Dual thermodynamic ensembles, relative entropy, and excess free energy

Gavin E. Crooks

2025-04-15

It has long been known that the entropy of a non-equilibrium thermodynamic ensemble relative to the canonical ensemble of the same system at thermodynamic equilibrium is the excess free energy, the difference in free energy between the corresponding equilibrium and non-equilibrium ensembles. Herein we show that the dual, or reverse, relative entropy of the equilibrium canonical ensemble to the non-equilibrium ensemble also has a physical interpretation as the excess free energy of a new ensemble that is thermodynamically dual to the original non-canonical ensemble.

We consider an ensemble B of a system thermally coupled to an idealized heat bath at constant inverse temperature $\beta = 1/k_BT$, with energy spectrum $E_B(x)$, and probabilities $p_B(x)$. This ensemble is out of thermal equilibrium, and therefore the probabilities are not given by the canonical ensemble of equilibrium statistical mechanics [1].

$$p_{B}(x) \neq \frac{e^{-\beta E_{B}(x)}}{Z_{B}}$$
(1)

We define the ensemble A as the equilibrium thermodynamic ensemble of the same system, coupled to the same heat bath, with the same energy spectrum $E_A(x) = E_B(x)$, but with canonical probabilities,

$$p_A(x) = \frac{e^{-\beta E_A(x)}}{Z_A} = e^{-\beta E_A(x) + \beta F_A}$$
. (2)

Here $Z_A = \sum_x \exp(-\beta E_A(x))$ is the partition function, and F_A is the free energy of the ensemble A [2], defined as $\beta F_A = \beta \langle E_A \rangle_A - S_A$, where S is the entropy of the ensemble [3–6], $S_A = -\sum p_A(x) \ln p_A(x)$, and $\langle E_A \rangle_A$ is the energy spectrum of A averaged over the distribution of A. For a canonical ensemble the partition function and free energy are related by $\beta F_A = -\ln Z_A$, but for a non-equilibrium non-canonical ensemble this relation no longer holds, $\beta F_B \neq -\ln Z_B$.

The relative entropy (KL-divergence) [7-9] of the ensemble B relative to the equilibrated ensemble A is the difference in free energy between them [10-21]. (For a

summary of the history of this result see [22])

$$D(B||A) = \sum_{x} p_{B}(x) \ln \frac{p_{B}(x)}{p_{A}(x)}$$
(3)
$$= \sum_{x} p_{B}(x) \ln p_{B}(x) - \sum_{x} p_{B}(x) \ln p_{A}(x)$$
$$= \sum_{x} p_{B}(x) \ln p_{B}(x) - \sum_{x} p_{B}(x) (-\beta E_{A}(x) + \beta F_{A})$$
$$= -S_{B} + \langle \beta E_{B} \rangle_{B} - \beta F_{A}$$
$$= \beta F_{B} - \beta F_{A}$$
$$= \beta F_{B}^{ex}$$

Here we define the free energy of a non-equilibrium ensemble analogously as for an equilibrium ensemble $\beta F_B = \beta \langle E_B \rangle_B - S_B$, Note that relative entropy is non-negative, being zero only if the ensembles are identical. Thus the excess free energy (the difference between non-equilibrium and corresponding equilibrium values) is positive, and the free energy is minimized in thermodynamic equilibrium, as expected.

The excess free energy is the amount of energy that would be dissipated if we allowed the system to relax to thermal equilibrium. It is also the reversible work, the amount of work that can be extracted by a thermodynamically reversible process connecting ensemble B and A. To show this it is useful to introduce a third ensemble, C, that has the same probabilities as B, $p_B(x) = p_C(x)$, but a different energy spectrum that ensures that C is also in thermodynamic equilibrium.

$$\beta E_{C}(x) = -\ln p_{B}(x) + \varepsilon_{c}$$
(4)

Here ε_c is an arbitrary energy offset that will be absorbed into the normalization constant. It does not effect the

probability distribution of C, but it does effect what we think the energy and free energy are. Since the maximum probability of any state is unity, ε_c is the lowest energy the ground state could have.

