How nanomechanical systems can minimize dissipation

Paolo Muratore-Ginanneschi[*] and Kay Schwieger[†]

Department of Mathematics and Statistics, University of Helsinki, P.O. Box 68, FIN-00014 Helsinki, Finland

Information processing machines at the nanoscales are unavoidably affected by thermal fluctuations. Efficient design requires understanding how nanomachines can operate at minimal energy dissipation. Here we focus on mechanical systems controlled by smoothly varying potential forces. We show that optimal control equations come about in a natural way if the energy cost to manipulate the potential is taken into account. When such a cost becomes negligible, an optimal control strategy can be constructed by transparent geometrical methods which recover the solution of optimal mass transport equations in the overdamped limit. Our equations are equivalent to hierarchies of kinetic equations of a form well known in the theory of dilute gases. From our results, optimal strategies for energy efficient nanosystems may be devised by established techniques from kinetic theory.

PACS number(s): 05.20.Dd, 02.30.Yy, 05.40.−a, 05.70.Ln

Recent experiments with colloidal particles, and integrated platforms of nanomagnetic memory and logic circuits, exhibited the possibility to design and control information processing machines on a molecular scale [1–3]. These experiments are a first experimental step towards low dissipation Brownian computers, a concept theoretically envisaged decades ago [4]. Furthermore, these experiments put to test new developments of nonequilibrium thermodynamics such as refinements of the second law stemming from fluctuation theorems (see, e.g., Refs. [5–9] and also Refs. [10,11]). Engineering nanomachines, however, remains technologically challenging [12] and demands a better theoretical understanding of how energy dissipation can be minimized.

The Langevin-Kramers dynamics is the reference model (cf. Ref. [13]) epitomizing effects in nanosystem mechanics: a kinetic-plus-potential Hamiltonian, mechanical friction by a Stokes drag force, and thermal noise. In the Langevin-Kramers framework, the second law of thermodynamics is amenable to a mathematical formulation in terms of an adapted *Schrödinger diffusion problem* [14]: Given the phase-space probability densities describing the state of the system at finite initial and final times, find the potential force which steers the initial into the final density while minimizing the average energy dissipation.

In this Rapid Communication we conceptualize the Schrödinger diffusion as an optimal stochastic control problem [15]. Our aim is to derive an optimal control strategy over the class of *smooth* potential forces for the problem adapted to the second law of thermodynamics. From an experimental slant, smooth potentials model macroscopic degrees of freedom of the system whose state is determined by external sources [16]. We show that our problem is well posed when regarded as the limit of a more general control problem which takes into account the energy cost of the control. The phase-space optimal control equations turn out to be amenable to the form of *Bogoljubov-Born-Green-Kirkwood-Yvon kinetic momentum hierarchies*, which are well known in the theory of dilute gases (see, e.g., Ref. [17]). This is an important observation as it renders an immediately available toolbox of mathematical

methods for the analysis of our optimal control equations [18]. In agreement with physical intuition, in the overdamped regime we recover the Monge-Ampère-Kantorovich equations [19], which were recently proved to govern minimal dissipation transitions in the Langevin-Smoluchowski modeling [20]. Moreover, elementary arguments show that the solution of the very same Monge-Ampère-Kantorovich equations yields a general lower bound for the average energy dissipation by smooth Langevin-Kramers dynamics.

Finally, we notice that our results are based on Pontryagin's principle (see Ref. [21] for a concise review) in a formulation inspired by Ref. [22]. Our formulation may be of relevance for control problems in other disciplines.

Model. We define the kinetic-plus-potential Langevin-Kramers dynamics by means of its scalar generator,

$$\mathfrak{L} = \frac{\boldsymbol{p}}{m} \cdot \partial_{\boldsymbol{q}} - \left(\frac{\boldsymbol{p}}{\tau} + \partial_{\boldsymbol{q}} U \right) \cdot \partial_{\boldsymbol{p}} + \frac{m}{\beta\tau} \partial_{\boldsymbol{p}}^2.$$

