An expression of excess work during the transition between nonequilibrium steady states

Tatsuro Yuge

Department of Physics, Osaka University, 1–1 Machikaneyama-Cho, Toyonaka, 560–0043, Japan

E-mail: yuge@acty.phys.sci.osaka-u.ac.jp

Received 10 June 2014, revised 7 October 2014
Accepted for publication 8 October 2014
Published 25 November 2014

Abstract
Excess work is a nondiverging part of the work during the transition between nonequilibrium steady states (NESSs). It is a central quantity in steady-state thermodynamics (SST), which is a candidate for nonequilibrium thermodynamics theory. We derive an expression of excess work during quasistatic transitions between NESSs by using the macroscopic linear response relation of the NESS. This expression is a line integral of a vector potential in the space of control parameters. We show a relationship between the vector potential and the response function of the NESS, and thus obtain a relationship between the SST and a macroscopic quantity. We also connect the macroscopic formulation to microscopic physics through a microscopic expression of the nonequilibrium response function, which gives a result that is consistent with previous studies.

Keywords: nonequilibrium steady state, excess work, response function
PACS numbers: 05.70.Ln, 05.60.-k

1. Introduction

The second law of thermodynamics gives a fundamental limit to thermodynamic operations on systems in equilibrium states. One of its formulations provides the lower bound for the work performed on a system during a thermodynamics operation that induces a transition between equilibrium states; the lower bound is given by the change in the free energy and is achieved for quasistatic operations. Establishing analogous thermodynamic theory for a nonequilibrium steady state (NESS) is one of the challenging problems in physics. Recent attempts [1–15] have focused on the relations to be satisfied by the work and the entropy production (or heat) for transitions between NESSs.
One of the candidates for nonequilibrium thermodynamics is the steady-state thermodynamics (SST) proposed in [16]. A central idea of SST is to use excess work and excess heat, which is defined as follows. In nonequilibrium states, work is continuously supplied to the system, so that the total work, $W_{\text{tot}}$, during the transition between NESSs diverges. Therefore, to construct a meaningful thermodynamic theory for the transition, it is necessary to take a finite part out of $W_{\text{tot}}$. To this end, the excess work, $W_{\text{ex}}$, is defined by subtracting the integral of the steady work flow in the instantaneous NESS at each point of the operation [16, 17] from $W_{\text{tot}}$.

Studies of this idea have been developed in [10–15, 18, 19]. In this paper, we focus our attention on quasistatic transitions. In the regime near equilibrium, the work version of the results in [10, 11] states that the excess work, $W_{\text{ex}}$, for quasistatic transitions is given by the change in a certain scalar potential. In the regime far from equilibrium, by contrast, the work version of [12, 15] states that, in general, it is not equal to the change in any scalar function, but rather is equal to a geometrical quantity (i.e., it is given by a line integral of a vector potential, $A^W$, in the operation parameter space). This suggests that $A^W$ plays an important role in SST.

Since these studies are based on microscopic dynamics, the relationship with the microscopic state has been developed. In contrast, the relationship with macroscopic quantities has been less clear, although some studies exist on this issue [8, 9, 20]. It is particularly important to clarify how the transport coefficients and response functions are treated in the framework of SST, because the transport coefficients are the main quantities that characterize the NESS, and they are accessible in experiments.

In this paper, we first derive an expression of $W_{\text{ex}}$ for quasistatic transitions in terms of the linear response function of NESSs. We show that $W_{\text{ex}}$ is equal to the line integral of a vector potential, $A^W$, and we clarify the relation between $A^W$ and the response function. Since the derivation relies only on the macroscopic phenomenological equation (linear response relation), the result is universally valid (independent of microscopic detail). We also show that in the regime near equilibrium, $W_{\text{ex}}$ is given by the change of a scalar potential, thanks to the reciprocal relation. Furthermore, we connect the macroscopic theory for $A^W$ to microscopic physics by using a microscopic expression of the response function, which is called the response-correlation relation (RCR) [21, 22]. We obtain a microscopic expression of $A^W$ that is consistent with the work version of the results in [12, 15].
2. Setup

We consider a system, $S$, that is in contact with multiple (say $n$) reservoirs. A schematic diagram of the setup is shown in figure 1. The $i$th reservoir is in the equilibrium state characterized by the chemical potential, $\mu_i$. We denote the set of the chemical potentials by $\mu$ (i.e., $\mu = \{\mu_i\}_{i=1}^n$). The temperatures of all the reservoirs are set to the same value. We denote the particle current between $S$ and the $i$th reservoir by $I_i$, where we take the sign of $I_i$ to be positive when it flows from the reservoir to $S$. More precisely, $I_i(t)$ is defined as $I_i(t) = -\frac{dN_i(t)}{dt}$, where $N_i$ is the particle number in the $i$th reservoir. We assume that the reservoirs are sufficiently large so that they are not affected by the change in $N_i$ and they remain in the equilibrium states on the time scale of interest. We also assume that for a fixed $\mu$, a stable NESS is realized in $S$ uniquely and independently of initial states after a relaxation time. We note that in the NESS, $\sum_i \langle I_i \rangle_{\mu}^{SS} = 0$ holds due to the particle number conservation in the total system ($S$ plus reservoirs), where $\langle I_i \rangle_{\mu}^{SS}$ is the expectation value of the current $I_i$ in the NESS characterized by $\mu$.

Examples of such a setup are seen in field-effect semiconductor devices. A typical one is the modulation-doped field-effect transistor [23, 24]. In this example, the system, $S$, is realized as a two-dimensional electron system, and the reservoirs are the electrodes (source, drain, and gate).

