Supplementary Information for

Energy Dissipation Bounds for Autonomous Thermodynamic Cycles

Samuel J. Bryant, Benjamin B. Machta

Benjamin B. Machta
E-mail: benjamin.machta@yale.edu

This PDF file includes:

Supplementary text
SI References

www.pnas.org/cgi/doi/10.1073/pnas.1915676117

Supporting Information Text

Main Derivation

Here we derive the main results (eq. 5 to eq. 14 in the main text) in greater detail. The system's energy is given by:

$$U(\boldsymbol{y}) = \lambda^\alpha(\theta)\phi_\alpha(\boldsymbol{x}) \tag{1}$$

Here $\boldsymbol{\lambda}$, $\boldsymbol{\phi}$ are indexed by α,β,γ and $\boldsymbol{x}$ is indexed by i,j,k. We use natural units and unit temperature so that $\beta = T = 1$ and $\mu = D = \gamma^{-1}$, where μ is the mobility, D is the diffusion, and γ is the resistance. When μ, D, γ and things like probability currents j appear without indices, they refer to θ. In particular we use ∂_i and ∂_j as shorthand for ∂_{x^i} and ∂_{x^j}. When they appear with roman indices (i,j,k), they refer to $\boldsymbol{x}$. The Langevin equations are:

$$\begin{aligned} \dot{\theta}(t) &= v + \sqrt{2D}\eta(t) & \langle \eta(t)\eta(t')\rangle &= \delta(t-t') \\ \dot{x}^i(t) &= D^{ij}F_j(\boldsymbol{y}) + \sqrt{2D^{ij}}\xi_j(t) & \langle \xi_i(t)\xi_j(t')\rangle &= \delta_{ij}\delta(t-t') \end{aligned} \tag{2}$$

where η and ξ are white noise functions and $F_j(\boldsymbol{y}) = -\partial_j U(\boldsymbol{y})$. We denote $p(\boldsymbol{y},t)$ to mean the probability of finding the system at state $\boldsymbol{y} = (\theta, \boldsymbol{x})$ at time t. The probability currents are given by:

$$\begin{aligned} j(\boldsymbol{y},t) &= [v - D\partial_\theta]\, p(\boldsymbol{y},t) = p(\boldsymbol{y},t)D\,[v\gamma - \partial_\theta \log p(\boldsymbol{y},t)] \\ j^i(\boldsymbol{y},t) &= D^{ij}\,[F_j(\boldsymbol{y}) - \partial_j]\, p(\boldsymbol{y},t) = p(\boldsymbol{y},t)D^{ij}\,[F_j(\boldsymbol{y}) - \partial_j \log p(\boldsymbol{y},t)] \end{aligned} \tag{3}$$

The total averaged entropy production rate is the sum of the production rate for the two individual variables:

$$\langle \dot{S}(t)\rangle = \langle \dot{S}^x + \dot{S}^\theta\rangle = \int \mathrm{d}\boldsymbol{y}\left[\frac{j^i(\boldsymbol{y},t)\gamma_{ij}j^j(\boldsymbol{y},t)}{p(\boldsymbol{y},t)} + \frac{j(\boldsymbol{y},t)\gamma j(\boldsymbol{y},t)}{p(\boldsymbol{y},t)}\right] \tag{4}$$

Starting with the θ contribution:

$$\begin{aligned} \langle \dot{S}^\theta\rangle &= \int \mathrm{d}\boldsymbol{y}\frac{j(\boldsymbol{y},t)\,\gamma\, j(\boldsymbol{y},t)}{p(\boldsymbol{y},t)} = \int \mathrm{d}\boldsymbol{y}\;[v\gamma - \partial_\theta \log p(\boldsymbol{y},t)]\, j(\boldsymbol{y},t) \\ &= v\gamma \int \mathrm{d}\boldsymbol{y}\, j(\boldsymbol{y},t) - \int \mathrm{d}\boldsymbol{y}\, j(\boldsymbol{y},t)\partial_\theta \log p(\boldsymbol{y},t) \\ &= v^2\gamma - v\int \mathrm{d}\boldsymbol{y}\, p(\boldsymbol{y},t)\,\partial_\theta \log p(\boldsymbol{y},t) - \int \mathrm{d}\boldsymbol{y}\, j(\boldsymbol{y},t)\,\partial_\theta \log p(\boldsymbol{y},t) \end{aligned} \tag{5}$$

