nu}_2 + \overline{\nu}_1\gamma_1\gamma_0]\, s_{n_1}(\tau)$$

$$+ \tfrac{1}{2}\overline{\nu}_1(1-\gamma_1)(1+\gamma_0)s_{n_1+1}(\tau) + \tfrac{1}{2}\overline{\nu}_1(1+\gamma_1)(1-\gamma_0)s_{n_1-1}(\tau) \Big\} \tag{4.44}$$

and similarly

$$P_{P_0}(n_2; \tau) = \mathrm{e}^{-\tau}\Big\{ c_{n_2}(\tau) + [\overline{\nu}_1 + \overline{\nu}_2\gamma_2\gamma_0]\, s_{n_2}(\tau)$$

$$+ \tfrac{1}{2}\overline{\nu}_2(1+\gamma_2)(1-\gamma_0)s_{n_2+1}(\tau) + \tfrac{1}{2}\overline{\nu}_2(1-\gamma_2)(1+\gamma_0)s_{n_2-1}(\tau) \Big\}. \tag{4.45}$$

From the identities $\sum_{n=-\infty}^{+\infty} c_n(\tau) = \cosh\tau$, and $\sum_{n=-\infty}^{+\infty} s_n(\tau) = \sinh\tau$, the probability that the total heat amount received from both thermostats is $\mathcal{Q}_1 + \mathcal{Q}_2 = (n_2 - n_1)\Delta e$ reads

$$P_{P_0}(n_2 - n_1 = 0; \tau) = \tfrac{1}{2}\left[1 + \gamma_\star\gamma_0\right] + \tfrac{1}{2}\left[1 - \gamma_\star\gamma_0\right]\mathrm{e}^{-2\tau}$$

$$P_{P_0}(n_2 - n_1 = \Delta n; \tau) \underset{\Delta n = \pm 1}{=} \tfrac{1}{4}\left\{1 - \gamma_\star\gamma_0 + \Delta n\left[\gamma_0 - \gamma_\star\right]\right\}\left[1 - \mathrm{e}^{-2\tau}\right]. \tag{4.46}$$

As a consequence

$$\langle[\mathcal{Q}_1 + \mathcal{Q}_2]_\tau\rangle_{P_0} = \tfrac{1}{2}\left[\gamma_0 - \gamma_\star\right]\left[1 - \mathrm{e}^{-2\tau}\right]\Delta e, \tag{4.47}$$

and we retrieve property (3.44).

We notice that all formulae are still valid in the limit where $T_1$ vanishes, namely where $\beta_1\Delta e$ goes to infinity.

### 4.3. Excess heats

The excess heats associated with the transition between two different steady states have been defined in (3.48). For the two-spin system they can be explicitly calculated. Indeed, when the system is initially prepared in the stationary state with distribution $P_0$ by contact with two thermostats at the inverse temperatures $\beta_1^0$ and $\beta_2^0$ and then is put at time $t = 0$ in contact with two heat baths at inverse temperatures $\beta_1$ and $\beta_2$, the mean heat amount received from the thermostat 1 between time $t = 0$ and time $t$ is given (with the convention (4.1)) by

$$\langle\mathcal{Q}_1(t)\rangle_{P_0} = -\Delta e \sum_{n_1 = -\infty}^{+\infty} n_1 \sum_{\sigma_1, \sigma_2} \sum_{\sigma_1', \sigma_2'} \sum_{\Delta n = 0, 1, -1} (\sigma_1', \sigma_2'|\mathbb{U}(n_1, n_1 + \Delta n; t)|\sigma_1, \sigma_2)$$

$$\times P_0(\sigma_1, \sigma_2). \tag{4.48}$$

By virtue of the definition (4.11) of the relevant generating function and the decomposition (4.13), the latter expression can be rewritten as

$$
\langle \mathcal{Q}_1(t) \rangle_{P_0} = -\Delta e \frac{\partial}{\partial z_1} \left( \sum_{\sigma_1, \sigma_2} \sum_{\sigma'_1, \sigma'_2} \sum_{\Delta n = 0, 1, -1} \left( \sigma'_1, \sigma'_2 | \widehat{\mathbb{U}}_{[\Delta n]}(z_1, z_2; t) | \sigma_1, \sigma_2 \right) \right.
$$
$$
\left. \times \; P_0(\sigma_1, \sigma_2) \right) \Bigg|_{z_1 = z_2 = 1}. \tag{4.49}
$$

From the explicit expressions (4.27) for the matrix elements and (4.22) for $\Delta_+(z)$, a straightforward calculation leads to

$$
\langle \mathcal{Q}_1(t) \rangle_{P_0} = \langle j_1 \rangle_{\text{st}} t + \left[ \overline{\nu}_1 \overline{\nu}_2 \left( \gamma_1 - \gamma_2 \right) - \overline{\nu}_1 \left( \gamma_1 - \gamma_\star^0 \right) \right] \frac{\Delta e}{2} [1 - \mathrm{e}^{-(\nu_1 + \nu_2)t}] \tag{4.50}
$$

where $\langle j_1 \rangle_{\text{st}} = -\langle j_2 \rangle_{\text{st}}$ is given in (3.27) and $\gamma_\star^0$ is a function of $\overline{\nu}_1$, $\overline{\nu}_2$, $\beta_1^0$ and $\beta_2^0$ written in (3.5). A similar calculation yields

$$
\langle \mathcal{Q}_2(t) \rangle_{P_0} = \langle j_2 \rangle_{\text{st}} t + \left[ -\overline{\nu}_1 \overline{\nu}_2 \left( \gamma_1 - \gamma_2 \right) - \overline{\nu}_2 \left( \gamma_2 - \gamma_\star^0 \right) \right] \frac{\Delta e}{2} [1 - \mathrm{e}^{-(\nu_1 + \nu_2)t}]. \tag{4.51}
$$

As a result, with the sign convention of definition (3.48), the excess heat given to the system by heat bath 1 in the present protocol reads

$$
Q_{\text{exc},1} |_{[\beta_1^0, \beta_2^0]}^{[\beta_1, \beta_2]} = - \left[ \overline{\nu}_1 \overline{\nu}_2 \left( \gamma_1 - \gamma_2 \right) - \overline{\nu}_1 \left( \gamma_1 - \gamma_\star^0 \right) \right] \frac{\Delta e}{2}. \tag{4.52}
$$

The excess heat given to the system by heat bath 2, $Q_{\text{exc},2} |_{[\beta_1^0, \beta_2^0]}^{[\beta_1, \beta_2]}$, is obtained from the latter expression by the exchanges $\overline{\nu}_1 \leftrightarrow \overline{\nu}_2$ and $\gamma_1 \leftrightarrow \gamma_2$. These two excess heats are not opposite to each other, and comparison with the expressions (3.14) for the mean energies in the initial and final stationary states shows that the sum of the excess heats coming from both thermostats indeed satisfy the identity (3.49).

### 4.4. Symmetry property for reversed heat transfers (when $T_1 \neq 0$) specific to the model

*4.4.1. Symmetry arising from modified detailed balance.* The symmetry properties for reversed heat transfers when $T_1 \neq 0$ are more conveniently exhibited after a change of variable in the complex plane where the integrals involved in $P_{P_0}(\mathcal{Q}_1, \mathcal{Q}_2, t)$ are defined. The relevant functions $c_n(\tau)$ and $s_n(\tau)$ are defined in (4.34) and (4.35), while $\Delta_+(z)$ is given in (4.22). The origin $z = 0$ is a singular point in $\Delta_+(z)$ and $z \Delta_+(z) = (A + B) P(z)$, where the second-order polynomial $P(z) = (z - z_+)(z - z_-)$ vanishes for the two roots $z_+$ and $z_-$. The product of the roots is equal to

$$
z_+ z_- = \frac{A - B}{A + B} = \frac{1}{\rho^2} \qquad \text{with } \rho \equiv \mathrm{e}^{(\beta_1 - \beta_2)\Delta e/2} \tag{4.53}
$$

and the sum of the roots is equal to $-(1 - 2A)/(A + B)$.

When $T_1 \neq 0$, $z_+ z_-$ does not vanish and, by using the variable change $\zeta = z/\sqrt{z_- z_+}$, namely

$$
\zeta = z\rho, \tag{4.54}
$$

the unit circle is changed into a circle with radius $\rho$, while the roots $z_+$ and $z_-$ are changed into $\zeta_-$ and $\zeta_+$, with $\zeta_-\zeta_+ = 1$. Then, for a function such as $\cosh(\sqrt{\Delta_+(z)})$ or $\sinh(\tau\sqrt{\Delta_+(z)})/\sqrt{\Delta_+(z)}$, each of which is in fact a function of $\Delta_+(z)$ denoted by $f(\Delta_+(z))$,

$$\oint_{|z|=1} \frac{\mathrm{d}z}{2\pi\mathrm{i}} \frac{1}{z^{n+1}} f(\Delta_+(z)) = \rho^n \oint_{|\zeta|=\rho} \frac{\mathrm{d}\zeta}{2\pi\mathrm{i}} \frac{1}{\zeta^{n+1}} f(\widetilde{\Delta}_+(\zeta)) \tag{4.55}$$

with $\widetilde{\Delta}_+(\zeta) \equiv \Delta_+(\zeta/\rho)$. $\widetilde{\Delta}_+(\zeta)$ is a symmetric function of $\zeta$ and $1/\zeta$,

$$\widetilde{\Delta}_+(\zeta) = b + a\frac{\zeta + \zeta^{-1}}{2} \tag{4.56}$$

with $a = 2\sqrt{A^2 - B^2}$ and $b = 1 - 2A$, namely

$$a = 2\overline{\nu}_1\overline{\nu}_2\sqrt{(1 - \gamma_1^2)(1 - \gamma_2^2)} \tag{4.57}$$

$$b = 1 - 2\overline{\nu}_1\overline{\nu}_2 (1 - \gamma_1\gamma_2). \tag{4.58}$$

Since the only singular points in the integrand in the right-hand side of (4.55) are $\zeta = 0$ and $\infty$, the circle $|\zeta| = \rho$ can be deformed into the unit circle and we get the identity

$$\oint_{|z|=1} \frac{\mathrm{d}z}{2\pi\mathrm{i}} \frac{1}{z^{n+1}} f(\Delta_+(z)) = \rho^n \oint_{|z|=1} \frac{\mathrm{d}z}{2\pi\mathrm{i}} \frac{1}{z^{n+1}} f(\widetilde{\Delta}_+(z)). \tag{4.59}$$

By inserting the latter identity in the definitions (4.34) and (4.35) we get the relations

$$\begin{aligned} c_n(\tau) &= \rho^n\widetilde{c}_n(\tau) \\ s_n(\tau) &= \rho^n\widetilde{s}_n(\tau) \end{aligned} \tag{4.60}$$

where

$$\widetilde{c}_n(\tau) \equiv \oint_{|z|=1} \frac{\mathrm{d}z}{2\pi\mathrm{i}} \frac{1}{z^{n+1}} \cosh\left(\tau\sqrt{\widetilde{\Delta}_+(z)}\right) \tag{4.61}$$

and

$$\widetilde{s}_n(\tau) \equiv \oint_{|z|=1} \frac{\mathrm{d}z}{2\pi\mathrm{i}} \frac{1}{z^{n+1}} \frac{\sinh(\tau\sqrt{\widetilde{\Delta}_+(z)})}{\sqrt{\widetilde{\Delta}_+(z)}}. \tag{4.62}$$

Since $\widetilde{\Delta}_+(z)$ is invariant under the exchange of $z$ and $1/z$,

$$\begin{aligned} \widetilde{c}_n(\tau) &= \widetilde{c}_{|n|}(\tau) \\ \widetilde{s}_n(\tau) &= \widetilde{s}_{|n|}(\tau). \end{aligned} \tag{4.63}$$

Therefore the functions involved in the matrix elements (4.36) of $\mathbb{U}(n_1, n_2; t)$ can be rewritten as

$$\begin{aligned} U^{(0)}(n, \sigma_1\sigma_2; \tau) &= \mathrm{e}^{-\tau}\rho^n \left[\widetilde{c}_{|n|}(\tau) + \gamma_\star\sigma_1\sigma_2\widetilde{s}_{|n|}(\tau)\right] \\ U^{(1)}(n, \Delta n; \tau) &= \mathrm{e}^{-\tau}\rho^n \left[\overline{\nu}_2 (1 - \gamma_2\Delta n) \widetilde{s}_{|n|}(\tau) + \overline{\nu}_1 (1 - \gamma_1\Delta n) \rho^{\Delta n}\widetilde{s}_{|n+\Delta n|}(\tau)\right] \end{aligned} \tag{4.64}$$

where $\rho$, defined in (4.53), also reads $\rho = \sqrt{(1 + \gamma_1)(1 - \gamma_2)/(1 - \gamma_1)(1 + \gamma_2)}$.

According to [16], the modified detailed balance implies some time-reversal symmetry property for histories, which itself entails some relation between the probabilities of forward and backward evolutions where the given initial and final configurations are exchanged (and the heat amounts are changed into their opposite values). In the spin model language, with the definitions (4.1), the symmetry exhibited in [16] for the probability that the system evolves from an initial configuration $\mathcal{C}_0 = (\sigma_1, \sigma_2)$ to a final configuration $\mathcal{C}_f = (\sigma_1', \sigma_2')$ while receiving the heat amount $\mathcal{Q}_1 = -n_1 \Delta e$ and $\mathcal{Q}_2 = n_2 \Delta e$ reads, for non-vanishing matrix elements,

$$\frac{(\sigma_1', \sigma_2'|\mathbb{U}(n_1, n_2; t)|\sigma_1, \sigma_2)}{(\sigma_1, \sigma_2|\mathbb{U}(-n_1, -n_2; t)|\sigma_1', \sigma_2')} = \mathrm{e}^{(n_1 \beta_1 - n_2 \beta_2)\Delta e}. \tag{4.65}$$

Comparison of the latter relation with the expressions (4.36) implies that

$$\frac{U^{(0)}(n, \sigma_1 \sigma_2; \tau)}{U^{(0)}(-n, \sigma_1 \sigma_2; \tau)} = \mathrm{e}^{[n\beta_1 - n\beta_2]\Delta e} \tag{4.66}$$

and

$$\frac{U^{(1)}(n, \Delta n; \tau)}{U^{(1)}(-n, -\Delta n; \tau)} = \mathrm{e}^{[n\beta_1 - (n+\Delta n)\beta_2]\Delta e}. \tag{4.67}$$

The latter relations can be checked from the explicit expressions (4.64).

We notice that the relation (4.65) can also be retrieved by noticing that the modified detailed balance entails that the matrix $\mathbb{A}(z_1, z_2)$, which rules the evolution of $\widehat{\mathbb{U}}(z_1, z_2; \tau)$ according to (4.16), obeys the symmetry $\mathbb{A}(z_1, z_2) = \mathbb{A}^{\mathrm{T}}(\mathrm{e}^{-\beta_1 \Delta e}/z_1, \mathrm{e}^{\beta_2 \Delta e}/z_2)$, where $\mathbb{A}^{\mathrm{T}}$ denotes the transposed matrix of $\mathbb{A}$. Therefore, after the variable change $z_1 = \zeta_1/\rho_1$ and $z_2 = \zeta_2/\rho_2$ with $\rho_1 = \exp(\beta_1 \Delta e/2)$ and $\rho_2 = \exp(-\beta_2 \Delta e/2)$, the matrix $\mathbb{A}(z_1, z_2)$ becomes the matrix $\widetilde{\mathbb{A}}(\zeta_1, \zeta_2) \equiv \mathbb{A}(z_1 = \zeta_1/\rho_1, z_2 = \zeta_2/\rho_2)$, which obeys the symmetry $\widetilde{\mathbb{A}}^{\mathrm{T}}(\zeta_1, \zeta_2) = \widetilde{\mathbb{A}}(1/\zeta_1, 1/\zeta_2)$. Then the derivation of the symmetry (4.65) is the following. First we make the variable change $z_1 = \zeta_1/\rho_1$ and $z_2 = \zeta_2/\rho_2$ in the integral representation (4.12) for $\mathbb{U}(n_1, n_2; t)$. Since $\widetilde{\mathbb{A}}(\zeta_1, \zeta_2)$ has no non-analyticity, apart from the $1/\zeta_1$ and $1/\zeta_2$ singular terms, the integrals on the circles with radii equal to $\rho_1$ and $\rho_2$ are equal to the integrals with the same integrands on the circles with radii equal to 1. If $\zeta$ is on the unit circle, $1/\zeta$ is also on this circle and we can make the variable change $\zeta_1 \to 1/\zeta_1$ and $\zeta_2 \to 1/\zeta_2$; then the symmetry of $\widetilde{\mathbb{A}}$ leads to the symmetry (4.65).

As shown in [16], the consequence (4.65) of the modified detailed balance (2.3) entails that, if the spin system is in an equilibrium state at inverse temperature $\beta_0$ at time $t = 0$ where it is put in contact with the two thermostats, then the ratio of the probabilities to measure some given heat amounts $\mathcal{Q}_1$ and $\mathcal{Q}_2$ or their opposite values obeys the fluctuation relation

$$\frac{P_{P_{\mathrm{can}}^{\beta_0}}(\mathcal{Q}_1, \mathcal{Q}_2, ; t)}{P_{P_{\mathrm{can}}^{\beta_0}}(-\mathcal{Q}_1, -\mathcal{Q}_2; t)} = \mathrm{e}^{(\beta_0 - \beta_1)\mathcal{Q}_1 + (\beta_0 - \beta_2)\mathcal{Q}_2}. \tag{4.68}$$

*4.4.2. A relation specific to the model.* The present model happens to obey a very specific relation for reversed heat transfers when the initial state of the system has an arbitrary probability distribution $P_0$. According to (4.36), after summation over the final

configuration,

$$
\begin{aligned}
\sum_{\sigma_1', \sigma_2'} (\sigma_1', \sigma_2' | \mathbb{U}(n, n; t) | \sigma_1, \sigma_2) &= U^{(0)}(n, \sigma_1 \sigma_2; \tau) \\
\sum_{\sigma_1', \sigma_2'} (\sigma_1', \sigma_2' | \mathbb{U}(n, n + \Delta n; t) | \sigma_1, \sigma_2) &\underset{\Delta n = \pm 1}{=} U^{(1)}(n, \Delta n; \tau)\, \delta_{\sigma_1 \sigma_2, \Delta n}.
\end{aligned}
\tag{4.69}
$$

Therefore, when the initial configurations are distributed with an arbitrary probability $P_0$ then

$$
\sum_{\sigma_1', \sigma_2'} (\sigma_1', \sigma_2' | \mathbb{U}(n, n; t) | P_0) = \sum_{\sigma_1, \sigma_2} U^{(0)}(n, \sigma_1 \sigma_2; \tau) P_0(\sigma_1, \sigma_2)
\tag{4.70}
$$

and

$$
\sum_{\sigma_1', \sigma_2'} (\sigma_1', \sigma_2' | \mathbb{U}(n, n + \Delta n; t) | P_0) \underset{\Delta n = \pm 1}{=} U^{(1)}(n, \Delta n; \tau) \times P_0(\sigma_1 \sigma_2 = \Delta n)
\tag{4.71}
$$

where $P_0(\sigma_1 \sigma_2 = \Delta n)$ denotes the probability that the product $\sigma_1 \sigma_2$ is equal to $\Delta n$ in the initial configuration.

