

# The length of time's arrow

Edward H. Feng

*College of Chemistry, University of California, Berkeley, Berkeley, California, 94720, USA*

Gavin E. Crooks

*Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, 94720, USA*

(Dated: August 29, 2008)

An unresolved problem in physics is how the thermodynamic arrow of time arises from an underlying time reversible dynamics. We contribute to this issue by developing a measure of time symmetry breaking, and by using the work fluctuation relations, we determine the time asymmetry of recent single molecule RNA unfolding experiments. We define time asymmetry as the Jensen-Shannon divergence between trajectory probability distributions of an experiment and its time-reversed conjugate. Among other interesting properties, the length of time's arrow bounds the average dissipation and determines the difficulty of accurately estimating free energy differences in non-equilibrium experiments.

PACS numbers: 05.70.Ln, 05.40.-a

In our everyday lives we have the sense that time flows inexorably from the past into the future; water flows downhill; mountains erode; we are born, grow old and die; we anticipate the future but remember the past. Yet almost all of the fundamental theories of physics – classical mechanics, electrodynamics, quantum mechanics, general relativity and so on – are symmetric with respect to time reversal. The only fundamental theory that picks out a preferred direction of time is the second law of thermodynamics, which asserts that the entropy of the universe increases as time flows towards the future [1]. This provides an orientation, or arrow of time, and it is generally believed that all other time asymmetries, such as our sense that future and past are different, are a direct consequence of this thermodynamic arrow [2, 3].

When the dissipation, or the total increase in entropy, is large, the orientation of time's arrow is self evident. If we watch a movie in which shards of pottery jump off the floor, assemble themselves into a cup, and land on a table, then clearly someone has threaded the film through the projector backwards. On the other hand, if the dissipation is microscopic, then the distinction between past and future becomes nebulous. This is because a more general statement the second law claims the dissipation is positive *on average*,  $\langle \Delta S_{\text{total}} \rangle \geq 0$  [8, 9]. If we repeat the same experiment many times, the entropy might increase or decrease on different occasions. Only the average dissipation must be positive. Thus, if we view a movie of a microscopic system undergoing a dissipative transformation, we cannot determine with certainty whether time moves forward or backwards.

Here, we seek a quantitative measure of time asymmetry in a driven microscopic system such as the single molecule RNA pulling experiments explained in Fig. 1. Naively, one might use the average dissipation for this quantification, but we will show that a large average dissipation can arise for dynamics that are essentially time

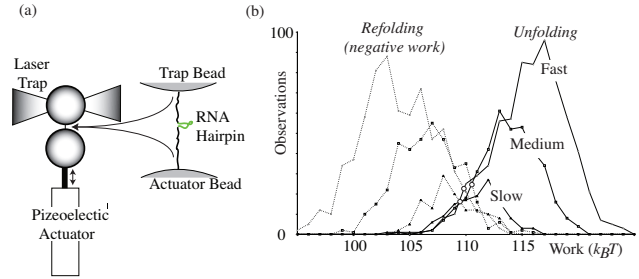


FIG. 1: In this Letter, we discuss the definition and measurement of time-asymmetry in microscopic systems. (a) As a concrete example, we analyze the time asymmetry of a single molecule experiment in which an RNA molecule is attached between two beads [4]. (a) One bead is captured in an optical laser trap that can measure the applied force on the bead. The other bead is fixed to a piezoelectric actuator. The controllable parameter  $\lambda$  is the distance between the fixed bead and the center of the laser trap. For the forward protocol, the RNA hairpin is initially in thermal equilibrium in the folded state with extension  $\lambda(a)$ . The extension is then increased to  $\lambda(b)$ , unfolding the RNA. In the conjugate, time reversed protocol, the RNA is initially in thermal equilibrium in the unfolded state, and the extension is lowered from  $\lambda(b)$  back to  $\lambda(a)$ , allowing the RNA to refold [4, 5, 6, 7]. (b) Histograms of work measurements for folding and unfolding an RNA hairpin at three different rates. Observations are binned into integers centered at  $1 k_B T$  intervals. Note that Eq. (6) predicts that the folding and unfolding work distributions cross at the free energy change.

symmetric. Instead, we develop a measure based on the Jensen-Shannon divergence between the forward and reverse probability distributions of trajectories for a microscopic system. We find that this measure of time's arrow has intuitive physical and information theoretic interpretations and constrains the minimum average dissipation. Moreover, recent advances in far-from-equilibrium statistical physics allow one to measure time's arrow in real

world experiments.

