PAPER

Entropy production in linear Langevin systems

View the article online for updates and enhancements.

Entropy production in linear Langevin systems

Gabriel T Landi, Tânia Tomé and Mário J de Oliveira

Instituto de Física da Universidade de São Paulo, 05314-970 São Paulo, Brazil

E-mail: gtlandi@gmail.com

Abstract

We study the entropy production rate in systems described by linear Langevin equations, containing mixed even and odd variables under time reversal. Exact formulas are derived for several important quantities in terms only of the means and covariances of the random variables in question. These include the total rate of change of the entropy, the entropy production rate, the entropy flux rate and the three components of the entropy production. All equations are cast in a way suitable for large-scale analysis of linear Langevin systems. Our results are also applied to different types of electrical circuits, which suitably illustrate the most relevant aspects of the problem.

PACS numbers: 68.43.De, 05.10.−a, 05.70.−a

(Some figures may appear in colour only in the online journal)

1. Introduction

Non-equilibrium systems have been the subject of intensive research for several decades. This is partly motivated by their broad range of applications in, e.g., physics, chemistry and biology. However, and most importantly, further progress in these areas is still hampered by more fundamental questions. Unlike equilibrium statistical mechanics, which is by now a well-established theoretical framework, in non-equilibrium statistical mechanics, several questions remain unanswered. Particularly challenging is the microscopic definition of *entropy production*. For, in the context of non-equilibrium thermodynamics [1], it provides the pathway through which irreversibility is described.

Suppose a certain system undergoes a change of state from state A to state B. If done reversibly, the total change in the entropy $\mathcal{S}$ of the system is $\Delta\mathcal{S} = \int dQ/T$, where T is the temperature and Q is the heat poured into the system. Hence $\int dQ/T$ is defined as minus the entropy flux from the system to the environment. If the process is irreversible, we have instead $\Delta\mathcal{S} \geqslant \int dQ/T$. The difference, being a positive quantity, is called the entropy production $\mathfrak{P}$:

$$\Delta\mathcal{S} - \int \frac{dQ}{T} = \mathfrak{P} \geqslant 0.$$

It is customary to divide by Δt and write, instead

$$\frac{\mathrm{d}\mathcal{S}}{\mathrm{d}t} = \Pi(t) - \Phi(t). \tag{1}$$

In this equation, $\Pi(t)$ is the entropy production rate of the system, which is always non-negative, and $\Phi(t)$ is the entropy flux rate *from* the system *to* the environment. For systems in a non-equilibrium steady state (NESS), we have $\mathrm{d}\mathcal{S}/\mathrm{d}t = 0$, which implies

$$\Pi_0 = \Phi_0 \geqslant 0.$$

It is only in thermodynamic equilibrium that the inequality in this equation becomes an equality.

Traditionally, non-equilibrium thermodynamics was founded on the basis of conservation equations. Nowadays, however, it has been realized that there are several advantages in using stochastic processes instead. For instance, by comparing forward and backward experiments, it enables one to relate the entropy production directly to the stochastic trajectories of the system [2–5]. When describing a non-equilibrium system in terms of stochastic processes, much of the focus has naturally been on Markovian dynamics, in particular using the master equation [2, 3, 6–8] or the Fokker–Planck approach [4, 9, 10], which will be the choice for this paper. We also note that non-Markovian dynamics have also been recently investigated [11].

Several formulas for the entropy production rate have been derived for both representations. In all cases, however, these are written in terms of integrals involving probability currents (cf section 3). Thus, they are not easily computed in most situations. In this paper, we will focus on linear Langevin systems; i.e., where all terms appearing in the stochastic differential equations are linear in the independent variables. First of all, one must always emphasize the importance of linear systems in view of the many circumstances in which they appear in nature. Moreover, we will show that, for such systems, it is possible to obtain exact formulas for the entropy production rate in terms of the means and variances of the independent variables. This enables one to study more complex situations, which are prohibitive for nonlinear systems. In fact, with the scheme to be derived below, it is possible to implement extremely efficient numerical procedures to study the entropy production even in large-scale systems.

The entropy production gives an insight into the properties of systems out of equilibrium. And with the formulas developed in this paper, it becomes simple to compute the entropy production even for the most complex linear systems. Moreover, these results are not restricted to the steady state as in most recent papers, but also naturally include the time dependence. We thus hope that these results are of value to deepen our understanding of non-equilibrium physics.

When discussing entropy production, it is paramount to distinguish between variables that are even and odd under time reversal. With this in mind, we shall illustrate our results by applying them to electrical circuits. These are excellent platforms for this type of problem, since they contain mixed even (charge and voltage) and odd (current) variables. These studies trace back to the works of Landauer [12], and the discussion about the differences between the minimum entropy principle and the maximum entropy principle. They were also recently discussed in [13] using the Onsager–Machlup Lagrangian [14].

The problem and the basic underlying equations will be stated in section 2. General remarks about the entropy production rate will be given in section 3 and the derivation of the formulas for linear systems will be carried out in section 4. Some of the lengthier calculations were postponed to the appendix. The applications in electrical circuits are contained in section 5 and the conclusions in section 6.

With the exception of section 5, we shall try to maintain the following notation. Vectors are denoted by lower-case letters, such as $x = (x_1, \ldots, x_n)$ and matrices by upper-case letters such as A and B. The exception will be random variables, such as X, which will also be in upper-case letters. All vectors are treated as column vectors, with x^{T} representing the corresponding row vector. Scalars and scalar functions are denoted by calligraphic letters, such as $\mathcal{S}$ or $\mathcal{P}$ (the only exception being Π and Φ in equation (1)). The gradient of $\mathcal{P}$ with respect to x is abbreviated as $\partial\mathcal{P}/\partial x$.

2. Systems described by Langevin equations

2.1. The Langevin equation

Let $X = (X_1, \ldots, X_n)$ be a vector of random variables satisfying

$$\dot{X} = f(X, t) + B\dot{\xi}(t). \tag{2}$$

In this equation, $f(x, t)$ is an arbitrary n-dimensional function of x and the time t. $\xi(t)$ are m-independent *standard* Wiener processes and, therefore, B is an $n \times m$ matrix[1]. In this paper, we shall focus mainly on linear systems, for which we write

$$f(x, t) = -Ax + b(t), \tag{3}$$

where A is $n \times n$ and $b(t)$ is an arbitrary n-dimensional function of time. We shall also make the reasonable assumption that all eigenvalues of A are in the open right-plane, thence guaranteeing the stability of the solutions.

The expectation of x is denoted by $\bar{x} = \langle X \rangle$, both notations being used interchangeably. The equation for the time evolution of $\bar{x}$ is obtained directly by taking the expectation of equation (2):

$$\frac{\mathrm{d}\bar{x}}{\mathrm{d}t} = f(\bar{x}, t) = -A\bar{x} + b(t). \tag{4}$$

Next, we obtain the equation describing the time evolution of the second moments. The latter can be constructed from the outer product $\langle XX^{\mathrm{T}} \rangle$, which gives a matrix whose (i, j)th entry is $\langle X_i X_j \rangle$. The result—obtained, for instance, by discretizing time, taking the outer product and then averaging—is

$$\frac{\mathrm{d}\langle XX^{\mathrm{T}} \rangle}{\mathrm{d}t} = Xf^{\mathrm{T}} + fX^{\mathrm{T}} + BB^{\mathrm{t}}. \tag{5}$$

For linear systems, it is more convenient to work with the covariance matrix:

$$\Theta = \langle XX^{\mathrm{T}} \rangle - \langle X \rangle \langle X \rangle^{\mathrm{T}}. \tag{6}$$

Note that Θ is symmetric by construction. Moreover, being a covariance matrix, it is also positive definite. Using equations (4) and (5) and assuming linearity, as in equation (3), we find that

$$\frac{\mathrm{d}\Theta}{\mathrm{d}t} = -(A\Theta + \Theta A^{\mathrm{T}}) + 2D \tag{7}$$

which is a matrix differential equation giving the time evolution of the covariance matrix. Here, D is the $n \times n$ diffusion tensor defined as

$$D = \tfrac{1}{2}BB^{\mathrm{T}}. \tag{8}$$

If B has full row rank, then D is positive definite; otherwise, it is positive semi-definite.