2.1. Transition between NESSs

Suppose that at the initial time, $t_0$, the system $S$ is in a NESS characterized by $\mu$. The difference, $\mu_i - \mu_j$, may be so large that the initial NESS is far from equilibrium. At $t_0 + \delta t$, we change the chemical potentials from $\mu$ to $\mu' = \{\mu_i + \delta \mu_i\}_{i=1}^n$, with small constants $\delta \mu_i$. Then the state of $S$ varies in time for $t > t_0$, and after a sufficiently long time it settles to a new NESS characterized by $\mu'$. In this paper, we investigate the work, $W$, done on $S$ during the transition between the NESSs.

For this purpose, we consider the expectation value of the current, $I_i$, from the macroscopic viewpoint of the linear response relation. To the linear order in $\delta \mu$, we can express the expectation value, $\langle I_i \rangle_{\mu}^{t}$, as

$$\langle I_i \rangle_{\mu}^{t} = \langle I_i \rangle_{\mu}^{SS} + \sum_{j=1}^{n} \Psi_{ij}(t-t_0) \delta \mu_j + O(\delta \mu^2),$$

where $\Psi_{ij}(\tau) \equiv \int_0^\tau d\tau' \Phi_{ij}(\tau')$, and $\Phi_{ij}(\tau)$ is the linear response function of the NESS [25–28], which satisfies the causality relation, $\Phi_{ij}(\tau < 0) = 0$. We note that the linear response relation (1) is a relation around a NESS (not equilibrium state) and is valid for a NESS even far from equilibrium if the NESS is stable with regard to perturbations. We can also express the expectation value of the current $I_i$ in the final NESS (characterized by $\mu'$) by the long-time limit of equation (1):

$$\langle I_i \rangle_{\mu'}^{\infty} = \langle I_i \rangle_{\mu'}^{SS} + \sum_{j} \phi_{ij} \delta \mu_j + O(\delta \mu^2),$$

where $\phi_{ij} \equiv \lim_{\tau \to \infty} \Psi_{ij}(\tau)$ is the transport coefficient (differential conductivity) of the initial NESS (characterized by $\mu$). We again note that the limit exists if the NESS is stable.
3. Excess Work

3.1. General case

In nonequilibrium states, work, $W$, is continuously supplied to the system $S$ from the reservoirs, accompanied by the particle current to $S$. We can use equilibrium thermodynamics to estimate the work done by the reservoirs since they are in equilibrium states. When the particle number in the $i$th reservoir increases by $\Delta N_i$, the work done by the $i$th reservoir is given by

$$W_i = \mu_i \Delta N_i.$$

Therefore, the unit-time work by the $i$th reservoir is

$$\langle W_i \rangle_t = \langle J_i^W \rangle_t / \langle J_i^\mu \rangle_t.$$

From equation (1), we obtain the average work flow,

$$\langle W \rangle_{t_0}^t = \sum_i \langle J_i^W \rangle_t / \langle J_i^\mu \rangle_t,$$

where we neglected the $O(\delta \mu^2)$ terms in the second line. The total work, $W_{\text{ex}}$, during the transition between the NESSs is given by the time-integral of equation (3). However, $W_{\text{tot}}$ is a diverging quantity because the system $S$ must still be supplied with the work from the reservoirs after it reaches the final NESS.

To extract a finite quantity intrinsic to the transition, we employ the idea of the SST [16]; we subtract from $W_{\text{tot}}$ the contribution of the steady work flow, $\langle W_{\text{ss}} \rangle_{t_0}$, in the final NESS:

$$W_{\text{ex}} = \int_{t_0}^{\infty} dt \left( \langle J_i^W \rangle_t - \langle J_i^W \rangle_{t_0} \right).$$

We refer to this quantity as the excess work. We note that $W_{\text{ex}}$ is related to the excess heat, $Q_{\text{ex}}$, as $W_{\text{ex}} + Q_{\text{ex}} = \Delta U$, where $\Delta U$ is the change in the energy of S between the NESSs due to the energy conservation in the transition between the NESSs and the energy balance in the steady states. Our definition (4) of $W_{\text{ex}}$ is consistent with the definition of $Q_{\text{ex}}$ in [10–12, 15]. We also note that the steady flow, $\langle J_i^W \rangle_{t_0}$, is equal to the long-time limit of equation (3):

$$\langle J_i^W \rangle_{t_0} = \sum_i \langle J_i \rangle_{t_0} (\mu_i + \delta \mu_i) + \mu_i \Phi_j \delta \mu_j.$$

By substituting equations (3) and (5) into equation (4), we obtain

$$W_{\text{ex}} = \sum_j A_j^W \delta \mu_j,$$

where the $j$th component, $A_j^W$, of the vector potential, $A^W$, is given by

$$A_j^W = \sum_i \mu_i \int_{t_0}^{\infty} dt \left[ \Phi_j \left( t - t_0 \right) - \Phi_j \right].$$

Equation (6) indicates that the excess work during quasistatic transitions between NESSs is not given by the difference of some scalar function, $F$, but rather is given by the geometrical quantity unless $A_j^W$ is equal to the $\mu_j$-derivative of $F$ for all $j$. This is consistent with the results in [12, 15]. Equation (7) relates the nonequilibrium linear response function, $\Phi$, with the vector potential $A^W$ in the expression (6). Therefore, $A^W$ can be experimentally determined in principle, because $\Phi$ is measurable.
The sufficient condition for \( A^W_j = \partial_j F \) for all \( j \) is that
\[
\partial_i A_j^W = \partial_j A_i^W
\]
(8)
holds for all \( i, j \), where \( \partial_j \) is the abbreviation of \( \partial/\partial \mu_j \).

3.2. Weakly nonequilibrium case

In the regime near equilibrium (linear response regime), we can use the response function, \( \Phi^{\text{eq}} \), of the equilibrium state in equation (7). Then equation (8) is valid because \( \Phi^{\text{eq}} \) is independent of \( \mu_i \) and the reciprocal relation, \( \Phi^{\text{eq}}_{ij} = \Phi^{\text{eq}}_{ji} \), holds. Therefore, the extension of the Clausius equality is possible in this regime, which is consistent with the results in [10, 11].