We consider a physical system driven from thermal equilibrium by an external perturbation. For such an experimental protocol,  $\Lambda$  denotes a set of controllable parameters  $\lambda(t)$  for  $t \in [a, b]$  which describe how the system is driven from the initial equilibrium at  $\lambda(a)$ . We are also interested in the conjugate time reversed protocol  $\tilde{\Lambda}$  in which the system begins in thermal equilibrium at  $\lambda(b)$  and the controllable parameters retrace the same series of changes, in reverse, back to  $\lambda(a)$ . In the single molecule experiments of Fig. 1, the distance between the center of the laser trap and the fixed bead plays the role of  $\lambda(t)$ . For each realization of the forward protocol  $\Lambda$ , the system travels along a trajectory  $\mathbf{x}$  which represents the states  $\mathbf{x}(t)$  for  $t \in [a, b]$ . We define a conjugate time reversed trajectory  $\tilde{\mathbf{x}}$  such that  $\tilde{\mathbf{x}}(t) = \mathbf{x}(t)$  for  $t \in [b, a]$ .

We quantify the intrinsic time asymmetry of a driven system as the distinguishability of conjugate forward and reverse experiments. Given a microscopic trajectory  $\mathbf{x}$ , can we tell if it was generated by the protocol  $\Lambda$ , or whether it is the time reversal of a trajectory generated by the reverse protocol  $\tilde{\Lambda}$ ? Specifically, we define the time asymmetry  $A$  as

$$A[\Lambda] \equiv \text{JS} \left( P[\mathbf{x}|\Lambda] ; P[\tilde{\mathbf{x}}|\tilde{\Lambda}] \right) \quad (1)$$

in which  $P[\mathbf{x}|\Lambda]$  and  $P[\tilde{\mathbf{x}}|\tilde{\Lambda}]$  are the probabilities of trajectories during the forward and reverse protocol respectively, and JS is the Jensen-Shannon divergence between two probability distributions [12, 13, 14, 15, 16]

$$\text{JS}(p; q) = \frac{1}{2} \sum_i p_i \ln \frac{p_i}{\frac{1}{2}(p_i + q_i)} + \frac{1}{2} \sum_i q_i \ln \frac{q_i}{\frac{1}{2}(p_i + q_i)}. \quad (2)$$

Each of the two summands is the relative entropy (or Kullback-Leibler divergence) between one of the distributions and the mean of the two distributions. Hence,  $\text{JS}(p; q) \geq 0$ , and is equal to zero only if the two distributions are identical:  $p_i = q_i$  for all  $i$  [17]. The Jensen-Shannon divergence reaches its maximum value of  $\ln 2$  nats [i.e. 1 bit], if the two distributions do not overlap,  $p_i q_i = 0$  for all  $i$ , and therefore are perfectly distinguishable.

The Jensen-Shannon divergence has a direct interpretation in terms of a Bayesian inference problem [14]. Suppose we are given a sample  $k$  taken from one of two probability distributions,  $p$  or  $q$ . With no other way to distinguish between the distributions, the prior probability for the distributions is  $P(s) = \{\frac{1}{2}, \frac{1}{2}\}$  in which  $s$  represents either  $p$  or  $q$ . The prior distribution of outcome  $k$  is therefore  $P(k) = \frac{1}{2}p_k + \frac{1}{2}q_k$  while the posterior distribution is

$$P(s|k) = \left\{ \frac{\frac{1}{2}p_k}{\frac{1}{2}p_k + \frac{1}{2}q_k}, \frac{\frac{1}{2}q_k}{\frac{1}{2}p_k + \frac{1}{2}q_k} \right\}. \quad (3)$$

The information gained about  $s$  from observing state  $k$  is the relative entropy between the posterior and prior

distribution. This information averaged over the prior distribution of outcomes is

$$\langle \Delta I \rangle = \sum_k P(k) \Delta I = \sum_k P(k) \sum_s P(s|k) \ln \frac{P(s|k)}{P(s)}. \quad (4)$$

Some elementary algebra reveals that this average information gain is equal to the Jensen-Shannon divergence,  $\langle \Delta I \rangle = \text{JS}(p; q)$ . Hence, the time asymmetry  $A[\Lambda]$  is the average gain in information about the orientation of time's arrow obtained from one realization of the experiment. Moreover, the square root of the Jensen-Shannon divergence is a metric between probability distributions [14, 15]. Consequently, the square root of the time asymmetry  $\sqrt{A[\Lambda]}$  measures the distance between the forward and reverse protocols in trajectory space, literally the length of time's arrow.

