IOP Publishing — Journal of Physics A: Mathematical and Theoretical

J. Phys. A: Math. Theor. **53** (2020) 125001 (11pp) https://doi.org/10.1088/1751-8121/ab73ae

Work fluctuation, entropy, and time's arrow in time-asymmetric engine cycles

Euijin Jeon[1] **and Juyeon Yi**[2,3]

[1] Graduate School of Nanoscience and Technology, Korea Advanced Institute of Science and Technology, Daejeon 34141, Republic of Korea
[2] Department of Physics, Pusan National University, Busan 46241, Republic of Korea

E-mail: jyi@pusan.ac.kr

Received 1 January 2020
Accepted for publication 5 February 2020
Published 3 March 2020

Abstract
We derive a thermodynamic uncertainty relation that governs the work yield and the entropy production of an engine operating with a time-asymmetric cycle such as the Carnot cycle. The relation shows an intercorrelation between one engine cycle and its time-reversed cycle, disclosing the role of the time-asymmetry of an engine cycle. It also illustrates that entropy production and the arrow of time appear as the central quantities controlling the average and the fluctuation of work.

Keywords: time-asymmetric engine cycles, thermodynamic uncertainty relation, entropy production, fluctuation of work

(Some figures may appear in colour only in the online journal)

1. Introduction

The efficiency of a macroscopic heat engine acting between two heat baths at different temperatures, T_H and T_C ($<T_H$), is limited by the second law of thermodynamics and has a fundamental upper bound called the Carnot efficiency $1 - T_C/T_H$. In a mesoscopic scale where fluctuations are considered, recent studies provide an elaborated proof of the fluctuation theorem of the following form:

$$\mathrm{e}^{-\Sigma} P_F(W, \Sigma) = P_R(-W, -\Sigma), \tag{1}$$

which is obtained for classical engines [1] and for quantum mechanical engines [2]. Here W is work performed by the working substance of an engine during one engine cycle. Work is defined by energy change of the working substance caused by time-dependent control of an external parameter such as volume in the Carnot cycle. See [3] for the definition of work

[3] Author to whom any correspondence should be addressed.

in more detail. The entropy production Σ is defined by $\Sigma \equiv \Delta E/T_C + Q_C/T_C + Q_H/T_H$ ($k_B \equiv 1$) where ΔE represents the difference in internal energy of the working substance at the beginning and end of the engine cycle, and $Q_{H,C}$ is heat absorbed by the heat bath at temperature $T_{H,C}$. The energy conservation law reads as $\Delta E = -W - Q_H - Q_C$ when energy cost to bring the working substance into thermal contact with heat baths is negligible. The engine is assumed to be cyclic in the sense that the probability distribution of the microstate returns to its initial distribution at the end of the cycle, so that the average value of ΔE is equal to zero.

Equation (1) relates two probability distribution functions: one is $P_F(W, \Sigma)$, the joint probability of observing W and Σ in an engine cycle (say, cycle F). The other joint probability, $P_R(W, \Sigma)$, is defined for the *reversed* operation of engine cycle F. To clarify the meaning of the term 'reversed', we let $\Lambda(t)$ to denote the external parameter at time t. Also, let $\Gamma_H(t)$ and $\Gamma_C(t)$ represent the coupling strengths of the system to a hot and to a cold bath, respectively. If engine cycle F runs in time interval, $0 \leqslant t \leqslant T$, in the reverse cycle R not only an external parameter control but also coupling to heat baths run backward in time, that is, $\Lambda(T-t)$, $\Gamma_{H,C}(T-t)$.

Integrating the relation (1) over Σ and W yields

$$\langle \mathrm{e}^{-\Sigma} \rangle_F \equiv \int \mathrm{d}W\mathrm{d}\Sigma P_F(W, \Sigma)\mathrm{e}^{-\Sigma} = 1. \tag{2}$$

From now on we use $\langle X \rangle_\alpha$ to denote the average of X, which is a function of W and Σ, taken with respect to $P_\alpha(W, \Sigma)$ with $\alpha = F, R$. The Jensen inequality $\mathrm{e}^{-\langle \Sigma \rangle_F} \leqslant \langle \mathrm{e}^{-\Sigma} \rangle_F$ together with equation (2) leads to $\langle \Sigma \rangle_F \geqslant 0$. Because $\langle \Delta E \rangle_F = 0$ for cyclic engine processes, $\langle \Sigma \rangle_F$ amounts to the average entropy produced in one engine cycle, and therefore, $\langle \Sigma \rangle_F \geqslant 0$ proves the second law of thermodynamics for a mesoscopic heat engine. On the other hand, using energy conservation, $\Delta E = -W - Q_H - Q_C$ the inequality $\langle \Sigma \rangle_F \geqslant 0$ for a heat engine ($\langle Q_H \rangle_F < 0$ and $\langle W \rangle_F > 0$) transforms into

$$\frac{\langle W \rangle_F}{|\langle Q_H \rangle_F|} \leqslant 1 - \frac{T_C}{T_H}. \tag{3}$$

This relation shows that the efficiency of a mesoscopic heat engine cannot exceed the Carnot efficiency.