Here β is the inverse of the temperature, m is the mass of a Brownian particle under a time dependent potential force $\partial_{\boldsymbol{q}} U \equiv (\partial_{\boldsymbol{q}} U)(\boldsymbol{q},t)$, and τ is the characteristic time of the Stokes drag. We suppose the dynamics to occur on an $2d$-dimensional Euclidean phase space with coordinates $\boldsymbol{x} \equiv [\boldsymbol{q}, \boldsymbol{p}]$, where $\boldsymbol{q}$ and $\boldsymbol{p}$ as usual denote positions and momenta. By the generator $\rho \equiv \rho(\boldsymbol{x},t)$ evolves according to the Fokker-Planck equation,

$$(\partial_t - \mathfrak{L}^{\dagger})\rho = 0, \tag{1}$$

where $\mathfrak{L}^{\dagger}$ is the $\mathbb{L}^2(\mathbb{R}^{2d})$ adjoint of $\mathfrak{L}$ with respect to the Lebesgue measure. Stochastic thermodynamics considerations (see, e.g., Refs. [9,23]) uphold the interpretation of

$$\mathcal{Q} = -\frac{dt_{\rm f}}{\beta\tau} + \int_0^{t_{\rm f}} dt \int_{\mathbb{R}^{2d}} d^{2d}x \, \rho \frac{\|\boldsymbol{p}\|^2}{m\tau}$$

as the mean heat release by the Brownian particle during the time interval $[0,t_{\rm f}]$. The mean heat release is given by the expected value of the line integral over the Stokes drag,

$$\mathcal{E} = \int_0^{t_{\rm f}} dt \int_{\mathbb{R}^{2d}} d^{2d}x \, \rho \frac{\|\boldsymbol{p}\|^2}{m\tau}, \tag{2}$$

minus its Maxwell-Boltzmann equilibrium equipartition value. Hence, finding a tight lower bound for the second law of

---

[*]paolo.muratore-ginanneschi@helsinki.fi
[†]kay.schwieger@helsinki.fi

thermodynamics over finite-time transitions governed by (1) and transforming $\rho(\boldsymbol{x},0) = \rho_\iota(\boldsymbol{x})$ into $\rho(\boldsymbol{x},t_\mathrm{f}) = \rho_\mathrm{f}(\boldsymbol{x})$ is equivalent to finding a Schrödinger diffusion process which minimizes (2) by an optimal choice of U. In this Rapid Communication, we will always consider boundary conditions compatible with thermal equilibrium: $\rho_\mathrm{j}(\boldsymbol{x}) = \mu_\mathrm{j}(\boldsymbol{q})\mu_\mathrm{MB}(\boldsymbol{p})$, $\mathrm{j} = \iota,\mathrm{f}$, with $\mu_\mathrm{MB}(\boldsymbol{p}) = [\beta/(2\pi m)]^{d/2}\exp\{-\frac{\beta\|\boldsymbol{p}\|^2}{2m}\}$ the Maxwell-Boltzmann momentum distribution.

Bounds. Elementary statistical moment inequalities immediately yield

$$\mathcal{E} \geqslant \int_0^{t_\mathrm{f}} \frac{dt}{\tau} \int_{\mathbb{R}^d} d^d q \, \mu m \|\boldsymbol{v}\|^2, \tag{3}$$

with the marginal density $\mu(\boldsymbol{q},t) \equiv \int_{\mathbb{R}^d} d^d p \, \rho(\boldsymbol{q},\boldsymbol{p},t)$ and the macroscopic velocity (or first-order kinetic cumulant) $\boldsymbol{v}(\boldsymbol{q},t) \equiv [\mu(\boldsymbol{q},t)\,m]^{-1} \int_{\mathbb{R}^d} d^d p \, \rho(\boldsymbol{q},\boldsymbol{p},t)\boldsymbol{p}$. A well-known result of kinetic theory [17] implies that μ obeys a continuity equation with respect to $\boldsymbol{v}$. Hence we obtain (see, e.g., Ref. [24] and below) a lower bound for the right hand side of (3) if we choose $\boldsymbol{v}$ to be solution of a Monge-Ampère-Kantorovich system (Burgers plus mass continuity) transporting μ_ι into μ_f in $[0,t_\mathrm{f}]$. Repeating analogous considerations on conditional position averages yields for equilibrium boundary conditions the bound $\mathcal{Q} \geqslant 0$. In general, there is no reason to expect these simple bounds to be tight. Moreover, knowledge of $\boldsymbol{v}$ does not specify an optimal control U.