In addition to its information theoretic interpretation, the time asymmetry can be measured in experiments due to recent advances in far-from-equilibrium statistical physics. In particular, the ratio of the probability of a trajectory during the forward protocol  $P[\mathbf{x}|\Lambda]$  and the probability of its conjugate trajectory on the reverse protocol  $P[\tilde{\mathbf{x}}|\tilde{\Lambda}]$  is [18]

$$\frac{P[\mathbf{x}|\Lambda]}{P[\tilde{\mathbf{x}}|\tilde{\Lambda}]} = e^{\beta W[\mathbf{x}|\Lambda] - \beta \Delta F[\Lambda]}, \quad (5)$$

in which  $\beta = 1/k_B T$ ,  $T$  is the temperature of the environment in natural units ( $k_B$  is the Boltzmann constant), and  $W[\mathbf{x}|\Lambda]$  is the work performed on the system during the forward protocol  $\Lambda$  [9, 18, 19], and  $\Delta F[\Lambda] = F_{\lambda(b)} - F_{\lambda(a)}$  is the difference in Helmholtz free energy between the initial and final ensembles. Eq. (5) is a direct consequence of the time reversal symmetry of the underlying dynamics [18] and implies the work fluctuation theorem

$$\frac{P(+W|\Lambda)}{P(-W|\Lambda)} = e^{\beta W - \beta \Delta F} \quad (6)$$

in which  $\Delta F \equiv \Delta F[\Lambda]$ . Moreover, Eq. (5) gives that the time asymmetry is

$$A[\Lambda] = \frac{1}{2} \left\langle \ln \frac{2}{1 + \exp(-\beta W[\mathbf{x}|\Lambda] + \beta \Delta F)} \right\rangle_{\Lambda} + \frac{1}{2} \left\langle \ln \frac{2}{1 + \exp(-\beta W[\tilde{\mathbf{x}}|\tilde{\Lambda}] - \beta \Delta F)} \right\rangle_{\tilde{\Lambda}}, \quad (7)$$

a non-linear average of the forward and reverse dissipation.

Time asymmetry is also closely related to the efficiency with which the free energy can be estimated from non-equilibrium measurements of the work. To determine  $\Delta F$  from experimental realizations of the forward and reverse protocols, Bennett's method gives the log likelihood of

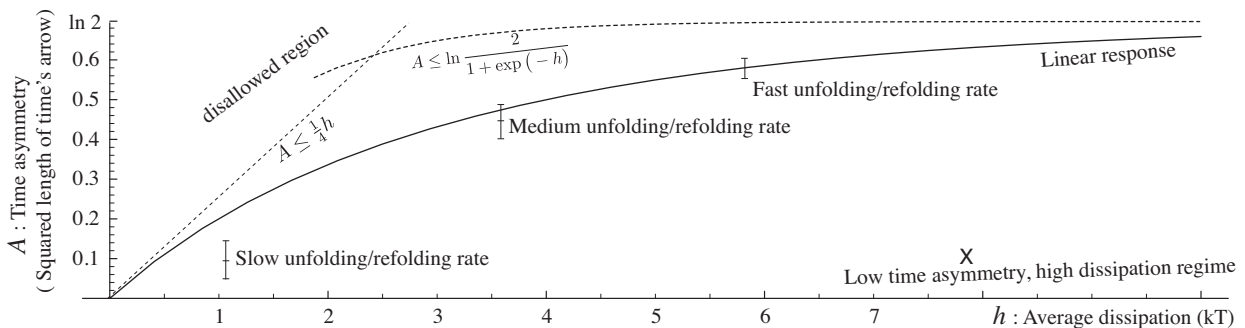


FIG. 2: The squared length of times arrow  $A$  [Eq. (1)] versus the hysteresis  $h$  [Eq. (10)], the dissipation averaged across a conjugate pair of forward and reversed experiments. The experiment is explained in Fig. 1. We equalize the number of data points between conjugate experiments, estimate the free energy from the data, obtain error bars by applying a Bayesian bootstrap [10], and apply a correction for experimental errors, as described in [11]. The slower the experiment is performed, the closer to thermodynamic reversibility, the lower the dissipation, and the lower the time-asymmetry. The slowest experiments are known to contain the greatest experimental error [11], which may explain the deviation of the slowest data from the linear response trend.