In addition to the average work, work fluctuation is relevant, in particular, to quantify how reliably a heat engine operates [4]. For a time symmetric control of an external parameter, the thermodynamic uncertainty relation shows the relationship between the relative fluctuation of work and the average entropy production as follows [5, 6]:

$$(\mathrm{e}^{\langle \Sigma \rangle} - 1)\mathrm{var}(W) \geqslant 2\langle W \rangle^2, \tag{4}$$

where $\langle W \rangle$ and $\mathrm{var}(W)$ denote the average and the mean square fluctuation of W defined accordingly in the studies [5, 6]. We emphasize here that in the derivation of equation (4), the coupling trajectory $\Gamma_{H,C}(t)$ is also time-symmetric, although not explicitly mentioned therein. One might presume that equation (4), perhaps with a minor modification to it, would be still valid even if an engine cycle is not symmetric in time. However, the Carnot engine disproves equation (4); in a finite-sized engine that operators according to the Carnot cycle, the average and the fluctuation of work are finite, whereas the average entropy production vanishes [7]. Therefore, the example of the Carnot cycle presents two points: first, the Carnot cycle is not time-symmetric, and second, for engines operating in time-asymmetric cycles, a different type of relationship from the equation must exist.

In this study, we prove that the work yield in engine cycle F intercorrelates with the work yield in the reverse cycle R through the relation,

$$\mathcal{B}_\Sigma[\mathrm{var}(W)_F + \mathrm{var}(W)_R] \geqslant (\langle W\rangle_F + \langle W\rangle_R)^2, \tag{5}$$

where $\mathrm{var}(W)_{\alpha=F,R}$ is the mean square fluctuation of W with respect to $P_\alpha(W,\Sigma)$, and $\mathcal{B}_\Sigma$ can be *any of following three functions of* $\langle\Sigma\rangle_s$, $\langle\Sigma\rangle_F$, *and* A,

$$\mathcal{B}_1 = \mathrm{e}^{\langle\Sigma\rangle_s} - 1 \tag{6}$$

$$\mathcal{B}_2 = 2(\mathrm{e}^{\langle\Sigma\rangle_F} - 1) \tag{7}$$

$$\mathcal{B}_3 = \frac{2}{2-\mathrm{e}^A}\left[\mathrm{e}^A - 1 - \log(2-\mathrm{e}^A)\right], \tag{8}$$

where $\mathcal{B}_2$ can be replaced with $2(\mathrm{e}^{\langle\Sigma\rangle_R} - 1)$, and we define the arithmetic mean, $\langle\Sigma\rangle_s = (\langle\Sigma\rangle_F + \langle\Sigma\rangle_R)/2$, of the entropy production in cycle F and R. We also prove that $\mathcal{B}_3$ is always greater than the other two functions. We can trivially obtain $\mathcal{B}_1$ from equation (4): in both equations (4) and (5) with $\mathcal{B}_\Sigma = \mathcal{B}_1$, every quantity is symmetrized, and the distinction between cycle F and cycle R becomes pointless. But the other two functions, $\mathcal{B}_2$ and $\mathcal{B}_3$, necessitate separate consideration. In equation (7), A is a quantity between 0 and ln 2, defined by the Jensen–Shannon divergence between two joint probabilities $P_F(W,\Sigma)$ and $P_R(-W,-\Sigma)$. As will be discussed in section 4, the quantity A represents the length of time's arrow and measures the distinguishability between cycle F realized in the forward time direction and cycle R realized in the backward time direction.

We organize this paper as follows: in section 2, we present a detailed derivation of equation (5). In section 3, we discuss what equation (5) conveys in regards to the time-symmetry of the engine cycle, and in section 4, the physical meaning of equation (5) is explained in terms of the arrow of time. Summary follows as section 5.