---

[1] Choosing B to be $n \times m$ is convenient because it includes the possibilities that a variable contains more than one source of noise or that the same noise is shared with more than one variable.

Equation (7) shows an important property of linear systems of Langevin equations, which is seldom discussed in the literature: all terms involving the external forcing term, $b(t)$, are dropped out. This means that the variability of the random variables in question is not influenced by external forces. In other words, we could say linear systems are not amenable to *synchronization*.

If $b(t) = b$, a constant, then the system will eventually reach equilibrium. (We are making the explicit assumption that A is stable.) The equilibrium value of $\bar{x}$ is read immediately from equation (4): $\bar{x}_0 = A^{-1}b$. Similarly, setting $\dot{\Theta}_0 = 0$ in equation (7), we obtain the matrix equation

$$A\Theta_0 + \Theta_0 A^{\mathrm{T}} = 2D. \tag{9}$$

This is a very important equation. It appears frequently in the literature of electrical engineering and control systems where it goes by the name of *continuous time Lyapunov equation*. It is also a particular case of the broader class of *Sylvester equations*.

It seems appropriate to stop now to briefly discuss the solution methods of equation (9). For algebraic solutions, we can transform it into a linear system of equations as follows. Let $A \otimes B$ denote the Kronecker product of two matrices A and B and define the operation $\mathrm{vec}(A)$ as that of creating a vector by stacking the columns of A. Then equation (9) can be written as

$$[(I \otimes A) + (A \otimes I)]\mathrm{vec}(\Theta_0) = 2\mathrm{vec}(D). \tag{10}$$

Given that Θ is symmetric, several equations will be repeated and, in the event that A is sparse, several equations will be of the form $0 = 0$, thence simplifying somewhat the computations. On the other hand, this approach should *never* be used for numerical calculations. The computational complexity of equation (10) is $O(n)^6$. However, specialized algorithms have been developed which reduce this to $O(n)^3$, a substantial improvement [15].

2.2. The Fokker–Planck equation

Let $\mathcal{P}(x, t)$ denote the probability density function corresponding to the vector of random variables X. The Fokker–Planck equation for $\mathcal{P}(x, t)$ associated with the general Langevin equation (2) reads

$$\frac{\partial \mathcal{P}}{\partial t} = -\sum_j \frac{\partial}{\partial x_j}[f_j(x, t)\mathcal{P}] + \sum_{j,k} \frac{\partial^2}{\partial x_j \partial x_k}[D_{jk}\mathcal{P}]. \tag{11}$$

It is also convenient to write this equation in the form of a continuity equation. For this, let us define the probability current

$$g(x, t) = f(x, t)\mathcal{P}(x, t) - D\frac{\partial \mathcal{P}(x, t)}{\partial x}. \tag{12}$$

Equation (11) can then be written as

$$\frac{\partial \mathcal{P}}{\partial t} = -\frac{\partial}{\partial x} \cdot g, \tag{13}$$

where the operation in the right-hand side denotes the divergence of g.

The Fokker–Planck equation for linear systems satisfies the very important property that its solution must be of the form of a multivariate normal distribution[2]:

$$\mathcal{P}(x, t) = \frac{1}{\sqrt{(2\pi)^n |\Theta|}} \exp\left\{-\frac{1}{2}(x - \bar{x})^{\mathrm{T}} \Theta^{-1}(x - \bar{x})\right\}, \tag{14}$$

---

[2] This can be understood intuitively by noting that, upon discretizing time in equation (2), $x_{t+\Delta t}$ is formed by summing normally distributed random variables ($\Delta\xi$ and x_t). But sums of normally distributed random variables must also be normally distributed and so will $x_{t+\Delta t}$.

where $\bar{x}$ and Θ are given by equations (4) and (7) (both being, in general, functions of time, which has been omitted for clarity). Here $|\Theta|$ denotes the determinant.

In the event that $b(t)$ is a constant, the system will reach a steady state whose distribution is

$$\mathcal{P}_0(x) = \frac{1}{\sqrt{(2\pi)^n |\Theta_0|}} \exp\left\{-\frac{1}{2}(x - \bar{x}_0)^{\mathrm{T}} \Theta_0^{-1}(x - \bar{x}_0)\right\}, \tag{15}$$

where $\bar{x}_0 = A^{-1}b$ and Θ_0 is given by equation (9).

Throughout this paper, we will make frequent use of the fact that both $\mathcal{P}(x, t)$ and $\partial \mathcal{P}/\partial x$ vanish at the boundaries of the probability space, so that cross terms appearing when integrating by parts can always be neglected.

2.3. Distinction between variables that are even and odd under time reversal

In studying the entropy production, it is important to distinguish between odd variables, which reverse sign under time reversal, and even variables, which do not. Examples of even variables include the position in mechanical systems and charges or voltages in circuits, their odd counterparts being velocities and currents. Following [4] and [16], let us define, for each variable x_i, a quantity ϵ_i, such that $\epsilon_i = \pm 1$ if x_i is even or odd, respectively. Moreover, let us define a diagonal matrix $E = \mathrm{diag}(\epsilon_1, \ldots, \epsilon_n)$. Then, time reversal is achieved by the operation $x \to Ex$.[3]

Let us also divide $f(x, t)$ in equation (2) into irreversible and reversible parts:

$$f(x, t) = f^{\mathrm{irr}}(x, t) + f^{\mathrm{rev}}(x, t), \tag{16}$$

where

$$\begin{aligned} f^{\mathrm{irr}}(x, t) &= \tfrac{1}{2}[f(x, t) + Ef(Ex, t)] = Ef^{\mathrm{irr}}(Ex, t) \\ f^{\mathrm{rev}}(x, t) &= \tfrac{1}{2}[f(x, t) - Ef(Ex, t)] = -Ef^{\mathrm{rev}}(Ex, t). \end{aligned} \tag{17}$$

It is convenient to separate even and odd variables by writing $x = (x_1, x_2)$, where it is agreed that x_1 contains all even variables and x_2 all odd ones (the dimensions of x_1 and x_2 depending on the problem in question). We then have that

$$Ex = (x_1, -x_2). \tag{18}$$

Let us also define

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \tag{19}$$

where the block matrices A_{ij} have dimensions compatible with those of x_1 and x_2. Then, according to equation (17), we may write

$$A = A^{\mathrm{irr}} + A^{\mathrm{rev}}, \tag{20}$$

where

$$A^{\mathrm{irr}} = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}, \qquad A^{\mathrm{rev}} = \begin{bmatrix} 0 & A_{12} \\ A_{21} & 0 \end{bmatrix}. \tag{21}$$

Finally, we divide the external forcing term $b(t)$ as $b(t) = (b_1(t), b_2(t))$ with b_1 and b_2 having the same dimensions as x_1 and x_2. It then follows again from equation (17) that

$$b^{\mathrm{irr}} = (b_1, 0), \qquad b^{\mathrm{rev}} = (0, b_2). \tag{22}$$

---

[3] The matrix E satisfies $E^{-1} = E$. Moreover, while this will not be necessary in this paper, it is interesting to note that the operators $\frac{1}{2}(I \pm E)$ may be used to select the even and odd variables, respectively.

Explicit examples of this separation are given in section 5 for several types of electric circuits.