Then from the expressions (4.70) and (4.71) for $P_{P_0}(\mathcal{Q}_1, \mathcal{Q}_2, t)$ in the cases $\mathcal{Q}_1 = -\mathcal{Q}_2$ and $\mathcal{Q}_1 \neq -\mathcal{Q}_2$ respectively, and by virtue of the consequences (4.66) and (4.67) of the modified detailed balance, we get the property

$$
\begin{aligned}
\frac{P_{P_0}(\mathcal{Q}_1, \mathcal{Q}_2, ; t)}{P_{P_0}(-\mathcal{Q}_1, -\mathcal{Q}_2; t)} &= \mathrm{e}^{-\beta_1 \mathcal{Q}_1 - \beta_2 \mathcal{Q}_2} \\
&\times \left[ \delta_{\mathcal{Q}_1 + \mathcal{Q}_2, 0} + \sum_{\epsilon = \pm 1} \delta_{\mathcal{Q}_1 + \mathcal{Q}_2, \epsilon \Delta e} \frac{P(\sigma_1 \sigma_2 = \epsilon; t = 0)}{1 - P(\sigma_1 \sigma_2 = \epsilon; t = 0)} \right].
\end{aligned}
\tag{4.72}
$$

The appearance of the initial probability for the sign of the spin product seemingly arises from the fact that, by virtue of energy conservation, the values of this sign in the final and initial states are related to the sum $\mathcal{Q}_2 + \mathcal{Q}_1$ by the constraint (4.5).

We stress that this relation is very specific to the present model. Its interest lies not in its precise form, but in that the right-hand side involves the initial distribution: it has sometimes been speculated that the (experimental) study of the ratio on the left-hand side for general systems could give some clues about the initial distribution. The above formula justifies this hope, but shows at the same time that even in the simple case at hand only partial information can be retrieved, suggesting that for more general systems even this partial information may be difficult to extract. In the cases where the initial distribution is equal either to an equilibrium state distribution at the inverse temperature $\beta_0$ or to the stationary state distribution which is a canonical distribution at the inverse temperature $\beta_\star$, the relation allows one to retrieve the generic relation (4.68).

### 4.5. Decaying property of joint probabilities for large heat exchanges

All quantities of interest involve the coefficients $c_m(\tau)$ and/or $s_m(\tau)$, computed via contour integrals in (4.34) and (4.35). The integrands involve functions which are holomorphic for $z \in \mathbb{C}^*$, the pointed complex plane: the quantity $\Delta_+(z)$ defined in (4.22) has this property and the functions $\cosh(w)$ and $\sinh(w)/w$ are entire even functions, so that the square roots in the composed functions $\cosh(\tau \sqrt{\Delta_+(z)})$ and $\sinh(\tau \sqrt{\Delta_+(z)})/\sqrt{\Delta_+(z)}$ do no harm.

Hence, in the formulae for $c_m(\tau)$ and $s_m(\tau)$, contours can be deformed. For $r \in ]0, +\infty[$, let $C(\tau, r) \equiv \sup_{|z|=r} |\cosh(\tau\sqrt{\Delta_+(z)})| < +\infty$ and $S(\tau, r) \equiv \sup_{|z|=r} |\sinh(\tau\sqrt{\Delta_+(z)})/\sqrt{\Delta_+(z)}| < +\infty$.

Taking $|z| = r$ as the integration contour, one gets immediately that, for each $r$, $c_m(\tau) \leq C(\tau, r) r^{-m}$ and $s_m(\tau) \leq S(\tau, r) r^{-m}$. This shows that $c_m(\tau)$ and $s_m(\tau)$ are $o(\mathrm{e}^{-K|m|})$ at large $|m|$ for any $K$.

With some efforts, we could get some explicit upper bounds for $C(\tau, r)$ and $S(\tau, r)$. Then we could extremize over $r$ to get a subexponential bound for $c_m(\tau)$ and $s_m(\tau)$, but we shall not need this refinement.

In the limit $T_1 = 0$, $\Delta_+(z)$ is in fact holomorphic for $z \in \mathbb{C}$ so that $c_m(\tau)$ and $s_m(\tau)$ vanish for $m = -1, -2, \ldots$.

## 5. Heat amount cumulants for any $T_1$ and $T_2$

### 5.1. Generic properties for a system with a finite number of configurations

*5.1.1. Characteristic function for the heat amount $\mathcal{Q}_2$.* The random variable $\mathcal{Q}_2$ can take only discrete values $n_2 \Delta e$, where $n_2$ is a positive or negative integer. Therefore its probability density $\Pi(\mathcal{Q}; t)$, defined as $\Pi(\mathcal{Q}; t) \, \mathrm{d}\mathcal{Q} = P(\mathcal{Q}_2 \in [\mathcal{Q}, \mathcal{Q} + \mathrm{d}\mathcal{Q}[; t)$, reads

$$\Pi(\mathcal{Q}; t) = \frac{1}{\Delta e} \sum_{n_2=-\infty}^{+\infty} \delta\left(\frac{\mathcal{Q}}{\Delta e} - n_2\right) P(n_2; t), \tag{5.1}$$

where $\delta$ stands for the Dirac distribution. Since $P(n_2; t)$ decays faster than $\exp(-K|n_2|)$ for any $K > 0$ when $|n_2|$ goes to infinity (see section 4.5), the Laplace transform $\widetilde{G}(\lambda; t)$ of $\Pi(\mathcal{Q})$, i.e. the characteristic function of the random variable $\mathcal{Q}_2$, is well defined for any $\lambda$,

$$\widetilde{G}(\lambda; t) \equiv \int_{-\infty}^{+\infty} \mathrm{d}\mathcal{Q} \, \mathrm{e}^{\lambda\mathcal{Q}} \Pi(\mathcal{Q}; t) = \langle \mathrm{e}^{\lambda\mathcal{Q}_2(t)} \rangle. \tag{5.2}$$

As a consequence, the properties of the probability density $\Pi(\mathcal{Q}; t)$ can be investigated through those of its Laplace transform, thanks to the inversion formula

$$\Pi(\mathcal{Q}; t) = \int_{-\mathrm{i}\infty}^{+\mathrm{i}\infty} \frac{\mathrm{d}\lambda}{2\pi\mathrm{i}} \mathrm{e}^{-\lambda\mathcal{Q}} \widetilde{G}(\lambda; t). \tag{5.3}$$

According to the property (5.1), $\widetilde{G}(\lambda)$ is a periodic function of $\lambda$ with period equal to $\mathrm{i}2\pi/\Delta e$, and the right-hand side of the latter formula can be written as

$$\frac{1}{\Delta e} \int_0^{2\pi} \frac{\mathrm{d}\theta}{2\pi} \mathrm{e}^{-\mathrm{i}\theta\mathcal{Q}/\Delta e} \widetilde{G}\left(\frac{\mathrm{i}\theta}{\Delta e}; t\right) \sum_{m=-\infty}^{+\infty} \mathrm{e}^{-\mathrm{i}2\pi m \mathcal{Q}/\Delta e}. \tag{5.4}$$

By virtue of the Poisson equality $\sum_{m=-\infty}^{+\infty} \mathrm{e}^{-\mathrm{i}2\pi m \mathcal{Q}/\Delta e} = \sum_{n_2=-\infty}^{+\infty} \delta(\mathcal{Q}/\Delta e - n_2)$, comparison with (5.1) leads to

$$P(n_2; t) = \int_0^{2\pi} \frac{\mathrm{d}\theta}{2\pi} \mathrm{e}^{-\mathrm{i}\theta n_2} \widetilde{G}\left(\frac{\mathrm{i}\theta}{\Delta e}; t\right). \tag{5.5}$$

The latter equality coincides with the inverse formula (analogous to (4.12)) in terms of the generating function, $G(z;t) = \sum_{n_2=-\infty}^{+\infty} z^{n_2} P(n_2;t)$,

$$P(n_2;t) = \oint_{|z|=1} \frac{\mathrm{d}z}{2\pi\mathrm{i}} \frac{1}{z^{n_2+1}} G(z;t), \tag{5.6}$$

where $G(z;t) \equiv \widetilde{G}((1/\Delta e)\ln z;t)$.

*5.1.2. Relation between long-time cumulants per unit time for $\mathcal{Q}_1$ and $\mathcal{Q}_2$.* The generic properties of the cumulants of $\mathcal{Q}_1$ and $\mathcal{Q}_2$ have been reviewed in [16]. We recall those which will be useful in the following. For a Markov process, the long-time behaviors of these cumulants are proportional to the time $t$ elapsed from the beginning of the measurements. The asymptotic behavior of the cumulants per unit time are given by the derivatives of

$$\alpha_a(\lambda) \equiv \lim_{t\to+\infty} \frac{1}{t} \ln\langle \mathrm{e}^{\lambda \mathcal{Q}_a(t)} \rangle \tag{5.7}$$

according to

$$\lim_{t\to+\infty} \frac{\kappa_{\mathcal{Q}_a}^{[p]}}{t} = \left.\frac{\partial^p \alpha_a(\lambda_a)}{\partial \lambda_a^p}\right|_{\lambda_a=0} \qquad \text{for } a = \{1,2\}. \tag{5.8}$$

Moreover, in the case of a system with a finite number of configurations, $\mathcal{Q}_1 + \mathcal{Q}_2 = \mathcal{E}(\mathcal{C}_\mathrm{f}) - \mathcal{E}(\mathcal{C}_0)$ is restricted to some finite interval and

$$\alpha_1(\lambda) = \alpha_2(-\lambda). \tag{5.9}$$

As a consequence the long-time cumulants per unit time obey the following relations

$$\lim_{t\to\infty} \frac{\kappa_{\mathcal{Q}_1}^{[p]}}{t} = (-1)^p \lim_{t\to\infty} \frac{\kappa_{\mathcal{Q}_2}^{[p]}}{t}. \tag{5.10}$$

## 5.2. Explicit formulæ for the cumulants per unit time

According to the relation (5.10) between the long-time cumulants per unit time for $\mathcal{Q}_1$ and $\mathcal{Q}_2$, we have only to consider the cumulants for the heat amount $\mathcal{Q}_2$ received from bath 2. For the two-spin system, where $\mathcal{Q}_2 = n_2 \Delta e$, it is convenient to introduce the cumulants $\kappa_{n_2}^{[p]}$ for the dimensionless variable $n_2$ and the associated characteristic function $\langle \mathrm{e}^{\overline{\lambda} n_2} \rangle$, where $\overline{\lambda}$ is a dimensionless variable. According to the relation (5.8) the long-time behavior of the cumulants per unit time are derived through the relation

$$\lim_{t\to+\infty} \frac{1}{t} \kappa_{n_2}^{[p]} = \left.\frac{\partial^p \alpha_2(\overline{\lambda})}{\partial \overline{\lambda}^p}\right|_{\overline{\lambda}=0} \tag{5.11}$$

with

$$\alpha_2(\overline{\lambda}) = \lim_{t\to+\infty} \frac{1}{t} \ln\langle \mathrm{e}^{\overline{\lambda} n_2} \rangle. \tag{5.12}$$

According to the definition of $\langle \mathrm{e}^{\overline{\lambda} n_2} \rangle$, the relation (4.9) between the probability $P(\sigma_1, \sigma_2, n_1, n_2; t)$ and the operator $\mathbb{U}(n_1, n_2; t)$, together with the definition (4.11) of $\widehat{\mathbb{U}}(z_1, z_2; t)$ and its evolution equation (4.16), the characteristic function may be expressed
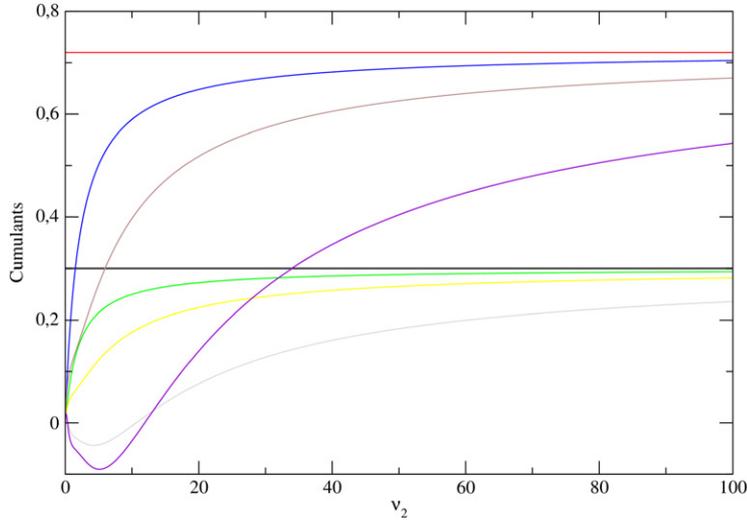
**Figure 1.** First infinite-time cumulants per unit time as functions of $\nu_2/\nu_1$. Illustration of the large $\nu_2/\nu_1$ behavior with the two asymptotes, $(\nu_1/2)(\gamma_1 - \gamma_2)$ for odd cumulants and $(\nu_1/2)(1 - \gamma_1\gamma_2)$ for even cumulants. The other model parameters are fixed to the sample values $\gamma_1 = 0.7$, $\gamma_2 = 0.4$, $\nu_1 = 2$.

as

$$\langle e^{\bar{\lambda} n_2} \rangle = \sum_{\sigma_1,\sigma_2} \sum_{\sigma_1',\sigma_2'} \left( \sigma_1, \sigma_2 | e^{((\nu_1+\nu_2)/2)t\,\mathbb{A}(z_1=1,z_2=e^{\bar{\lambda}})} | \sigma_1', \sigma_2' \right) P(\sigma_1', \sigma_2'; t=0). \tag{5.13}$$

According to (4.17), $\mathbb{A}(z_1 = 1, z_2 = e^{\bar{\lambda}}) + (1 + \gamma_\star)\mathbb{I}_4$ is a real positive matrix and the Perron–Frobenius theorem can be applied. Henceforth, $\alpha_2(\bar{\lambda})$ coincides with the eigenvalue of the matrix $[(\nu_1 + \nu_2)/2]\mathbb{A}(z_1 = 1, z_2 = e^{\bar{\lambda}})$ with the largest modulus (and which is necessarily real). The four eigenvalues to consider are the $\mu^{(\epsilon,\eta)}(z = e^{\bar{\lambda}})$ which are given by the expression (4.26), with $\epsilon = \pm$ and $\eta = \pm$. The one with the largest modulus corresponds to $(\epsilon, \eta) = (+, +)$ and reads

$$\alpha_2(\bar{\lambda}) = \frac{\nu_1 + \nu_2}{2} \left[ -1 + \sqrt{1 - 2A + (A + B)e^{\bar{\lambda}} + (A - B)e^{-\bar{\lambda}}} \right], \tag{5.14}$$

where $A$ and $B$ are defined in (4.23).

It is simple to calculate a number of cumulants per unit time in the infinite-time limit from (5.11). Their behavior as a function of the kinetic parameter $\nu_2/\nu_1$ exhibits some interesting features. For large $\nu_2/\nu_1$, the cumulants go to a limit which is the same for all odd and for all even cumulants, as will be explained in section 7. Figure 1 illustrates this convergence, which gets slower and slower for higher moments. The first six cumulants are represented. This figure also shows some oscillations at finite $\nu_2/\nu_1$. These oscillations become more and more visible on higher cumulants. Figure 2 illustrates this phenomenon. Cumulants from the fifth to the ninth are represented. In both figures, the other model parameters are fixed to the sample values $\gamma_1 = 0.7$, $\gamma_2 = 0.4$, $\nu_1 = 2$.

Only the first cumulants have analytic expressions simple enough to fit on a line. For the sake of conciseness, the results are first expressed in terms of the dimensionless time
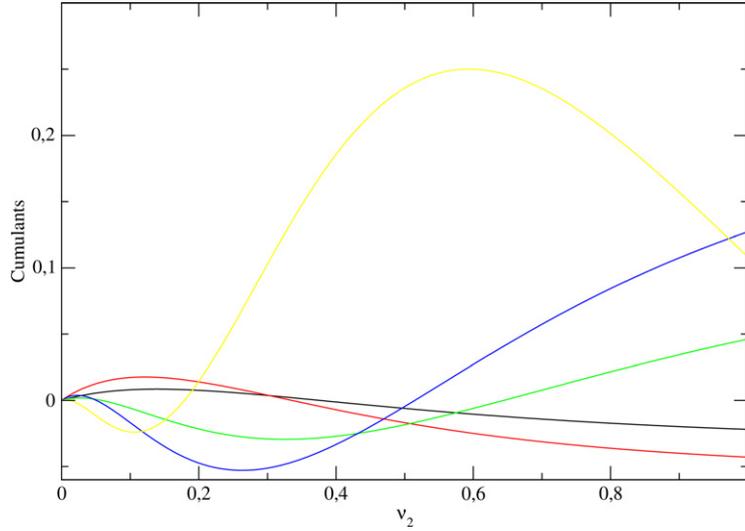
**Figure 2.** Higher cumulants. Illustration of the oscillations at finite $\nu_2/\nu_1$. The other model parameters are fixed to the sample values $\gamma_1 = 0.7$, $\gamma_2 = 0.4$, $\nu_1 = 2$.