4. Connection to microscopic physics

Until now, our formulation has been closed on the macroscopic level. Now we connect our formulation to microscopic physics. In this paper, we assume that the microscopic dynamics of the system S is governed by the quantum master equation (QME) [29]:
\[
\frac{\partial \hat{\rho}}{\partial t} = \mathcal{K} \hat{\rho}.
\]
(9)
Here, \( \hat{\rho} \) is the density matrix of S, and the generator, \( \mathcal{K} \), is written as \( \mathcal{K} = [\hat{H}, \hat{\rho}]/i\hbar + \sum_j \mathcal{L}_j \), where \( \hat{H} \) is the Hamiltonian of S and \( \mathcal{L}_j \) is the dissipator induced by the interaction with the \( j \)th reservoir. As in the previous sections of this paper, we assume that a unique steady state exists in the QME. The steady-state density matrix, \( \hat{\rho}_{\text{ss}} \), satisfies \( \mathcal{K} \hat{\rho}_{\text{ss}} = 0 \).

First, we consider the connection to microscopic physics through the response function, \( \Phi \), of the NESS. For this purpose, we employ the RCR [21, 22], which is a microscopic expression of \( \Phi \). In the framework of the QME and for the response of the current, \( I_i \), from the \( i \)th reservoir into S, the RCR reads
\[
\mathcal{K} \hat{\rho} = \int_{-\infty}^{\infty} \hat{L}_t(\tau) \hat{\rho} \left( \int_{-\infty}^{\tau} \hat{L}_t(\tau) \hat{\rho} \right) d\tau + \int_{-\infty}^{\infty} \hat{L}_t(\tau) \hat{\rho} \left( \int_{-\infty}^{\tau} \hat{L}_t(\tau) \hat{\rho} \right) d\tau.
\]
(10)
where \( \hat{L}_t(\tau) = \hat{N} \hat{L}_t(\tau) \hat{N} \) with \( \hat{N} \) being the particle number operator in S. See appendix A for the derivation of equation (10). Note that \( \hat{L}_t \) can be regarded as the particle current operator from the \( i \)th reservoir into S because it satisfies the continuity equation: \( \frac{\partial}{\partial t} \text{Tr} [\hat{N} \hat{\rho}(t)] = \sum_i \text{Tr} [\hat{L}_t \hat{\rho}(t)] \). Substituting equation (10) into equation (7), we obtain
\[
A_j^W = \sum_i \mu_i \int_{-\infty}^{\infty} d\tau \left\{ \int_{-\infty}^{\tau} \text{Tr} \left[ \hat{L}_t e^{\mathcal{K}(\tau-t)} \left( \partial_j \mathcal{K} \right) \hat{\rho}_{\text{ss}} \right] + \text{Tr} \left[ \hat{L}_t R \left( \partial_j \mathcal{K} \right) \hat{\rho}_{\text{ss}} \right] \right\}.
\]
(11)
Note that the contribution from the first term in equation (10) vanishes. Here, we defined
\[
\mathcal{R} \equiv -\lim_{T \to \infty} \int_{-\infty}^{T} dt' e^{\mathcal{K}(T-t')} Q_0,
\]
(12)
and \( Q_0 = 1 - \hat{P}_0 \), where the projection superoperator, \( \hat{P}_0 \), is defined such that \( \hat{P}_0 \hat{X} = \hat{P}_0 \text{Tr} \hat{X} \) holds for any linear operator, \( \hat{X} \). See appendix B for the fact that \( \mathcal{R} \) is a well-defined superoperator. To rewrite equation (11) further, we note the following relation:
\[ \frac{d}{dt} \text{Tr} \left[ \hat{I}_i e^{\mathcal{K}(t-t')} \mathcal{R} \left( \partial_j \mathcal{K} \right) \hat{\rho}_{\alpha} \right] \]

\[ = - \text{Tr} \left[ \hat{I}_i e^{\mathcal{K}(t-t')} \mathcal{K} \mathcal{R} \left( \partial_j \mathcal{K} \right) \hat{\rho}_{\alpha} \right] \]

\[ = - \text{Tr} \left[ \hat{I}_i e^{\mathcal{K}(t-t')} \hat{Q}_0 \left( \partial_j \mathcal{K} \right) \hat{\rho}_{\alpha} \right] \]

\[ = - \text{Tr} \left[ \hat{I}_i e^{\mathcal{K}(t-t')} \left( \partial_j \mathcal{K} \right) \hat{\rho}_{\alpha} \right] + \text{Tr} \left[ \hat{I}_i \hat{\rho}_{\alpha} \right] \text{Tr} \left[ \left( \partial_j \mathcal{K} \right) \hat{\rho}_{\alpha} \right] \]

\[ = - \text{Tr} \left[ \hat{I}_i e^{\mathcal{K}(t-t')} \left( \partial_j \mathcal{K} \right) \hat{\rho}_{\alpha} \right]. \quad (13) \]

In the third line, we used \( \mathcal{R} \mathcal{K} = \mathcal{K} \mathcal{R} = \hat{Q}_0 \). \quad (14)

See appendix B for the derivation of equation (14). In the last line of equation (13), we used

\[ \text{Tr} \left[ \left( \partial_j \mathcal{K} \right) \hat{\rho}_{\alpha} \right] = 0 \] for any \( \hat{X} \) (trace-preserving property of the QME). Integrating equation (13), we can rewrite the first term on the right-hand side of equation (11) as

\[ \int_{t_0}^t \frac{d}{dt'} \text{Tr} \left[ \hat{I}_i e^{\mathcal{K}(t-t')} \left( \partial_j \mathcal{K} \right) \hat{\rho}_{\alpha} \right] = \text{Tr} \left[ \hat{I}_i e^{\mathcal{K}(t-t')} \mathcal{R} \left( \partial_j \mathcal{K} \right) \hat{\rho}_{\alpha} \right] \]

\[ - \text{Tr} \left[ \hat{I}_i \mathcal{R} \left( \partial_j \mathcal{K} \right) \hat{\rho}_{\alpha} \right]. \]