It will also be important to distinguish the irreversible and reversible parts of the probability current:

$$g = g^{\mathrm{irr}} + g^{\mathrm{rev}}, \tag{23}$$

where

$$g^{\mathrm{irr}}(x, t) = f^{\mathrm{irr}}(x, t)\mathcal{P}(x, t) - D\frac{\partial \mathcal{P}(x, t)}{\partial x} \tag{24}$$

$$g^{\mathrm{rev}}(x, t) = f^{\mathrm{rev}}(x, t)\mathcal{P}(x, t). \tag{25}$$

2.4. Conditions for equilibrium

In the steady state, $\dot{\mathcal{P}}_0 = 0$ and thence $\frac{\partial}{\partial x} \cdot g_0 = 0$. Moreover, the probability currents $g_0(x)$ should be such that even components change sign under time reversal, whereas odd components do not, i.e.,

$$g_0(x) = -Eg_0(Ex). \tag{26}$$

Using the definitions in equations (23)–(25), we find

$$g_0(Ex) = [f^{\mathrm{rev}}(Ex) + f^{\mathrm{irr}}(Ex)]\mathcal{P}_0(Ex) - D\frac{\partial \mathcal{P}_0(Ex)}{\partial Ex}.$$

With the aid of equation (17), this may be written as

$$g_0(Ex) = E[-f^{\mathrm{rev}}(x) + f^{\mathrm{irr}}(x)]\mathcal{P}_0(Ex) - DE\frac{\partial \mathcal{P}_0(Ex)}{\partial x}.$$

Whence,

$$\begin{aligned} g_0(x) + Eg_0(Ex) &= f^{\mathrm{rev}}(x)[\mathcal{P}_0(x) - \mathcal{P}_0(Ex)] + f^{\mathrm{irr}}(x)[\mathcal{P}_0(x) + \mathcal{P}_0(Ex)] \\ &\quad - D\frac{\partial \mathcal{P}_0(x)}{\partial x} - EDE\frac{\partial \mathcal{P}_0(Ex)}{\partial x} \\ &= 0. \end{aligned} \tag{27}$$

If $\mathcal{P}_0(Ex) = \mathcal{P}_0(x)$ (the distribution is an even function of the odd variables), the first term in equation (27) vanishes. Moreover, if $EDE = D$ (which is generally true if D is diagonal or block-diagonal), then by comparing equation (27) with definition (24), we conclude that

$$g_0^{\mathrm{irr}} \equiv f^{\mathrm{irr}}(x)\mathcal{P}_0(x) - D\frac{\partial \mathcal{P}_0(x)}{\partial x} = 0,$$

i.e., the irreversible portion of the current is zero in the steady state. This condition has already been discussed in the context where D is diagonal [16]. For the more general case presented here, we see that the two conditions for this to happen are that $\mathcal{P}_0(Ex) = \mathcal{P}_0(x)$ and $EDE = D$.

Being this the case, we find from $\frac{\partial}{\partial x} \cdot g_0 = 0$ that

$$\frac{\partial}{\partial x} \cdot g_0^{\mathrm{rev}} = \left[\frac{\partial}{\partial x} \cdot f^{\mathrm{rev}}(x)\right]\mathcal{P}_0 + [f^{\mathrm{rev}}(x)]^{\mathrm{T}}\frac{\partial \mathcal{P}_0}{\partial x} = 0.$$

In most physical systems, the reversible part of the force, $f^{\mathrm{rev}}(x)$, is divergence less:

$$\frac{\partial}{\partial x} \cdot f^{\mathrm{rev}}(x) = 0. \tag{28}$$

In this case, we may use equation (24) to write $\frac{\partial \mathcal{P}_0}{\partial x} = D^{-1}f^{\mathrm{irr}}(x)\mathcal{P}_0(x)$ so as to finally arrive at

$$[f^{\mathrm{rev}}(x)]^{\mathrm{T}}D^{-1}[f^{\mathrm{irr}}(x)] = 0. \tag{29}$$

This is the required condition for equilibrium. It says that the vectors f^{rev} and f^{irr} should be orthogonal in a space whose metric is D^{-1}. In the event that D is semi-definite, then by D^{-1} we mean the pseudo-inverse of D.

6

3. Entropy production: general statement

The following calculations follow closely those of [9], but generalize it in two ways. First, it includes systems with arbitrary combinations of even and odd variables and second, it includes non-diagonal Ds.

The entropy is defined as

$$S = - \int \mathcal{P} \log \mathcal{P} \, dx. \tag{30}$$

Differentiating with respect to time, using equation (13) and integrating by parts, we find

$$\frac{dS}{dt} = - \int \frac{\partial \mathcal{P}}{\partial t} \log \mathcal{P} \, dx = - \int g^{\mathrm{T}} \frac{\partial \mathcal{P}}{\partial x} \frac{dx}{\mathcal{P}}. \tag{31}$$

Separating g as in equation (23), we find for the term involving g^{rev} after integrating by parts

$$- \int (g^{\mathrm{rev}})^{\mathrm{T}} \frac{\partial \mathcal{P}}{\partial x} \frac{dx}{\mathcal{P}} = \int \left[\frac{\partial}{\partial x} \cdot f^{\mathrm{rev}} \right] \mathcal{P} \, dx = 0,$$

since we are assuming f^{rev} is divergence less [cf equation (28)]. Hence, equation (31) becomes

$$\frac{dS}{dt} = - \int g^{\mathrm{irr\,T}} \frac{\partial \mathcal{P}}{\partial x} \frac{dx}{\mathcal{P}}. \tag{32}$$

Next, we use equation (24) to write

$$\frac{\partial \mathcal{P}}{\partial x} = D^{-1}[f^{\mathrm{irr}}(x, t)\mathcal{P} - g^{\mathrm{irr}}(x, t]$$

and obtain

$$\frac{dS}{dt} = \int g^{\mathrm{irr\,T}} D^{-1} g^{\mathrm{irr}} \frac{dx}{\mathcal{P}} - \int g^{\mathrm{irr\,T}} D^{-1} f^{\mathrm{irr}} \, dx. \tag{33}$$

The first term is a quadratic form. Since, D is positive definite, it is always non-negative and can thence be identified with the entropy production rate:

$$\Pi(t) = \int g^{\mathrm{irr\,T}} D^{-1} g^{\mathrm{irr}} \frac{dx}{\mathcal{P}}. \tag{34}$$

Consequently, the entropy flux rate is found to be

$$\Phi(t) = - \int g^{\mathrm{irr\,T}} D^{-1} f^{\mathrm{irr}} \, dx. \tag{35}$$

Using a different approach, equation (34) was further separated in [4] into three terms. These, unlike equation (34), hold *only* for diagonal D. They also assume that a steady-state, time-independent, configuration exists. Usually, we will simply take this to mean that $b(t) = b$, a constant. In our present notation, the formulas are

$$\Pi_1(t) = - \int \frac{\partial \mathcal{P}}{\partial t} \log \frac{\mathcal{P}(x, t)}{\mathcal{P}_0(x)} \, dx \tag{36}$$

$$\Pi_2(t) = \int \left[g_0^{\mathrm{irr}}(Ex) \right]^{\mathrm{T}} D^{-1} \left[g_0^{\mathrm{irr}}(Ex) \right]^{\mathrm{T}} \frac{\mathcal{P}(x, t)}{[\mathcal{P}_0(Ex)]^2} \, dx \tag{37}$$

$$\Pi_3(t) = - \int \frac{\partial \mathcal{P}}{\partial t} \log \frac{\mathcal{P}_0(x)}{\mathcal{P}_0(Ex)} \, dx, \tag{38}$$

where $\mathcal{P}_0$ is given by equation (15) and g_0^{irr} is obtained from g^{irr} in equation (24) by replacing $\mathcal{P}$ with $\mathcal{P}_0$. The first contribution, termed nonadiabatic, is related to the relaxation processes of the system and should be zero at a NESS. On the other hand, the second contribution, termed adiabatic, is directly related to the absence of detailed balance in the system and will be the sole contribution to Π when a NESS is reached. Both Π_1 and Π_2 have been shown [4] to obey integral fluctuation theorems. However, the same is not true of the third term, Π_3.