Optimal control. To tackle the optimal control problem we construct from (2) the *energy cost* functional

$$\mathcal{A} = \mathcal{E} + \frac{2\,g}{\beta}\mathcal{D}(\rho\|\bar\rho) - \int_0^{t_\mathrm{f}} dt \int_{\mathbb{R}^{2d}} d^{2d}x \, V(\partial_t - \mathfrak{L}^\dagger)\rho, \tag{4}$$

where $g \geqslant 0$ and, upon denoting the local equilibrium potential by $S(\boldsymbol{q},t) = -\ln\frac{\tau^d \mu(\boldsymbol{q},t)}{\beta^{d/2}\,m^{d/2}}$, we define

$$\mathcal{D}(\rho|\bar\rho) \equiv \frac{\tau\beta}{2m} \int_0^{t_\mathrm{f}} dt \int_{\mathbb{R}^{2d}} d^{2d}x \, \rho \left\| \partial_{\boldsymbol{q}}\left(U - \frac{k}{\beta}S\right) \right\|^2,$$

with a real constant k.

The last term in (4) encapsulates *Pontryagin's principle.* Namely, the term vanishes if ρ satisfies the Fokker-Planck equation (1). Thus, the dynamics is enforced by the Lagrange multiplier $V \equiv V(\boldsymbol{x},t)$, which we interpret as the value function of Bellman's formulation of optimal control [15]. The parameter $g \geqslant 0$ couples the energy dissipation $\mathcal{E}$ to a term $\mathcal{D}$ modeling the energy cost of the control. Namely, $\mathcal{D}(\rho\|\bar\rho)$ is the Kullback-Leibler divergence between ρ of the

process and the density $\bar\rho$ of a Langevin-Kramers process driven by the potential $\frac{k}{\beta}S$. As $\frac{1}{\beta}S$ is the local equilibrium potential for the position marginal density, $\mathcal{D}(\rho\|\bar\rho)$ vanishes at equilibrium for $k = 1$. The case $k = 0$ describes instead the relative entropy between the controlled and the uncontrolled ($U = 0$) Langevin-Kramers dynamics.

Following Pontryagin's principle, we look for an extremal value of the energy cost functional (4) versus the triple (ρ,V,U). The calculation is straightforward but relies on three crucial observations. First, for any $g > 0$, the energy cost is convex in the control $\partial_{\boldsymbol{q}}U$ and hence coercive [15]. Second, the energy cost depends nonlinearly upon ρ via the local equilibrium potential S. Third, since U depends only upon position variables, we can average out momenta from energy cost variations with respect to U. We thus arrive at the system of three extremal equations formed by the following: first, the Fokker-Planck equation (1); second, the dynamic programming equation

$$(\partial_t + \mathfrak{L})V + \frac{\|\boldsymbol{p}\|^2}{m\tau} + \frac{2g}{\beta}\mathcal{D}'_\rho = 0, \tag{5}$$

with $\mathcal{D}'_\rho \equiv \frac{\beta\tau}{2m}\{\|\partial_{\boldsymbol{q}}(U - \frac{k}{\beta}S)\|^2 - \frac{2k}{\beta}\partial_{\boldsymbol{q}}^2(U - \frac{k}{\beta}S)\}$; and third, the equation for the control potential,

$$\nabla_{\boldsymbol{q}}^S \cdot \left[\frac{m}{2\tau}\mathsf{V}^{(1)} - g\,\partial_{\boldsymbol{q}}\left(U - \frac{k}{\beta}S\right)\right] = 0, \tag{6}$$

where we write $\nabla_{\boldsymbol{q}}^S \equiv \partial_{\boldsymbol{q}} - (\partial_{\boldsymbol{q}}S)$ and define

$$\mathsf{V}^{(1)}(\boldsymbol{q},t) = \int_{\mathbb{R}^d} d^d p \frac{\rho(\boldsymbol{q},\boldsymbol{p},t)}{\mu(\boldsymbol{q},t)}\partial_{\boldsymbol{p}}V(\boldsymbol{q},\boldsymbol{p},t). \tag{7}$$

These three equations, the optimal control system, is the first result of this work. We now analyze its main physical consequences.