The second term on the right-hand side of this equation cancels out the second term on the right-hand side of equation (11). We thus rewrite equation (11) as

\[ A_j^W = - \sum_i \mu_i \text{Tr} \left[ \hat{I}_i \mathcal{R}^2 \left( \partial_j \mathcal{K} \right) \hat{\rho}_{\alpha} \right]. \quad (15) \]

where we used \( \int_{t_0}^{\infty} d\tau e^{\mathcal{K}(t-t')} = \hat{Q}_0 - \mathcal{R} \). This is a microscopic expression of the vector potential, \( A^W \).

Next, we investigate the consistency of equation (15) with the results in [12, 15]. In a manner almost the same as those in [12, 15], we can derive another microscopic expression of \( A^W \) without relying on equation (7):

\[ A_j^W = - \text{Tr} \left( \hat{\rho}_{\alpha} \partial_j \right). \quad (16) \]

where \( \hat{\rho}_{\alpha} \equiv \partial \hat{\rho}_{\alpha} / \partial (i \hat{\chi}) \big|_{\hat{\chi}=0} \). Here, \( \hat{\chi} \) is the counting field in the full counting statistics (FCS) of the work, \( W \), from the reservoirs, and \( \hat{\rho}_{\alpha} \) is the left eigenvector of \( \mathcal{K} \) corresponding to the eigenvalue, \( \lambda_\alpha \), that has the maximum real part. \( \mathcal{K} = [\hat{H}, 1/i\hbar + \sum_j \mathcal{L}_j] \) is the \( \chi \)-modified generator, which is introduced for the FCS in the framework of the QME [30]. See appendix C for the details and derivation. We note that \( \hat{\chi}_{\alpha=0} = \hat{I} \) (identity operator), \( \lambda_{\alpha=0} = 0 \), and \( \partial \hat{\chi} / \partial (i \hat{\chi}) \big|_{\hat{\chi}=0} = \langle J^W \rangle_\chi \). In the following, we rewrite equation (16) to show its equivalence to equation (15).

We first rewrite \( \hat{\rho}_{\alpha} \) in equation (16). By differentiating the left eigenvalue equation, \( \mathcal{K} \hat{\rho}_{\alpha}^{(\alpha)} = (\lambda_\alpha)^{\alpha}_{\chi=0} \hat{\rho}_{\alpha}^{(\alpha)} \), with respect to \( i \hat{\chi} \) and setting \( \hat{\chi} = 0 \), we obtain

\[ \mathcal{K} \hat{\rho}_{\alpha}^{(\alpha)} - (\mathcal{K}^*)^\dagger \hat{1} = - (\mathcal{K}^*)^\dagger \hat{1}. \quad (17) \]

Here, \( \mathcal{K} = \partial \mathcal{K} / \partial (i \hat{\chi}) \big|_{\hat{\chi}=0} \) and the adjoint, \( \mathcal{O}^* \), of a superoperator, \( \mathcal{O} \), is defined by \( \text{Tr} \left[ (\mathcal{O}^* \hat{X}_1) \hat{X}_2 \right] = \text{Tr} \left[ \hat{X}_1 \mathcal{O}^* \hat{X}_2 \right] \) for any pair \( (\hat{X}_1, \hat{X}_2) \) of linear operators. By operating on the both sides of equation (17) with \( \mathcal{R}^\dagger \), we obtain
\[ \hat{t}_0' = -R^j(\mathbf{\mathcal{K}'}')^i + c \hat{1}, \]  

where we used equation (14) and \( c = -(J^\mu s \omega + \text{Tr} \{ \delta \hat{\mathcal{K}} \hat{t}_0 \} \). Substituting equation (18) into equation (16) and using \( \text{Tr} \{ \partial_j \hat{\mathcal{K}} \} = \partial_j \text{Tr} \{ \hat{\mathcal{K}} \} = 0 \), we have \( A^j_W = \text{Tr} \{ (\mathbf{\mathcal{K}'}')^i \} R \delta \beta_{s} \}. \) Furthermore, we can show \( (\mathbf{\mathcal{K}'}')^i \hat{1} = \sum_i \mu_i \mathcal{L}_i \hat{N} = \sum_i \mu_i \hat{N} \). With this equation, we have

\[ A^j_W = \sum_i \mu_i \text{Tr} \{ \hat{1} R \partial_j \beta_{s} \}. \]

We next rewrite \( \partial_j \beta_{s} \). By differentiating the steady-state equation, \( \mathcal{K} \beta_{s} = 0 \), with respect to \( \mu_j \) and operating on it with \( R \), we have \( \partial_j \beta_{s} = -R \langle \partial_j \mathbf{\mathcal{K}} \rangle \beta_{s} \}. \) Substituting this equation into equation (19), we obtain

\[ A^j_W = -\sum_i \mu_i \text{Tr} \{ \hat{1} R^2 \{ \partial_j \mathbf{\mathcal{K}} \} \beta_{s} \}. \]

This is the same as equation (15). We thus show that equation (15), and therefore equation (7), is consistent with the results in [12, 15].

5. Concluding remarks

We have derived an expression of the excess work for quasistatic transitions between NESSs in particle transport systems on the basis of the linear response relation. We have related the vector potential, \( A^W \), in the expression with the response function. We note that it is possible to extend our formulation to situations where other control parameters for transition between NESSs are varied. In particular, we can obtain a similar result in heat conducting systems, where the temperatures of heat reservoirs are changed. Our final remarks are as follows.