the free energy difference as [11, 20, 21, 22]

$$\begin{aligned} \ell(\Delta F) &= \sum_{i=1}^K \ln \frac{1}{1 + e^{-\beta W[x_i|\Lambda] + \beta \Delta F}} \\ &+ \sum_{j=1}^K \ln \frac{1}{1 + e^{-\beta W[\tilde{x}_j|\tilde{\Lambda}] - \beta \Delta F}}. \end{aligned} \quad (8)$$

Here,  $W[x_i|\Lambda]$  and  $W[\tilde{x}_j|\tilde{\Lambda}]$  are the work measured during the  $i$  and  $j$  realizations of the forward and reverse protocols respectively. The maximum likelihood estimate  $\Delta \hat{F} = \text{argmax} \ell(\Delta F)$  has the minimum variance among all estimators of  $\Delta F$  [11, 21, 22]. Comparing Eqs. (7) and (8), the time asymmetry can be estimated with the Bennett likelihood in the large sample limit, [23, 24]

$$A[\Lambda] \approx \frac{1}{2K} \ell(\Delta F) + \ln 2. \quad (9)$$

Thus, we can simultaneously estimate  $\Delta F$  and the time asymmetry by maximizing  $A[\Lambda]$  with respect to  $\Delta F$ .

It is enlightening to contrast the length of time's arrow with the hysteresis, the average dissipation of the forward and reverse protocols,

$$h[\Lambda] = \frac{1}{2} \beta \langle W[x|\Lambda] \rangle_{\Lambda} + \frac{1}{2} \beta \langle W[\tilde{x}|\tilde{\Lambda}] \rangle_{\tilde{\Lambda}}. \quad (10)$$

Because of Eq. (5), the hysteresis is also a divergence between the forward and reverse trajectory distributions [25, 26, 27, 28, 29]

$$h[\Lambda] = \frac{1}{2} \text{Jeffreys} \left( P[x|\Lambda]; P[\tilde{x}|\tilde{\Lambda}] \right) \quad (11)$$

in which

$$\text{Jeffreys}(p; q) = \sum_i p_i \log \frac{p_i}{q_i} + \sum_i q_i \log \frac{q_i}{p_i} \quad (12)$$

is the Jeffreys J-divergence (or symmetrized Kullback-Leibler divergence) [30, 31]. In Fig. 2, we plot time asymmetry against the hysteresis for single molecule RNA pulling experiments at three different rates (see Fig. 1). Both the hysteresis and length of time's arrow increases as the pulling rate increases; however, we will show that this need not always be the case. For comparison, we also display the time asymmetry  $A[\Lambda]$  in the linear response regime. The work distributions are normal with variance twice the average dissipation [9].

The relative values of time asymmetry and hysteresis are bounded by several inequalities. Taneja demonstrated that  $A \leq h/4$  using convexity arguments [32]. Figure (2) shows that this bound is obeyed by the linear response calculation and the experimental data. For large values of the hysteresis, we can derive a tighter bound than Taneja. Since the function  $f(x) = \ln(1 + e^{-x})$  is convex, Jensen's inequality [17] implies that  $\langle \ln(1 + e^{-x}) \rangle \geq \ln(1 + e^{-\langle x \rangle})$ . Thus,

$$\begin{aligned} \text{JS}(p; q) &= \frac{1}{2} \sum_i p_i \ln \frac{2}{1 + e^{\ln \frac{q_i}{p_i}}} + \frac{1}{2} \sum_i q_i \ln \frac{2}{1 + e^{\ln \frac{p_i}{q_i}}} \\ &\leq \frac{1}{2} \ln \frac{2}{1 + \exp(-D(p||q))} + \frac{1}{2} \ln \frac{2}{1 + \exp(-D(q||p))} \\ &\leq \ln \frac{2}{1 + \exp(-\frac{1}{2} \text{Jeffreys}(p; q))} \end{aligned} \quad (13)$$

in which last line follows by a second application of Jensen's inequality. Hence,  $A \leq \ln[2/(1 + e^{-h})]$  which we show in Fig. 2 for large  $h$ .