Kinetic hierarchies and controllability. In general, we can turn the energy cost (4) into an average restricted to configuration space by introducing conditional momentum cumulants of any order n:

$$\mathsf{F}_{\boldsymbol{i}_n}^{(n)}(\boldsymbol{q},t) \equiv \left.\frac{\partial_{\bar{p}_{i_1}}\cdots\partial_{\bar{p}_{i_n}}}{m^n\,\iota^n}\ln\int_{\mathbb{R}^d} d^d p \, e^{\iota\,\bar{\boldsymbol{p}}\cdot\boldsymbol{p}}\frac{\rho(\boldsymbol{q},\boldsymbol{p},t)}{\mu(\boldsymbol{q},t)}\right|_{\bar{\boldsymbol{p}}=0},$$

together with dual tensors $\{\mathsf{V}_{\boldsymbol{i}_n}^{(n)}\}_{n=0}^\infty$. Here we denote by $\boldsymbol{i}_n \equiv [i_1,\ldots,i_n]$ the n-tuple of Euclidean indices of any rank-n cumulant or dual tensor. If we apply Pontryagin's principle by looking for stationary variations of $\{\mathsf{F}_{\boldsymbol{i}_n}^{(n)},\mathsf{V}_{\boldsymbol{i}_n}^{(n)}\}_{n=0}^\infty$ and μ, then instead of the Fokker-Planck equation (1) and the dynamic programming (5), we obtain two coupled hierarchies of kinetic equations for the $\mathsf{F}^{(n)}$'s and the $\mathsf{V}^{(n)}$'s and the local equilibrium potential S:

$$\left(\partial_t + \boldsymbol{v}\cdot\partial_{\boldsymbol{q}} + \frac{n}{\tau}\right)\mathsf{F}_{\boldsymbol{i}_n}^{(n)} + \frac{1}{m}\sum_{l=2}^n \binom{n}{l}\underset{\boldsymbol{i}_{n-l},\boldsymbol{i}_l}{\mathrm{Sym}}\mathsf{F}_{\boldsymbol{i}_{n-l},j}^{(n-l+1)}\partial_{q_j}\mathsf{F}_{\boldsymbol{i}_l}^{(l)} + \frac{1}{m}\nabla_{q_j}^S\mathsf{F}_{j,\boldsymbol{i}_n}^{(n+1)} = \delta_{n,0}\partial_t S - \delta_{n,1}\partial_{q_i}U + \delta_{n,2}\frac{2m\,\delta_{\boldsymbol{i}_2}}{\beta\,\tau}, \tag{8}$$

$$\left(\partial_t + \boldsymbol{v}\cdot\partial_{\boldsymbol{q}} - \frac{n}{\tau}\right)\mathsf{V}_{\boldsymbol{i}_n}^{(n)} - \frac{1}{m}\sum_{l\geqslant 2}\binom{n+l-1}{l-1}\left(\frac{n}{l}\underset{\boldsymbol{i},\boldsymbol{i}_{n-1}}{\mathrm{Sym}}(\partial_{q_i}\mathsf{F}_{j_l}^{(l)})\mathsf{V}_{\boldsymbol{j}_l,\boldsymbol{i}_{n-1}}^{(n+l-1)} - \nabla_{q_j}^S\mathsf{F}_{\boldsymbol{j}_{l-1},\boldsymbol{i}_n}^{(l)}\mathsf{V}_{\boldsymbol{j}_{l-1},\boldsymbol{i}_n}^{(n+l-1)}\right)$$

$$+ \frac{1}{m}\underset{i,\boldsymbol{i}_{n-1}}{\mathrm{Sym}}(\partial_{q_i}\mathsf{V}_{\boldsymbol{i}_{n-1}}^{(n-1)} - n(\partial_{q_i}v_j)\mathsf{V}_{j,\boldsymbol{i}_{n-1}}^{(n)}) = -\frac{\delta_{\boldsymbol{i}_2}\delta_{n,2}}{\beta\,m\,\tau} - \frac{2\,v_i\,\delta_{n,1}}{\beta\,\tau} - \delta_{n,0}\left(\frac{\mathsf{F}_{j,j}^{(2)}}{m\,\tau} + \frac{m\|\boldsymbol{v}\|^2}{\tau} + \frac{2\,g\mathcal{D}'_\rho}{\beta}\right). \tag{9}$$