$\tau = \frac{1}{2}(\nu_1 + \nu_2)t$ as

$$
\begin{aligned}
\lim_{\tau \to +\infty} \frac{1}{\tau}\kappa_{n_2}^{[1]} &= B \\
\lim_{\tau \to +\infty} \frac{1}{\tau}\kappa_{n_2}^{[2]} &= A - B^2 \\
\lim_{\tau \to +\infty} \frac{1}{\tau}\kappa_{n_2}^{[3]} &= B[1 - 3A + 3B^2] \\
\lim_{\tau \to +\infty} \frac{1}{\tau}\kappa_{n_2}^{[4]} &= A - 3A^2 + B^2[-4 + 18A - 15B^2].
\end{aligned}
\tag{5.15}
$$

All odd cumulants are proportional to $B$, because all odd powers of $\overline{\lambda}$ in the expansion of the expression (5.14) for $\alpha_2(\overline{\lambda})$ are proportional to $B$. The first three cumulants are rewritten in terms of the model parameters as

$$
\begin{aligned}
\lim_{t \to +\infty} \frac{\langle \mathcal{Q}_2 \rangle_{\mathrm{st}}}{t} &= \overline{\nu}_1 \overline{\nu}_2 (\gamma_1 - \gamma_2) \frac{(\nu_1 + \nu_2)\Delta e}{2} = \langle j_2 \rangle_{\mathrm{st}} \\
\lim_{t \to +\infty} \frac{\langle \mathcal{Q}_2^2 \rangle_{\mathrm{st}} - \langle \mathcal{Q}_2 \rangle_{\mathrm{st}}^2}{t} &= \overline{\nu}_1 \overline{\nu}_2 \left[ 1 - \gamma_1 \gamma_2 - \overline{\nu}_1 \overline{\nu}_2 (\gamma_1 - \gamma_2)^2 \right] \frac{(\nu_1 + \nu_2)(\Delta e)^2}{2} \\
\lim_{t \to +\infty} \frac{\langle \mathcal{Q}_2^3 \rangle_{\mathrm{st}}^{\mathrm{c}}}{t} &= \overline{\nu}_1 \overline{\nu}_2 (\gamma_1 - \gamma_2) \left[ 1 - 3\overline{\nu}_1 \overline{\nu}_2 (1 - \gamma_1 \gamma_2) + 3\overline{\nu}_1^2 \overline{\nu}_2^2 (\gamma_1 - \gamma_2)^2 \right] \frac{(\nu_1 + \nu_2)(\Delta e)^3}{2}.
\end{aligned}
\tag{5.16}
$$

$\langle \mathcal{Q}_2^3 \rangle^{\mathrm{c}}$ is the third cumulant, which is equal to the third centered moment, namely $\langle \mathcal{Q}_2^3 \rangle^{\mathrm{c}} \equiv \langle [\mathcal{Q}_2 - \langle \mathcal{Q}_2 \rangle]^3 \rangle$.

At equilibrium $\gamma_1 = \gamma_2$, so that $B = 0$: then, by virtue of the remark after (5.15), the long-time behavior of all odd cumulants of $\mathcal{Q}_2$ is subdominant with respect to the elapsed time $t$, and in the long-time limit $P(\mathcal{Q}_2; t)$ becomes an even function of $\mathcal{Q}_2$ at leading order in time $t$. The fourth cumulant of the cumulated heat $\mathcal{Q}_2$ received from the thermostat 2 per unit time does not vanish: $\lim_{t \to +\infty} 1/t \ln \langle e^{\lambda \mathcal{Q}_2} \rangle_{\mathrm{eq}}$ is not quadratic in $\lambda$, and even in the long-time limit the variable $\mathcal{Q}_2$ has a non-Gaussian distribution, contrary

to the variable $\left[\mathcal{Q}_2 - \langle\mathcal{Q}_2\rangle\right]/\sqrt{t}$ (for which all cumulants of order larger than three vanish in the infinite-time limit). The first two even cumulants per unit time read

$$
\begin{aligned}
&\lim_{t\to+\infty} \frac{\langle\mathcal{Q}_2^2\rangle_{\mathrm{eq}} - \langle\mathcal{Q}_2\rangle_{\mathrm{eq}}^2}{t} = \overline{\nu}_1\overline{\nu}_2\left(1-\gamma^2\right)\frac{(\nu_1+\nu_2)(\Delta e)^2}{2}\\
&\lim_{t\to+\infty} \frac{\langle\mathcal{Q}_2^4\rangle_{\mathrm{eq}}^{\mathrm{c}}}{t} = \overline{\nu}_1\overline{\nu}_2\left(1-\gamma^2\right)\left[1-3\overline{\nu}_1\overline{\nu}_2(1-\gamma^2)\right]\frac{(\nu_1+\nu_2)(\Delta e)^4}{2}.
\end{aligned}
\tag{5.17}
$$

$\langle\mathcal{Q}_2^4\rangle^{\mathrm{c}}$ denotes the fourth cumulant, which can be expressed as $\langle\mathcal{Q}_2^4\rangle^{\mathrm{c}} = \langle[\mathcal{Q}_2 - \langle\mathcal{Q}_2\rangle]^4\rangle - 3\langle[\mathcal{Q}_2 - \langle\mathcal{Q}_2\rangle]^2\rangle$.

For a system weakly out-of-equilibrium the Einstein–Green–Kubo relation, namely

$$
\lim_{(\beta_1,\beta_2)\to(\beta,\beta)} \frac{\langle j_2\rangle_{\mathrm{st}}}{\beta_1 - \beta_2} = \frac{1}{2}\lim_{t\to+\infty} \frac{\langle\mathcal{Q}_2^2\rangle_{\mathrm{eq}} - (\langle\mathcal{Q}_2\rangle_{\mathrm{eq}})^2}{t},
\tag{5.18}
$$

is indeed obeyed by the system, as it should be. This can be checked by comparing the expression (5.17) with the limit obtained when $(\beta_1,\beta_2) \to (\beta,\beta)$ for the ratio $\langle j_2\rangle_{\mathrm{st}}/(\beta_1 - \beta_2)$, which, by virtue of (3.27), reads

$$
\frac{\langle j_2\rangle_{\mathrm{st}}}{\beta_1 - \beta_2} = \overline{\nu}_1\overline{\nu}_2\frac{\gamma_1 - \gamma_2}{(\beta_1 - \beta_2)\Delta e}\frac{(\nu_1+\nu_2)(\Delta e)^2}{2}.
\tag{5.19}
$$

When the system is far from equilibrium, comparison of the latter expression for $\langle j_2\rangle_{\mathrm{st}}/(\beta_1 - \beta_2)$ with the expression (5.16) for the long-time limit of the second cumulant per unit time shows that $\langle j_2\rangle_{\mathrm{st}}/(\beta_1-\beta_2) \neq \lim_{t\to+\infty}[\langle\mathcal{Q}_2^2\rangle_{\mathrm{st}} - \langle\mathcal{Q}_2\rangle_{\mathrm{st}}^2]/t$ , as it should be (see subsection 5.3 of [16]). Indeed, by virtue of equation (5.14), $\alpha_2(\overline{\lambda})$ obeys the symmetry relation $\alpha_2(\overline{\lambda}) = \alpha_2(-\overline{\mathcal{F}} - \overline{\lambda})$ with $\overline{\mathcal{F}} = \ln(A + B)/(A - B) = (\beta_1 - \beta_2)\Delta e$, but $\alpha_2(\overline{\lambda})$ is not a quadratic function of $\overline{\lambda}$, i.e. $\mathcal{Q}_2$ has a non-Gaussian distribution in the long-time limit.

## 6. Large deviation function for the cumulative heat current $\mathcal{Q}_2/t$

In this section, we derive the large deviation function for the cumulative heat current $\mathcal{Q}_2/t$ by three methods. The first one is based on the general theory of large deviations for the definition of large deviation functions and uses one of its cornerstones, the Gärtner–Ellis theorem. The second and the third rely on the fact that $\mathcal{Q}_2$ takes discrete values in a $t$-independent set, and uses an ad hoc definition of large deviation functions (see appendix E of [16]). Though the general theory of large deviations and the ad hoc definition for discrete exchanged quantities do not have to be the same, the ad hoc definition is nevertheless a sensible definition of large deviations. Physically, the general and the ad hoc definition are expected to yield the same result in a case as simple as the two-spin system, and our explicit computations can be seen as a proof of this fact. A natural tool to compute the ad hoc large deviation function is via a contour integral representation, but, as we shall see below, this method is surprisingly tricky even for the simple two-spin system at hand. In contrast with the general theory of large deviations, the contour integral method is the basis of a systematic expansion at large times. However, corrections are less universal than the dominant term.

The cumulative heat current received from heat bath 2 during the time interval $t$ takes the values $\mathcal{J} = \mathcal{Q}_2/t$, with $\mathcal{Q}_2 = n\Delta e$, $n$ integer. By dimensional analysis, the large

deviation function $f_{\mathcal{Q}_2}(\mathcal{J})$, which has the dimension of an inverse time, must be a function of

$$\widetilde{\jmath} = \frac{\mathcal{J}}{\Delta e} = \frac{n}{t}, \tag{6.1}$$

and we shall often consider the expressions of

$$\widetilde{f}_{\mathcal{Q}_2}(\widetilde{\jmath}) \equiv f_{\mathcal{Q}_2}(\mathcal{J}) \tag{6.2}$$

rather than those of $f_{\mathcal{Q}_2}(\mathcal{J}) = \widetilde{f}_{\mathcal{Q}_2}(\mathcal{J}/\Delta e)$. Moreover, the explicit calculations are more conveniently dealt with if, instead of considering $\widetilde{\jmath}$, we introduce the dimensionless current $\overline{\jmath}$ associated with the dimensionless time $\tau = [(\nu_1 + \nu_2)/2]t$,

$$\overline{\jmath} = \frac{n}{\tau} = \frac{2}{(\nu_1 + \nu_2)}\widetilde{\jmath}. \tag{6.3}$$

The dimensionless large deviation function $\overline{f}_{\mathcal{Q}_2}$ of $\overline{\jmath}$ is such that $t\widetilde{f}_{\mathcal{Q}_2}(\widetilde{\jmath}) = \tau\overline{f}_{\mathcal{Q}_2}(\overline{\jmath})$, and the expression of $\widetilde{f}_{\mathcal{Q}_2}(\widetilde{\jmath})$ can be retrieved from that for $\overline{f}_{\mathcal{Q}_2}(\overline{\jmath})$ through

$$\widetilde{f}_{\mathcal{Q}_2}(\widetilde{\jmath}) = \frac{\nu_1 + \nu_2}{2}\,\overline{f}_{\mathcal{Q}_2}\left(\frac{2}{(\nu_1 + \nu_2)}\widetilde{\jmath}\right). \tag{6.4}$$

We notice that large deviation functions for other cumulative quantities are related to $f_{\mathcal{Q}_2}$. Indeed, in a system with a finite number of configurations, $\mathcal{Q}_1 + \mathcal{Q}_2$ is bounded and, as a consequence of the general theory of large deviations (see e.g. [16]),

$$f_{\mathcal{Q}_1}(\mathcal{J}) = f_{\mathcal{Q}_2}(-\mathcal{J}). \tag{6.5}$$

In the same vein, as $\Delta_{\mathrm{exch}}S = \beta_1\mathcal{Q}_1 + \beta_2\mathcal{Q}_2 = -(\beta_1 - \beta_2)\mathcal{Q}_2 + \beta_1(\mathcal{Q}_1 + \mathcal{Q}_2)$, with $\mathcal{Q}_1 + \mathcal{Q}_2$ bounded, the large deviation function for $\Delta_{\mathrm{exch}}S$ and that for $\mathcal{Q}_2$ satisfy the simple relation

$$f_{\Delta_{\mathrm{exch}}S}(\mathcal{J}) = f_{\mathcal{Q}_2}\left(-\frac{\mathcal{J}}{\beta_1 - \beta_2}\right). \tag{6.6}$$

### 6.1. Derivation from Gärtner–Ellis theorem

*6.1.1. Method.* By analogy with (5.12), we introduce the dimensionless function

$$\overline{\alpha}_2(\overline{\lambda}) \equiv \lim_{\tau \to +\infty} \frac{1}{\tau} \ln\langle e^{\overline{\lambda}n_2}\rangle. \tag{6.7}$$

A simplified version of the Gärtner–Ellis theorem (see e.g. the review for physicists [23] or the mathematical point of view [24]) states that, if $\overline{\alpha}_2(\overline{\lambda})$ exists and is differentiable for all $\overline{\lambda}$ in $\mathbb{R}$, then the large deviation function of the current $\overline{\jmath} = n_2/\tau$ exists and it can be calculated as the Legendre–Fenchel transform of $\overline{\alpha}_2(\overline{\lambda})$, namely, with the signs chosen in the definitions used in the present paper,

$$\overline{f}_{\mathcal{Q}_2}(\overline{\jmath}) = \min_{\overline{\lambda} \in \mathbb{R}}\{\overline{\alpha}_2(\overline{\lambda}) - \overline{\lambda}\overline{\jmath}\}. \tag{6.8}$$

As a consequence, since $\overline{\alpha}_2(\overline{\lambda})$ obeys the symmetry $\overline{\alpha}_2(\overline{\lambda}) = \overline{\alpha}_2(-(\beta_1 - \beta_2)\Delta e - \overline{\lambda})$ (as can be checked from (5.14)), $\overline{f}(\overline{\jmath})$ obeys the fluctuation relation $\overline{f}(\overline{\jmath}) - \overline{f}(-\overline{\jmath}) = (\beta_1 - \beta_2)\Delta e \times \overline{\jmath}$. Moreover, the cumulant generating function $\ln\langle e^{\overline{\lambda}n_2}\rangle$ is necessarily convex (downward). In

the present case $\overline{\alpha}_2(\overline{\lambda})$ is strictly convex and continuously differentiable for all real $\overline{\lambda}$, so that the minimum in the definition of the Legendre–Fenchel transform can be readily calculated by using the Legendre transform,

$$\overline{f}_{\mathcal{Q}_2}(\overline{\jmath}) = \overline{\alpha}_2\left(\overline{\lambda}_{\mathrm{c}}(\overline{\jmath})\right) - \overline{\jmath}\overline{\lambda}_{\mathrm{c}}(\overline{\jmath}) \qquad \text{with } \frac{\mathrm{d}\overline{\alpha}_2}{\mathrm{d}\overline{\lambda}}(\overline{\lambda}_{\mathrm{c}}) = \overline{\jmath}. \tag{6.9}$$

*6.1.2. Various expressions for $f_{\mathcal{Q}_2}$ and its properties.* From the relation $\alpha_2(\overline{\lambda}) = [(\nu_1 + \nu_2)/2]\overline{\alpha}_2(\overline{\lambda})$ and the expression (5.14) for $\alpha_2(\overline{\lambda})$, when $T_1 \neq 0$ ($\gamma_1 \neq 1$), $A \neq B$ and we get

$$\overline{f}_{\mathcal{Q}_2}(\overline{\jmath}) = \overline{\jmath}\ln\sqrt{\frac{A+B}{A-B}} - |\overline{\jmath}|\cosh^{-1}\left[\frac{Y(\overline{\jmath})}{\sqrt{A^2-B^2}}\right] - 1 + \sqrt{1-2A+2Y(\overline{\jmath})}. \tag{6.10}$$

$\cosh^{-1}x$ denotes the positive real whose hyperbolic cosine is equal to $x$, namely $\cosh^{-1}x = \ln\left[x + \sqrt{x^2-1}\right]$, and

$$Y(\overline{\jmath}) = \overline{\jmath}^2 + \sqrt{\overline{\jmath}^4 + (1-2A)\overline{\jmath}^2 + A^2 - B^2}. \tag{6.11}$$

The expression for $\overline{f}_{\mathcal{Q}_2}(\overline{\jmath})$ involves the combinations of the model parameters

$$\ln\sqrt{\frac{A+B}{A-B}} = (\beta_1 - \beta_2)\frac{\Delta e}{2}, \qquad A^2 - B^2 = \overline{\nu}_1^2\overline{\nu}_2^2\left(1-\gamma_1^2\right)\left(1-\gamma_2^2\right), \tag{6.12}$$
$$A = \overline{\nu}_1\overline{\nu}_2(1-\gamma_1\gamma_2).$$

The expression (6.10) for $\overline{f}_{\mathcal{Q}_2}(\overline{\jmath})$ can be rewritten in two different forms according to the sign of $\overline{\jmath}$. By using the identity $\cosh^{-1}x = \ln\left[x + \sqrt{x^2-1}\right]$, the $\cosh^{-1}$ term in (6.10) can be split into two contributions and, according to the sign of $\overline{\jmath}$, we get

$$\overline{f}_{\mathcal{Q}_2}(\overline{\jmath}) \underset{\overline{\jmath}<0}{=} -\overline{\jmath}\ln(A-B) + \overline{\jmath}\ln\left[Y(\overline{\jmath}) + \sqrt{Y^2(\overline{\jmath}) - (A^2-B^2)}\right] - 1 + \sqrt{1-2A+2Y(\overline{\jmath})}, \tag{6.13}$$

while

$$\overline{f}_{\mathcal{Q}_2}(\overline{\jmath}) \underset{\overline{\jmath}>0}{=} +\overline{\jmath}\ln(A+B) - \overline{\jmath}\ln\left[Y(\overline{\jmath}) + \sqrt{Y^2(\overline{\jmath}) - (A^2-B^2)}\right] - 1 + \sqrt{1-2A+2Y(\overline{\jmath})}. \tag{6.14}$$

In the limit where $T_1$ vanishes ($A \to B$), the latter expressions yield the results discussed in section 8.