2. Proof of equation (5)

Derivation of equation (5) consists of three steps: we first introduce a trial probability distribution $P_\lambda(W,\Sigma)$, parameterized by λ. Second, we obtain the bound of the variance of W from the Cauchy–Schwarz inequality,

$$\mathrm{var}(W)_{\lambda=W_s}\mathrm{var}(Y)_{\lambda=W_s} \geqslant |\mathrm{cov}(W,Y)|^2_{\lambda=W_s} \tag{9}$$

with Y taken as

$$Y = d\ln P_\lambda(W,\Sigma)/\mathrm{d}\lambda,$$

where $\mathrm{var}(X)_\lambda$ is the variance of X with respect to $P_\lambda(W,\Sigma)$, and $2W_s = \langle W\rangle_F + \langle W\rangle_R$. Third, we find upper bound of $\mathrm{var}(Y)_{\lambda=W_s}$ and the functions in equations (6)–(8). In addition to this, the inequalities between the three upper bounds are derived in the last subsection.

2.1. Trial probability distribution

We introduce the trial probability distribution as

$$P_\lambda(W,\Sigma) \equiv P_s(W,\Sigma)\left[\frac{\lambda}{W_s} + \frac{W_s-\lambda}{W_s}Nf(\Sigma)\right] \tag{10}$$

with $\lambda \leqslant W_s$ for the positivity of $P_\lambda(W,\Sigma)$. Here the function $f(\Sigma)$ comes into play when λ deviates from W_s, it is defined as

$$f(\Sigma) \equiv 2/(1+\mathrm{e}^{\Sigma}), \tag{11}$$

and N is introduced to normalize $P_\lambda(W,\Sigma)$ as

$$N^{-1} = \langle f(\Sigma)\rangle_s. \tag{12}$$

In equation (10), $P_s(W,\Sigma) = [P_F(W,\Sigma) + P_R(W,\Sigma)]/2$ is the symmetrized probability that satisfies,

$$P_s(W,\Sigma) = \mathrm{e}^{\Sigma} P_s(-W,-\Sigma). \tag{13}$$

This form is equivalent to the fluctuation theorem (1) when the cycle F is indistinguishable from its reversed cycle R. Note also that $P_s(W,\Sigma)$ can be seen as the joint probability of acquiring W and Σ in a time-symmetric engine cycle C, by defining cycle C as a probabilistic mixture of cycle F and cycle R. There is no distinction between the forward and reverse operation of cycle C since the engine runs randomly in cycle F or cycle R with probability 1/2 in both operations. In what follows, for ease of writing, we use simplified notations for the averages

$$\int_{-\infty}^{\infty} \mathrm{d}W\mathrm{d}\Sigma X(W,\Sigma) P_\alpha(W,\Sigma) = \langle X\rangle_\alpha, \ \alpha = s, \lambda,$$

in the same manner defining $\langle X\rangle_{F,R}$. At $\lambda = W_s$, the trial probability distribution is identical to the symmetrized distribution:

$$P_{\lambda=W_s}(W,\Sigma) = P_s(W,\Sigma). \tag{14}$$

We mention, in order, two properties of $P_\lambda(W,\Sigma)$ important in our derivation. First, the average of W with respect to $P_\lambda(W,\Sigma)$ is λ:

$$\langle W\rangle_\lambda = \lambda + \frac{W_s - \lambda}{W_s} N\langle Wf(\Sigma)\rangle_s = \lambda. \tag{15}$$

Here the average $\langle Wf(\Sigma)\rangle_s$ vanishes, for $f(\Sigma)P_s(W,\Sigma) = f(-\Sigma)P_s(-W,-\Sigma)$, which can be proven by the use of equations (13) and (11). Second, the Fisher information of $P_\lambda(W,\Sigma)$, defined as

$$I(\lambda) = \left\langle \left[\frac{d\ln P_\lambda(W,\Sigma)}{\mathrm{d}\lambda}\right]^2 \right\rangle_\lambda, \tag{16}$$

becomes, when $\lambda = W_s$,

$$I(W_s) = (N-1)/W_s^2. \tag{17}$$

We can show equation (17), inserting equation (10) into equation (16) and setting $\lambda = W_s$:

$$I(W_s) = -\frac{1}{W_s^2} + \frac{N^2}{W_s^2}\langle f^2(\Sigma)\rangle_s. \tag{18}$$

Further exploiting the identity,

$$\begin{aligned}\langle f^2(\Sigma)\rangle_s &= 2\int \mathrm{d}W\mathrm{d}\Sigma \left[\frac{P_s(W,\Sigma)}{(1+\mathrm{e}^{\Sigma})^2} + \frac{P_s(-W,-\Sigma)}{(1+\mathrm{e}^{-\Sigma})^2}\right] \\ &= \langle f(\Sigma)\rangle_s = N^{-1},\end{aligned}$$

where equations (13) and (12) are used, we reach equation (17).