4. Entropy production for linear systems

The last section summarized a series of results valid for general Langevin systems. In this section, we shall specialize them to linear systems, as in equation (3), and obtain formulas for all relevant quantities. The major advantage of considering linear systems is that all results can be expressed in terms of $\bar{x}$ and Θ, thus avoiding the necessity of performing the several integrals involved in those expressions.

4.1. Rate of the change of the total entropy

We begin with the total rate of change of the entropy. Substituting equation (14) for $\log \mathcal{P}$ in equation (30), we find

$$\mathcal{S} = \frac{1}{2} \int [(x - \bar{x})^{\mathrm{T}} \Theta^{-1} (x - \bar{x}) + \log(2\pi)^n |\Theta|] \mathcal{P} \, \mathrm{d}x.$$

The last term is simply $\log(2\pi)^n |\Theta|$. To compute the first term, we proceed as follows. Since Θ is symmetric positive definite, we can construct its square, triangular Cholesky factorization $\Theta = QQ^{\mathrm{T}}$. Now, consider the transformation $x = Qz + \bar{x}$. It transforms $(x - \bar{x})^{\mathrm{T}} \Theta^{-1} (x - \bar{x}) = z^{\mathrm{T}} z$. Whence, we see that this transformation produces a standard multivariate normal, where all variables are statistically independent. Hence, the first term acquires the form $\langle z^{\mathrm{T}} z \rangle = n$. We thus obtain

$$\mathcal{S}(t) = \frac{1}{2} \log |\Theta(t)| + \frac{n}{2} \log(2\pi e). \tag{39}$$

Differentiating equation (39) with respect to time and using a known formula of matrix calculus yields

$$\frac{\mathrm{d}\mathcal{S}}{\mathrm{d}t} = \frac{1}{2} \mathrm{tr}\left(\Theta^{-1} \frac{\mathrm{d}\Theta}{\mathrm{d}t}\right). \tag{40}$$

From equation (9), we have

$$\Theta^{-1} \frac{\mathrm{d}\Theta}{\mathrm{d}t} = -[\Theta^{-1}A\Theta + A^{\mathrm{T}}] + 2\Theta^{-1}D.$$

The matrices A^{T} and $\Theta^{-1}A\Theta$ are similar, and thence share the same trace. Thus, equation (40) becomes

$$\frac{\mathrm{d}\mathcal{S}}{\mathrm{d}t} = \mathrm{tr}(\Theta^{-1}D - A), \tag{41}$$

where Θ^{-1} is generally a function of time. This is the required formula. It gives the rate of change of the total entropy of the system, which is seen to be a function of only the inverse of the covariance matrix, Θ^{-1}. Note, as before, that $\dot{S}$ is entirely independent of $b(t)$. (The same will not be true for Π and Φ.)

In the steady state, we set $\dot{\mathcal{S}} = 0$ in equation (41) to find

$$\mathrm{tr}\left(\Theta_0^{-1}D - A\right) = 0. \tag{42}$$

This relation can also be derived directly from the steady-state solution, equation (9), by multiplying both sides by Θ_0^{-1} and then taking the trace.

4.2. Entropy production rate and entropy flux rate

We now derive formulas for Π and Φ. The formulas for Π_1, Π_2 and Π_3 are derived by similar methods in the appendix. Before we proceed, let us establish a convenient mathematical

operation which will be used extensively, specially in the appendix. We will frequently encounter quantities of the following form:

$$\int (Fx + u)^{\mathrm{T}}(Gx + v)\mathcal{P}(x, t)\, \mathrm{d}x$$

which correspond to an average over $\mathcal{P}(x, t)$ of a quadratic form; here, F, G, u and v are arbitrary matrices and vectors. When we expand the product, there will be only one term that is of second order in x, which will have the form $x^{\mathrm{T}}F^{\mathrm{T}}Gx$. Taking the expectation of this term yields

$$\langle x^{\mathrm{T}}F^{\mathrm{T}}Gx \rangle = \sum_{ij}(F^{\mathrm{T}}G)_{ij}\langle x_i x_j \rangle.$$

Next, we use equation (6) to write $\langle x_i x_j \rangle = \Theta_{ij} + \bar{x}_i \bar{x}_j$, which results in

$$\langle x^{\mathrm{T}}F^{\mathrm{T}}Gx \rangle = \mathrm{tr}(F^{\mathrm{T}}G\Theta) + \bar{x}^{\mathrm{T}}F^{\mathrm{T}}G\bar{x}.$$

The last term is again a quadratic form and may thus be reincorporated into the original product. We thus arrive at the following relation:

$$\int (Fx + u)^{\mathrm{T}}(Gx + v)\mathcal{P}(x, t)\, \mathrm{d}x = \mathrm{tr}(F^{\mathrm{T}}G\Theta) + (F\bar{x} + u)^{\mathrm{T}}(G\bar{x} + v). \quad (43)$$

That is, whenever we average a quadratic form, there will be a trace term containing Θ and the original term inside the integral, with x replaced by $\bar{x}$.

With this result, it is now straightforward to compute the entropy production rate in equation (34). From equation (14), we have

$$\frac{\partial \mathcal{P}}{\partial x} = -\Theta^{-1}(x - \bar{x})\mathcal{P}(x, t). \quad (44)$$

Using this in equation (24), we find

$$\begin{aligned}
g^{\mathrm{irr}} &= (-A^{\mathrm{irr}}x + b^{\mathrm{irr}})\mathcal{P} - D\frac{\partial \mathcal{P}}{\partial x} \\
&= [(D\Theta^{-1} - A^{\mathrm{irr}})x + b^{\mathrm{irr}} - D\Theta^{-1}\bar{x}]\mathcal{P} \\
&= (Fx + u)\mathcal{P},
\end{aligned} \quad (45)$$

where we have defined $F = D\Theta^{-1} - A^{\mathrm{irr}}$ and $u = b^{\mathrm{irr}} - D\Theta^{-1}\bar{x}$. By applying the method described in equation (43), we obtain

$$\begin{aligned}
\Pi(t) &= \int (Fx + u)^{\mathrm{T}}D^{-1}(Fx + u)\mathcal{P}\, \mathrm{d}x \\
&= \mathrm{tr}(F^{\mathrm{T}}D^{-1}F\Theta) + (F\bar{x} + u)^{\mathrm{T}}D^{-1}(F\bar{x} + u).
\end{aligned}$$

Using that $F\bar{x} + u = b^{\mathrm{irr}} - A^{\mathrm{irr}}\bar{x}$ and expanding the matrix product inside the trace, we finally find that

$$\Pi(t) = \mathrm{tr}(D\Theta^{-1} - A^{\mathrm{irr}}) + \mathrm{tr}(A^{\mathrm{irr}\,\mathrm{T}}D^{-1}A^{\mathrm{irr}}\Theta - A^{\mathrm{irr}}) + (A^{\mathrm{irr}}\bar{x} - b^{\mathrm{irr}})^{\mathrm{T}}D^{-1}(A^{\mathrm{irr}}\bar{x} - b^{\mathrm{irr}}), \quad (46)$$

which is the required result.

We can identify the first term in equation (46) as being simply $\mathrm{d}\mathcal{S}/\mathrm{d}t$ in equation (41). Hence, from equation (1), we immediately find

$$\Phi(t) = \mathrm{tr}(A^{\mathrm{irr}\,\mathrm{T}}D^{-1}A^{\mathrm{irr}}\Theta - A^{\mathrm{irr}}) + (A^{\mathrm{irr}}\bar{x} - b^{\mathrm{irr}})^{\mathrm{T}}D^{-1}(A^{\mathrm{irr}}\bar{x} - b^{\mathrm{irr}}). \quad (47)$$

This result can also be easily derived from equation (35) using the same method.