The thermodynamical and kinetic parameters of the heat baths are disentangled if, in place of $A$ and $B$, we consider the parameters

$$p_+ = \tfrac{1}{2}(1+\gamma_1)(1-\gamma_2) \qquad \text{and} \qquad p_- = \tfrac{1}{2}(1-\gamma_1)(1+\gamma_2). \tag{6.15}$$

The relations with $A$ and $B$ are $A = \overline{\nu}_1\overline{\nu}_2\left(p_+ + p_-\right)$ and $B = \overline{\nu}_1\overline{\nu}_2\left(p_+ - p_-\right)$. Therefore, $(A+B)/(A-B) = p_+/p_-$, $\sqrt{A^2-B^2} = 2\overline{\nu}_1\overline{\nu}_2\sqrt{p_+p_-}$. Then, by virtue of the relation (6.4)
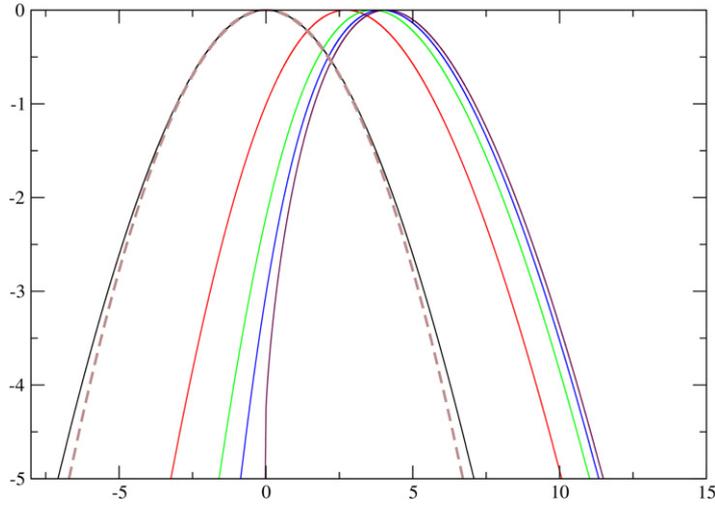
**Figure 3.** Effect of increasing the non-equilibrium driving parameter $\gamma_1 - \gamma_2$. A few large deviation functions are represented for various values of $\gamma_1$, when the other parameters are fixed to the sample values $\gamma_2 = 0.1$, $\nu_1 = 100$, $\nu_2 = 10$. The leftmost curve (for $\gamma_1 = 0.1$) is the equilibrium large deviation function and the dashed curve is the quadratic with the same curvature at the origin. The rightmost curve is for $\gamma_1 = 1$, i.e. heat bath 1 at zero temperature. For the intermediate curves, from left to right, $\gamma_1$ takes the values 0.7, 0.9, 0.966 6667.

and the expression (6.10) for $\overline{f}_{\mathcal{Q}_2}(\overline{\jmath})$, $\widetilde{f}_{\mathcal{Q}_2}(\widetilde{\jmath})$ reads

$$
\widetilde{f}_{\mathcal{Q}_2}(\widetilde{\jmath}) = \widetilde{\jmath} \ln \sqrt{\frac{p_+}{p_-}} - |\widetilde{\jmath}| \cosh^{-1}\left[\frac{Z(\widetilde{\jmath})}{\sqrt{p_+ p_-}}\right]
$$
$$
+ \frac{\nu_1 + \nu_2}{2}\left[-1 + \sqrt{1 - 2\overline{\nu}_1 \overline{\nu}_2 \left[p_+ + p_- - 2Z(\widetilde{\jmath})\right]}\right], \tag{6.16}
$$

where, with the definition $Y(\overline{\jmath})/\sqrt{A^2 - B^2} \equiv Z(\widetilde{\jmath})/\sqrt{p_+ p_-}$,

$$
Z(\widetilde{\jmath}) = \frac{1}{\nu_1 \nu_2}\left[2\widetilde{\jmath}^2 + \sqrt{4\widetilde{\jmath}^4 + \left[(\nu_1 + \nu_2)^2 - 2\nu_1 \nu_2 (p_+ + p_-)\right]\widetilde{\jmath}^2 + (\nu_1 \nu_2)^2 p_+ p_-}\right]. \tag{6.17}
$$

By virtue of the definitions (6.15) of $p_+$ and $p_-$, the thermodynamic parameters of the thermal baths appear in $\widetilde{f}_{\mathcal{Q}_2}(j)$ through the following combinations

$$
\sqrt{\frac{p_+}{p_-}} = e^{(\beta_1 - \beta_2)\Delta e/2}, \qquad \sqrt{p_+ p_-} = \tfrac{1}{2}\sqrt{(1 - \gamma_1^2)(1 - \gamma_2^2)}, \qquad p_+ + p_- = 1 - \gamma_1 \gamma_2. \tag{6.18}
$$

At equilibrium the large deviation function is even. As $\gamma_1 - \gamma_2$ increases, the large deviation function becomes more and more asymmetric. In the zero-temperature limit $\gamma_1 = 1$, the large deviation function becomes infinite for $\widetilde{\jmath} < 0$. Figure 3 illustrates the changes in the shape of the large deviation function, with increasing departure from equilibrium.

Some generic properties of a large deviation function can be checked in the case of the above explicit formulas. By virtue of (5.15) $\langle \overline{\jmath}\rangle_{\mathrm{st}} = B$ and one checks that $\overline{f}_{\mathcal{Q}_2}(\langle \overline{\jmath}\rangle_{\mathrm{st}}) = 0$,

$\overline{f}'_{\mathcal{Q}_2}(\langle \overline{\jmath} \rangle_{\mathrm{st}}) = 0$ and $\overline{f}''_{\mathcal{Q}_2}(\langle \overline{\jmath} \rangle_{\mathrm{st}}) = -1/(A - B^2)$, namely

$$\overline{f}''_{\mathcal{Q}_2}(\langle \overline{\jmath} \rangle_{\mathrm{st}}) = - \left[ \lim_{\tau \to +\infty} \frac{\kappa_{n_2}^{[2]}}{\tau} \right]^{-1}. \tag{6.19}$$

The expression (6.10) for $\overline{f}_{\mathcal{Q}_2}(\overline{\jmath})$ is the sum of a term $\overline{\jmath}(\beta_1 - \beta_2)\Delta e/2$ and an even function of $\overline{\jmath}$. As a consequence, we check again that $\overline{f}_{\mathcal{Q}_2}(\overline{\jmath}) - \overline{f}_{\mathcal{Q}_2}(-\overline{\jmath}) = \overline{\jmath}(\beta_1 - \beta_2)\Delta e$, namely, by virtue of (6.4), $f_{\mathcal{Q}_2}(\mathcal{J})$ obeys the fluctuation relation $f_{\mathcal{Q}_2}(\mathcal{J}) - f_{\mathcal{Q}_2}(-\mathcal{J}) = (\beta_1 - \beta_2)\mathcal{J}$. Moreover, the absolute value of $\overline{\jmath}$ in the expression (6.10) for $\overline{f}_{\mathcal{Q}_2}(\overline{\jmath})$ is responsible for a (rather mild) singularity in the curve $f_{\mathcal{Q}_2}(\mathcal{J})$ at $\mathcal{J} = 0$: a jump in the third derivative.

We notice that the large current behavior of $\overline{f}_{\mathcal{Q}_2}(\overline{\jmath})$ in the present model reads

$$\overline{f}_{\mathcal{Q}_2}(\overline{\jmath}) \underset{|\overline{\jmath}| \to +\infty}{\sim} -2|\overline{\jmath}| \ln |\overline{\jmath}|. \tag{6.20}$$

### 6.2. Derivation from a saddle-point method

Before embarking on the derivation, let us explain why the saddle-point method is not straightforward for this model.

The saddle-point approximation or expansion is well-suited for the asymptotic study of integrals of the form $\int_\Gamma \mathrm{d}z\, \Psi(z) \mathrm{e}^{\tau \Phi(z)}$, where $\tau$ is some large real parameter and the functions $\Psi(z)$, $\Phi(z)$ are holomorphic in a domain large enough such that the initial integration contour $\Gamma$ can be deformed to a steepest descent path while remaining within the holomorphicity domain during the deformation. One may also have to encircle some singularities when deforming the contour, and then one must keep track of their contributions, which may or may not dominate the saddle-point contribution. This can of course be generalized to a finite sum $\sum_i \int_{\Gamma_i} \mathrm{d}z\, \Psi_i(z) \mathrm{e}^{\tau \Phi_i(z)}$ when each individual term satisfies the hypotheses above. Note, however, that to get the leading behavior one may have to take into account possible destructive interferences between different pieces, for instance if the real parts of saddle-point values are the same for several $\Psi_i$, or if the saddle points for certain terms compete with encircled singularities for other terms.

In our case, we deal with an integral of the type $\int_\Gamma \mathrm{d}z \sum_i \Psi_i(z) \mathrm{e}^{\tau \Phi_i(z)}$ where there is a single integration contour, and the sum[4] $\sum_i \Psi_i(z) \mathrm{e}^{\tau \Phi_i(z)}$ has nice holomorphicity properties that allow one to deform contours (almost) freely, but each term in itself has singularities and cuts. So we have to face a kind of dilemma: either we want to keep holomorphicity, then the large parameter does not appear in an exponential—and to our knowledge no straightforward constant phase technique applies—or we look at each pure exponential piece individually, and then some branch cuts may prevent one from deforming the contour purely as a constant phase steepest descent path: the steepest descent path is not closed, some parts of the original path are deformed along the cuts and they may dominate the saddle-point. But, also, the contribution of the pure exponential pieces may interfere. In our case, we have managed to show that in fact the interferences between contributions of one pure exponential and cut contributions from another pure exponential are destructive (with remainder terms controlled explicitly), leaving the contribution of only a single saddle-point (not one saddle-point for each pure exponential). But our argument relies on some tricks and features that appear to us at this stage as coincidences: we have not been

---

[4] Which in our case consists of only two terms.

able to identify a general framework avoiding our tedious analysis. And indeed, examples are known [25, 26] where (depending possibly on parameters) the cut contributions do or do not dominate the saddle-point.

To conclude these comments, let us mention one general direction that seems worth pursuing, though we have not been able to use it to simplify significantly our argument even in our simple case. In physical problems, the functions $\Phi_i(z)$ will often be closely related to the different branches of a single algebraic function. For instance, the functions $\Phi_i(z)$ are often closely related to the eigenvalues of some $z$-dependent matrix. So a natural route would be to regard the integrals not in the $z$ plane, but on the appropriate uniformizing Riemann surface, in our case an elliptic curve.

We now turn to the detailed analysis.

*6.2.1. Method.* The current probability density $\Pi(\bar{\jmath}; \tau)$ is related to the probability $P\left(\mathcal{Q}_2/(\tau\Delta e) \in [\bar{\jmath}, \bar{\jmath} + d\bar{\jmath}[; \tau\right)$ by the definition $P\left(\mathcal{Q}_2/(\tau\Delta e) \in [\bar{\jmath}, \bar{\jmath} + d\bar{\jmath}[; \tau\right) \equiv \Pi(\bar{\jmath}; \tau)d\bar{\jmath}$. Since $\mathcal{Q}_2/\Delta e$ can take only integer values, the density distribution $\Pi(\bar{\jmath}; \tau)$ is a sum of Dirac distributions

$$\Pi(\bar{\jmath}; \tau) = \sum_{n=-\infty}^{+\infty} \delta\left(\bar{\jmath} - \frac{n}{\tau}\right) \tau P\left(\frac{\mathcal{Q}_2}{\Delta e} = n; \tau\right). \tag{6.21}$$

In the long-time limit

$$\Pi(\bar{\jmath}; \tau) \underset{\tau \to +\infty}{\sim} \sum_{n=-\infty}^{+\infty} \delta\left(\bar{\jmath} - \frac{n}{\tau}\right) \tau P^{\mathrm{as}}(\bar{\jmath}; \tau), \tag{6.22}$$

where $P^{\mathrm{as}}(\bar{\jmath}; \tau)$ is a function of the continuous parameter $\bar{\jmath}$ that we shall compute below, and which is such that the following asymptotic behavior holds:

$$P\left(\frac{\mathcal{Q}_2}{\Delta e} = \tau\bar{\jmath}; \tau\right)\bigg|_{\tau\bar{\jmath}\,\mathrm{integer}} \underset{\tau \to +\infty}{\sim} P^{\mathrm{as}}(\bar{\jmath}; \tau). \tag{6.23}$$

The notation $g(\bar{\jmath}, \tau)|_{\tau\bar{\jmath}\,\mathrm{integer}}$ is a reminder of the rule that if the function $g(\bar{\jmath}, \tau)$ is given by an integral representation, the latter must be calculated in the case where $\tau\bar{\jmath}$ is an integer. By using one of the ad hoc definitions of the large deviation function introduced in appendix E.2 of [16], the function $P^{\mathrm{as}}(\bar{\jmath}; \tau)$ can be rewritten as

$$P^{\mathrm{as}}(\bar{\jmath}; \tau) = A(\bar{\jmath}, \tau)\mathrm{e}^{\tau\bar{f}_{\mathcal{Q}_2}(\bar{\jmath})} \qquad \mathrm{with} \quad \lim_{\tau \to +\infty} \frac{1}{\tau} \ln A(\bar{\jmath}, \tau) = 0. \tag{6.24}$$

When one is interested only in the large deviation function, the only information to be retained from the latter equation is merely

$$\bar{f}_{\mathcal{Q}_2}(\bar{\jmath}) = \lim_{\tau \to +\infty} \frac{1}{\tau} \ln P(\mathcal{Q}_2 = \tau\bar{\jmath}\Delta e; \tau)\bigg|_{\tau\bar{\jmath}\,\mathrm{integer}}. \tag{6.25}$$

Consequently, $\bar{f}_{\mathcal{Q}_2}(\bar{\jmath})$ can be investigated by means of a saddle-point method applied to the representation of $P(\mathcal{Q}_2 = \tau\bar{\jmath}\Delta e; \tau)|_{\tau\bar{\jmath}\,\mathrm{integer}}$ in the complex $z$ plane given by (4.45). In the latter expression $P(\mathcal{Q}_2 = n\Delta e; \tau)$ is equal to $\mathrm{e}^{-\tau}$ times a linear combination of $c_n(\tau)$, $s_n(\tau)$, $s_{n+1}(\tau)$ and $s_{n-1}(\tau)$. When $T_1 = 0$ the expressions (4.34) and (4.35) of the latter functions are convenient for studying the large $\tau$ behavior of $c_{n=\tau\bar{\jmath}}(\tau)$, $s_{n=\tau\bar{\jmath}}(\tau)$. When

$T_1 \neq 0$ the study is slightly more complicated and it is more conveniently performed by considering the related coefficients defined by $c_n(\tau) = \rho^n \widetilde{c}_n(\tau)$ and $s_n(\tau) = \rho^n \widetilde{s}_n(\tau)$, where the expression (4.53) of $\rho$ is finite when $T_1 \neq 0$. We present the details in the case where $T_1 \neq 0$.

When $T_1 \neq 0$, in the long-time limit, we have to consider the behaviors of the functions

$$K_c(\bar{\jmath}; \tau) \equiv \widetilde{c}_{n=\tau\bar{\jmath}}(\tau) = \oint_{|z|=1} \frac{\mathrm{d}z}{2\pi\mathrm{i}} \frac{1}{z^{\tau\bar{\jmath}+1}} \cosh\left(\tau\sqrt{\widetilde{\Delta}_+(z)}\right) \tag{6.26}$$

and

$$K_s^{(\Delta n)}(\bar{\jmath}; \tau) \equiv \widetilde{s}_{\tau\bar{\jmath}+\Delta n}(\tau) = \oint_{|z|=1} \frac{\mathrm{d}z}{2\pi\mathrm{i}} \frac{1}{z^{\Delta n}} \frac{1}{z^{\tau\bar{\jmath}+1}} \frac{\sinh\left(\tau\sqrt{\widetilde{\Delta}_+(z)}\right)}{\sqrt{\widetilde{\Delta}_+(z)}} \tag{6.27}$$

with $\Delta n = 0, 1, -1$. It is sufficient to exhibit the derivation of the long-time behavior of $K_c(\bar{\jmath}; \tau)$, because the calculation of the long-time behavior of $K_s^{(\Delta n)}(\bar{\jmath}; \tau)$ follows the same lines. Moreover, according to the property $\widetilde{c}_n(\tau) = \widetilde{c}_{|n|}(\tau)$, we have to consider only the case where $\bar{\jmath} > 0$.

For the study of the large $\tau$ limit, the cosh function in the integrand of $K_c(\bar{\jmath}; \tau)$ is split into two exponentials, and $K_c(\bar{\jmath}; \tau)$ appears as the sum of two integrals

$$2K_c(\bar{\jmath}; \tau) = \oint_{|z|=1} \frac{\mathrm{d}z}{2\pi\mathrm{i}z} \mathrm{e}^{\tau\Phi^{(+)}(z)} + \oint_{|z|=1} \frac{\mathrm{d}z}{2\pi\mathrm{i}z} \mathrm{e}^{\tau\Phi^{(-)}(z)} \tag{6.28}$$

where

$$\Phi^{(\pm)}(z) = -\bar{\jmath}\ln z \pm \sqrt{b + \frac{a}{2}\left(z + \frac{1}{z}\right)}. \tag{6.29}$$

We notice that, since $\tau\bar{\jmath}$ is in fact an integer, $\exp(-\tau\bar{\jmath}\ln z)$ is single valued and there is no cut in the complex plane $z$ associated with the logarithmic function. However, since the cosh function has been split into two exponentials, we have to consider the two cuts associated with $\sqrt{\widetilde{\Delta}_+(z)}$. These cuts are

$$]-\infty, -x_>] \qquad \text{and} \qquad [-x_<, 0], \tag{6.30}$$

where $-x_>$ and $-x_<$ are the two negative real roots of the second-order polynomial $z\widetilde{\Delta}_+(z)$, where $\widetilde{\Delta}_+(z)$ is given in (4.56). The roots are such that $0 \leq x_< < 1 < x_>$.