First, the relationship between the excess work and the response function suggests that the response functions can be calculated within the framework of the SST. We expect that this expression becomes a first step for understanding how transport phenomena are treated in the SST.

Second, as was mentioned below equation (4), our definition of the excess work is consistent with the definition of the excess heat in [10–12, 15]. However, the definition of the excess work and heat is not unique; for example, there are Hatano–Sasa-type [5, 9, 18] and Maes–Netočný-type [13] approaches. Recently, [9] provided evidence that the Hatano–Sasa approach is appropriate for the definition. Since the Hatano–Sasa approach relies on microscopic information such as steady-state distribution and transition rate, the connection to macroscopic quantities is not clear. It is therefore important to investigate the definition from the viewpoint of the response function in future research.

Third, in recent years the nonequilibrium response function has been a hot topic in statistical physics [25–28]. One of the foci of recent research is the decomposition of the response function [27, 28]. We expect that the application of these results to the expression of the excess work will lead to a further decomposition of the work that is appropriate for the construction of the SST.

Acknowledgments

The author thanks K Akiba and M Yamaguchi for helpful discussions. This work was supported by a JSPS Research Fellowship for Young Scientists (No. 24-1112), KAKENHI (No. 26287087), and ImPACT Program of the Council for Science, Technology and Innovation (Cabinet Office, Government of Japan).
Appendix A. Linear response function of a NESS in the quantum master equation approach

Here we derive equation (10), the RCR in QME. We consider the QME (9) that depends on multiple parameters, \( \alpha = (\alpha_1, \alpha_2, \alpha_3, \ldots) \), like chemical potentials (i.e., we assume that the generator of the QME depends on parameters \( \mathcal{K} = \mathcal{K}(\alpha) \)).

Suppose that at time \( t \leq t_0 \), the system S is in the NESS with \( \alpha = \alpha^0 \) (i.e., \( \rho = \rho^0 \) for \( t \leq t_0 \)). For \( t > t_0 \), we weakly modulate the parameters in time: \( \alpha(t) = \alpha(t) + \delta_t \), where \( \delta_t \) is much smaller than a typical value of \( \alpha_l \). Then we can expand the generator \( \mathcal{K} \) around \( \alpha^0 \) in terms of \( \delta_t \):

\[
\mathcal{K}(\alpha(t)) \approx \mathcal{K}^0 + \sum_l \delta_t \partial_l \mathcal{K}^0,
\]

where \( \partial_l \mathcal{K}^0 \equiv \partial \mathcal{K}(\alpha)/\partial \alpha_l |_{\alpha^0}^{\alpha} \).

To solve the QME (9) with the weakly time-dependent \( \alpha \) and the initial condition \( \rho(t_0) \), we transform the QME into an 'interaction picture.' That is, we introduce \( \hat{\rho}(t) = e^{-\mathcal{K}_{t}(t-t_0)}\hat{\rho}(t) \). Then, from the QME (9), we have the equation of motion for \( \hat{\rho}(t) \) as

\[
\frac{\partial \hat{\rho}(t)}{\partial t} = \sum_l \delta_t(t) e^{-\mathcal{K}_{t}(t-t_0)} \left( \partial_l \mathcal{K}^0 \right) e^{\mathcal{K}_{t}(t-t_0)} \hat{\rho}(t),
\]

where we used equation (A.1). By integrating this equation from \( t_0 \) to \( t \) with \( \rho(t_0) = \hat{\rho}_{\alpha^0} \), we obtain

\[
\hat{\rho}(t) = \hat{\rho}_{\alpha^0} + \sum \int_{t_0}^{t} dt' \delta_t(t') e^{-\mathcal{K}_{t}(t-t_0)} \left( \partial_l \mathcal{K}^0 \right) e^{\mathcal{K}_{t}(t-t_0)} \hat{\rho}(t') \approx \hat{\rho}_{\alpha^0} + \sum \int_{t_0}^{t} dt' \delta_t(t') e^{-\mathcal{K}_{t}(t-t_0)} \left( \partial_l \mathcal{K}^0 \right) \hat{\rho}_{\alpha^0}.
\]

In going from the first line to the second, we approximately replaced \( \hat{\rho}(t') \) in the integral with \( \hat{\rho}_{\alpha^0} \). This approximation corresponds to the first-order, time-dependent perturbation theory in quantum mechanics. Going back to the Schrödinger picture, we have

\[
\hat{\rho}(t) = \hat{\rho}_{\alpha^0} + \sum \int_{t_0}^{t} dt' \delta_t(t') e^{\mathcal{K}_{t}(t-t') \left( \partial_l \mathcal{K}^0 \right)} \hat{\rho}_{\alpha^0}.
\]

We thus obtain the time dependence of the expectation value of a quantity, \( \hat{X} \), that is independent of \( \alpha \):

\[
\langle X \rangle_{\alpha^0} = \text{Tr} \left[ \hat{X} \hat{\rho}(t) \right] = \langle X \rangle_{\alpha^0} + \sum \int_{t_0}^{t} dt' \text{Tr} \left[ \hat{X} e^{\mathcal{K}_{t}(t-t') \left( \partial_l \mathcal{K}^0 \right)} \hat{\rho}_{\alpha^0} \right] \delta_t(t') \]

\[
= \langle X \rangle_{\alpha^0} + \sum \int_{t_0}^{t} dt' \text{Tr} \left[ \left\{ \partial \mathcal{K}(\alpha(t)) \right\} e^{\mathcal{K}_{t}(t-t') \hat{X}} \right] \hat{\rho}_{\alpha^0} \delta_t(t').
\]

Equations (A.5) and (A.6) give the RCR in the QME. We note that equation (A.6) reduces to the Kubo formula if \( \hat{\rho}_{\alpha^0} \) is an equilibrium state (i.e., when we consider the response of an equilibrium state) [21].