Although B is a non-equilibrium system we can, in principle, instantaneously stabilize B by changing the energy levels of B to match the energy levels of C [16, 17] in a single instantaneous jump. This requires work, whose average value is also the difference in free energy between C and B (since the entropies are the same, the difference is entirely energetic).

$$\begin{split} \beta \langle W_{B \to C}^{inst} \rangle &= \sum_{x} p_{B}(x) \big(\beta E_{C}(x) - \beta E_{B}(x) \big) \\ &= \langle \beta E_{C} \rangle_{C} - \langle \beta E_{B} \rangle_{B} \\ &= \beta F_{C} - \beta F_{B} \end{split} \tag{5}$$

In a sense this instantaneous transformation is thermodynamically reversible, since there is no change in entropy, provided that we do not allow any time for the system to relax towards equilibrium and dissipate energy.

We can now convert the equilibrium ensemble C to A by a quasi-static thermodynamically reversible transformation, for which the work done during this reversible transformation is equal to the difference in free energies.

$$\beta W_{C \to A}^{rev} = \beta F_A - \beta F_C \tag{6}$$

Combining the instantaneous stabilization of B to C, followed by the quasi-static transformation of C to A, we see that the excess free energy is indeed the total reversible work of this process, justifying extending the definition of free energy to non-equilibrium ensembles.

Another approach to understanding that the excess free energy bounds the work is via the Jarzynski identity [23– 25]

$$\left\langle e^{-\beta W} \right\rangle_{\mathrm{A},\Lambda} = e^{-\beta \Delta \mathsf{F}}$$
 (7)

Here *W* is the work done on a system that is initially in the equilibrium A, and ΔF is the change in the equilibrium free energy induced by the protocol Λ . The Jarzynski identity implies that the average excess work is nonnegative (which essentially a statement of the second law) via an application of Jensen's inequality.

$$\langle \beta W \rangle_{A,\Lambda} \ge \beta \Delta F_{\Lambda}$$
 (8)

The Jarzynski identity assumes that the system starts in thermal equilibrium. But we can make a small change and obtain a similar relation where we start the process from a non-equilibrium ensemble. First we write the Jarzynski relation more explicitly.

$$\sum_{\mathbf{x}_0, \mathbf{X}} \mathbf{p}_{\mathbf{A}}(\mathbf{x}_0) \ \mathbf{p}(\mathbf{X}|\mathbf{\Lambda}, \mathbf{x}_0) \ e^{-\beta W[\mathbf{X}, \mathbf{\Lambda}]} = e^{-\beta \Delta \mathbf{F}_{\mathbf{\Lambda}}} \tag{9}$$

Here x_0 is the initial state of the system, and X is the trajectory of the system during the driving process.

We can now substitute in a different initial probability distribution over states, moving extra terms to the exponential

$$\sum_{x_0, X} p_B(x_0) p(X|x_0) \ e^{-\beta W - \ln \frac{p_B(x_0)}{p_A(x_0)}} = e^{-\beta \Delta F_A}$$
(10)

which leads to a Jarzynski-like expression that is valid for any initial distribution.

$$\left\langle e^{-\beta W - \ln \frac{p_{B}(x_{0})}{p_{A}(x_{0})}} \right\rangle_{B,\Lambda} = e^{-\beta \Delta F_{\Lambda}}$$
(11)

Then by Jensen's inequality we have

$$\langle \beta W \rangle_{B,\Lambda} \ge \beta \Delta F_{\Lambda} - D(B||A) ,$$
 (12)

$$\geq \beta \Delta F_{\Lambda} - \beta F_{B}^{ex} . \tag{13}$$

If we now further assume that the initial and final equilibrium ensembles associated with the protocol Λ are both A, then we have that the work done, starting from a non-equilibrium ensemble, is bounded by the free energy difference between the equilibrium and non-equilibrium ensembles.

$$\langle \beta W \rangle_{B,\Lambda} \ge \beta F_A - \beta F_B$$
 (14)