2.2. Cauchy–Schwarz inequality

Now consider the Cauchy–Schwarz inequality (9). Note that the average of Y defined below equation (9) vanishes because

$$\left\langle \frac{d \ln P_\lambda(W,\Sigma)}{\mathrm{d}\lambda} \right\rangle_\lambda = \frac{d}{\mathrm{d}\lambda} \int \mathrm{d}W\mathrm{d}\Sigma P_\lambda(W,\Sigma) = 0 \tag{19}$$

for the normalized distribution P_λ. Therefore, the covariance between W and Y is given by

$$\begin{aligned} \mathrm{cov}(W,Y) &= \left\langle W \frac{d \ln P_\lambda(W,\Sigma)}{\mathrm{d}\lambda} \right\rangle_\lambda \\ &= \frac{d}{\mathrm{d}\lambda}\langle W\rangle_\lambda = 1. \end{aligned} \tag{20}$$

The last equality is obtained, using equation (15). Also because $\langle Y\rangle_\lambda = 0$, the variance of Y is identical to the Fisher information (16), namely,

$$\mathrm{var}(Y)_{\lambda=W_s} = I(W_s) = (N-1)/W_s^2. \tag{21}$$

Inserting equations (20) and (21) into equation (9), we obtain

$$\mathrm{var}(W)_{\lambda=W_s} \geqslant W_s^2/(N-1). \tag{22}$$

Finally using $2\mathrm{var}(W)_{\lambda=W_s} = \mathrm{var}(W)_F + \mathrm{var}(W)_R$ and $2W_s = \langle W\rangle_F + \langle W\rangle_R$, we arrive at

$$\mathrm{var}(W)_F + \mathrm{var}(W)_R \geqslant \frac{(\langle W\rangle_F + \langle W\rangle_R)^2}{2(N-1)}, \tag{23}$$

which indicates that the fluctuation of work is bounded from below by the sum, $\langle W\rangle_F + \langle W\rangle_R$, and also by the factor N. The remaining problem is to find N and, in particular, its relation with entropy production $\langle\Sigma\rangle_F$ and the Jensen–Shannon divergence A.

2.3. Bound of N

We now attain a few inequalities that N satisfies and obtain $\mathcal{B}_2$ and $\mathcal{B}_3$. Recall that equation (12) determines N as

$$N^{-1} = \frac{1}{2}\langle f(\Sigma)\rangle_F + \frac{1}{2}\langle f(\Sigma)\rangle_R.$$

For nonnegative $\langle f(\Sigma)\rangle_{F,R}$, using $\langle f(\Sigma)\rangle_F = \langle f(\Sigma)\rangle_R$ because of the fluctuation relation (1), we can rewrite equation (12) as

$$N^{-1} = \frac{2-q}{2}\langle f(\Sigma)\rangle_F + \frac{q}{2}\langle f(\Sigma)\rangle_R.. \tag{24}$$

Here q can any arbitrary constants, but the particular choices of q enable us to relate the upper bound of N with the average entropy production and the arrow of time, as will be shown below.

(i) Obtaining $\mathcal{B}_2$: for $q = 0$, using Jensen's inequality, $\ln\langle f(\Sigma)\rangle_F \geqslant \langle \ln f(\Sigma)\rangle_F$, we obtain, from equation (24),

$$-\ln N = \ln\langle f(\Sigma)\rangle_F \geqslant \langle \ln f(\Sigma)\rangle_F \geqslant -\langle\Sigma\rangle_F. \tag{25}$$

The last inequality comes from considering

$$\langle \ln f(\Sigma)\rangle_F = -\langle\Sigma\rangle_F + \langle \ln(\mathrm{e}^{\Sigma} f(\Sigma))\rangle_F \geqslant -\langle\Sigma\rangle_F. \tag{26}$$

Therefore, N satisfies $N \leqslant \mathrm{e}^{\langle\Sigma\rangle_F}$, which combined with equation (23) proves that

$$2(\mathrm{e}^{\langle\Sigma\rangle_F} - 1)[\mathrm{var}(W)_F + \mathrm{var}(W)_R] \geqslant (\langle W\rangle_F + \langle W\rangle_R)^2.$$

This relation shows that $\mathcal{B}_\Sigma$ in equation (5) is $\mathcal{B}_2 = 2(\mathrm{e}^{\langle\Sigma\rangle_F} - 1)$ as given in equation (7). Choosing $q = 2$ replaces $\langle\Sigma\rangle_F$ with $\langle\Sigma\rangle_R$ only.