In the steady state, making use of equation (42), we find

$$\Pi_0 = \Phi_0 = \mathrm{tr}(A^{\mathrm{irr}\,\mathrm{T}}D^{-1}A^{\mathrm{irr}}\Theta_0 - A^{\mathrm{irr}}) + (A^{\mathrm{irr}}\bar{x}_0 - b^{\mathrm{irr}})^{\mathrm{T}}D^{-1}(A^{\mathrm{irr}}\bar{x}_0 - b^{\mathrm{irr}}). \quad (48)$$

We can also use the same approach to derive formulas for the three contributions to the entropy production rate defined in equations (36)–(38). These calculations, however, are somewhat lengthier. Thus, we will simply state the results here and postpone the details to the appendix. We emphasize, once again, that these results are valid only for diagonal D and constant b.

First, the result for $\Pi_1(t)$ is

$$\Pi_1(t) = \mathrm{tr}(D\Theta^{-1} - A) + \mathrm{tr}\big(A^{\mathrm{T}}\Theta_0^{-1}\Theta - A\big) + (\bar{x} - \bar{x}_0)^{\mathrm{T}}\Theta_0^{-1}D\Theta_0^{-1}(\bar{x} - \bar{x}_0). \tag{49}$$

The nature of Π_1, as being related to the relaxation of the system, becomes quite visible from the structure of this formula. Next, we obtain for Π_2

$$\Pi_2(t) = \mathrm{tr}(A^{\mathrm{irr}\,\mathrm{T}}D^{-1}A^{\mathrm{irr}}\Theta) - \mathrm{tr}(A^{\mathrm{T}}E\Theta_0^{-1}E\Theta) + \big(A^{\mathrm{irr}}\bar{x} - b^{\mathrm{irr}}\big)^{\mathrm{T}}D^{-1}(A^{\mathrm{irr}}\bar{x} - b^{\mathrm{irr}})$$
$$- (A\bar{x} - b)^{\mathrm{T}}E\Theta_0^{-1}(E\bar{x} - \bar{x}_0). \tag{50}$$

Finally, the result for Π_3 is

$$\Pi_3(t) = \mathrm{tr}\big(A^{\mathrm{T}}E\Theta_0^{-1}E\Theta - A\big) - \mathrm{tr}\big(A^{\mathrm{T}}\Theta_0^{-1}\Theta - A\big) + (A\bar{x} - b)^{\mathrm{T}}E\Theta_0^{-1}(E\bar{x} - \bar{x}_0) \tag{51}$$

$$- (\bar{x} - \bar{x}_0)^{\mathrm{T}}\Theta_0^{-1}D\Theta_0^{-1}(\bar{x} - \bar{x}_0). \tag{52}$$

5. Application to electrical circuits

Henceforth, we will soften our previous notation of vectors and matrices and use the usual nomenclature of electrical circuits. We shall consider electrical circuits connected to different heat baths. The coupling constant is, in this case, the resistance R. The fluctuations will be assumed to correspond to white noise with spectral density given by the Jonhson–Nyquist formula $\sqrt{2RT}$, where T is the temperature of the bath.

5.1. RL circuit in series

We begin by considering a simple, yet very instructive, example of a resistor R and an inductor L in series with a battery providing a constant emf $\mathcal{E}$. The independent variable is the current I through the resistor, which is odd under time reversal. The equation for the time evolution of I is obtained from Kirchhoff's voltage law:

$$\frac{\mathrm{d}I}{\mathrm{d}t} = -\frac{R}{L}I + \frac{\mathcal{E}}{L} + \sqrt{\frac{2RT}{L^2}}\dot{\xi}. \tag{53}$$

Making reference to the notation of section 2.1 (all matrices are now 1×1), we have $A = R/L$, $b = \mathcal{E}/L$, $B = \sqrt{2RT/L^2}$ and $D = (1/2)BB^{\mathrm{T}} = RT/L^2$. For simplicity, we will assume $I(t = 0) = 0$. The expectation of I as a function of time is obtained by solving equation (4):

$$\bar{I}(t) = \frac{\mathcal{E}}{R}(1 - \mathrm{e}^{-Rt/L}). \tag{54}$$

The steady-state current is clearly seen to be $\bar{I}_0 = \mathcal{E}R$. The covariance 'matrix' is $\Theta = \langle I^2 \rangle - \bar{I}^2$ and satisfies the differential equation [cf equation (7)]

$$\frac{\mathrm{d}\Theta}{\mathrm{d}t} = -\frac{2R}{L}\Theta + \frac{2RT}{L^2} \tag{55}$$

whose solution is
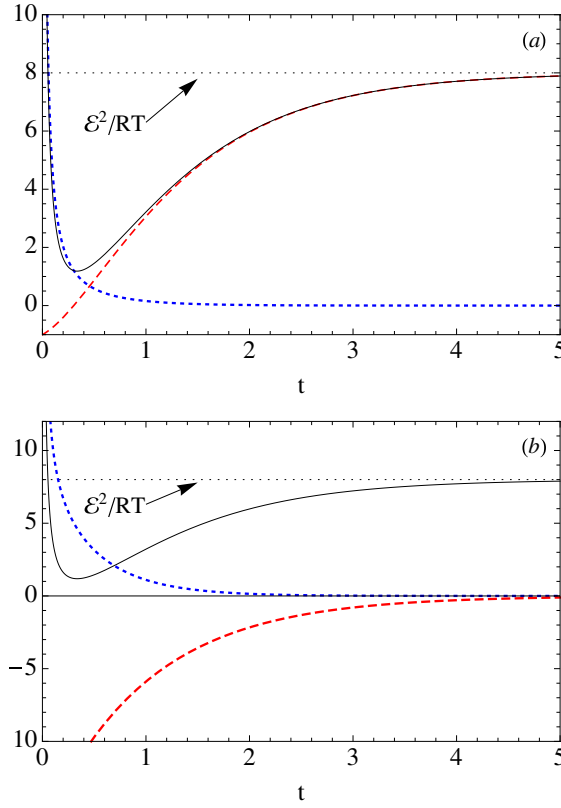
$$\Theta(t) = \frac{T}{L}(1 - \mathrm{e}^{-2Rt/L}). \tag{56}$$

Figure 1. Entropy production as a function of time for an RL circuit with $\mathcal{E} = 2$V, $R = 1\Omega$, $L = 1$H and $k_B T = 1/2$J. (*a*) $d\mathcal{S}/dt$ (blue, dotted), $\Phi(t)$ (red, dashed) and $\Pi(t)$ (black); the dotted line indicate the NESS value $\Pi_0 = \Pi_2 = \mathcal{E}^2/RT$. (*b*) The different contributions to the entropy production rate, Π_1 (blue, dotted), Π_3 (red, dashed) and Π (black). Again, Π_2 is given by the dotted horizontal line.

The current is an odd variable under time reversal. Thence, we have from equation (21) that $A^{\text{irr}} = A = R/L$ and from equation (22), $b^{\text{irr}} = 0$. With this in mind, the entropy production rate is readily computed from equation (46):

$$\Pi(t) = \frac{\mathcal{E}^2}{RT}(1 - e^{-Rt/L})^2 + \frac{R}{L}\frac{e^{-2Rt/L}}{e^{2Rt/L} - 1}. \tag{57}$$

If we take the limit $t \to \infty$, we obtain for the NESS entropy production rate

$$\Pi_0 = \Phi_0 = \frac{\mathcal{E}^2}{RT} = \frac{R\bar{I}_0^2}{T}. \tag{58}$$

This formula shows that, in the NESS, there is a constant production of entropy due to the dissipation in the resistor.