*6.2.2. Deformation of contours.* The large $\tau$ behavior of $K_c(\bar{\jmath}; \tau)$ can be investigated by applying the saddle-point method to the contribution from the integral involving $\Phi^{(+)}(z)$. For that purpose we have to find a way to deform the unit circle into a contour that goes through a saddle-point along a constant phase path where $\Phi^{(+)}(z)$ is maximum at the saddle-point. It can be easily found that the function $\Phi^{(+)}(z)$ has two real saddle points where $\Phi^{(+)}(z)$ as well as its second derivative are real, but only one of them corresponds to a maximum of $\Phi^{(+)}(z)$ when the real axis is crossed perpendicularly. The latter saddle-point is $x_c = \exp\left[\cosh^{-1}(x_+(\bar{\jmath}))\right]$, namely by using $\cosh^{-1}(x) = \ln[x + \sqrt{x^2 - 1}]$,

$$x_c = x_+(\bar{\jmath}) + \sqrt{x_+^2(\bar{\jmath}) - 1} \tag{6.31}$$
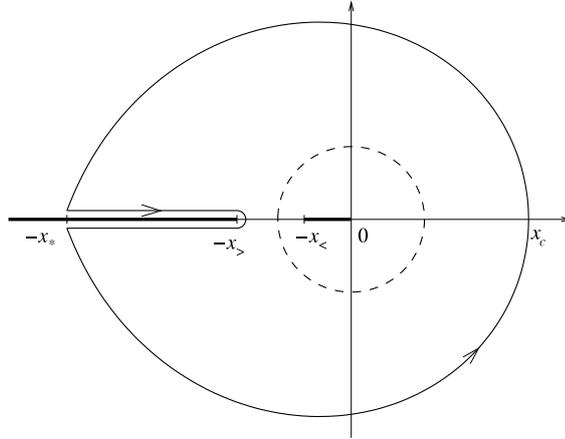
**Figure 4.** Deformed contour for $\mathrm{e}^{\tau\Phi^{(+)}(z)}$, which decomposes as a steepest descent contour $\Gamma^\star$ which goes through the saddle-point $x_\mathrm{c}$ in the anti-clockwise sense and a piece which circumvents part of the cut $]-\infty, -x_>]$ in the clockwise sense. The other cut $[-x_<, 0]$ lies inside the unit circle represented by a dotted line.

with

$$x_+(\bar{\jmath}) \equiv \frac{2}{a}\left[\bar{\jmath}^2 + \sqrt{\bar{\jmath}^4 + b\bar{\jmath}^2 + \frac{a^2}{4}}\right]. \tag{6.32}$$

The constant phase contour $\Gamma^\star$ which crosses the real axis perpendicularly at $x_\mathrm{c}$ can be looked for in the form $z_\star(\theta) = \mathrm{e}^{\lambda_\star(\theta)+\mathrm{i}\theta}$. It proves to be

$$z_\star(\theta) = \mathrm{e}^{\cosh^{-1}(y_\star(\theta;\bar{\jmath}))+\mathrm{i}\theta}, \tag{6.33}$$

where $\theta \in \,]-\pi, \pi[$ and

$$y_\star(\theta;\bar{\jmath}) = u(\theta;\bar{\jmath})\cos\theta + \sqrt{1 + \frac{2b}{a}u(\theta;\bar{\jmath}) + u^2(\theta;\bar{\jmath})} \tag{6.34}$$

with

$$u(\theta;\bar{\jmath}) = \frac{\theta^2}{\sin^2\theta}\frac{2\bar{\jmath}^2}{a}. \tag{6.35}$$

The contour $\Gamma^\star$ crosses the negative real axis at the point $z_\star(\theta = \pi) \equiv -x_\star(\bar{\jmath})$ with

$$-x_\star(\bar{\jmath}) = -\exp\left[\cosh^{-1}\left(\frac{b}{a} + \pi^2\frac{\bar{\jmath}^2}{a}\right)\right] \tag{6.36}$$

which lies on the cut $]-\infty, -x_>]$, because $-x_> = -\exp\left[-\cosh^{-1}(b/a)\right]$. As a consequence, the unit circle can be deformed into the contour $\Gamma^\star$ and a contour $\mathcal{C}^{(+)}_{[-x_\star, -x_>]}$ that goes around the cut $]-\infty, -x_>]$ between the points $-x_\star(\bar{\jmath})$ and $-x_>$ in the clockwise sense (see figure 4)

$$\oint_{|z|=1}\frac{\mathrm{d}z}{2\pi\mathrm{i}z}\mathrm{e}^{\tau\Phi^{(+)}(z)} = \oint_{\Gamma^\star}\frac{\mathrm{d}z}{2\pi\mathrm{i}z}\mathrm{e}^{\tau\Phi^{(+)}(z)} + \oint_{\mathcal{C}^{(+)}_{[-x_\star, -x_>]}}\frac{\mathrm{d}z}{2\pi\mathrm{i}z}\mathrm{e}^{\tau\Phi^{(+)}(z)}. \tag{6.37}$$
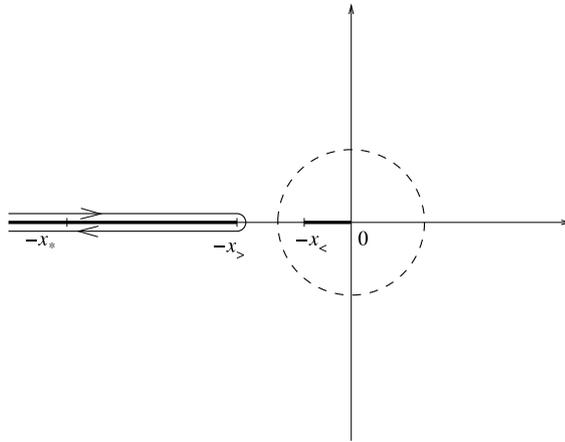
**Figure 5.** Deformed contour for $\mathrm{e}^{\tau\Phi^{(-)}(z)}$, consisting of the circle at infinity, not represented here, and a path which circumvents the whole cut $]-\infty, -x_>]$ in the clockwise sense.

On the other hand, in the integral involving $\Phi^{(-)}(z)$ the unit circle can be deformed into a circle, minus the point on the negative real axis, with radius $R$ that goes to infinity and a path around the cut $]-\infty, -x_>]$. By using the parametrization $z = R\mathrm{e}^{\mathrm{i}\theta}$, with $\theta \neq \pi$, we get the following large $|z|$ behavior: $|\exp\left[\tau\Phi^{(-)}(z)\right]|\Big|_{|z|\to+\infty} \sim \exp\left[-\tau\sqrt{aR}\cos(\theta/2)\right]$, so that the contribution of the integral along a circle of radius $R$ vanishes in the limit where $R$ goes to infinity. Consequently, (see figure 5),

$$\oint_{|z|=1} \frac{\mathrm{d}z}{2\pi\mathrm{i}z}\mathrm{e}^{\tau\Phi^{(-)}(z)} = \oint_{\mathcal{C}^{(+)}_{]-\infty,-x_>]}} \frac{\mathrm{d}z}{2\pi\mathrm{i}z}\mathrm{e}^{\tau\Phi^{(-)}(z)}. \tag{6.38}$$

The crucial point is then to notice that the sum $\mathrm{e}^{\tau\Phi^{(+)}(z)} + \mathrm{e}^{\tau\Phi^{(-)}(z)}$ is an analytic function of $z$, which has no cut along the interval $]-\infty, -x_>]$. As a consequence, the integral along the contour $\mathcal{C}^{(+)}_{]-\infty,-x_>]}$ with $\Phi^{(-)}(z)$ can be replaced by the opposite of the same integral with $\Phi^{(+)}(z)$ in place of $\Phi^{(-)}(z)$, and the equality (6.38) becomes

$$\oint_{|z|=1} \frac{\mathrm{d}z}{2\pi\mathrm{i}z}\mathrm{e}^{\tau\Phi^{(-)}(z)} = \oint_{\mathcal{C}^{(-)}_{]-\infty,-x_>]}} \frac{\mathrm{d}z}{2\pi\mathrm{i}z}\mathrm{e}^{\tau\Phi^{(+)}(z)}, \tag{6.39}$$

where $\mathcal{C}^{(-)}_{]-\infty,-x_>]}$ is the contour which goes around the cut $]-\infty, -x_>]$ in the anti-clockwise sense.

When the contributions (6.37) and (6.39) from the integrals involving respectively $\Phi^{(+)}$ and $\Phi^{(-)}$ are summed according to the definition (6.28) of $K_c(\bar{\jmath};\tau)$, we get

$$2K_c(\bar{\jmath};\tau) = \oint_{\Gamma^\star} \frac{\mathrm{d}z}{2\pi\mathrm{i}z}\mathrm{e}^{\tau\Phi^{(+)}(z)} + \oint_{\mathcal{C}^{(-)}_{]-\infty,-x_\star]}} \frac{\mathrm{d}z}{2\pi\mathrm{i}z}\mathrm{e}^{\tau\Phi^{(+)}(z)}. \tag{6.40}$$

We stress that $\exp\left[\tau\Phi^{(+)}(z)\right]$ diverges when $|z|$ goes to infinity, except on the negative real axis, so that the contour integral along the cut $]-\infty, -x_\star]$ cannot be closed at the point $z = -\infty$. The expression (6.40) corresponds to integrating $\mathrm{e}^{\tau\Phi^{(+)}(z)}/(2\pi\mathrm{i}z)$ along the contour in figure 6.
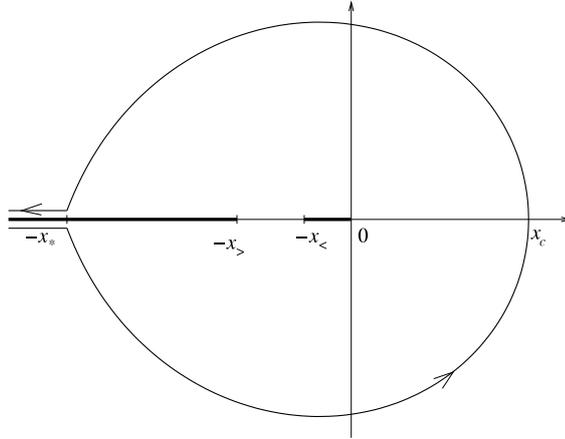
**Figure 6.** Contour of integration for $\mathrm{e}^{\tau \Phi^{(+)}(z)}$ in the integral representation (6.40) of $K_{\mathrm{c}}(\bar{\jmath}; \tau)$.

On the contour $\mathcal{C}_{]-\infty, -x_\star]}$, $z = \mathrm{e}^{\lambda + \mathrm{i}\sigma\pi}$, where $\sigma = +1$ if $z$ is above the cut and $\sigma = -1$ otherwise. Since $\tau\bar{\jmath} = n$, where $n$ is an integer,

$$\oint_{\mathcal{C}^{(-)}_{]-\infty, -x_\star]}} \frac{\mathrm{d}z}{2\pi \mathrm{i}z} \mathrm{e}^{\tau \Phi^{(+)}(z)} = (-1)^n \int_{\ln x_\star}^{+\infty} \frac{\mathrm{d}\lambda}{\pi} \mathrm{e}^{-\tau\bar{\jmath}\lambda} \sin\left(\tau\sqrt{a\cosh\lambda - b}\right). \quad (6.41)$$

The sign of this contribution changes for two consecutive values of $\bar{\jmath}$, but its absolute value is bounded,

$$\left| \oint_{\mathcal{C}^{(-)}_{]-\infty, -x_\star]}} \frac{\mathrm{d}z}{2\pi \mathrm{i}z} \mathrm{e}^{\tau \Phi^{(+)}(z)} \right| \leq \frac{1}{\pi\tau\bar{\jmath}} \mathrm{e}^{-\tau\bar{\jmath}\ln x_\star}. \quad (6.42)$$

*6.2.3. Large $\tau$ behavior.* According to the saddle-point formula

$$\oint_{\Gamma^\star} \frac{\mathrm{d}z}{2\pi \mathrm{i}z} \mathrm{e}^{\tau \Phi^{(+)}(z)} \underset{\tau \to +\infty}{\sim} \frac{1}{\sqrt{\tau}} \frac{1}{x_{\mathrm{c}}\sqrt{2\pi \times \mathrm{d}^2\Phi^{(+)}/\mathrm{d}z^2|_{z=x_{\mathrm{c}}}}} \mathrm{e}^{\tau \Phi^{(+)}(x_{\mathrm{c}})}. \quad (6.43)$$

By using the inequalities, $-\ln x_\star < -\ln x_{\mathrm{c}}$, derived from the expression (6.31) and (6.32) and (6.36), and $-\bar{\jmath}\ln x_{\mathrm{c}} \leq \Phi^{(+)}(x_{\mathrm{c}})$, derived from (6.29), the bound exhibited in (6.42) implies that

$$\oint_{\mathcal{C}^{(-)}_{]-\infty, -x_\star]}} \frac{\mathrm{d}z}{2\pi \mathrm{i}z} \mathrm{e}^{\tau \Phi^{(+)}(z)} = o\left(\mathrm{e}^{\tau \Phi^{(+)}(x_{\mathrm{c}})}\right) \quad (6.44)$$

where $o\left(\mathrm{e}^{\tau g(j)}\right)$ denotes a function which decays faster than $\mathrm{e}^{\tau g(j)}$ when $\tau$ goes to $+\infty$.

Eventually, the definition (6.26) of $K_{\mathrm{c}}(\bar{\jmath}; \tau)$ and the decomposition (6.40) together with the behaviors (6.43) et (6.44) lead to

$$\widetilde{c}_{\tau\bar{\jmath}}(\tau)\big|_{\tau\bar{\jmath}\,\mathrm{integer}} \underset{\tau \to +\infty}{\sim} A_{\widetilde{c}} \times \mathrm{e}^{\tau \Phi^{(+)}(x_{\mathrm{c}})}, \quad (6.45)$$

where

$$\Phi^{(+)}(x_{\mathrm{c}}) = -\bar{\jmath}\ln[x_+(\bar{\jmath}) + \sqrt{x_+^2(\bar{\jmath}) - 1}] + \sqrt{b + ax_+(\bar{\jmath})} \quad (6.46)$$

and

$$A_{\widetilde{c}} = \frac{1}{\sqrt{\tau}} \frac{1}{2\sqrt{2\pi}} \frac{1}{x_{\mathrm{c}} \sqrt{\mathrm{d}^2 \Phi^{(+)}/\mathrm{d}z^2|_{z=x_{\mathrm{c}}}}} \tag{6.47}$$

with $x_{\mathrm{c}} = x_+(\bar{\jmath}) + \sqrt{x_+^2(\bar{\jmath}) - 1}$ and

$$\mathrm{d}^2 \Phi^{(+)}/\mathrm{d}z^2|_{z=x_{\mathrm{c}}} = \frac{2}{a} \bar{\jmath} \frac{\sqrt{\bar{\jmath}^4 + b\bar{\jmath}^2 + a^2/4}}{[x_+(\bar{\jmath}) + \sqrt{x_+^2(\bar{\jmath}) - 1}\,]^2 \sqrt{x_+^2(\bar{\jmath}) - 1}}. \tag{6.48}$$

The same argument can be performed for $K_{\mathrm{s}}^{(\Delta n)}(\bar{\jmath}; \tau)$ defined in (6.27), with the result

$$\widetilde{s}_{\tau\bar{\jmath}+\Delta n}|_{\tau\bar{\jmath}\,\text{integer}}(\tau) \underset{\tau\to+\infty}{\sim} A_{\widetilde{s}} \times \mathrm{e}^{\tau\Phi^{(+)}(x_{\mathrm{c}})} \qquad \text{with } A_{\widetilde{s}} = \left[ x_{\mathrm{c}}^{\Delta n} \sqrt{\widetilde{\Delta}_+(x_{\mathrm{c}})} \right]^{-1} A_{\widetilde{c}}. \tag{6.49}$$

As a consequence,

$$\lim_{\tau\to+\infty} \frac{1}{\tau} \ln \left[ \mathrm{e}^{-\tau} \rho^{\tau\bar{\jmath}} \widetilde{c}_{\tau\bar{\jmath}}(\tau) \right] = \overline{f}(\bar{\jmath}) \tag{6.50}$$

and

$$\lim_{\tau\to+\infty} \frac{1}{\tau} \ln \left[ \mathrm{e}^{-\tau} \rho^{\tau\bar{\jmath}} \widetilde{s}_{\tau\bar{\jmath}+\Delta n}(\tau) \right] = \overline{f}(\bar{\jmath}) \tag{6.51}$$

with $\overline{f}(\bar{\jmath}) = -1 + \bar{\jmath} \ln \rho + \Phi^{(+)}(x_{\mathrm{c}})$. By virtue of (4.53), $\ln \rho = \ln \sqrt{(A+B)/(A-B)}$ and, according to the definitions in (4.56), $\cosh^{-1}((1-b)/a) = \cosh^{-1}(A/(\sqrt{A^2-B^2})) = \ln \sqrt{(A+B)/(A-B)}$. Therefore $\overline{f}(\bar{\jmath})$ reads

$$\overline{f}(\bar{\jmath}) = -1 + \bar{\jmath} \cosh^{-1}\left(\frac{1-b}{a}\right) - |\bar{\jmath}| \ln \left[ x_+(\bar{\jmath}) + \sqrt{x_+^2(\bar{\jmath}) - 1} \right] + \sqrt{b + ax_+(\bar{\jmath})}, \tag{6.52}$$

where $x_+(\bar{\jmath})$ is defined in (6.32).

Eventually, according to (4.45), $P(\mathcal{Q}_2/\Delta e = n; \tau)$ is a finite linear combination of functions of $n$ plus a finite increment $\Delta n$, which can be rewritten as

$$P\left(\frac{\mathcal{Q}_2}{\Delta e} = n; \tau\right) \underset{n=\tau\bar{\jmath}}{=} \sum_{\Delta n = 0, 1, -1} b_{\Delta n} \, g_{\Delta n}(\tau\bar{\jmath} + \Delta n; \tau), \tag{6.53}$$

and all functions $g_{\Delta n}$ prove to have the same 'large deviation function' $\overline{f}(\bar{\jmath})$ in the sense of definition (6.25),

$$g_{\Delta n}(\tau\bar{\jmath} + \Delta n; \tau) \underset{\substack{\tau\to+\infty \\ \tau\bar{\jmath}\,\text{integer}}}{\sim} A_{\Delta n}(\bar{\jmath}, \tau) \mathrm{e}^{\tau\overline{f}(\bar{\jmath})}. \tag{6.54}$$

Therefore, by comparison with (6.23) and (6.24) we get

$$\overline{f}_{\mathcal{Q}_2}(\bar{\jmath}) = \overline{f}(\bar{\jmath}) \qquad \text{and} \qquad A(\bar{\jmath}, \tau) = \tau \sum_{\Delta n = 0, 1, -1} b_{\Delta n} A_{\Delta n}(\bar{\jmath}, \tau), \tag{6.55}$$

where the expression (6.52) of $\overline{f}(\bar{\jmath})$ indeed coincides with the result (6.10).