In (8) and (9) we use the conventions that repeated (multi)indices are contracted and that "Sym" denotes symmetrization of the free indices in the underscript. A straightforward calculation, which will be reported elsewhere, also shows that (8) and (9) can be also derived directly from (1) and (5). Finally, the hierarchies are coupled by (6), which continues to hold with the interpretation of a constraint relating the dual tensor $\mathsf{V}^{(1)}$ to the control potential. The advantage of introducing (8) and (9) is that we can use them to approximate the solution of our optimal control problem by means of *realizable closures*, i.e., finite order truncations of the hierarchies preserving the probabilistic interpretation of the cumulants [18]. We notice that truncating (8) and (9) at any finite order n yields a controllable system of equations in the sense that the number of time derivatives in the hierarchies equals that of the boundary conditions for S and $\{\mathsf{F}^{(l)}\}_{l=1}^{n}$. Furthermore, increasing the order of truncation imposes more constraints on the control strategy. The solution of a sequence of realizable truncations therefore yields more and more refined lower bounds to the energy cost (4). For instance, the estimate (3) corresponds to the truncation of lowest order.

Limit of vanishing g. If we set *a priori* $g = 0$, by Pontryagin's principle we need to replace (6) with

$$U_\star(\boldsymbol{q},t) = \arg\inf_U \left\{ -\int_{\mathbb{R}^d} d^d q \,\mu\, \mathsf{V}^{(1)} \cdot \partial_{\boldsymbol{q}} U \right\},$$

reminiscent of singular optimal control [15]. Qualitatively, this equation suggests the decomposition of phase space into a "no-action region" where $\mathsf{V}^{(1)}$ vanishes and a "push region" where $\mathsf{V}^{(1)} \cdot \partial_{\boldsymbol{q}} U > 0$ and where $\|\partial_{\boldsymbol{q}} U\|$ is constrained only by the boundary conditions. Inspecting (8) and (9) shows, however, that requiring the condition $\mathsf{V}^{(1)} = 0$ to hold and be preserved by the dynamics enslaves the macroscopic velocity to the remaining cumulants by the equation

$$\frac{2\boldsymbol{v}_i}{\tau} = \sum_{l \geqslant 2} \frac{\left(\partial_{\boldsymbol{q}_i}\mathsf{F}^{(l)}_{j_l}\right)\mathsf{V}^{(l)}_{j_l} - l\,\nabla^S_{\boldsymbol{q}_j}\mathsf{F}^{(l)}_{j,j_{l-1}}\mathsf{V}^{(l)}_{j_{l-1},i}}{m} - \frac{\partial_{\boldsymbol{q}_i}\mathsf{V}^{(0)}}{m}.$$
(10)

As the boundary values of the dual tensors $\{\mathsf{V}^{(n)}_{\boldsymbol{i}_n}\}_{n\neq 1}$ are determined by the boundary conditions imposed on μ and $\{\mathsf{F}^{(n)}_{\boldsymbol{i}_n}\}_{n\geqslant 2}$, we see that in general (10) cannot satisfy independent boundary conditions on $\boldsymbol{v}$. In this sense, for $g = 0$ the system is not controllable. The consideration of an exactly solvable case indicates, however, that in a weaker sense $\mathsf{V}^{(1)} = 0$ still governs the optimal control strategy at $g = 0$.

Evolution between Gaussian states. Let us consider the transition between Gaussian densities at finite initial and final times. Physically, this is a stylized model of a moving laser trap of changing size (see Ref. [25], and references therein). For these boundary conditions and fixed $g > 0$, the optimal control equations admit a solution in terms of a probability density which stays Gaussian in the entire control horizon, a value function which is a quadratic polynomial in $\boldsymbol{p}$ and $\boldsymbol{q}$, and a potential which is quadratic in $\boldsymbol{q}$. Correspondingly, the hierarchies (8) and (9) reduce to a system of $2d\,(2d+3)$ first-order differential equations accompanied by the same number of boundary conditions for the initial and final cumulants. For
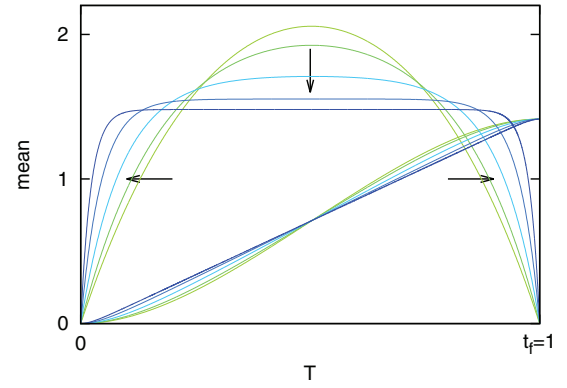