The formula for the total rate of change of the entropy is

$$\frac{d\mathcal{S}}{dt} = \frac{R}{L}\frac{1}{e^{2Rt/L} - 1}.$$

The entropy flux rate is then simply $\Phi = \Pi - d\mathcal{S}/dt$. These quantities are illustrated in figure 1(*a*) for arbitrary values of the parameters.
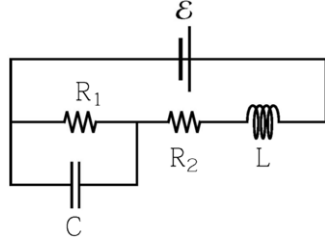
Figure 2. Scheme of the circuit studied in section 5.2. Each resistor is connected to a heat bath at a different temperature.

If we had instead studied an RC circuit then the independent variable would be the voltage through the resistor, which is even under time reversal. In this case, the formulas would be quite similar, except that now the emf would be a part of b^{irr}. The resulting expression for the entropy production would now go to zero as $t \to \infty$, since, in an RC circuit, the current (which is responsible for the dissipation) goes to zero as $t \to \infty$.

Continuing further, we may now compute formulas for the three parts of the entropy production rate, equations (49)–(51):

$$\Pi_1(t) = \frac{\mathcal{E}^2}{RT} e^{-2Rt/L} + \frac{R}{L} \frac{e^{-2Rt/L}}{e^{2Rt/L} - 1}$$

$$\Pi_2(t) = \frac{\mathcal{E}^2}{RT}$$

$$\Pi_3(t) = -2\frac{\mathcal{E}^2}{RT} e^{-Rt/L}.$$

These formulas are physically very rich. Note that Π_2, which is exactly the contribution related to the absence of detailed balance, is simply the steady-state entropy production rate Π_0 in equation (58). Graphs for these quantities are illustrated in figure 1(*b*).

5.2. Mixed RC and RL circuits

Let us now turn to the circuit denoted in figure 2, which was also studied in [13]. We now have, as independent variables, the voltage U through resistor R_1 and the current I through the resistor R_2; i.e., we have mixed even and odd variables. Using Kirchhoff's voltage law, we find the equations

$$\dot{U} = \frac{I}{C} - \frac{U}{R_1 C} + \sqrt{\frac{2T_1}{R_1 C^2}} \dot{\xi}_1 \tag{59}$$

$$\dot{I} = \frac{\mathcal{E} - U - R_2 I}{L} + \sqrt{\frac{2R_2 T_2}{L^2}} \dot{\xi}_2. \tag{60}$$

In our usual matrix notation, we have from equation (21)

$$A^{irr} = \begin{bmatrix} 1/CR_1 & 0 \\ 0 & R_2/L \end{bmatrix}, \qquad A^{rev} = \begin{bmatrix} 0 & -1/C \\ 1/L & 0 \end{bmatrix}$$

and from equation (22)

$$b^{irr} = 0, \qquad b^{rev} = b = \begin{bmatrix} 0 \\ \mathcal{E}/L \end{bmatrix}.$$

The diffusion matrix, equation (8), is

$$D = \begin{bmatrix} T_1/R_1C^2 & 0 \\ 0 & R_2T_2/L^2 \end{bmatrix}.$$

The formulas for the time dependence of the entropy production rates are now somewhat cumbersome. Thus, we will illustrate their time dependence numerically and provide formulas only for the steady-state quantities. The equilibrium values of U and I are

$$\bar{U}_0 = \frac{R_1}{R_1 + R_2}\mathcal{E}, \qquad \bar{I}_0 = \frac{1}{R_1 + R_2}\mathcal{E}.$$

The steady-state values of the covariance matrix are obtained by solving equation (9) (the notation Var and Cov meaning variance and covariance, respectively):

$$\mathrm{Var}(U) = \frac{\alpha}{C}[L(R_1 + R_2)T_1 + CR_1R_2(R_2T_1 + R_1T_2)]$$
$$\mathrm{Var}(I) = \frac{\alpha}{L}[L(R_1T_1 + R_2T_2) + CR_1R_2(R_1 + R_2)T_2]$$
$$\mathrm{Cov}(U, I) = \alpha[R_1R_2(T_2 - T_1)]$$
$$\alpha := [(R_1 + R_2)(L + CR_1R_2)]^{-1}.$$

From these results, we may now compute the steady-state entropy production rate from equation (48):

$$\Pi_0 = \Phi_0 = \frac{\bar{U}_0}{R_1T_1} + \frac{R_2\bar{I}_0}{T_2} + \frac{R_1R_2(T_2 - T_1)^2}{(R_1 + R_2)(L + CR_1R_2)T_1T_2}. \tag{61}$$

The structure of this result is worth noticing. The first two terms are similar to those appearing in equation (58), each one corresponding to a resistor. They stem from the absence of detailed balance, related to the presence of the batteries. In addition to these, however, there is now a second term related to the lack of detailed balance due to the presence of two heat baths. We see that it depends only on $(T_1 - T_2)^2$. It is thus zero when both temperatures are equal and shows that it is irrelevant which temperature is the largest.

It is worth comparing equation (61) with a similar result obtained for the same circuit in [13]. There, the third term was not present. This follows because, in the formalism of the Onsager–Machlup Lagrangian [14], the entropy production rate is related to the extremum of the Lagrangian with respect to the independent variables. Since the last term in equation (61) is independent of U and I, it does not appear in that formalism.

In figure 3(a), we show the total rate of change of the entropy, the entropy production rate and the entropy flux rate as a function of time, for arbitrary choices of the parameters. Figure 3(b) then shows the three contributions to the entropy production rate.

5.3. Inductively coupled RL circuits

Next, we turn to the interesting example of two inductively coupled RL circuits, each connected to a different heat bath. Let the suffixes 1 and 2 denote the quantities pertaining to the two circuits and let m denote the mutual inductance. Then, the corresponding dynamical equations are

$$L_1\dot{I}_1 + m\dot{I}_2 = -R_1I_1 + \mathcal{E}_1 + \sqrt{2R_1T_1}\dot{\xi}_1$$

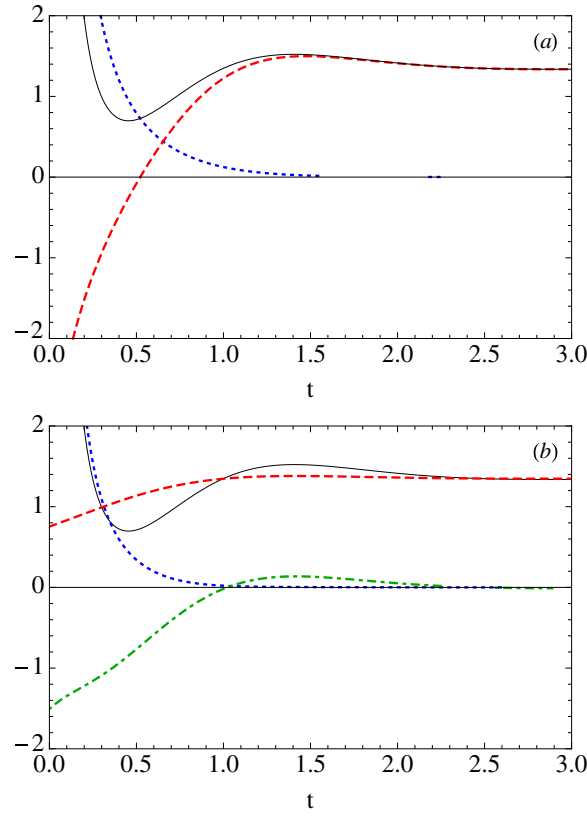$$m\dot{I}_1 + L_2\dot{I}_2 = -R_2I_2 + \mathcal{E}_2 + \sqrt{2R_2T_2}\dot{\xi}_1.$$

Figure 3. Entropy production as a function of time for the circuit in figure 2, with $\mathcal{E} = 2\,\mathrm{V}, R = 2\,\Omega$, $R_2 = 1\,\Omega$, $L = 1\,\mathrm{H}$, $C = 0.2\,\mathrm{F}$, $k_B T_1 = 1\,\mathrm{J}$ and $k_B T_2 = 2\,\mathrm{J}$. (*a*) $d\mathcal{S}/dt$ (blue, dotted), $\Phi(t)$ (red, dashed) and $\Pi(t)$ (black). (*b*) The different contributions to the entropy production rate, Π_1 (blue, dotted), Π_2 (red, dashed) and Π_3 (green, dash–dotted); $\Pi(t)$ is again shown in black.