### 6.3. Derivation by Laplace's method on a discrete sum

As the reader may have noticed, the computation of the large deviation function via contour integrals is a bit tricky and clumsy due to the cuts, and relies on some compensations which are not totally obvious to foresee.

In the case at hand it is possible to derive the large deviation function via Laplace's method applied to a discrete sum of non-negative contributions. We illustrate this briefly in the case of $K_{\mathrm{s}}^{(\Delta n)}$.

The key is an explicit formula for $\sin(\tau \sqrt{\widetilde{\Delta}_+(z)})/\sqrt{\widetilde{\Delta}_+(z)}$ as a Laurent series in $z$. From the symmetry $z \leftrightarrow 1/z$ we can concentrate on positive powers of $z$. We start with

$$\frac{\sin\left(\tau \sqrt{\widetilde{\Delta}_+(z)}\right)}{\sqrt{\widetilde{\Delta}_+(z)}} = \sum_{k=0}^{+\infty} \frac{\tau^{2k+1}}{(2k+1)!}[\widetilde{\Delta}_+(z)]^k \tag{6.56}$$

and expand $[\widetilde{\Delta}_+(z)]^k$ as a Laurent polynomial in $z$,

$$\left(b + \frac{a}{2}\left(z + \frac{1}{z}\right)\right)^k = \sum_{\substack{l,m \geq 0 \\ l+m \leq k}} z^{l-m}\left(\frac{a}{2}\right)^{l+m} b^{k-l-m} \frac{k!}{l!m!(k-l-m)!}. \tag{6.57}$$

So for $n \geq 0$ one gets (take $l = m + n$ above)

$$\oint_{|z|=1} \frac{\mathrm{d}z}{2\pi\mathrm{i}} \frac{1}{z^{n+1}} \frac{\sin\left(\tau \sqrt{\widetilde{\Delta}_+(z)}\right)}{\sqrt{\widetilde{\Delta}_+(z)}} = \sum_{m \geq 0} \sum_{k \geq 2m+n} \frac{\tau^{k+1}}{(2k+1)!}\left(\frac{a}{2}\right)^{2m+n}$$

$$\times\, b^{k-2m-n} \frac{k!}{m!(m+n)!(k-2m-n)!}. \tag{6.58}$$

As $\tau$, $a$ and $b$ are $> 0$, this is a (double) sum of positive terms, and we are interested in the limit

$$\tau \to +\infty \qquad n = \tau \bar{\jmath} + \Delta n \qquad \text{with } \Delta n \in \{0, -1, 1\}. \tag{6.59}$$

It is straightforward to check that in this limit the maximal term in the (double) sum is in the bulk (i.e. not for $m = 0$ or $k = 2m + n$) and such that $k$ and $m$ scale linearly with $\tau$. One can use the Stirling approximation for all factorials and obtain the large deviation function straightforwardly, the most painful part of the computation being the location of the maximal term. We omit all details.

## 7. Dependence upon typical time scales of the thermostats

In this section we are interested in the influence of the typical time scales in the limit where the heat exchanges with one of the two thermostats become infinitely fast. We only explicitly consider the limit when $\nu_2/\nu_1$ goes to infinity, namely when heat bath 2 exchanges heat with the two-spin system far faster than heat bath 1 does. Indeed, the opposite limit when $\nu_2/\nu_1$ vanishes involves similar calculations (where the roles of $\mathcal{Q}_1^d \equiv -\mathcal{Q}_1$ and $\mathcal{Q}_2$ are interchanged), and the results in both limits are essentially the

same when they are stated in terms of quantities pertaining either to the 'fast' heat bath or to the 'slow' heat bath. The main results are summarized with the latter terminology in section 1.2.

### 7.1. Stationary mean values in the infinite $\nu_2/\nu_1$ limit

The evolution of the probability distribution $P(\sigma_1, \sigma_2; t)$ is given in (4.39), where the time scales involve the parameter $\alpha = \sqrt{1 - 4A}$ with $A = \overline{\nu}_1 \overline{\nu}_2 (1 - \gamma_1 \gamma_2)$. When the ratio of inverse time scales $\nu_2/\nu_1$ goes to $+\infty$, $\lim_{\nu_2/\nu_1 \to +\infty} \frac{1}{2}(\nu_1 + \nu_2)[1 + \alpha] = \nu_2$ and $\lim_{\nu_2/\nu_1 \to +\infty} \frac{1}{2}(\nu_1 + \nu_2)[1 - \alpha] = \nu_1(1 - \gamma_1\gamma_2)$, so that the probability distribution $P(\sigma_1, \sigma_2; t)$ goes exponentially fast to its stationary value over the time scale $1/[\nu_1(1 - \gamma_1\gamma_2)]$. Moreover, since $\gamma_\star \equiv \overline{\nu}_1 \gamma_1 + \overline{\nu}_2 \gamma_2$, $\lim_{\nu_2/\nu_1 \to +\infty} \gamma_\star = \gamma_2$, and, according to (3.6), the stationary probability distribution $P_{\mathrm{st}}(\sigma_1, \sigma_2)$ coincides with the canonical distribution at the inverse temperature $\beta_2$ of the fast thermostat, namely

$$\lim_{\nu_2/\nu_1 \to +\infty} P_{\mathrm{st}}(\sigma_1, \sigma_2) = \tfrac{1}{4}(1 + \gamma_2 \sigma_1 \sigma_2). \tag{7.1}$$

As a consequence the heat capacities $C_{\mathrm{st}}^{[a]}(T_1, T_2)$ corresponding to a variation of the temperature $T_a$ of bath $a$ (for $a = 1, 2$) and given in (3.22), become $\lim_{\nu_2/\nu_1 \to +\infty} C_{\mathrm{st}}^{[1]}(T_1, T_2) = 0$ and $\lim_{\nu_2/\nu_1 \to +\infty} C_{\mathrm{st}}^{[2]}(T_1, T_2) = C_{\mathrm{eq}}(T_2)$ respectively, while the heat capacity $C_{\mathrm{st}}(T_1, T_2)$ corresponding to equal variations of both bath temperatures, and given at the end of section 3.2.3, becomes $\lim_{\nu_2/\nu_1 \to +\infty} C_{\mathrm{st}}(T_1, T_2) = C_{\mathrm{eq}}(T_2)$.

On the other hand the time scale of the mean currents of exchanged quantities is that of the slow thermostat. Indeed, according to (3.27),

$$\lim_{\nu_2/\nu_1 \to +\infty} \langle j_2 \rangle_{\mathrm{st}} = (\gamma_1 - \gamma_2) \frac{\nu_1 \Delta e}{2}, \tag{7.2}$$

and the thermal conductivity is determined by the typical time scale $1/\nu_1$ of the slow thermostat. Similarly the housekeeping entropy flow (3.47), which is equal to the opposite of the mean exchange entropy flow in the stationary state $\mathrm{d}_{\mathrm{exch}}S/\mathrm{d}t|_{\mathrm{st}}$, becomes

$$\lim_{\nu_2/\nu_1 \to +\infty} \sigma_{\mathrm{hk}}[P_{\mathrm{st}}] = \nu_1 (\gamma_1 - \gamma_2)(\beta_1 - \beta_2) \frac{\Delta e}{2}. \tag{7.3}$$

### 7.2. Various probabilities in the infinite $\nu_2/\nu_1$ limit

The stationary probability that the heat received from bath 1 is equal to $\mathcal{Q}_1 = -n_1 \Delta e$ is given by (4.44), where $\gamma_0$ is to be replaced by $\lim_{\nu_2/\nu_1 \to +\infty} \gamma_\star = \gamma_2$, with the result

$$\lim_{\nu_2/\nu_1 \to +\infty} P_{\mathrm{st}}(n_1; t) = \lim_{\nu_2/\nu_1 \to +\infty} \mathrm{e}^{-\tau} [c_{n_1}(\tau) + s_{n_1}(\tau)]. \tag{7.4}$$

According to the definitions (4.34) and (4.35), the integral representations of $c_n(\tau)$ and $s_n(\tau)$ in the complex plane involve the discriminant $\Delta_+(z)$. In order to discuss the dependence upon the inverse time scales $\nu_1$ and $\nu_2$, it is convenient to rewrite the expression (4.22) of $\Delta_+(z)$ in terms of the parameters $p_+$ and $p_-$ defined in (6.15). The correspondence with $A$ and $B$ reads $A = \overline{\nu}_1 \overline{\nu}_2 (p_+ + p_-)$ and $B = \overline{\nu}_1 \overline{\nu}_2 (p_+ - p_-)$, and

$$\Delta_+(z) = 1 + 2\overline{\nu}_1 \overline{\nu}_2 \left[ -(p_+ + p_-) + p_+ z + p_- \frac{1}{z} \right]. \tag{7.5}$$

Therefore $\lim_{\nu_2/\nu_1 \to +\infty} \Delta_+(z) = 1$ and the leading order of $\frac{1}{2}(\nu_1 + \nu_2)[-1 + \sqrt{\Delta_+(z)}]$ is merely $\frac{1}{2}\nu_1[-(p_+ + p_-) + p_+ z + p_-(1/z)]$. Consequently $e^{-\tau} c_{n_1}(\tau)$ and $e^{-\tau} s_{n_1}(\tau)$ have the same asymptotic behavior and

$$\lim_{\nu_2/\nu_1 \to +\infty} e^{-\tau} [c_n(\tau) + s_n(\tau)] = \oint_{|z|=1} \frac{\mathrm{d}z}{2\pi \mathrm{i}} \frac{1}{z^{n+1}} G_{\mathrm{RW}}(z; \nu_1 t), \tag{7.6}$$

where

$$G_{\mathrm{RW}}(z; \nu_1 t) = \exp\left\{ -\frac{1}{2}(p_+ + p_-)\nu_1 t + \frac{1}{2}\left[ p_+ z + p_- \frac{1}{z} \right] \nu_1 t \right\}. \tag{7.7}$$

This expression can be interpreted as $G_{\mathrm{RW}}(z; \nu_1 t) = \sum_{n=-\infty}^{+\infty} z^n P_{\mathrm{RW}}(n; \nu_1 t)$, namely $G_{\mathrm{RW}}(z; \nu_1 t)$ is the generating function of the probability $P_{\mathrm{RW}}(n; \nu_1 t)$ for the continuous-time random walk, also referred to as the 'randomized' random walk (see for instance page 59 of [27]), which is determined by the Markov evolution equation

$$\frac{\mathrm{d}P_{\mathrm{RW}}(n; \nu_1 t)}{\mathrm{d}t} = \frac{\nu_1}{2} \left[ -(p_+ + p_-) P_{\mathrm{RW}}(n; \nu_1 t) + p_+ P_{\mathrm{RW}}(n - 1; \nu_1 t) + p_- P_{\mathrm{RW}}(n + 1; \nu_1 t) \right], \tag{7.8}$$

and the initial condition $P_{\mathrm{RW}}(n; t = 0) = \delta_{n,0}$. By virtue of the identity which defines the generating function of modified Bessel functions $I_n(x)$,

$$e^{(1/2)[p_+ z + p_-(1/z)]\nu_1 t} = \sum_{n=-\infty}^{+\infty} \left( z\sqrt{\frac{p_+}{p_-}} \right)^n I_n \left( \nu_1 t \sqrt{p_+ p_-} \right), \tag{7.9}$$

where

$$I_n(x) \equiv \int_0^{2\pi} \frac{\mathrm{d}\theta}{2\pi} e^{-\mathrm{i}n\theta} e^{x\cos\theta}. \tag{7.10}$$

As can be derived from the latter integral representation, the modified Bessel function $I_n(x)$ is an even function of $n$, $I_{-n} = I_n$. Therefore the series representation of the expression (7.7) for the generating function $G_{\mathrm{RW}}(z; \nu_1 t)$ yields

$$P_{\mathrm{RW}}(n; \nu_1 t) = \left( \sqrt{\frac{p_+}{p_-}} \right)^n I_{|n|} \left( \sqrt{p_+ p_-} \times \nu_1 t \right) e^{-(1/2)(p_+ + p_-)\nu_1 t}. \tag{7.11}$$

Eventually the limit in (7.4) reads

$$\lim_{\nu_2/\nu_1 \to +\infty} P_{\mathrm{st}}(n_1; t) = P_{\mathrm{RW}}(n_1; \nu_1 t), \tag{7.12}$$

where the thermodynamic parameters of the thermal baths appear through the combinations of $p_+$ and $p_-$ explicitly given in (6.18). On the other hand the probability that at time $t$ the system has received a heat amount $\mathcal{Q}_2 = n_2 \Delta e$ from bath 2 is given by (4.45). Since $\lim_{\nu_2/\nu_1 \to +\infty} e^{-\tau} c_n(\tau) = \frac{1}{2} P_{\mathrm{RW}}(n; \nu_1 t)$ and $\lim_{\nu_2/\nu_1 \to +\infty} e^{-\tau} s_n(\tau) = \frac{1}{2} P_{\mathrm{RW}}(n; \nu_1 t)$,

$$\lim_{\nu_2/\nu_1 \to +\infty} P_{\mathrm{st}}(n_2; t) = \frac{1 + \gamma_2^2}{2} P_{\mathrm{RW}}(n_2; \nu_1 t) + \frac{1 - \gamma_2^2}{4} [P_{\mathrm{RW}}(n_2 + 1; \nu_1 t) + P_{\mathrm{RW}}(n_2 - 1; \nu_1 t)]. \tag{7.13}$$

We notice that at equilibrium, namely in the stationary state where $\beta_1 = \beta_2 = \beta$, $p_+ = p_-$ according to (6.18), and the limit (7.12) reads

$$\lim_{\nu_2/\nu_1 \to +\infty} P_{\mathrm{eq}}(n_1; t) = I_{|n_1|}(\tfrac{1}{2}(1 - \gamma^2)\nu_1 t)\mathrm{e}^{-(1/2)(1-\gamma^2)\nu_1 t}. \tag{7.14}$$

Similar formulae hold for the three contributions in $\lim_{\nu_2/\nu_1 \to +\infty} P_{\mathrm{eq}}(n_2; t)$, which is derived from (7.13).

### 7.3. Interpretation: mean-field regime

When the initial state is distributed according to the stationary measure, the probability that at time $t$ the system is in a configuration where $\sigma_1\sigma_2$ is equal to $\pm 1$ and that the system has received a heat amount $\mathcal{Q}_1 = -n_1\Delta e$ from bath 1 is given by (4.42) and (4.43), where $\gamma_\star$ is to be replaced by $\lim_{\nu_2/\nu_1 \to +\infty}\gamma_\star = \gamma_2$, namely

$$\lim_{\nu_2/\nu_1 \to +\infty} P_{\mathrm{st}}(\sigma_1\sigma_2 = \pm 1, n_1; t) = \tfrac{1}{2}[1 \pm \gamma_2] \lim_{\nu_2/\nu_1 \to +\infty} \mathrm{e}^{-\tau}[c_{n_1}(\tau) + s_{n_1}(\tau)]. \tag{7.15}$$

Comparison with (7.1) and (7.4) shows that the latter equation can be interpreted as

$$P_{\mathrm{st}}(\sigma_1\sigma_2 = \pm 1, n_1; t) \underset{\nu_2/\nu_1 \to +\infty}{\sim} P_{\mathrm{st}}(\sigma_1\sigma_2 = \pm 1) \times P_{\mathrm{st}}(n_1; t). \tag{7.16}$$

This is a mean-field property: between two flips of spin $\sigma_1$, spin $\sigma_2$ is flipped so many times by thermostat 2 that, when spin $\sigma_1$ is flipped again, the sign of $\sigma_1\sigma_2$ is no longer correlated to its value when the previous flip of $\sigma_1$ occurred. Therefore the variation of $n_1$, which is generated by the flip of $\sigma_1$ and the value of which is determined by the sign of $\sigma_1\sigma_2$, is no longer correlated to the sign which $\sigma_1\sigma_2$ had when the previous variation of $n_1$ occurred: the probability distributions of $\sigma_1\sigma_2$ and $n_1$ are independent of each other.

On the other hand, the probability that at time $t$ the system is in a configuration where $\sigma_1\sigma_2$ is equal to 1 and that the system has received a heat amount $\mathcal{Q}_2 = n_2\Delta e$ from bath 2 has an expression given by the remark after (4.42) and (4.43). We get

$$\lim_{\nu_2/\nu_1 \to +\infty} P_{\mathrm{st}}(\sigma_1\sigma_2 = +1, n_2; t) = \frac{(1+\gamma_2)^2}{4}P_{\mathrm{RW}}(n_2; \nu_1 t) + \frac{1-\gamma_2^2}{4}P_{\mathrm{RW}}(n_2 + 1; \nu_1 t), \tag{7.17}$$

while

$$\lim_{\nu_2/\nu_1 \to +\infty} P_{\mathrm{st}}(\sigma_1\sigma_2 = -1, n_2; t) = \frac{(1-\gamma_2)^2}{4}P_{\mathrm{RW}}(n_2; \nu_1 t) + \frac{1-\gamma_2^2}{4}P_{\mathrm{RW}}(n_2 - 1; \nu_1 t). \tag{7.18}$$

Comparison with (7.13) shows that there is no factorization similar to (7.16). In other words the variables $n_2$ and $\sigma_1\sigma_2$ are still correlated.

### 7.4. Symmetry property specific to the probability of $\mathcal{Q}_1$

The probability $P_{\mathrm{RW}}(n; \nu_1 t)$ for the continuous-time random walk, recalled in (7.11), obeys the symmetry

$$\ln \frac{P_{\mathrm{RW}}(n; \nu_1 t)}{P_{\mathrm{RW}}(-n; \nu_1 t)} = n \ln \frac{p_+}{p_-} \qquad \text{at any time } t. \tag{7.19}$$

As a consequence, according to (7.12) and the relations (6.18), the probability distribution for the heat amount dissipated towards the slow bath, $\mathcal{Q}_1^{\mathrm{d}} = -n_1\Delta e$, obeys the finite-time

symmetry

$$\ln \frac{P_{\text{st}}(\mathcal{Q}_1^{\text{d}};t)}{P_{\text{st}}(-\mathcal{Q}_1^{\text{d}};t)} \underset{\nu_2/\nu_1 \to +\infty}{=} (\beta_1 - \beta_2)\mathcal{Q}_1^{\text{d}}. \tag{7.20}$$

However, by virtue of (7.13), there is no similar finite-time symmetry property for $P_{\text{st}}(\mathcal{Q}_2;t)$.