FIG. 1. (Color online) Means for Gaussian boundary conditions with parameters $\tau = 1$, $m = 1$, $\beta = 1$, $t_{\mathrm{f}} = 1$, $k = 1$. The initial and final means of $\boldsymbol{x} \equiv [\boldsymbol{q},\boldsymbol{p}]$ are $(0,0)$ and $(\sqrt{2},0)$, respectively. The parameter g varies logarithmically from 1.28×10^{-1} (green/light) to 1.25×10^{-4} (blue/dark). The arrows indicate the behavior for decreasing g.

dimension $d = 1$ the typical behavior of the solution is shown in Figs. 1 and 2 for different values of g.

As $g \to 0$ the limit behavior of the cumulants is described by the "slow manifold" specified by the condition $\mathsf{V}^{(1)} = 0$ and the evolution law (10). Only in a layer close to the boundaries of the control horizon do cumulants get away from the slow manifold along exponentially stable and unstable directions with rates of the order $O(1/\sqrt{g})$ in order to satisfy the boundary conditions. The description of singular boundary value problems in terms of invariant manifolds is well known in the theory of dynamical systems [26]. We refer the interested reader to Ref. [27] for details of the multiscale expansion [28], proving the foregoing qualitative picture.

Overdamped limit. The relevance of the "slow manifold" condition $\mathsf{V}^{(1)} = 0$ appears from the fact that it permits one to recover directly the "overdamped" limit of the Langevin-Kramers dynamics at $g = 0$. The overdamped regime
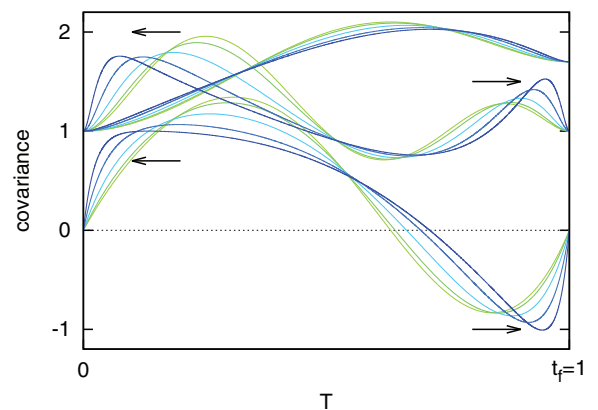

FIG. 2. (Color online) Covariances for Gaussian boundary conditions with parameters as in Fig. 1. The initial and final covariance matrices of $\boldsymbol{x} \equiv [\boldsymbol{q},\boldsymbol{p}]$ are given by $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1.7 & 0 \\ 0 & 1 \end{pmatrix}$, respectively.

corresponds to the assumption of a wide scale separation between the control horizon $[0, t_{\mathrm{f}}]$ and the Stokes time τ and between the characteristic length scale L of the configuration space boundary data $\mu_\iota(\boldsymbol{q}/L)$, $\mu_{\mathrm{f}}(\boldsymbol{q}/L)$, and a typical length scale $\ell = \tau/\sqrt{\beta m}$ of the uncontrolled process. In the overdamped regime, the quantifier of the scale separation is the Stokes number $\varepsilon = \tau/t_{\mathrm{f}} = \ell^2/L^2 \ll 1$. Under these hypotheses, we can look for an asymptotic solution of (1) and (5) by expanding around a Maxwell-Boltzmann momentum equilibrium distribution $\mu_{\mathrm{MB}}(\boldsymbol{p})$ perturbed at large scales, $(\tilde{\boldsymbol{q}}, \tilde{t}) \equiv (\sqrt{\varepsilon}\,\boldsymbol{q}, \varepsilon t)$, by the action of a control potential of the form $U(\boldsymbol{q}, t) \equiv U_0(\sqrt{\varepsilon}\,\boldsymbol{q}, \varepsilon t) + O(\sqrt{\varepsilon})$. From now on we specify the order of the perturbative expansion by an underscript. Upon setting $\partial_{\tilde{\boldsymbol{q}}} S_0 = -\partial_{\tilde{\boldsymbol{q}}} \ln \mu_0$, and applying standard homogenization techniques (see, e.g., Ref. [28]; see also Ref. [24]) we get for the solution of the Fokker-Planck equation (1)