Now we consider the current, \( \hat{I}_i = \mathcal{L}_i \hat{X} \), from the \( i \)th reservoir into the system S. We note that \( \hat{I}_i \) is dependent on \( \alpha \) because \( \mathcal{L}_i \) is too. Therefore, we have the average current at time, \( t \),
as
\[
\{\hat{L}^I\}_{a(t)} = \{\hat{L}^I\}_{a_0}
\]
\[
= \text{Tr} \left[ \hat{L}(\mathbf{a}(t)) \hat{\rho}(t) \right] - \text{Tr} \left[ \hat{L}(\mathbf{a}_0) \hat{\rho}_m^0 \right]
\]
\[
= \sum_i \text{Tr} \left[ \left( \partial_\mathbf{a} \hat{L}^0 \right) \hat{\rho}_m^0 \right] f_i(t) + \text{Tr} \left[ \hat{L}_t^0 \hat{\rho}(t) \right] - \text{Tr} \left[ \hat{L}_t^0 \hat{\rho}_m^0 \right]
\]
\[
= \sum_i \text{Tr} \left[ \left( \partial_\mathbf{a} \hat{L}^0 \right) \hat{\rho}_m^0 \right] f_i(t) + \sum_t \int_{t_0}^t dt' \text{Tr} \left[ \hat{L}_t^0 e^{\mathbf{K}(t-t')} (\partial_\mathbf{a} \mathbf{K}) \hat{\rho}_m^0 \right] f_i(t'),
\]
where \( \hat{L}_t^0 = \hat{L}(\mathbf{a}_0) \) and \( \partial_\mathbf{a} \hat{L}^0 = \partial_\mathbf{a} \hat{L}(\mathbf{a}) / \partial_\mathbf{a} \mathbf{a} \mid_\mathbf{a} = \mathbf{a}_0 \). We used equation (A.5) in the third line. Finally, by performing the functional differentiation with respect to \( f_i(t') \), we obtain
\[
\Phi_{ij}(t - t') = \frac{\delta \{\hat{L}^I\}_{a(t)}}{\delta f_j(t')}
\]
\[
= \text{Tr} \left[ \left( \partial_\mathbf{a} \hat{L}^0 \right) \hat{\rho}_m^0 \right] \delta(t - t') + \text{Tr} \left[ \hat{L}_t^0 e^{\mathbf{K}(t-t')} (\partial_\mathbf{a} \mathbf{K}) \hat{\rho}_m^0 \right].
\]
This is equivalent to equation (10).

**Appendix B. Inverse-like superoperator in the quantum master equation approach**

First, we show that \( \mathbf{R} \) in equation (12) is well defined. To this end, we denote the eigenvalue and the corresponding left and right eigenvectors of \( \mathbf{K} \) as \( \lambda_m, \ell_m ^\dagger \), and \( r_m \), respectively. We assign the steady state of \( \mathbf{K} \) to the index, \( m = 0 \) (i.e., \( \lambda_0 = 0, \ell_0 = 1, \) and \( r_0 = \hat{\rho}_m^0 \)). By the assumption of the unique existence of the stable steady state, \( \text{Re} \lambda_m < 0 \) for \( m \neq 0 \). Then, for any linear operator \( \hat{X} \), we obtain the following equation:
\[
\mathbf{R} \hat{X} = - \int_0^\infty dr e^{\mathbf{K}r} \text{Tr} \left[ \hat{X} \right]
\]
\[
= - \int_0^\infty dr e^{\mathbf{K}r} \sum_{m \neq 0} \text{Tr} \left[ \hat{X} \ell_m ^\dagger \right] \ell_m
\]
\[
= - \int_0^\infty dr \sum_{m \neq 0} e^\lambda_m r \text{Tr} \left[ \hat{X} \ell_m ^\dagger \right] \ell_m
\]
\[
= \sum_{m \neq 0} \frac{\text{Tr} \left[ \hat{X} \ell_m ^\dagger \right] \ell_m}{\lambda_m}.
\]
Since this is not diverging, \( \mathbf{R} \) is well defined.

We show here that equation (14) holds for the generator, \( \mathbf{K} \), of the QME. We first note that \( \mathbf{R} \mathbf{K} = \mathbf{K} \mathbf{R} \) follows from \( \mathbf{Q}_0 \mathbf{K} = \mathbf{K} \mathbf{Q}_0 = \mathbf{K} \), which we can derive from the fact that
\[
\mathbf{R}_0 \mathbf{K} \hat{X} = \hat{\rho}_m \text{Tr} \left[ \mathbf{K} \hat{X} \right] = 0,
\]
\[
\mathbf{K} \mathbf{R}_0 \hat{X} = \mathbf{K} \hat{\rho}_m \text{Tr} \hat{X} = 0,
\]
hold for any linear operator, $\hat{X}$. Equation (B.2) follows from $\text{Tr} \{ \mathcal{K} \hat{X} \} = 0$ (trace-preserving property of QME), and equation (B.3) follows from $\mathcal{K} \hat{\rho}_0 = 0$ (steady-state equation). Then we can show equation (14) as follows:

$$R \mathcal{K} = \mathcal{K} R = \lim_{T \to \infty} \int_{t_0}^T dt \frac{d}{dt} e^{\mathcal{K}(T-t)} Q_0$$

$$= \left( 1 - \lim_{T \to \infty} e^{\mathcal{K}(T-t_0)} \right) Q_0$$

$$= (1 - P_0) Q_0$$

$$= Q_0.$$  (B.4)

Here, the third line follows from the convergence theorem of the Markov process, which we can derive from the fact that for any linear operator, $\hat{X}$, the following equation holds:

$$\mathcal{K} \rho_\mu = \hat{\rho}_{ss}$$

This gives the third line in equation (B.4). We note that equation (B.4) leads to $R \mathcal{K} R \mathcal{K} = \mathcal{K}$. This implies that $R$ satisfies one of the conditions for the Moore–Penrose pseudoinverse of $\mathcal{K}$.