Let us define the matrix

$$M = \begin{bmatrix} L_1 & m \\ m & L_2 \end{bmatrix}.$$

Then, in the notation of equation (2), we have

$$A = M^{-1} \begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix}$$

and

$$D = M^{-1} \begin{bmatrix} R_1 T_1 & 0 \\ 0 & R_2 T_2 \end{bmatrix} M^{-1}.$$

In this case, we see that the diffusion matrix D is not diagonal, so formulas (49)–(51) are not applicable. We will thus restrict our discussion to the steady-state expression for the entropy production rate. The steady-state currents are simply

$$\bar{I}_{10} = \frac{\mathcal{E}_1}{R_1}, \qquad \bar{I}_{20} = \frac{\mathcal{E}_2}{R_2}.$$

The steady-state covariance matrix is computed from equation (9). It reads

$$\Theta_0 = \begin{bmatrix} T_1 & \frac{mR_2(T_1-T_2)}{L_2R_1+L_1R_2} \\ \frac{mR_1(T_2-T_1)}{L_2R_1+L_1R_2} & T_2 \end{bmatrix} M^{-1}.$$

14

On noting that both variables are odd under time reversal, we obtain from equation (48)

$$\Pi_0 = \frac{\mathcal{E}_1^2}{R_1 T_1} + \frac{\mathcal{E}_2^2}{R_2 T_2} \tag{62}$$

$$+ \frac{m^2 R_1 R_2}{(L_1 L_2 - m^2)(L_2 R_1 + L_1 R_2)} \frac{(T_1 - T_2)^2}{T_1 T_2}. \tag{63}$$

Again, we arrive at a structure compatible with the two sources of disorder: the first two terms come from the batteries and the last from the presence of two heat baths. As expected, the last term goes to zero when $m \to 0$, since the mutual inductance is the sole mechanism of coupling in the system.

Finally, it is also illustrative to consider three inductively coupled RL circuits. For simplicity, we will make the assumption that all mutual inductances and self-inductances are the same, being m and L, respectively. The result is then

$$\Pi_0 = \frac{\mathcal{E}_1^2}{R_1 T_1} + \frac{\mathcal{E}_2^2}{R_2 T_2} + \frac{\mathcal{E}_3^2}{R_3 T_3} + \frac{m^2}{W}[\alpha_1 T_1 (T_2 - T_3)^2$$
$$+ \alpha_2 T_2 (T_1 - T_3)^2 + \alpha_3 T_3 (T_1 - T_2)^2] \frac{1}{T_1 T_2 T_3},$$

where

$$W = (L - m)(L + 2m)[2m^2 R_1 R_2 R_3 + L(L + m)(R_1 + R_2)(R_2 + R_3)(R_1 + R_3)]$$

and

$$\alpha_1 = R_2 R_3 \big[2m R_1^2 + L(R_1 + R_2)(R_2 + R_3) \big]$$

with similar expressions for α_2 and α_3. Again, the results depend only on the temperature differences, as expected.

6. Conclusions

In conclusion, we have studied the entropy production rate in systems of linear Langevin equations. Linearity and its consequences enabled us to compute formulas for (i) the total rate of change of the entropy, (ii) the entropy production rate *per se*, (iii) the entropy flux rate and (iv) the three contributions to the entropy production stemming from different physical sources. All these formulas were expressed in terms of the mean and covariance matrix of the random variables in question. This makes their implementation in large-scale systems quite easy. Our results were applied to electrical circuits of various types. For instance, we have shown that two circuits which are coupled only via their mutual inductance have a steady-state entropy production rate related to the difference in temperature between the two circuits.

Acknowledgments

Appendix. The three parts of the entropy production rate

In this section, we describe the steps to obtain formulas (49)–(51) from equations (36)–(38).

A.1. Equation for $\Pi_1(t)$

We begin with Π_1 in equation (36), which can be written as

$$\Pi_1(t) = \frac{d\mathcal{S}}{dt} + \Gamma_1, \tag{A.1}$$

where $d\mathcal{S}/dt$ is given by equation (41) and

$$\Gamma_1(t) = \int \frac{\partial \mathcal{P}}{\partial t} \log \mathcal{P}_0(x)\, dx. \tag{A.2}$$

Using the FP equation (13) and integrating by parts, we find

$$\Gamma_1 = \int g^{\mathrm{T}} \frac{\partial \mathcal{P}_0(x)}{\partial x} \frac{dx}{\mathcal{P}_0(x)}.$$

Similarly to equation (44), we have $\partial \mathcal{P}_0(x)/\partial x = -\Theta_0^{-1}(x - \bar{x}_0)\mathcal{P}_0(x)$. Using also equation (45), but with g instead of g^{irr}; i.e., with $F = D\Theta^{-1} - A$ and $u = b - D\Theta^{-1}\bar{x}$, we find

$$\Gamma_1 = -\int (Fx + u)^{\mathrm{T}} \Theta_0^{-1}(x - \bar{x}_0) P(x, t)\, dx.$$

Using the method of equation (43), we obtain

$$\Gamma_1 = -\mathrm{tr}(F^{\mathrm{T}}\Theta_0\Theta) - (F\bar{x} + u)^{\mathrm{T}}\Theta_0(\bar{x} - \bar{x}_0).$$

This can be simplified to read

$$\Gamma_1(t) = \mathrm{tr}\big(A^{\mathrm{T}}\Theta_0^{-1}\Theta - A\big) + (A\bar{x} - b)^{\mathrm{T}}\Theta_0^{-1}(\bar{x} - \bar{x}_0). \tag{A.3}$$

The last term in equation (A.3) can also be modified as follows. Multiplying both sides of equation (9) by Θ_0^{-1}, we find

$$\Theta_0^{-1}A + A^{\mathrm{T}}\Theta_0^{-1} = 2\Theta_0^{-1}D\Theta_0^{-1}. \tag{A.4}$$

Now, let y be an arbitrary vector. Noting that $y^{\mathrm{T}}My = y^{\mathrm{T}}M^{\mathrm{T}}y$, we find

$$y^{\mathrm{T}}\Theta_0^{-1}Ay + y^{\mathrm{T}}A^{\mathrm{T}}\Theta_0^{-1}y = 2y^{\mathrm{T}}A^{\mathrm{T}}\Theta_0^{-1}y$$

$$= 2y^{\mathrm{T}}\Theta_0^{-1}D\Theta_0^{-1}y.$$

Using this and recalling that $b = A\bar{x}_0$, we may write

$$(A\bar{x} - b)^{\mathrm{T}}\Theta_0^{-1}(\bar{x} - \bar{x}_0) = (\bar{x} - \bar{x}_0)^{\mathrm{T}}A^{\mathrm{T}}\Theta_0^{-1}(\bar{x} - \bar{x}_0)$$

$$= (\bar{x} - \bar{x}_0)^{\mathrm{T}}\Theta_0^{-1}D\Theta_0^{-1}(\bar{x} - \bar{x}_0).$$

Using this result in equation (A.3) and substituting in equation (A.1), we finally obtain equation (49).