### 7.5. Cumulants per unit time for $\mathcal{Q}_1^{\text{d}}$ and $\mathcal{Q}_2$

At any time, according to (7.7), the characteristic function $G_{\text{RW}}(\text{e}^{\overline{\lambda}};\nu_1 t)$ for the continuous-time random walk takes the very simple form

$$G_{\text{RW}}(\text{e}^{\overline{\lambda}};\nu_1 t) = \exp[t\alpha_{\text{RW}}(\overline{\lambda};\nu_1)], \tag{7.21}$$

where

$$\alpha_{\text{RW}}(\overline{\lambda};\nu_1) = \frac{\nu_1}{2}\left[-(p_+ + p_-) + p_+\text{e}^{\overline{\lambda}} + p_-\text{e}^{-\overline{\lambda}}\right]. \tag{7.22}$$

On the other hand, according to (7.12), $\sum_{n_1=-\infty}^{+\infty}\text{e}^{\overline{\lambda}n_1}\lim_{\nu_2/\nu_1 \to +\infty}P_{\text{st}}(n_1;t) = G_{\text{RW}}(\text{e}^{\overline{\lambda}};\nu_1 t)$. Therefore, the cumulants for $\mathcal{Q}_1^{\text{d}}$ are given at any finite time by the formulæ $\kappa_{n_1}^{[q]} = \partial^q \ln G_{\text{RW}}(\text{e}^{\overline{\lambda}};\nu_1 t)/\partial\lambda^q|_{\lambda=0}$ and the cumulants per unit time read for $p \geq 0$

$$\frac{1}{t}\kappa_{n_1}^{[2p+1]} \underset{\nu_2/\nu_1 \to +\infty}{=} \frac{\nu_1}{2}(p_+ - p_-) = \frac{\nu_1}{2}(\gamma_1 - \gamma_2) \tag{7.23}$$

and for $p \geq 1$

$$\frac{1}{t}\kappa_{n_1}^{[2p]} \underset{\nu_2/\nu_1 \to +\infty}{=} \frac{\nu_1}{2}(p_+ + p_-) = \frac{\nu_1}{2}(1 - \gamma_1\gamma_2). \tag{7.24}$$

In the case of $\mathcal{Q}_2$, by virtue of (7.13)

$$\sum_{n_2=-\infty}^{+\infty}\text{e}^{\overline{\lambda}n_2}P_{\text{st}}(n_2;t) \underset{\nu_2/\nu_1 \to +\infty}{=} \tfrac{1}{2}\left[1 + \gamma_2^2 + (1 - \gamma_2^2)\cosh\overline{\lambda}\right]G_{\text{RW}}(\text{e}^{\overline{\lambda}};\nu_1 t). \tag{7.25}$$

The cumulants per unit time of $\mathcal{Q}_2$ coincide with the cumulants per unit time of $\mathcal{Q}_1^{\text{d}}$ only in the long-time limit, in agreement with (5.10) and because $\mathcal{Q}_1^{\text{d}} = -\mathcal{Q}_1 = n_1\Delta e$,

$$\lim_{t \to +\infty}\frac{1}{t}\kappa_{n_2}^{[p]} = \lim_{t \to +\infty}\frac{1}{t}\kappa_{n_1}^{[p]}. \tag{7.26}$$

When the system is at equilibrium $\gamma_1 = \gamma_2$ and the long-time behavior of all odd cumulants per unit time vanish, as already noticed in section 5.2.

We notice that the previous results can also be retrieved directly from the expression for the generating function $\alpha_2(\overline{\lambda})$ of the long-time cumulants per unit time. (The limits $\nu_2/\nu_1 \to +\infty$ and $t \to +\infty$ do commute with each other.) The expression of $\alpha_2(\overline{\lambda})$ is given in (5.14), and, according to the relations after (6.15), it reads

$$\alpha_2(\overline{\lambda}) = \tfrac{1}{2}\left\{-(\nu_1 + \nu_2) + \sqrt{(\nu_1 + \nu_2)^2 + 2\nu_1\nu_2\left[-(p_+ + p_-) + p_+\text{e}^{\overline{\lambda}} + p_-\text{e}^{-\overline{\lambda}}\right]}\right\}. \tag{7.27}$$

$\alpha_2(\overline{\lambda})$ is a symmetric function of $\nu_1$ and $\nu_2$. In the limit $\nu_2/\nu_1 \to +\infty$ the generating function of the long-time cumulants per unit time becomes

$$\lim_{\nu_2/\nu_1 \to +\infty} \alpha_2(\overline{\lambda}) = \alpha_{\mathrm{RW}}(\overline{\lambda}; \nu_1), \tag{7.28}$$

where $\alpha_{\mathrm{RW}}(\overline{\lambda}; \nu_1)$ is given in (7.22).

### 7.6. Long-time current distribution in the infinite $\nu_2/\nu_1$ limit

According to the definition (7.10) of the modified Bessel function

$$I_{n=t\widetilde{\jmath}}(\alpha t) = \int_0^{2\pi} \frac{\mathrm{d}\theta}{2\pi} \mathrm{e}^{tg(\theta;\widetilde{\jmath})} \qquad \text{with } g(\theta; \widetilde{\jmath}) = -\mathrm{i}\widetilde{\jmath}\theta + \alpha\cos\theta, \tag{7.29}$$

where $\alpha$ denotes some parameter. In the complex plane where the affix reads $z = \theta + \mathrm{i}\theta'$, $g(z; \widetilde{\jmath})$, the analytic continuation of $g(\theta; \widetilde{\jmath})$, is a periodic function of $z$ with period $2\pi$ when $t\widetilde{\jmath}$ is equal to an integer. Therefore $I_{n=t\widetilde{\jmath}}$ can be rewritten as $I_{n=t\widetilde{\jmath}} = \int_{[-\pi,\pi]}(\mathrm{d}z/2\pi)\exp[tg(z;\widetilde{\jmath})]$ and, by applying a saddle-point method to the latter integral, with a deformation of the initial contour in order to exhibit the constant phase path which goes through the saddle-point in the direction where it is indeed a maximum, as done in section 6.2, one obtains (with the relevant saddle-point $z_{\mathrm{c}} = -\mathrm{i}\ln[(|\widetilde{\jmath}|/\alpha) + \sqrt{(\widetilde{\jmath}/\alpha)^2 + 1}]$) that

$$I_{n=t\widetilde{\jmath}}(\alpha t) \underset{t \to +\infty}{\sim} \frac{1}{\sqrt{2\pi\alpha t\sqrt{(\widetilde{\jmath}/\alpha)^2 + 1}}}$$

$$\times \exp\left[t\left(\alpha\sqrt{\left(\frac{\widetilde{\jmath}}{\alpha}\right)^2 + 1} - |\widetilde{\jmath}|\ln\left[\frac{|\widetilde{\jmath}|}{\alpha} + \sqrt{\left(\frac{\widetilde{\jmath}}{\alpha}\right)^2 + 1}\right]\right)\right]. \tag{7.30}$$

(The latter asymptotic behavior can also be directly read at page 378 of [28]). The long-time behavior of a current density $\Pi(\widetilde{\jmath}; t)$ is given in terms of $P(n = t\widetilde{\jmath}; t)$ by (6.22) and (6.23). From the expression (7.11) for $P_{\mathrm{RW}}(n = t\widetilde{\jmath}; \nu_1 t)$ with $\alpha = \nu_1\sqrt{p_+p_-}$, we get

$$P_{\mathrm{RW}}^{\mathrm{as}}(\widetilde{\jmath}, t; \nu_1) = A_{\mathrm{RW}}(\widetilde{\jmath}, t; \nu_1)\mathrm{e}^{tf_{\mathrm{RW}}(\widetilde{\jmath};\nu_1)} \tag{7.31}$$

with $A_{\mathrm{RW}}(\widetilde{\jmath}, t; \nu_1) = 1/\sqrt{2\pi\nu_1 t\sqrt{p_-p_- + (\widetilde{\jmath}/\nu_1)^2}}$ and

$$f_{\mathrm{RW}}(\widetilde{\jmath}; \nu_1) = -\frac{\nu_1}{2}(p_+ + p_-) + \widetilde{\jmath}\ln\sqrt{\frac{p_+}{p_-}} + |\widetilde{\jmath}|\ln\sqrt{p_+p_-} + \nu_1\sqrt{\frac{\widetilde{\jmath}^2}{\nu_1^2} + p_+p_-}$$

$$- |\widetilde{\jmath}|\ln\left[\frac{|\widetilde{\jmath}|}{\nu_1} + \sqrt{\frac{\widetilde{\jmath}^2}{\nu_1^2} + p_+p_-}\right], \tag{7.32}$$

where $f_{\mathrm{RW}}(\widetilde{\jmath}; \nu_1)$ is the large deviation function for the randomized random walk described by the Markov equation (7.8). We notice that, as predicted by large deviation theory and in particular the Gärtner–Ellis theorem, $f_{\mathrm{RW}}(\widetilde{\jmath}; \nu_1)$ can also be retrieved as the inverse Legendre transform of the generating function of the long-time cumulants per unit time $\alpha_{\mathrm{RW}}(\overline{\lambda}; \nu_1)$ written in (7.22).

The probability density of the cumulative heat current $\mathcal{Q}_1^{\mathrm{d}}/t$ in the long-time limit is given by (6.22), (7.12) and (7.31), with the result

$$\Pi^{\mathrm{as}}\left(\frac{\mathcal{Q}_1^{\mathrm{d}}}{t\Delta e} = \widetilde{\jmath}; t\right) \underset{\nu_2/\nu_1 \to +\infty}{=} \sum_{n=-\infty}^{+\infty} \delta\left(\widetilde{\jmath} - \frac{n}{t}\right) t A_{\mathrm{RW}}(\widetilde{\jmath}, t; \nu_1) \mathrm{e}^{t f_{\mathrm{RW}}(\widetilde{\jmath}; \nu_1)}. \qquad (7.33)$$

Similarly, according to (7.13), by an argument similar to that leading to (6.55), the probability density of the cumulative heat current $\mathcal{Q}_2/t$ in the long-time limit is shown to read

$$\Pi^{\mathrm{as}}\left(\frac{\mathcal{Q}_2}{t\Delta e} = \widetilde{\jmath}; t\right) \underset{\nu_2/\nu_1 \to +\infty}{=} \sum_{n=-\infty}^{+\infty} \delta\left(\widetilde{\jmath} - \frac{n}{t}\right) t A_2(\widetilde{\jmath}, t; \nu_1) \mathrm{e}^{t f_{\mathrm{RW}}(\widetilde{\jmath}; \nu_1)}, \qquad (7.34)$$

with $A_2(\widetilde{\jmath}, t; \nu_1) \neq A_{\mathrm{RW}}(\widetilde{\jmath}, t; \nu_1)$. We retrieve that the cumulative heat $\mathcal{Q}_1^{\mathrm{d}} = n_1 \Delta e$ and $\mathcal{Q}_2 = n_2 \Delta e$ have the same large deviation function, and, more precisely,

$$\lim_{\nu_2/\nu_1 \to +\infty} f_{\mathcal{Q}_2}(\mathcal{J}) = f_{\mathrm{RW}}\left(\frac{\mathcal{J}}{\Delta e}; \nu_1\right). \qquad (7.35)$$

We notice that the expression (7.32) for the large deviation function $f_{\mathrm{RW}}(\widetilde{\jmath}; \nu_1)$ agrees with the limit of the expression (6.16) for $\widetilde{f}_{\mathcal{Q}_2}(\widetilde{\jmath})$, when $\nu_2/\nu_1$ goes to infinity and $\widetilde{\jmath}/\nu_1$ is fixed. Indeed, in the expression (6.16), which is valid when $T_1 \neq 0$, the function $Z(\widetilde{\jmath})$ given in (6.17) is such that

$$\lim_{\substack{\nu_2/\nu_1 \to +\infty \\ \widetilde{\jmath}/\nu_1 \text{ fixed}}} Z(\widetilde{\jmath}) = \sqrt{\frac{\widetilde{\jmath}^2}{\nu_1^2} + p_+ p_-}, \qquad (7.36)$$

while $(\nu_1 + \nu_2)\left(-1 + \sqrt{1 - 2\overline{\nu}_1 \overline{\nu}_2 [p_+ + p_- - 2Z(\widetilde{\jmath})]}\right) \sim -\nu_1 [p_+ + p_- - 2Z(\widetilde{\jmath})].$

## 8. Case where $T_1 = 0$: pure energy dissipation towards thermal bath 1

### 8.1. Microscopic irreversibility

When the temperature $T_1$ of the colder bath vanishes, in the sense that $\beta_1 \Delta e$ goes to infinity, the microscopic reversibility (2.2) is broken,

$$(-\sigma, \sigma|\mathbb{W}|\sigma, \sigma) = 0 \qquad \text{whereas} \quad (\sigma, \sigma|\mathbb{W}|-\sigma, \sigma) \neq 0 \qquad (8.1)$$

by virtue of the expression (2.9) for the transition rates when $\gamma_1 = 1$. In other words the thermal bath at zero temperature cannot provide energy to the system, i.e., it cannot flip spin $\sigma_1$ if the flip corresponds to an increase of the two-spin system energy. There is only energy dissipation towards the zero-temperature bath.

When $\beta_1 \Delta e = +\infty$, the ratio in the modified detailed balance (2.3) vanishes or is infinite when two configurations differ from each other by the sign of $\sigma_1$; while when spin $\sigma_1$ is flipped by thermostat 1, the corresponding variation of the thermostat entropy, $\delta S_1^{\mathrm{TH}}(\mathcal{C}' \leftarrow \mathcal{C}) \equiv -\beta_1 \delta q_1(\mathcal{C}' \leftarrow \mathcal{C})$ with definition (3.26), is infinite. All direct consequences of the modified detailed balance (2.3) are no longer valid.

However, the Markov matrix (2.10) of the configuration evolution is still irreducible (see the definition after (2.10)) because histories such as

$$(\sigma, \sigma) \rightarrow (\sigma, -\sigma) \rightarrow (-\sigma, -\sigma) \rightarrow (-\sigma, \sigma) \tag{8.2}$$

correspond to a succession of flips with non-vanishing transition rates. Therefore, according to the Perron–Frobenius theorem, there still exists a single stationary distribution, and in the latter distribution every configuration has a non-vanishing weight. The stationary probability given by (3.6) is still a canonical distribution with an effective inverse temperature $\beta_*^0 = (2/\Delta e) \tanh^{-1} \gamma_\star^0$, with $\gamma_\star^0 = \overline{\nu}_1 + \overline{\nu}_2 \gamma_2$,

$$\lim_{\beta_1 \Delta e \rightarrow +\infty} P_{st}(\sigma_1, \sigma_2) = \tfrac{1}{4}[1 + \gamma_\star^0 \sigma_1 \sigma_2]. \tag{8.3}$$

According to (3.27), the mean current is finite,

$$\lim_{\beta_1 \Delta e \rightarrow +\infty} \langle j_2 \rangle_{st} = \frac{\nu_1 \nu_2}{\nu_1 + \nu_2} (1 - \gamma_2) \frac{\Delta e}{2}. \tag{8.4}$$

Since $\beta_1 \Delta e = +\infty$, the stationary exchange entropy flow (3.34) is infinitely negative in the stationary state

$$\lim_{\beta_1 \Delta e \rightarrow +\infty} \frac{d_{exch} S}{dt} \bigg|_{st} = -\infty, \tag{8.5}$$

while the rate of entropy production, which has the opposite value in the stationary state, is infinitely positive, $\lim_{\beta_1 \Delta e \rightarrow +\infty} d_{int} S^{SG}/dt|_{st} = +\infty$. We also notice that the heat capacity with respect to a variation of the temperature $T_1$ from the zero value, $C_{st}^{[1]}(T_1 = 0, T_2)$, defined in (3.19), vanishes according to (3.22).

As a consequence of the fact that the thermal bath at zero temperature cannot give energy to the system, $\mathcal{Q}_1 = -n_1 \Delta e$ is necessarily negative and

$$P(\mathcal{Q}_1, \mathcal{Q}_2; t) = 0 \qquad \text{if } \mathcal{Q}_1 > 0. \tag{8.6}$$

This can be checked on the explicit expressions of section 4.2 as follows. The probability that the system is in configuration $(\sigma_1, \sigma_2)$ at time $t_0 = 0$, in configuration $(\sigma_1', \sigma_2')$ at time $t$ and receives $\mathcal{Q}_1 = -n\Delta e$ and $\mathcal{Q}_2 = (n + \Delta n)\Delta e$ during the time interval $[0, t]$ is $(\sigma_1', \sigma_2'|\mathbb{U}(n, n + \Delta n; t)|\sigma_1, \sigma_2)$. According to (4.36) and (4.37), where $\gamma_1$ is to be set equal to 1, the latter matrix elements involve the functions $c_n(\tau)$ and $s_n(\tau)$ defined in (4.34) and (4.35). When $T_1$ vanishes, $\gamma_1$ tends to 1, $A - B = \overline{\nu}_1 \overline{\nu}_2 (1 - \gamma_1)(1 + \gamma_2)$ goes to zero and, by virtue of (4.22), $\lim_{\beta_1 \Delta e \rightarrow +\infty} \Delta_+(z) = 1 - 2A + 2Az$. As a consequence, $\cosh(\tau \sqrt{\Delta_+(z)}) = \sum_{p=0}[1/(2p)!]\tau^{2p}(1 - 2A + 2Az)^p$ contains only positive powers of $z$, and so does $\sinh(\tau \sqrt{\Delta_+(z)})/\sqrt{\Delta_+(z)}$. Consequently $c_n(\tau)$ and $s_n(\tau)$ vanish for $n < 0$ and $\lim_{\beta_1 \Delta e \rightarrow +\infty} (\sigma_1', \sigma_2'|\mathbb{U}(n, n + \Delta n; t)|\sigma_1, \sigma_2) = 0$ for any $n < 0$.