Since $p(\boldsymbol{y},t)\partial_\theta \log p(\boldsymbol{y},t) = \partial_\theta p(\boldsymbol{y},t)$ and θ is periodic, after integrating over θ, the middle term disappears leaving:

$$\langle \dot{S}^\theta\rangle = \frac{v^2}{D} - \int \mathrm{d}\boldsymbol{y}\, j(\boldsymbol{y},t)\partial_\theta \log p(\boldsymbol{y},t) \tag{6}$$

Next we consider the $\boldsymbol{x}$ contribution:

$$\begin{aligned} \langle \dot{S}^x\rangle &= \int \mathrm{d}\boldsymbol{y}\,\frac{j^i(\boldsymbol{y},t)\,\gamma_{ij}\, j^j(\boldsymbol{y},t)}{p(\boldsymbol{y},t)} = \int \mathrm{d}\boldsymbol{y}\, j^i(\boldsymbol{y},t)(F_i(\boldsymbol{y}) - \partial_i \log p(\boldsymbol{y},t)) \\ &= -\int \mathrm{d}\boldsymbol{y}\, j^i(\boldsymbol{y},t)\partial_i U(\boldsymbol{y}) - \int \mathrm{d}\boldsymbol{y}\, j^i(\boldsymbol{y},t)\partial_i \log p(\boldsymbol{y},t) \end{aligned} \tag{7}$$

Giving us:

$$\langle \dot{S}\rangle = \frac{v^2}{D} - \int \mathrm{d}\boldsymbol{y}\;\left[(j(\boldsymbol{y},t)\partial_\theta + j^i(\boldsymbol{y},t)\partial_i)\log p(\boldsymbol{y},t) + j^i(\boldsymbol{y},t)\partial_i U(\boldsymbol{y})\right] \tag{8}$$

We can rewrite $U = \mathcal{F} - \log p_{eq}$ where $\mathcal{F}(\theta)$ is the free energy for the system for a fixed θ and $p_{eq}(\boldsymbol{x}|\theta) = e^{\mathcal{F}(\theta) - U(\theta,\boldsymbol{x})}$ is the equilibrium Boltzmann distribution for fixed θ. Then $\partial_i U = -\partial_i \log p_{eq}$ since $\mathcal{F}$ is a function of θ only:

$$\langle \dot{S}\rangle = \frac{v^2}{D} - \int \mathrm{d}\boldsymbol{y}\;\left[(j(\boldsymbol{y},t)\partial_\theta + j^i(\boldsymbol{y},t)\partial_i)\log p(\boldsymbol{y},t) - j^i(\boldsymbol{y},t)\partial_i \log p_{eq}(\boldsymbol{x}|\theta)\right] \tag{9}$$

Now we may integrate by parts on each of the terms in the integral:

$$\langle \dot{S}\rangle = \frac{v^2}{D} + \int \mathrm{d}\boldsymbol{y}\;\left[(\partial_\theta j(\boldsymbol{y},t) + \partial_i j^i(\boldsymbol{y},t))\log p(\boldsymbol{y},t) - (\partial_i j^i(\boldsymbol{y},t))\log p_{eq}(\boldsymbol{x}|\theta)\right] \tag{10}$$

where the boundary terms disappear because $p(\boldsymbol{y},t)$ goes to zero at the boundaries of x.