(This argument from the Jarzynski identity is similar to those found in [15, 26])

These preceding arguments show that it is proper to define the free energy of a non-equilibrium ensemble as $\beta F = S - \beta \langle E \rangle$, and that this free energy is related to the relative entropy between non-equilibrium and corresponding equilibrium ensembles (3). Howsoever, we are intending to give a thermodynamic interpretation for the reverse, or dual, relative entropy D(A||B). To this end we introduce a fourth ensemble D. This ensemble has the same probability distribution as A, and the same energy spectrum as C.

$$p_D(x) = p_A(x) = \frac{e^{-\beta E_B(x)}}{Z_B}$$
(15)

$$\beta E_D(x) = \beta E_C(x) = -\ln p_B(x) + \varepsilon_c$$
(16)

This new ensemble D is thermodynamically dual to ensemble B in that the roles of entropy and energy have been interchanged: the probabilities of D are defined by the energies of B, and the energies of D are defined by the probabilities of B.

The interrelation of these four ensembles is summarized by the following diagram.



We are now in a position to give a thermodynamic interpretation of D(A||B). Since the probability distribution of A is the same as D, and that of B the same as C, it follows that D(A||B) = D(D||C). But since C is the canonical equilibrium that the non-canonical ensemble D would relax to (they have the same energy spectrum and bath temperature), it follows that D(D||C) is the excess free energy of the dual ensemble D.

$$D(A||B) = D(D||C) = \beta F_D - \beta F_C = \beta F_D^{ex}$$
(17)

To make the concept of a thermodynamically dual ensemble more concrete, we can introduce the notation $B^* = D$. The corresponding equilibrated ensemble can be notated as $B^{eq} = A$, and the instantaneously stabilized ensemble as $B^{inst} = B^{*eq} = C$.

$$p_{B^{\star}}(x) = \frac{e^{-\beta E_B(x)}}{Z_B}$$
(18)

$$\beta \mathsf{E}_{\mathsf{B}^{\star}}(\mathsf{x}) = -\ln \mathsf{p}_{\mathsf{B}}(\mathsf{x}) + \beta \mathsf{F}_{\mathsf{B}^{\mathsf{eq}}} \tag{19}$$

We have also fixed the energy offset $\varepsilon_{B^*} = \beta F_{B^{eq}}$. This is convenient (but not thermodynamically necessary) as it ensures that that the free energies of the equilibrated and stabilized ensembles are the same, $\beta F_{B^{inst}} = \beta F_{B^{eq}}$, and therefore that the thermodynamic duality is an involution $B^{**} = B$. Ensembles in thermodynamically equilibrium are self-dual $B^{eq*} = B^{eq}$.

References

- [1] J. Willard Gibbs. *Elementary principles in statistical mechanics*. Charles Scribner's Sons, New York (1902).
- [2] Hermann von Helmholtz. Physical memoirs, selected and translated from foreign sources, volume 1, chapter On the Thermodynamics of Chemical Processes (1882), pages 43–97. Taylor & Francis (1888).