$$
\begin{aligned}
\rho(\boldsymbol{p}, \tilde{\boldsymbol{q}}, \tilde{t}) = {} & \mu_{\mathrm{MB}}(\boldsymbol{p}) \mu_0(\tilde{\boldsymbol{q}}, \tilde{t}) \\
& \times \left(1 + \sqrt{\varepsilon}\frac{\tau}{m}\boldsymbol{p} \cdot \partial_{\tilde{\boldsymbol{q}}}(S_0 - \beta U_0)(\tilde{\boldsymbol{q}}, \tilde{t}) + O(\varepsilon) \right).
\end{aligned}
\tag{11}
$$

For the solution of the dynamic programming (5) we obtain

$$
\begin{aligned}
V(\boldsymbol{p}, \tilde{\boldsymbol{q}}, t, \tilde{t}) = {} & \frac{(t_{\mathrm{f}} - t)d}{\beta\,\tau} + \frac{\|\boldsymbol{p}\|^2}{2\,m} + V_0(\tilde{\boldsymbol{q}}, \tilde{t}) + \sqrt{\varepsilon}\bigg(V_1\left(\tilde{\boldsymbol{q}}, \tilde{t}\right) \\
& + \frac{\tau}{m}\,\boldsymbol{p} \cdot \partial_{\tilde{\boldsymbol{q}}}(V_0 - U_0)\left(\tilde{\boldsymbol{q}}, \tilde{t}\right) \bigg) + O(\varepsilon).
\end{aligned}
\tag{12}
$$

The functions S_0 in (11) and V_0 in (12), respectively, obey the local equilibrium potential equation

$$
\partial_{\tilde{t}} S_0 - \frac{\tau}{m}\big((\partial_{\tilde{\boldsymbol{q}}} S_0) \cdot \partial_{\tilde{\boldsymbol{q}}} - \partial_{\tilde{\boldsymbol{q}}}^2\big)\left(U_0 - \frac{1}{\beta} S_0 \right) = 0,
$$

and the dynamic programming equation

$$
\partial_{\tilde{t}} V_0 - \frac{\tau}{m}\left((\partial_{\tilde{\boldsymbol{q}}} U_0) \cdot \partial_{\tilde{\boldsymbol{q}}} - \frac{1}{\beta}\partial_{\tilde{\boldsymbol{q}}}^2\right)(V_0 - U_0) = 0.
$$

These equations specifying two of the three optimal control equations governing the minimal heat release by a Langevin-Smoluchowski dynamics between μ_ι and μ_{f} [20]. In order to recover the third condition, we use (11) and (12) to evaluate

$$
\mathsf{V}^{(1)}\left(\tilde{\boldsymbol{q}}, \tilde{t}\right) = -\sqrt{\varepsilon}\frac{\tau}{m}\partial_{\tilde{\boldsymbol{q}}}\left(2U_0 - \frac{S_0}{\beta} - V_0 \right)(\tilde{\boldsymbol{q}}, \tilde{t}) + O(\varepsilon).
$$

Then the condition $\mathsf{V}^{(1)} = 0$ yields exactly the relation between U_0, V_0, S_0 that allows us to recover the very same Monge-Ampère-Kantorovich equations of Ref. [20]:

$$
\partial_{\tilde{t}} \tilde{U} - \frac{\tau}{2\,m}\partial_{\tilde{\boldsymbol{q}}} \tilde{U} \cdot \partial_{\tilde{\boldsymbol{q}}} \tilde{U}\, \partial_{\tilde{t}} S_0 - \frac{\tau}{m}(\partial_{\tilde{\boldsymbol{q}}} S_0) \cdot \partial_{\tilde{\boldsymbol{q}}} \tilde{U} + \frac{\tau}{m}\partial_{\tilde{\boldsymbol{q}}}^2 \tilde{U} = 0,
$$

with $\tilde{U} \equiv U_0 - S_0/\beta$.