Now we assume that the system reaches a steady state so all time dependences drop out of the problem. In particular, the Fokker-Plank equation yields $\partial_\theta j(\boldsymbol{y}) + \partial_i j^i(\boldsymbol{y}) = -\partial_t p = 0$, allowing us to remove the first term in the integral and swap $\partial_i j^i(\boldsymbol{y}) = -\partial_\theta j(\boldsymbol{y})$ in the second:

$$\langle \dot{S} \rangle = \frac{v^2}{D} + \int \mathrm{d}\boldsymbol{y}\ [\partial_\theta j(\boldsymbol{y})] \log p_{eq}(\boldsymbol{x}|\theta) \tag{11}$$

Next we expand $\partial_\theta j(\boldsymbol{y}) = v\partial_\theta p(\boldsymbol{y}) - D\partial_\theta^2 p(\boldsymbol{y})$ and $\log p_{eq}(\boldsymbol{x}|\theta) = \mathcal{F}(\theta) - U(\boldsymbol{y})$, giving:

$$\langle \dot{S} \rangle = \frac{v^2}{D} + \int \mathrm{d}\boldsymbol{y}\ \left[v\partial_\theta p(\boldsymbol{y}) - D\partial_\theta^2 p(\boldsymbol{y})\right] (\mathcal{F}(\theta) - U(\boldsymbol{y})) \tag{12}$$

Now we integrate by parts to move the ∂_θ, ∂_θ^2 to the other term. Since θ is periodic, we can always neglect the boundary terms:

$$\langle \dot{S} \rangle = \frac{v^2}{D} - \int \mathrm{d}\boldsymbol{y}\, p(\boldsymbol{y}) \left[v\partial_\theta + D\partial_\theta^2\right] [\mathcal{F}(\theta) - U(\boldsymbol{y})] \tag{13}$$

We now need to compute the derivatives of U and $\mathcal{F}$ with respect to θ. Recall we denote $\langle \cdots \rangle_\theta$ as an average over all microstates $\boldsymbol{x}$ with non-equilibrium weight $p(\boldsymbol{x}|\theta)$ for a fixed value of θ. We also denote $\langle \cdots \rangle_{\mathrm{eq},\theta}$ to indicate an average over all $\boldsymbol{x}$ with equilibrium Boltzmann weight $p_{\mathrm{eq}}(\boldsymbol{x}|\theta)$ for a fixed value of θ:

$$\partial_\theta U(\theta, \boldsymbol{x}) = \frac{\partial U}{\partial \lambda^\alpha}\frac{\partial \lambda^\alpha}{\partial \theta} = \phi_\alpha(\boldsymbol{x})\frac{\partial \lambda^\alpha}{\partial \theta}(\theta) \tag{14}$$

$$\partial_\theta^2 U(\theta, \boldsymbol{x}) = \phi_\alpha(\boldsymbol{x})\frac{\partial^2 \lambda^\alpha}{\partial \theta^2}(\theta) \tag{15}$$

$$\partial_\theta \mathcal{F}(\theta) = \frac{\partial \mathcal{F}}{\partial \lambda^\alpha}\frac{\partial \lambda^\alpha}{\partial \theta} = \langle \phi_\alpha \rangle_{\mathrm{eq},\theta}\frac{\partial \lambda^\alpha}{\partial \theta}(\theta) \tag{16}$$

$$\partial_\theta^2 \mathcal{F}(\theta) = \langle \phi_\alpha \rangle_{\mathrm{eq},\theta}\frac{\partial^2 \lambda^\alpha}{\partial \theta^2}(\theta) - g^\lambda_{\alpha\beta}\frac{\partial \lambda^\alpha}{\partial \theta}\frac{\partial \lambda^\beta}{\partial \theta} \tag{17}$$

where $g^\lambda_{\alpha\beta} = -\frac{\partial}{\partial \lambda^\beta}\langle \phi_\alpha \rangle_{\mathrm{eq},\theta}$ is the equilibrium Fisher information metric in the λ basis. Plugging these in gives:

$$\dot{S}(\boldsymbol{y}) = \frac{v^2}{D} + \left[v\frac{\partial \lambda^\alpha}{\partial \theta}(\theta) + D\frac{\partial^2 \lambda^\alpha}{\partial \theta^2}(\theta)\right]\delta\phi_\alpha(\boldsymbol{y}) + Dg^\lambda_{\alpha\beta}(\theta)\frac{\partial \lambda^\alpha}{\partial \theta}(\theta)\frac{\partial \lambda^\beta}{\partial \theta}(\theta) \tag{18}$$