- [3] Rudolf Clausius. Ueber verschiedene für die anwendung bequeme formen der hauptgleichungen der mechanischen wärmetheorie. *Annalen der Physik und Chemie*, 201(7):353–400 (1865). doi:10.1002/andp. 18652010702.
- [4] Ludwig Boltzmann. Weitere studien über das Wärmegleichgewicht unter Gasmolekülen (Further studies on the thermal equilibrium of gas molecules). *Sitzungsberichte Akad. Wiss., Vienna, part II,* 66:275– 370 (1872).
- [5] Claude E. Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27:379–423, 623–656 (1948). doi:10.1002/j.1538-7305.1948.tb01338.x.
- [6] Edwin T. Jaynes. Information theory and statistical mechanics. *Phys. Rev.*, 106:620–630 (1957). doi:10. 1103/PhysRev.106.620.
- [7] Solomon Kullback and Richard A. Leibler. On information and sufficiency. *Ann. Math. Statist.*, 22:79–86 (1951). doi:10.1214/aoms/1177729694.
- [8] Thomas M. Cover and Joy A. Thomas. *Elements of information theory*. Wiley, New York, 2nd edition (2006).
- [9] Gavin E. Crooks. On measures of entropy and information. Technical report, BITS (2015-2021). Tech. Note 009 v0.8 http://threeplusone.com/info.
- [10] R. B. Bernstein and R. D. Levine. Entropy and chemical change. I. Characterization of product (and reactant) energy distributions in reactive molecular collisions: Information and entropy deficiency. *J. Chem. Phys.*, 57(1):434–449 (1972). doi:10.1063/1.1677983.
- [11] Robert Shaw. *The Dripping Faucet as a Model Chaotic System*. Aerial Press, Santa Cruz (1984).
- [12] Bernard Gaveau and L. S. Schulman. A general framework for non-equilibrium phenomena: the master equation and its formal consequences. *Phys. Lett. A*, 229(6):347–353 (1997). doi:10.1016/ S0375-9601(97)00185-0.
- [13] Hong Qian. Relative entropy: free energy associated with equilibrium fluctuations and nonequilibrium deviations. *Phys. Rev. E*, 63:042103 (2001). doi:10. 1103/PhysRevE.63.042103.
- [14] Takahiro Hatano and Shin-ichi Sasa. Steadystate thermodynamics of Langevin systems. *Phys. Rev. Lett.*, 86(16):3463–3466 (2001). doi:10.1103/ PhysRevLett.86.3463.

- [15] S. Vaikuntanathan and Christopher Jarzynski. Dissipation and lag in irreversible processes. *Europhys. Lett.*, 87:60005 (2009). doi:10.1209/0295-5075/87/ 60005.
- [16] H.-H. Hasegawa, J. Ishikawa, K. Takara, and D.J. Driebe. Generalization of the second law for a nonequilibrium initial state. *Phys. Lett. A*, 374(8):1001–1004 (2010). doi:10.1016/j.physleta.2009. 12.042.
- [17] K. Takara, H.-H. Hasegawa, and D.J. Driebe. Generalization of the second law for a transition between nonequilibrium states. *Phys. Lett. A*, 375(2):88–92 (2010). doi:10.1016/j.physleta.2010.11.002.
- [18] Massimiliano Esposito and Christian Van den Broeck. Second law and Landauer principle far from equilibrium. *Europhys. Lett.*, 95(4):40004 (2011). doi:10.1209/0295-5075/95/40004.
- [19] Gavin E. Crooks and David A. Sivak. Measures of trajectory ensemble disparity in nonequilibrium statistical dynamics. J. Stat. Mech.: Theor. Exp., page P06003 (2011). doi:10.1088/1742-5468/2011/ 06/P06003.
- [20] David A. Sivak and Gavin E. Crooks. Nearequilibrium measurements of nonequilibrium free energy. *Phys. Rev. Lett.*, 108(15):150601 (2012). doi: 10.1103/PhysRevLett.108.150601.
- [21] Sebastian Deffner and Eric Lutz. Information free energy for nonequilibrium states. arXiv:1201.3888.
- [22] Gavin E. Crooks. On the Fluctuations of Dissipation: An annotated bibliography on the Fluctuation Theorem, Stochastic Thermodynamics and related subjects (2010-2025). Tech. Note 005 http: //threeplusone.com/ftbib.
- [23] Christopher Jarzynski. Nonequilibrium equality for free energy differences. *Phys. Rev. Lett.*, 78(14):2690– 2693 (1997). doi:10.1103/PhysRevLett.78.2690.
- [24] Christopher Jarzynski. Equilibrium free-energy differences from nonequilibrium measurements: A master-equation approach. *Phys. Rev. E*, 56(5):5018– 5035 (1997). doi:10.1103/PhysRevE.56.5018.
- [25] Gavin E. Crooks. Nonequilibrium measurements of free energy differences for microscopically reversible Markovian systems. J. Stat. Phys., 90(5-6):1481–1487 (1998). doi:10.1023/A:1023208217925.
- [26] ? [citation needed]. ?