However, there is no lower bound to the time asymmetry given the hysteresis. A system can be almost time symmetric, but exhibit a large average dissipation. To illustrate this situation, imagine that occasionally, while gently unfolding an RNA hairpin, the RNA becomes stuck in a tangled configuration that resists being pulled apart by force. While most repetitions of the experiment

give a work measurement very close to the free energy change  $W \approx \Delta F$ , for the rare instances of entanglement the work is very large  $W \gg \Delta F$ . The normalized distribution of dissipation  $\mathcal{D} = W - \Delta F$  for the forward process may well be approximated as

$$P(\mathcal{D}|\Lambda) = (1-p)N(\bar{\mathcal{D}}_0, \sigma_0^2) + pN(\bar{\mathcal{D}}_1, \sigma_1^2) \quad (14)$$

in which  $N(\mu, \sigma^2)$  denotes a normal distribution with mean  $\mu$  and variance  $\sigma^2$ . The second term corresponds to the rare events, so  $p \ll 1$  and  $\bar{\mathcal{D}}_1 \gg 1$ . The work fluctuation theorem, Eq. (6) implies that the dissipation of the reverse process is

$$P(\mathcal{D}|\tilde{\Lambda}) = (1-q)N(\sigma_0^2 - \bar{\mathcal{D}}_0, \sigma_0^2) + qN(\sigma_1^2 - \bar{\mathcal{D}}_1, \sigma_1^2) \quad (15)$$

in which  $q = p \exp(-\bar{\mathcal{D}}_1 + \sigma_1^2/2)$  and normalization requires that

$$(1-p) \exp(-\bar{\mathcal{D}}_0 + \sigma_0^2/2) + p \exp(-\bar{\mathcal{D}}_1 + \sigma_1^2/2) = 1. \quad (16)$$

The linear response regime corresponds to  $p = 0$  with  $\bar{\mathcal{D}}_0 = \sigma_0^2/2$ . For small  $p$  and relatively small variance, we almost never see rare trajectories on the reverse protocol with a negative dissipation  $-\bar{\mathcal{D}}_1$  since  $q$  will be exponentially smaller than  $p \ll 1$ . The values  $p = 0.1$ ,  $\bar{\mathcal{D}}_1 = 150$ ,  $\bar{\mathcal{D}}_0 = 0.2$  and  $\sigma_0^2 = 1$  give  $h = 7.8$  and  $A = 0.1$  (Marked in Fig 2), so a system with small time asymmetry can have a large hysteresis. While Eq. (5) shows that dissipation measures the time symmetry breaking of individual trajectories [25, 29, 33], the average dissipation is sensitive to unusual events and is not a reliable measure of time asymmetry for the entire system.

One interpretation of relative entropy, and therefore Jeffreys divergence, is that it represents an encoding cost [17]. If we encode messages using an optimal code for the message probability distribution  $q_i$ , but the messages actually arrive with probabilities  $p_i$ , then each message, on average, will require an additional  $D(p||q)$  bits to encode compared to the optimal encoding. Analogously, the hysteresis represents a cost, the entropy lost to dissipation. Thus, the time asymmetry  $A$  measures the extent of time-symmetry breaking and the average dissipation measures the price paid.

The interrelation between time asymmetry and dissipation may be important for molecular motors and other macromolecular biological machinery. One of the central imperatives of any life form is to make tomorrow look different from today. On the molecular level, this requires rectifying the ever present thermal fluctuations. Since the horizontal-axis in Fig. 2 represents dissipation in units of  $k_B T$ , it takes about 4-8  $k_B T$  of free energy per cycle to ensure that a machine mostly advances forward in time, assuming it stays in the linear regime. This is a substantial fraction of the energy budget available from the hydrolysis of an ATP molecule, about 20  $k_B T$ .