Where $\delta\phi_\alpha = \phi_\alpha(\boldsymbol{x}) - \langle \phi_\alpha \rangle_{\mathrm{eq},\theta}$ is the deviation of ϕ from its equilibrium value for fixed system state θ. If we identify $g^\theta \equiv \frac{\partial \lambda^\alpha}{\partial \theta} g^\lambda_{\alpha\beta} \frac{\partial \lambda^\beta}{\partial \theta}$ as the Fisher information metric on θ inherited from λ, we get:

$$\dot{S}(\boldsymbol{y}) = \frac{v^2}{D} + \left[v\frac{\partial \lambda^\alpha}{\partial \theta}(\theta) + D\frac{\partial^2 \lambda^\alpha}{\partial \theta^2}(\theta\right]\delta\phi_\alpha(\boldsymbol{y}) + Dg^\theta(\theta) \tag{19}$$

To find the average dissipation rate when the system is at the point θ, we average over all $\boldsymbol{x}$ with weight $p(x|\theta)$. This immediately yields the multidimensional analog of (13):

$$\langle \dot{S} \rangle_\theta = \frac{v^2}{D} + Dg^\theta(\theta) + \left[v\frac{\partial \lambda^\alpha}{\partial \theta}(\theta) + D\frac{\partial^2 \lambda^\alpha}{\partial \theta^2}(\theta)\right]\langle \delta\phi_\alpha \rangle_\theta \tag{20}$$

Again, this is an exact expression. Next we use linear response to approximate $\langle \delta\phi_\alpha \rangle_\theta$ assuming that v and D are small.

Linear Response Approximation. Our starting point is the following expression (1) which gives the average linear response of ϕ at time t_0 to a specific control trajectory $\theta(\tau)$:

$$\langle \delta\phi_\alpha(t_0) \rangle_{\theta(\tau)} = \int_{-\infty}^{0} \mathrm{d}t' \left[\frac{\mathrm{d}C^{\theta(t_0)}_{\alpha\beta}}{\mathrm{d}t'}\right] \left[\boldsymbol{\lambda}(t_0) - \boldsymbol{\lambda}(t_0 + t')\right]^\beta \tag{21}$$

with the multidimensional autocorrelation function for $\boldsymbol{\phi}$:

$$C^{\theta(t_0)}_{\alpha\beta}(t) = \langle \delta\phi_\alpha(0)\delta\phi_\beta(t) \rangle_{\mathrm{eq},\theta(t_0)} = \langle \phi_\alpha(0)\phi_\beta(t_0) \rangle_{\mathrm{eq},\theta(t_0)} - \langle \phi_\alpha(0) \rangle_{\mathrm{eq},\theta(t_0)} \langle \phi_\beta(t) \rangle_{\mathrm{eq},\theta(t_0)} \tag{22}$$

This is the linear response function to a *single* path.

The expression we need is $\langle \delta\phi_\alpha(t_0) \rangle_{\theta_0}$, the average of the above expression over *all paths* $\theta(\tau)$ such that $\theta(t_0) = \theta_0$ for a specific point (θ_0, t_0). The important point is that only $\boldsymbol{\lambda}(t + t')$ is trajectory dependent, the other parts of the expression only depend on the value of the trajectory at the moment t_0. Therefore we may write:

$$\langle \delta\phi_\alpha(t_0) \rangle_{\theta_0} = \int_{-\infty}^{0} \mathrm{d}t' \left[\frac{\mathrm{d}C^{\theta_0}_{\alpha\beta}}{\mathrm{d}t'}\right] \langle \lambda^\beta(t_0) - \lambda^\beta(t_0 + t') | \theta_0, t_0 \rangle \tag{23}$$

where here $\langle \cdots | \theta_0, t_0 \rangle$ represents an average over all possible control paths $\theta(\tau)$ such that $\theta(t_0) = \theta_0$ weighted by their probability as given by the Langevin dynamics of θ. By integrating by parts we are left with:

$$\langle \delta\phi_\alpha(t_0) \rangle_{\theta_0} = \int_{-\infty}^{0} \mathrm{d}t' C^{\theta_0}_{\alpha\beta}(t') \frac{\mathrm{d}}{\mathrm{d}t'} \langle \lambda^\beta(t_0 + t') | \theta_0, t_0 \rangle \tag{24}$$