### Appendix C. Derivation of equation (16)

For completeness, we now derive equation (16), the work version of the results in [12, 15].

First, we note that we can measure the work, $W$, while varying the chemical potentials, $\mu = \{ \mu_i \}_i$, with a time interval, $\tau$, as follows. At the initial time $t = t_0$, we perform a projection measurement of reservoir particle numbers, $\{ \hat{N}_i \}_i$, to obtain measurement outcomes, $\{ \hat{N}_i(t_0) \}_i$. For $t > t_0$, we vary $\mu$, where the system evolves by interacting with the reservoirs. At $t = t_0 + \tau$, we again perform a measurement of $\hat{N}_i$ to obtain outcomes, $\{ \hat{N}_i(t_0 + \tau) \}_i$. The difference of the outcomes gives the work, $W = \sum_i [\mu_i(t_0 + \tau) N_i(t_0 + \tau) - \mu_i(t_0) N_i(t_0)]$. By repeating the measurements, we obtain a probability distribution, $p_W(W)$. The average work is given by $\langle W \rangle = \int dW p_W(W) W$, and the average work flow in a NESS is given by $J_W = \lim_{\tau \to \infty} \langle W \rangle / \tau$, with $\mu$ being fixed.

In the following, we calculate the average work by $\langle W \rangle = \partial G(\chi) / \partial (\chi 1)_{\chi=0}$, where $G(\chi) \equiv \ln \int dW p_W(W) e^{\chi W}$ is the cumulant generating function and $\chi$ is the counting field. By using the full counting statistics [30], we can calculate $G(\chi)$ by

$$G_\tau(\chi) = \ln \text{Tr}_{\alpha} \hat{\rho}^{\hat{X}}(\chi).$$  (C.1)

Here, $\hat{\rho}^{\hat{X}}$ is the solution of the generalized quantum master equation (GQME):

$$\frac{\partial \hat{\rho}^{\hat{X}}(t)}{\partial t} = \mathcal{K}^{\hat{X}}(\mu(t)) \hat{\rho}^{\hat{X}}(t),$$  (C.2)

where the generalized generator is given by $\mathcal{K}^{\hat{X}} = [\hat{H}, 1/\hbar + \sum_j \mathcal{L}_j^{\hat{X}}]$, with the generalized dissipator $\mathcal{L}_j^{\hat{X}} \hat{\rho} \equiv -(1/\hbar^2) \int_0^\infty dt' \text{Tr} \{ \hat{H}_{\hat{S}_j} [-t'], \hat{\rho} \otimes \hat{\rho}_{\hat{S}_j}(t') \} \mathcal{L}_j^{\hat{X}} \hat{\rho} \otimes \hat{\rho}_{\hat{S}_j}(t') \text{Tr}_{\hat{S}_j}$ is the trace over the $j$th reservoir, $\hat{H}_{\hat{S}_j}$ is the interaction Hamiltonian between the system and the $j$th reservoir, $\hat{\rho}_{\hat{S}_j}$
is its interaction picture, \( \hat{\rho}_i(\mu) \) is the thermal equilibrium state of the \( j \)th reservoir with the chemical potential \( \mu_j \). \( [\hat{O}_1, \hat{O}_2]_\mu \equiv \hat{O}_1 \hat{O}_2 - \hat{O}_2 \hat{O}_1 \), and \( \hat{O}_\chi^x \equiv e^{-i\chi} \hat{\Sigma}_\mu \Sigma^{1/2} \hat{\Sigma}_\mu \Sigma^{1/2} \).

Note that, if we set \( \chi = 0 \), the GQME (C.2) reduces to the original QME (9), and \( K^x, \hat{\rho}_m^x \), and \( \hat{\rho}_m^\mu \) also reduce to \( K, \hat{\rho}_m^\mu \), and \( \hat{\rho}_m^\mu \), respectively.

For fixed \( \mu \), we can define the left and right eigenvectors of \( K^x(\mu) \) corresponding to the eigenvalue \( \lambda^x_m(\mu) \), which are respectively denoted by \( \hat{\lambda}_m^x(\mu) \) and \( \hat{\rho}_m^x(\mu) \). They are normalized as \( \text{Tr}(\hat{\rho}_m^x \hat{\rho}_n^x) = \delta_{mn} \). We assign the label for the eigenvalue with the maximum real part to \( m = 0 \). It is known that \( \lim_{\tau \to \infty} G_\tau(x)/\tau = \lambda^x_0 \) holds [30]. Therefore, the average work flow, \( J_w \), in the NESS can be calculated by

\[
J_w(\mu) = \left. \frac{\partial \lambda^x_m(\mu)}{\partial (i\chi)} \right|_{\chi=0}. \tag{C.3}
\]

We now derive equation (16). We first note that the excess work can be written as \( W_{ex} = \partial G_{\chi}(\chi)/\partial (i\chi) |_{\chi=0} \), where \( G_{\chi}(\chi) \equiv G_\tau(\chi) - A_\tau^\chi(\tau) \) and \( A_\tau^\chi(\tau) \equiv \int_{t_0}^{t_0+\tau} dt' \lambda^x_m(\mu(t')) \). This is because \( (W)_{\tau} = \partial G_{\chi}(\chi)/\partial (i\chi) |_{\chi=0} \) and equation (C.3). To calculate \( G_{\chi}(\chi) \), we solve the GQME (C.2). For this purpose, we expand \( \hat{\rho}^x(\tau) \) as

\[
\hat{\rho}^x(\tau) = \sum_m \hat{\rho}_m^x(\mu) e^{A_\tau(\tau)\hat{\rho}_m^x(\mu)}. \tag{C.4}
\]