This research was supported by the U.S. Dept. of Energy, under contracts DE-AC02-05CH11231. E.H.F.

thanks the Miller Institute for Basic Research in Science for financial support. We thank Felix Ritort for providing the experimental data used in our analysis.

- 
- [1] R. Clausius, *Annalen der Physik und Chemie* **201**, 353 (1865).
  - [2] A. S. Eddington, *The nature of the physical world* (Cambridge University Press, Cambridge, 1928).
  - [3] H. Price, *Time's arrow and Archimedes' point* (Oxford University Press, New York, 1996).
  - [4] D. Collin, F. Ritort, C. Jarzynski, S. B. Smith, I. Tinoco Jr., and C. Bustamante, *Nature* **437**, 231 (2005).
  - [5] J. Liphardt, S. Dumont, S. B. Smith, I. Tinoco Jr., and C. Bustamante, *Science* **296**, 1832 (2002).
  - [6] C. Bustamante, J. Liphardt, and F. Ritort, *Phys. Today* **58**, 43 (2005).
  - [7] F. Ritort, *J. Phys.: Condens. Matter* **18**, R531 (2006).
  - [8] D. J. Evans, E. G. D. Cohen, and G. P. Morriss, *Phys. Rev. Lett.* **71**, 3616 (1993), Errata 71(21):2401.
  - [9] C. Jarzynski, *Phys. Rev. Lett.* **78**, 2690 (1997).
  - [10] D. B. Rubin, *Ann. Stat.* **9**, 130 (1981).
  - [11] P. Maragakis, F. Ritort, M. Karplus, C. Bustamante, and G. E. Crooks, *J. Chem. Phys.* **129**, 024102 (2008).
  - [12] J. Lin, *IEEE Trans. Info. Theory* **37**, 145 (1991).
  - [13] F. Topsøe, *IEEE Trans. Info. Theory* **46**, 1602 (2000).
  - [14] D. M. Endres and J. E. Schindelin, *IEEE Trans. Info. Theory* **49**, 1858 (2003).
  - [15] F. Österreicher and I. Vajda, *Ann. Inst. Statist. Math.* **55**, 639 (2003).
  - [16] A. Majtey, P. W. Lamberti, M. T. Martin, and A. Plastino, *Eur. Phys. J. D* **32**, 413 (2005).
  - [17] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 1991).
  - [18] G. E. Crooks, *J. Stat. Phys.* **90**, 1481 (1998).
  - [19] L. Peliti, *J. Stat. Mech.: Theor. Exp.* p. P05002 (2008).
  - [20] C. H. Bennett, *J. Comput. Phys.* **22**, 245 (1976).
  - [21] M. R. Shirts, E. Bair, G. Hooker, and V. S. Pande, *Phys. Rev. Lett.* **91**, 140601 (2003).
  - [22] P. Maragakis, M. Spichty, and M. Karplus, *Phys. Rev. Lett.* **96**, 100602 (2006).
  - [23] G. E. Crooks, *Phys. Rev. Lett.* **99**, 100602 (4) (2007).
  - [24] E. Feng and G. E. Crooks (2008), arXiv:0807.0621.
  - [25] P. Gaspard, *J. Chem. Phys.* **120**, 8898 (2004).
  - [26] C. Jarzynski, *Phys. Rev. E* **73**, 046105 (2006).
  - [27] R. Kawai, J. M. R. Parrondo, and C. Van den Broeck, *Phys. Rev. Lett.* **98**, 080602 (2007).
  - [28] D. Andrieux, P. Gaspard, S. Ciliberto, N. Garnier, S. Joubaud, and A. Petrosyan, *J. Stat. Mech.: Theor. Exp.* p. P01002 (2008).
  - [29] A. Gomez-Marin, J. M. R. Parrondo, and C. Van den Broeck (2008), arXiv:0805.4703v1.
  - [30] H. Jeffreys, *Theory of probability* (Clarendon Press, 1948), 2nd ed.
  - [31] S. Kullback and R. A. Leibler, *Ann. Math. Stat.* **22**, 79 (1951).
  - [32] I. J. Taneja, *Aust. J. Math. Anal. Appl.* **2**, 8 (2005).
  - [33] C. Maes and K. Netočný, *J. Stat. Phys.* **110**, 269 (2003).