A.2. Equation for $\Pi_3(t)$

Let us next compute Π_3 in equation (38). It can be written as

$$\Pi_3(t) = \Gamma_2(t) - \Gamma_1(t), \tag{A.5}$$

where $\Gamma_1(t)$ is given by equation (A.2) or (A.3) and

$$\Gamma_2(t) = \int \frac{\partial \mathcal{P}(x, t)}{\partial t} \log \mathcal{P}_0(Ex)\, dx$$

$$= \int g(x, t)^{\mathrm{T}} \frac{\partial \mathcal{P}_0(Ex)}{\partial x} \frac{dx}{\mathcal{P}_0(Ex)}. \tag{A.6}$$

We now have

$$\frac{\partial \mathcal{P}_0(Ex)}{\partial x} = -E\Theta_0^{-1}(Ex - \bar{x}_0)\mathcal{P}_0(Ex).$$

Using again the definitions $F = D\Theta^{-1} - A$ and $u = b - D\Theta^{-1}\bar{x}$, we find

$$\Gamma_2 = -\int (Fx + u)^{\mathrm{T}} E\Theta_0^{-1}(Ex - \bar{x}_0) P(x, t)\, \mathrm{d}x.$$

Using the method of equation (43), we obtain

$$\Gamma_2 = -\mathrm{tr}(F^{\mathrm{T}} E\Theta_0 E\theta) - (F\bar{x} + u)^{\mathrm{T}} E\Theta_0(E\bar{x} - \bar{x}_0).$$

The first term reads

$$-\mathrm{tr}(F^{\mathrm{T}} E\Theta_0 E\theta) = \mathrm{tr}\big(A^{\mathrm{T}} E\Theta_0^{-1} E\Theta - \Theta^{-1}DE\Theta_0^{-1}E\Theta\big).$$

For the last term, we may use the cyclic property of the trace to write it as $\mathrm{tr}(EDE\Theta_0^{-1})$. However, since D is diagonal, we have that $EDE = D$. Whence, the last term is simply $\mathrm{tr}(D\Theta_0^{-1}) = \mathrm{tr}(A)$, according to equation (42).

Collecting our results, we finally conclude that

$$\Gamma_2(t) = \mathrm{tr}\big(A^{\mathrm{T}} E\Theta_0^{-1} E\Theta - A\big) + (A\bar{x} - b)^{\mathrm{T}} E\Theta_0^{-1}(E\bar{x} - \bar{x}_0). \tag{A.7}$$

Using this and equation (A.3) in equation (A.5), we finally obtain equation (51).

A.3. Equation for $\Pi_2(t)$

Finally, we turn to Π_2 in equation (37). We begin by writing

$$\begin{aligned} g_0^{\mathrm{irr}}(Ex) &= \big[\big(D\Theta_0^{-1}E - A^{\mathrm{irr}}E\big)x + b^{\mathrm{irr}} - D\Theta_0^{-1}\bar{x}_0\big]\mathcal{P}_0(Ex) \\ &= (Fx + u)\mathcal{P}_0(Ex). \end{aligned} \tag{A.8}$$

We then have

$$\begin{aligned} \Pi_2 &= \int (Fx + u)^{\mathrm{T}} D^{-1}(Fx + u)\mathcal{P}(x, t)\, \mathrm{d}x \\ &= \mathrm{tr}(F^{\mathrm{T}} D^{-1} F\Theta) + (F\bar{x} + u)^{\mathrm{T}} D^{-1}(F\bar{x} + u). \end{aligned} \tag{A.9}$$

Let us first simplify the trace term

$$\mathrm{tr}(F^{\mathrm{T}} D^{-1} F\Theta) = \mathrm{tr}\big[E\Theta_0^{-1}\big(D\Theta_0^{-1} - 2A^{\mathrm{irr}}\big)E\Theta\big] + \mathrm{tr}(EA^{\mathrm{irr}^{\mathrm{T}}} D^{-1} A^{\mathrm{irr}} E\Theta). \tag{A.10}$$

The first term can be simplified as follows. First, we note the following relation stemming from equations (19)–(21):

$$A^{\mathrm{irr}} = \tfrac{1}{2}(A + EAE). \tag{A.11}$$

Using this in the first term of equation (A.10), we find

$$\mathrm{tr}\big[E\Theta_0^{-1}\big(D\Theta_0^{-1} - 2A^{\mathrm{irr}}\big)E\Theta\big] = \mathrm{tr}\big[E\Theta_0^{-1}\big(D\Theta_0^{-1} - A\big)E\Theta\big] - \mathrm{tr}\big(E\Theta_0^{-1}EA\Theta\big).$$

Similarly to the first term in equation (A.7), the last term in this equation can be written as $\mathrm{tr}(A^{\mathrm{T}} E\Theta_0^{-1} E\Theta)$. Moreover, the first term is zero, which may be shown by substituting

$$D\Theta_0^{-1} = \tfrac{1}{2}\big(A + \Theta_0 A^{\mathrm{T}}\Theta_0^{-1}\big).$$

Finally, the last term in equation (A.10) can be written as $\mathrm{tr}(A^{\mathrm{irr}^{\mathrm{T}}} D^{-1} A^{\mathrm{irr}}\Theta)$. This is a direct consequence of the structure of A^{irr} in equation (21) and the fact that we are assuming D to be diagonal. In fact, for a non-diagonal D, this simplification would not occur and this term would not correctly match its counterpart appearing in equation (46). Thus, we conclude that

$$\mathrm{tr}(F^{\mathrm{T}} D^{-1} F\Theta) = \mathrm{tr}(A^{\mathrm{irr}^{\mathrm{T}}} D^{-1} A^{\mathrm{irr}}\Theta) - \mathrm{tr}\big(A^{\mathrm{T}} E\Theta_0^{-1} E\Theta\big). \tag{A.12}$$

For the second term in equation (A.9), we now have

$$F\bar{x} + u = D\Theta_0^{-1}(E\bar{x} - \bar{x}_0) + b^{\mathrm{irr}} - A^{\mathrm{irr}}E\bar{x}.$$

Expanding the quadratic form, using equation (A.11) and simplifying finally yields the required corresponding term in equation (50).

References

[1] de Groot S R and Mazur P 1961 *Non-Equilibrium Thermodynamics* 1st edn (Amsterdam: North-Holland)
[2] Esposito M 2010 *Phys. Rev. Lett.* **104** 090601
[3] Verley G and Lacoste D 2012 *Phys. Rev.* E **86** 051127
[4] Spinney R and Ford I 2012 *Phys. Rev.* E **85** 051113
[5] Crooks G 2000 *Phys. Rev.* E **61** 2361
[6] Ford I and Spinney R 2012 *Phys. Rev.* E **86** 021127
[7] Tomé T and de Oliveira M J 2012 *Phys. Rev. Lett.* **108** 020601
[8] Schnakenberg J 1976 *Rev. Mod. Phys.* **48** 571
[9] Tomé T and de Oliveira M J 2010 *Phys. Rev.* E **82** 021120
[10] Van den Broeck C and Esposito M 2010 *Phys. Rev.* E **82** 011144
[11] García-García R 2012 *Phys. Rev.* E **86** 031117
[12] Landauer R 1975 *J. Stat. Phys.* **13** 1
[13] Bruers S, Maes C and Netočný K 2007 *J. Stat. Phys.* **129** 725
[14] Onsager L and Machlup S 1953 *Phys. Rev.* **91** 1505
[15] Bartels R and Stewart G 1972 *Commun. ACM* **15** 820
[16] Risken H 1989 *The Fokker–Planck Equation* 2nd edn (Berlin: Springer)