Conclusion. We showed how Pontryagin's principle can be used to derive refined bounds for the second law of thermodynamics in the case of nanomechanical systems. We also established a relation between optimal control and kinetic theory, which renders available ideas and tools of dilute gas [17,18] and optimal transport theory [19] to the construction of optimal protocols implementing, at the nanoscale, information processing operations such as the erasure of a bit [10].

[1] S. Toyabe, T. Sagawa, M. Ueda, E. Muneyuki, and M. Sano, Nat. Phys. **6**, 988 (2010).

[2] B. Lambson, D. Carlton, and J. Bokor, Phys. Rev. Lett. **107**, 010604 (2011).

[3] A. Bérut, A. Arakelyan, A. Petrosyan, S. Ciliberto, R. Dillenschneider, and E. Lutz, Nature (London) **483**, 187 (2012).

[4] C. H. Bennett, Int. J. Theor. Phys. **21**, 905 (1982).

[5] D. J. Evans and D. J. Searles, Phys. Rev. E **50**, 1645 (1994).

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

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

[8] G. E. Crooks, J. Stat. Phys. **90**, 1481 (1997).

[9] D.-Q. Jiang, M. Qian, and M.-P. Qian, *Mathematical Theory of Nonequilibrium Steady States*, Lecture Notes in Mathematics Vol. 1833 (Springer, Berlin, 2004), p. 276.

[10] R. Dillenschneider and E. Lutz, Phys. Rev. Lett. **102**, 210601 (2009).

[11] E. Aurell, K. Gawędzki, C. Mejía-Monasterio, R. Mohayaee, and P. Muratore-Ginanneschi, J. Stat. Phys. **147**, 487 (2012).

[12] C. Van den Broeck, Nat. Phys. **6**, 937 (2010).

[13] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd ed. (Elsevier, Amsterdam, 2007), p. 463.

[14] R. Aebi, *Schrödinger Diffusion Processes* (Birkhäuser, Basel, 1996), p. 186.

[15] W. H. Fleming and M. H. Soner, *Controlled Markov Processes and Viscosity Solutions*, 2nd ed., Springer Series in Stochastic Modeling and Applied Probability Vol. 25 (Springer, Berlin, 2006), p. 428.

[16] A. Alemany, M. Ribezzi, and F. Ritort, in *Nonequilibrium Statistical Physics Today: Proceedings of the 11th Granada Seminar on Computational and Statistical Physics*, edited by P. L. Garrido, J. Marro, and F. de los Santos, AIP Conf. Proc. No. 1332 (AIP, Melville, NY, 2011), p. 96.

[17] C. Cercignani, R. Illner, and M. Pulvirenti, *The Mathematical Theory of Dilute Gases*, Springer Series in Applied Mathematical Sciences Vol. 106 (Springer, New York, 1994), p. 349.

[18] C. D. Levermore, J. Stat. Phys. **83**, 1021 (1996).

[19] C. Villani, *Optimal Transport: Old and New*, Grundlehren der mathematischen Wissenschaften Vol. 338 (Springer, Berlin, 2009), p. 973.

[20] E. Aurell, C. Mejía-Monasterio, and P. Muratore-Ginanneschi, Phys. Rev. Lett. **106**, 250601 (2011).

[21] U. Boscain and B. Piccoli, in *Controle Non lineaire et Applications*, Le cours du CIMPA, edited by T. Sari (Hermann, Paris, 2005), pp. 19–66.

[22] F. Guerra and L. M. Morato, Phys. Rev. D **27**, 1774 (1983).

[23] R. Chétrite and K. Gawędzki, Commun. Math. Phys. **282**, 469 (2008).

[24] P. Muratore-Ginanneschi, J. Stat. Mech.: Theory Exp. (2014) P05013.

[25] A. Gomez-Marin, T. Schmiedl, and U. Seifert, J. Chem. Phys. **129**, 024114 (2008).

[26] S.-K. Tin, N. Kopell, and C. K. R. T. Jones, SIAM J. Numer. Anal. **31**, 1558 (1994).

[27] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.90.060102 for the explicit treatment of the cumulant expansion of the Gaussian model in the limit of vanishing g.

[28] G. A. Pavliotis and A. M. Stuart, *Multiscale Methods: Averaging and Homogenization*, Springer Texts in Applied Mathematics Vol. 53 (Springer, Berlin, 2008), p. 307.