## 8.2. Long-time behavior

The explicit values of the infinite-time limit for the cumulants per unit time of the heats $\mathcal{Q}_1^d$ and $\mathcal{Q}_2$ can be calculated as in section 5.2. When $T_1$ vanishes, $\gamma_1 = 1$, $A$ tends to

$A^0 = \overline{\nu}_1\overline{\nu}_2(1 - \gamma_2)$, and $A - B$ vanishes. Then the expression (5.14) of $\alpha_2(\overline{\lambda})$ is reduced to

$$\lim_{\beta_1\Delta e \to +\infty} \alpha_2(\overline{\lambda}) = \tfrac{1}{2}\left\{-(\nu_1 + \nu_2) + \sqrt{(\nu_1 + \nu_2)^2 + 2\nu_1\nu_2(1 - \gamma_2)\left[-1 + e^{\overline{\lambda}}\right]}\right\}. \tag{8.7}$$

The expressions of the first three cumulants can be retrieved by setting $\gamma_1 = 1$ in the expressions (5.16).

As in section 6, the large deviation function can be derived either as the Legendre transform of $\alpha_2(\overline{\lambda})$ or by a saddle-point method similar to that performed in section 6.2 for $P_{P_0}(n_2; t)$ given by (4.45), which also provides the amplitude of the probability, or it can be retrieved directly by taking the limit $\gamma_1 \to 1$ in the expressions (6.13) and (6.14) for the large deviation function $\overline{f}_{\mathcal{Q}_2}(\overline{\jmath})$, as follows. In the limit where $T_1$ vanishes, so does $A - B$, and $Y(\overline{\jmath})$ tends to $Y^0(\overline{\jmath}) = \overline{\jmath}\left[\overline{\jmath} + \sqrt{\overline{\jmath}^2 + 1 - 2A^0}\right]$, where $A^0 = \overline{\nu}_1\overline{\nu}_2(1 - \gamma_2)$, while, according to (3.27) and (6.3),

$$\langle\overline{\jmath}\rangle_{\mathrm{st}}^0 = \overline{\nu}_1\overline{\nu}_2(1 - \gamma_2). \tag{8.8}$$

As a result,

$$\lim_{\beta_1\Delta e \to +\infty} f_{\mathcal{Q}_2}(\mathcal{J}) \underset{\mathcal{J}<0}{=} -\infty$$

$$\lim_{\beta_1\Delta e \to +\infty} f_{\mathcal{Q}_2}(\mathcal{J}) \underset{\mathcal{J}>0}{=} \frac{\nu_1 + \nu_2}{2}\left(-1 + \left[1 + \ln\langle\overline{\jmath}\rangle_{\mathrm{st}}^0\right]\overline{\jmath}\right.$$

$$\left. + \sqrt{\overline{\jmath}^2 + 1 - 2\langle\overline{\jmath}\rangle_{\mathrm{st}}^0} - \overline{\jmath}\ln\overline{\jmath}\left[\overline{\jmath} + \sqrt{\overline{\jmath}^2 + 1 - 2\langle\overline{\jmath}\rangle_{\mathrm{st}}^0}\right]\right) \tag{8.9}$$

where $\overline{\jmath} = [2/(\nu_1 + \nu_2)]\mathcal{J}/\Delta e$.

### 8.3. Limit where $\nu_2/\nu_1$ becomes infinite

*8.3.1. Finite-time behaviors.* The discussion can be performed along the same lines as in section 7. We only point out the features which are qualitatively different when $\gamma_1 = 1$. The discrepancies are due to the fact that, according to (6.15), when $\gamma_1 = 1$, though $p_+$ remains finite,

$$p_- = 0, \tag{8.10}$$

and the random walk process associated with the variation of the heat amounts $\mathcal{Q}_1^{\mathrm{d}}$ or $\mathcal{Q}_2$ can have only positive increments. As shown below, a Poisson process appears as a randomized random walk which can proceed only in the sense of increasing positive $n_2$.

Since $p_-$ vanishes, while $p_+ = 1 - \gamma_2$, the expression (7.5) for $\Delta_+(z)$ becomes $\lim_{\beta_1\Delta e \to +\infty}\Delta_+(z) = 1 + 2\overline{\nu}_1\overline{\nu}_2(1 - \gamma_2)[-1 + z]$ and the generating function $G_{\mathrm{RW}}(z; \nu_1 t)$ which appears in section 7.2 (see (7.7)) is to be replaced by

$$G_{\mathrm{Pois}}(z; \nu_1 t) = e^{[-1+z]p_+\nu_1 t} \tag{8.11}$$

with

$$p_+\nu_1 = \tfrac{1}{2}(1 - \gamma_2)\nu_1 = \lim_{\nu_2/\nu_1 \to +\infty}\langle\overline{\jmath}\rangle_{\mathrm{st}}^0 \equiv \langle\overline{\jmath}\rangle_{\mathrm{st}}^{0\,\star}. \tag{8.12}$$

$G_{\mathrm{Pois}}(z; \nu_1 t)$ is the generating function $G_{\mathrm{Pois}}(z; \nu_1 t) = \sum_{n=0}^{+\infty} z^n P_{\mathrm{Pois}}(n; \nu_1 t)$ for the Poisson process ruled by the Markov evolution equation

$$\frac{\mathrm{d}P_{\mathrm{Pois}}(n; \nu_1 t)}{\mathrm{d}t} = p_+ \nu_1 \left[ -P_{\mathrm{Pois}}(n; \nu_1 t) + P_{\mathrm{Pois}}(n-1; \nu_1 t) \right], \tag{8.13}$$

and the initial condition $P_{\mathrm{Pois}}(n; t = 0) = \delta_{n,0}$. The solution reads

$$P_{\mathrm{Pois}}(n; \nu_1 t) = \frac{(p_+ \nu_1 t)^n}{n!} \mathrm{e}^{-p_+ \nu_1 t}. \tag{8.14}$$

Eventually the probability that at time $t$ the system has dissipated a heat amount $\mathcal{Q}_1^{\mathrm{d}} = n_1 \Delta e$ towards bath 1 at zero temperature reads

$$\lim_{\nu_2/\nu_1 \to +\infty} \lim_{\beta_1 \Delta e \to +\infty} P_{\mathrm{st}}(n_1; t) = P_{\mathrm{Pois}}(n_1; \nu_1 t). \tag{8.15}$$

A formula similar to (7.13), where $P_{\mathrm{RW}}(n_2; \nu_1 t)$ is to be replaced by $P_{\mathrm{Pois}}(n_2; \nu_1 t)$, holds for the probability that at time $t$ the system has received a heat amount $\mathcal{Q}_2 = n_2 \Delta e$ from bath 2.

*8.3.2. Long-time behavior.* The infinite-time cumulants per unit time for the heat $\mathcal{Q}_2$ can be obtained by noticing that, when $\nu_1/\nu_2$ vanishes, the expression (8.7) for $\lim_{\beta_1 \Delta e \to +\infty} \alpha_2(\overline{\lambda})$ becomes

$$\lim_{\nu_2/\nu_1 \to +\infty} \lim_{\beta_1 \Delta e \to +\infty} \alpha_2(\overline{\lambda}) = \alpha_{\mathrm{Pois}}\left(\overline{\lambda}; \nu_1\right), \tag{8.16}$$

where

$$\alpha_{\mathrm{Pois}}\left(\overline{\lambda}; \nu_1\right) = \frac{\nu_1}{2}(1 - \gamma_2)\left[-1 + \mathrm{e}^{\overline{\lambda}}\right] \tag{8.17}$$

is the cumulant generating function for a Poisson process with average $\nu_1 p_+ = (\nu_1/2)(1 - \gamma_2)$. We notice that the two limits can be taken in the reverse order: by virtue of (7.28) $\lim_{\beta_1 \Delta e \to +\infty} \lim_{\nu_2/\nu_1 \to +\infty} \alpha_2(\overline{\lambda}) = \lim_{\beta_1 \Delta e \to +\infty} \alpha_{\mathrm{RW}}(\overline{\lambda}; \nu_1) = \alpha_{\mathrm{Pois}}(\overline{\lambda}; \nu_1)$. The long-time cumulants per unit time are all equal

$$\lim_{t \to +\infty} \frac{1}{t} \left( \lim_{\nu_2/\nu_1 \to +\infty} \lim_{\beta_1 \Delta e \to +\infty} \kappa_{n_2}^{[p]} \right) = \frac{\nu_1}{2}(1 - \gamma_2). \tag{8.18}$$

The large deviation function can be retrieved

- either as the Legendre transform of $\alpha_2(\overline{\lambda}) = \alpha_{\mathrm{Pois}}(\overline{\lambda}; \nu_1)$,
- or by a saddle-point method applied to the expression (8.14) of $P_{\mathrm{Pois}}(n; \nu_1 t)$ and similar to that performed for $P_{\mathrm{RW}}(n; \nu_1 t)$, with the result,

$$\left. P_{\mathrm{Pois}}(t\widetilde{\jmath}; \nu_1 t) \right|_{t\widetilde{\jmath} \text{ integer}} \underset{t \to +\infty}{\sim} \frac{1}{\sqrt{2\pi t\widetilde{\jmath}}} \mathrm{e}^{t f_{\mathrm{Pois}}(\widetilde{\jmath}; \nu_1)} \tag{8.19}$$

with

$$f_{\mathrm{Pois}}(\widetilde{\jmath}; \nu_1) = -\langle\widetilde{\jmath}\rangle_{\mathrm{st}}^{0\star} + \widetilde{\jmath} - \widetilde{\jmath} \ln \frac{\widetilde{\jmath}}{\langle\widetilde{\jmath}\rangle_{\mathrm{st}}^{0\star}}, \tag{8.20}$$

where $\langle\widetilde{\jmath}\rangle_{\mathrm{st}}^{0\star}$ is defined in (8.12),

- or directly by taking the limit $\nu_2/\nu_1 \to +\infty$ with $\widetilde{\jmath}/\nu_1$ fixed in the expression for the large deviation function $\lim_{\beta_1 \Delta e \to +\infty} f_{\mathcal{Q}_2}(\mathcal{J})$ given in (8.9) (and by noticing that $\langle \widetilde{\jmath} \rangle_{\mathrm{st}}^0$ is of order $\nu_1/\nu_2$, while $\overline{\jmath}^2$ is of order $(\widetilde{\jmath}/\nu_1)^2 \times (\nu_1/\nu_2)^2$),

- or by taking first the expression for $\mathcal{J} > 0$ of $\lim_{\nu_2/\nu_1 \to +\infty} f_{\mathcal{Q}_2}(\mathcal{J})$ given by (7.35) and (7.32), and then taking the limit $\beta_1 \Delta e \to +\infty$, namely $p_- \to 0$.

Eventually, the large deviation function takes the simple form

$$\lim_{\nu_2/\nu_1 \to +\infty} \lim_{\beta_1 \Delta e \to +\infty} f_{\mathcal{Q}_2}(\mathcal{J}) \underset{\mathcal{J} < 0}{=} -\infty$$

$$\lim_{\nu_2/\nu_1 \to +\infty} \lim_{\beta_1 \Delta e \to +\infty} f_{\mathcal{Q}_2}(\mathcal{J}) \underset{\mathcal{J} > 0}{=} f_{\mathrm{Pois}}\left( \frac{\mathcal{J}}{\Delta e}; \nu_1 \right) = \frac{1}{\Delta e}\left[ -\langle \mathcal{J}_2 \rangle_{\mathrm{st}}^{0\,\star} + \mathcal{J} - \mathcal{J} \ln \frac{\mathcal{J}}{\langle \mathcal{J}_2 \rangle_{\mathrm{st}}^{0\,\star}} \right] \quad (8.21)$$

where $\langle \mathcal{J}_2 \rangle_{\mathrm{st}}^{0\,\star} = (\nu_1/2)(1 - \gamma_2)\Delta e$.

The expression of $\langle \mathcal{J}_2 \rangle_{\mathrm{st}}^{0\,\star}$ can be interpreted as follows. Since bath 1 is at zero temperature, spin $\sigma_1$ may be flipped only when it is opposite to spin $\sigma_2$. Moreover, since $\nu_2 \gg \nu_1$, once spin 1 has been flipped so that $\sigma_1\sigma_2 = 1$, on average spin 2 is flipped a great odd number of times with a net energy transfer $\Delta e$ from heat bath 2 until spin $\sigma_1$ is again flipped with an energy transfer $\Delta e$ to heat bath 1 so that $\sigma_1\sigma_2 = 1$ again. As a consequence, the mean energy current through the spin system is equal to $\Delta e$ times the typical inverse time $\nu_1$ between two possible flips induced by thermal bath 1 times the probability that $\sigma_2$ is opposite to $\sigma_1$, namely $(1 - \gamma_2)/2$.

### 8.4. Limit where $\nu_2/\nu_1$ tends to zero

In the reverse limit, where $\nu_2 \ll \nu_1$, the roles of the two heat baths in the discussion of section 7 are interchanged (see comment after (4.42) and (4.43), as well as comparison of (4.44) and (4.45)). The slow thermostat is heat bath 2 and the evolution of $\mathcal{Q}_2$ is a Poisson process with the kinetic parameter $\nu_2$.

The stationary mean heat current received by the system is now $\langle \mathcal{J}_2 \rangle_{\mathrm{st}}^{0\,\star} = (\nu_2/2)(1 - \gamma_2)\Delta e$. The interpretation of the latter expression is the following. Since $\nu_2 \ll \nu_1$, as soon as spin $\sigma_2$ is flipped to a value opposite to spin $\sigma_1$ with an energy transfer $\Delta e$ from heat bath 2, bath 1 flips spin $\sigma_1$ so that $\sigma_1\sigma_2 = 1$ and an energy $\Delta e$ is transferred to heat bath 1. The next flip can be only a flip of spin $\sigma_2$, and its probability per time unit is the value of the transition rate of spin 2 when $\sigma_1\sigma_2 = 1$, namely $(1/2)\nu_2(1 - \gamma_2)$. As a consequence the mean energy current through the spins system is equal to $\Delta e$ times $(\nu_2/2)(1 - \gamma_2)$.

### Acknowledgment

### References

[1] Jarzynski C and Wójcik D K, *Classical and quantum fluctuation theorems for heat exchange*, 2004 *Phys. Rev. Lett.* **92** 230602

[2] Derrida B, *Non-equilibrium steady states: fluctuations and large deviations of the density and of the current*, 2007 *J. Stat. Mech.* P07023

[3] Bodineau T and Derrida B, *Cumulants and large deviations of the current through non-equilibrium steady states*, 2007 *C. R. Physique* **8** 540

[4] Harris R J and Schütz G M, *Fluctuation theorems for stochastic dynamics*, 2007 *J. Stat. Mech.* P07020

[5] Seifert U, *Stochastic thermodynamics, fluctuation theorems and molecular machines*, 2012 *Rep. Prog. Phys.* **75** 126001

[6] Van den Broeck C, *Stochastic thermodynamics: a brief introduction*, 2013 *Physics of Complex Colloids, Proc. Int. School of Physics 'Enrico Fermi'* ed C Bechinger, F Sciortino and P Ziherl (Amsterdam: IOS) p 155

[7] Lecomte V, Rácz Z and van Wijland F, *Energy flux distribution in a two-temperature Ising model*, 2005 *J. Stat. Mech.* P02008

[8] Zia R K P and Schmittmann B, *Probability currents as principal characteristics in the statistical mechanics of non-equilibrium steady states*, 2007 *J. Stat. Mech.* P07012

[9] Lavrentovich M O and Zia R K P, *Energy flux near the junction of two Ising chains at different temperatures*, 2010 *Europhys. Lett.* **91** 50003

[10] Corberi F, Gonnella G and Piscitelli A, *Heat exchanges in coarsening systems*, 2011 *J. Stat. Mech.* P10022

[11] Derrida B and Brunet É, 2005 *Einstein Aujourd'hui* (Les Ulis: EDP Sciences) chapter (Le mouvement Brownien et le Théorème de Fluctuation–Dissipation) p 203

[12] Visco P, *Work fluctuations for a Brownian particle between two thermostats*, 2006 *J. Stat. Mech.* P06006

[13] Sekimoto K, *Kinetic characterization of heat bath and the energetics of thermal ratchet models*, 1997 *J. Phys. Soc. Japan* **66** 1234

[14] Hatano T and Sasa S, *Steady-state thermodynamics of Langevin systems*, 2001 *Phys. Rev. Lett.* **86** 3463

[15] Gomez-Marin A and Sancho J M, *Heat fluctuations in Brownian transducers*, 2006 *Phys. Rev.* E **73** 045101(R)

[16] Cornu F and Bauer M, *Thermal Contact I. Symmetries ruled by Exchange Entropy Variations*, 2013 arXiv:1302.4538 [cond-mat.stat-mech]

[17] Glauber R J, *Time-Dependent Statistics of the Ising Model*, 1963 *J. Math. Phys.* **4** 294

[18] Komatsu T S, Nakagawa N, Sasa S and Tasaki H, *Entropy and nonlinear nonequilibrium thermodynamic relation for heat conducting steady states*, 2011 *J. Stat. Phys.* **142** 127

[19] Oono Y and Paniconi M, *Steady state thermodynamics*, 1998 *Prog. Theor. Phys. Suppl.* **130** 29

[20] Sekimoto K, *Langevin equation and thermodynamics*, 1998 *Prog. Theor. Phys. Suppl.* **130** 17

[21] Speck T and Seifert U, *Integral fluctuation theorem for the housekeeping heat*, 2005 *J. Phys. A: Math. Gen.* **38** 581

[22] Boksenbojm E, Maes C, Netočný K and Pešek J, *Heat capacity in nonequilibrium steady states*, 2011 *Europhys. Lett.* **96** 40001

[23] Touchette H, *The large deviation approach to statistical mechanics*, 2009 *Phys. Rep.* **478** 1

[24] Dembo A and Zeitouni O, 1998 *Large Deviations Techniques and Applications* 2nd edn (New York: Springer)

[25] Farago J, *Injected power fluctuations in Langevin equation*, 2002 *J. Stat. Phys.* **107** 781

[26] van Zon R and Cohen E G D, *Extended heat-fluctuation theorems for a system with deterministic and stochastic forces*, 2004 *Phys. Rev.* E **69** 056121

[27] Feller W, 1971 *An Introduction to Probability Theory and Its Applications* 2nd edn, vol II (New York: Wiley)

[28] Abramowitz M and Stegun I A (ed), 1972 *Handbook of Mathematical Functions* (New York: Dover)