To compute this, we need an expression for the expectation value of λ^β for this set of trajectories, which we may formally write as:

$$\langle \lambda^\beta(t_0 + t') | \theta_0, t_0 \rangle = \int_{-\infty}^{\infty} \mathrm{d}\theta' \, p(\theta_0 + \theta', t_0 + t' | \theta_0, t_0) \lambda^\beta(\theta_0 + \theta') \tag{25}$$

Note that despite the fact that θ is periodic, the integration bounds here are not. This is because $\theta' = 2\pi$ in this context refers to the control parameter making a full cycle in time t' which is not the same as it not moving ($\theta' = 0$).

Since the stochasticity of θ is driven by Gaussian noise, it's trivial to write down p:

$$p(\theta_0 + \theta', t_0 + t' | \theta_0, t_0) = \frac{1}{\sqrt{4\pi D|t'|}} e^{-\frac{(\theta' - vt')^2}{4D|t'|}} \tag{26}$$

which is a Gaussian that diffuses away from a Dirac delta function as $|t'| > 0$ (see Fig 2 from the main text). We also Taylor expand $\lambda^\beta(\theta_0 + \theta')$ about θ_0:

$$\lambda^\beta(\theta_0 + \theta') = \lambda^\beta(\theta_0) + \theta' \partial_\theta \lambda^\beta(\theta_0) + \frac{1}{2!}(\theta')^2 \partial_\theta^2 \lambda^\beta(\theta_0) + \cdots \tag{27}$$

Putting these together gives:

$$\langle \lambda^\beta(t_0 + t') | \theta_0, t_0 \rangle = \sum_{k=0} \frac{\partial_\theta^k \lambda^\beta(\theta_0)}{k! \sqrt{4\pi D|t'|}} \int_{-\infty}^{\infty} \mathrm{d}\theta' \, e^{-\frac{(\theta' - vt')^2}{4D|t'|}} (\theta')^k \tag{28}$$

Solving this integral for each value of k will yield terms of the form:

$$(D|t'|)^m (vt')^n \partial_\theta^{2m+n} \lambda^\beta(\theta_0) \tag{29}$$

An underlying assumption of this work is that we are in the near-equilibrium regime where the perturbation on the system is weak. We treat v and D as small parameters and only keep terms of order $m + n = 1$ in v, D:

$$\langle \lambda^\beta(t_0 + t') | \theta_0, t_0 \rangle = \lambda^\beta(\theta_0) + vt' \partial_\theta \lambda^\beta(\theta_0) + D|t'| \partial_\theta^2 \lambda^\beta(\theta_0) + \mathcal{O}(2) \tag{30}$$

We can see that to leading order, the time-derivative of this expression will be time-independent. Thus, to leading order we can replace $\frac{\mathrm{d}}{\mathrm{d}t'} \langle \lambda^\beta(t + t') | \theta_0, t_0 \rangle$ with its value at $t' = 0$. However, due to the presence of the absolute value terms $|t'|$, the time derivative at $t' = 0$ is discontinuous. This is more than a minor detail; it's connected with the choice of whether to define derivatives in stochastic calculus using left sided limits or right sided limits analogous to choices in Riemann sums (2).

We introduce the following notation to distinguish between the left and right-sided limits of the time derivative of the conditional expectation value of λ:

$$\left\langle \frac{\mathrm{d}\lambda^\beta}{\mathrm{d}t^\pm} \right\rangle_{\theta_0} \equiv \lim_{t' \to 0^\pm} \frac{\mathrm{d}}{\mathrm{d}t'} \langle \lambda^\beta(t_0 + t') | \theta_0, t_0 \rangle \approx v \partial_\theta \lambda^\beta(\theta_0) \pm D \partial_\theta^2 \lambda^\beta(\theta_0) \tag{31}$$

As one may expect, this is exactly the formula for the net drift of λ^β at $t' = 0$ given by the Ito formula under the Ito $(+)$ and reverse-Ito $(-)$ conventions (2).