Substituting this expansion into equation (C.2) and taking the Hilbert–Schmidt inner product with \( \hat{\rho}_m^x(\mu) \), we obtain

\[
\frac{d\hat{\rho}_m^x(t)}{dt} = -\sum_m \hat{\lambda}_m^x(\mu) e^{A_\tau(\tau)\hat{\rho}_m^x(\mu)} \text{Tr}[\hat{\rho}_m^x(\mu) \hat{\rho}^x(\tau)].
\]

If the time scale of varying \( \mu \) is sufficiently slower than the relaxation time of the system, we can approximate the sum on the right-hand side solely by the contribution from the term with \( m = 0 \) (adiabatic approximation). By solving the approximate equation, we obtain

\[
c_0(t_0 + \tau) = c_0(t_0) \exp \left\{ -\int_C \text{Tr}[\hat{\rho}_0^x(\mu) d\hat{\rho}^x(\mu)] \right\}, \tag{C.5}
\]

where \( C \) is a path connecting \( \mu(t_0) \) and \( \mu(t_0 + \tau) \) in the parameter space and \( d\hat{\rho}^x(\mu) \equiv \sum_j (\partial\hat{\rho}^x(\mu)/\partial \mu_j) d\mu_j \). If \( \hat{\rho}_m^x(t_0) \equiv \hat{\rho}_m^x(\mu(t_0)) \), then \( c_0(t_0) = \text{Tr}[\hat{\rho}_0^x(\mu(t_0)) \hat{\rho}^x(\tau)] \).

For a long time, only the \( m = 0 \) term remains in equation (C.4) since \( \lambda^x_m(\mu) \) has the maximum real part. Therefore, we obtain

\[
\hat{\rho}^x(t_0 + \tau) = c_0(t_0 + \tau) e^{A_\tau(\tau)\hat{\rho}_m^x(\mu(t_0 + \tau))}. \tag{C.6}
\]

Substituting equation (C.5) into this equation, we obtain an expression for

\[
G_{\chi}(\chi) = \text{InTr} e^{A_\tau(\tau)\hat{\rho}_m^x(\mu(t_0 + \tau))} \lambda^x_m(\mu(t_0))
\]

and

\[
G_{\chi}(\chi) = \text{InTr} e^{A_\tau(\tau)\hat{\rho}_m^x(\mu(t_0))} \lambda^x_m(\mu(t_0)) \tag{C.7}
\]

Finally, by differentiating equation (C.7) with respect to \( i\chi \) and setting \( \chi = 0 \), we obtain an expression for the excess work:
\[ W_{\text{ex}} = -\int_C \text{Tr}_S \left[ \hat{\mathcal{L}}_0^+(\mu) \hat{\Phi}_S(\mu) \right]. \] (C.8)

We thus obtain equation (16).

References

[1] Esposito M, Harbola U and Mukamel S 2007 Phys. Rev. E 76 031132
[2] Esposito M and van den Broeck C 2010 Phys. Rev. Lett. 104 090601
[3] Deffner S and Lutz E 2010 Phys. Rev. Lett. 105 170402
[4] Takara K, Hasegawa H H and Driebe D J 2010 Phys. Lett. A 375 88
[5] Deffner S and Lutz E 2012 arXiv:1201.3888
[6] Nakagawa N 2012 Phys. Rev. E 85 051115
[7] Verley G and Lacoste D 2012 Phys. Rev. E 86 051127
[8] Boksenbojm E, Maes C, Netočný K and Pešek J 2011 Europhys. Lett. 96 40001
[9] Mandal D 2013 Phys. Rev. E 88 062135
[10] Komatsu T S, Nakagawa N, Sasa S and Tasaki H 2008 Phys. Rev. Lett. 100 230602
[11] Saito K and Tasaki H 2011 J. Stat. Phys. 145 1275
[12] Sagawa T and Hayakawa H 2011 Phys. Rev. E 84 051110
[13] Maes C and Netočný K 2014 J. Stat. Phys. 154 188
[14] Bertini L, Gabrielli D, Jona-Lasinio G and Landim C 2013 Phys. Rev. Lett. 110 020601
[15] Yuge T, Sagawa T, Sugita A and Hayakawa H 2013 J. Stat. Phys. 153 412
[16] Oono Y and Paniconi M 1998 Prog. Theor. Phys. Suppl. 130 29
[17] Landauer R 1978 Phys. Rev. A 18 255
[18] Hatano N and Sasa S 2001 Phys. Rev. Lett. 86 3463
[19] Sasa S and Tasaki H 2006 J. Stat. Phys. 125 125
[20] Maes C 2014 J. Stat. Phys. 154 705
[21] Shimizu A and Yuge T 2010 J. Phys. Soc. Japan 79 013002
[22] Yuge T 2010 Phys. Rev. E 82 051130
[23] Sze S M 2002 Semiconductor Devices: Physics and Technology vol 2 (New York: Wiley)
[24] Davies J H 1998 The Physics of Low-Dimensional Semiconductors: An Introduction (Cambridge: Cambridge University Press)
[25] Marconi U M B, Puglisi A, Rondoni L and Vulpiani A 2008 Phys. Rep. 461 111
[26] Chetrite R and Gupta S 2011 J. Stat. Phys. 143 543
[27] Baiesi M, Maes C and Wynants B 2009 Phys. Rev. Lett. 103 010602
[28] Baiesi M and Maes C 2013 New J. Phys. 15 013004
[29] Breuer H P and Petruccione F 2002 The Theory of Open Quantum Systems (Oxford: Oxford University Press)
[30] Esposito M, Harbola U and Mukamel S 2009 Rev. Mod. Phys. 81 1665