Plugging this result into eq. **24** and using the definition of thermodynamic friction:

$$\tilde{g}^\lambda_{\alpha\beta} = \tau_{\alpha\beta} \circ g^\lambda_{\alpha\beta} = \int_{-\infty}^{0} \mathrm{d}t' \, C^\theta_{\alpha\beta}(t') \tag{32}$$

yields the time-independent expression for the average linear response:

$$\langle \delta\phi_\alpha \rangle_{\theta_0} = \tilde{g}^\lambda_{\alpha\beta}(\theta_0) \left\langle \frac{\mathrm{d}\lambda^\beta}{\mathrm{d}t^-} \right\rangle_{\theta_0} \tag{33}$$

which in turn gives an average entropy production rate for the system when it's at the state θ_0:

$$\langle \dot{S} \rangle_{\theta_0} = \left\langle \frac{\mathrm{d}\lambda^\alpha}{\mathrm{d}t^+} \right\rangle_{\theta_0} \tilde{g}^\lambda_{\alpha\beta}(\theta_0) \left\langle \frac{\mathrm{d}\lambda^\beta}{\mathrm{d}t^-} \right\rangle_{\theta_0} + \frac{v^2}{D} + D g^\theta(\theta_0) \tag{34}$$

What about the higher order terms in (30)? We can neglect them under the assumption that the speed of control is small compared to the excitation timescale τ of the system at equilibrium. However, an explicit mathematical statement of this requirement is challenging because of the unknown form of $\lambda(\theta)$. Since C is a decay function, we expect roughly that:

$$\int_{-\infty}^{0} \mathrm{d}t'\,(t')^k C^{\theta}_{\alpha\beta}(t') \sim g^{\lambda}_{\alpha\beta}\tau^{k+1} \qquad [35]$$

Thus in keeping terms of order $m+n>1$ we would generate additional contributions to (33) of the form:

$$(D\tau)^m (v\tau)^n \left[g^{\lambda}_{\alpha\beta}\partial_\theta^{2m+n}\lambda^\beta(\theta)\right] \qquad [36]$$

Comparing these to the leading order terms $(D\tau)g^{\lambda}_{\alpha\beta}\partial^2_\theta\lambda^\beta$ and $(v\tau)g^{\lambda}_{\alpha\beta}\partial_\theta\lambda^\beta$, we can see that for a reasonably behaved function $\lambda(\theta)$ with a characteristic length scale L we can summarize our assumption via the requirements:

$$v\tau/L \ll 1 \qquad\qquad D\tau/L^2 \ll 1 \qquad [37]$$

Essentially this is the requirement that $\left\langle \frac{\mathrm{d}\lambda^\alpha}{\mathrm{d}t}(t)\right\rangle$ (under the reverse-Ito convention) remains relatively constant over the system relaxation timescale τ. This is the natural stochastic generalization of the constraint imposed by [3].

However, because we have not explicitly given a definition for L, this requirement is admittedly a little vague. One major complication in doing so arises from the fact that because we only care about the *total* dissipation bounds, we only want to keep the terms in (34) that contribute to the leading order behavior of the *integral* of (34). Thus, while it may be the case that for a specific point θ_0, the second order term dominates: $(v\tau)^2\partial^2_\theta\lambda(\theta_0) \gg (v\tau)\partial_\theta\lambda(\theta_0)$, we still want to drop the higher order term because its contribution to the total integral is subleading. The actual formal constraints dictating when this approximation is appropriate is further complicated by the unconstrained behavior of $\tilde{g}(\theta)$ and $\lambda(\theta)$. However, it should be clear that as we approach equilibrium behavior (D, $v\to 0$), the kept terms dominate over the dropped terms. We feel that the constraint given in (37) satisfactorily captures this idea.

References

1. Zwanzig R (2001) *Nonequilibrium statistical mechanics.* (Oxford University Press, Oxford ; New York).
2. Gardiner CW (2004) *Handbook of stochastic methods for physics, chemistry, and the natural sciences*, Springer series in synergetics. (Springer-Verlag, Berlin ; New York), 3rd ed edition.