(ii) Obtaining $\mathcal{B}_3$: from equation (24) for $q = 1$, we can reexpress N in terms of absolute value of entropy $|\Sigma|$ as

$$\begin{aligned} N^{-1} &= 2\langle f(\Sigma)\rangle_s - \langle f(\Sigma)\rangle_s = 2\langle f(\Sigma)\rangle_s - \langle f^2(\Sigma)\rangle_s \\ &= \langle f(\Sigma) f(-\Sigma)\rangle_s \\ &= \langle f_N(|\Sigma|)\rangle_s \end{aligned} \tag{27}$$

with the function $f_N(x) = f(x)f(-x)$. On the other hand, the Jensen–Shannon divergence A is defined by

$$A = \frac{1}{2} KL[P_F(\xi)||M(\xi)] + \frac{1}{2} KL[P_R(-\xi)||M(\xi)] \tag{28}$$

with $\xi = W, \Sigma$ and $2M(\xi) = 2M(W, \Sigma) = P_F(W, \Sigma) + P_R(-W, -\Sigma)$, and in terms of Σ, it is written as

$$A = \langle f_A(|\Sigma|)\rangle_s \tag{29}$$

with the function $f_A(x)$ defined by

$$f_A(x) = \frac{1}{2} f(x) \ln f(x) + \frac{1}{2} f(-x) \ln f(-x). \tag{30}$$

Since $f_A(x)$ is a monotonically increasing function for $x \geqslant 0$, we can define the inverse function f_A^{-1} satisfying $f_A^{-1} \circ f_A(x) = x$. Then the composite function $f_N \circ f_A^{-1}$ becomes a convex function (the graph of $f_N \circ f_A^{-1}$ is presented in figure 1(a)), and the Jensen's inequality gives

$$N^{-1} = \langle f_N \circ f_A^{-1}(f_A(|\Sigma|))\rangle_s \geqslant f_N \circ f_A^{-1}(A). \tag{31}$$

From this, we obtain another lower bound of N^{-1} less than $f_N \circ f_A^{-1}(A)$ but written in more manageable form

$$N^{-1} \geqslant \frac{2 - \mathrm{e}^A}{1 - \log(2 - \mathrm{e}^A)}. \tag{32}$$

The comparison between these two bounds is illustrated in figure 1(a). Combining this with equation (23), the inequality

$$\frac{2}{2 - \mathrm{e}^A}\left[\mathrm{e}^A - 1 - \log(2 - \mathrm{e}^A)\right][\mathrm{var}(W)_F + \mathrm{var}(W)_R] \geqslant (\langle W\rangle_F + \langle W\rangle_R)^2$$

is proven as desired.

2.4. Inequalities between $\mathcal{B}_1$, $\mathcal{B}_2$, and $\mathcal{B}_3$

Here, we prove that $\mathcal{B}_3$ is less than or equal to $\mathcal{B}_1$ and $\mathcal{B}_2$. To demonstrate the inequality between $\mathcal{B}_1$ and $\mathcal{B}_3$, we express the symmetrized entropy in terms of $|\Sigma|$

$$\begin{aligned}\langle\Sigma\rangle_s &= \frac{1}{2}\langle\Sigma f(\Sigma)\rangle_s + \frac{1}{2}\langle\Sigma f(-\Sigma)\rangle_s \\ &= \frac{1}{2}\langle\Sigma[f(-\Sigma) - f(\Sigma)]\rangle_s \\ &= \langle f_s(|\Sigma|)\rangle_s \end{aligned} \tag{33}$$

with a monotonic increasing function $f_s(x) = x[f(x) - f(-x)]/2$ of $x \geqslant 0$, where the second equality comes from equation (13). Considering a monotonic increasing functions $b_1(x)$ defined by

$$b_1(x) = \mathrm{e}^x - 1 \tag{34}$$

for $x > 0$ and $b_3(x)$ defined by

$$b_3(x) = \frac{2}{2 - \mathrm{e}^x}\left[\mathrm{e}^x - 1 - \ln(2 - \mathrm{e}^x)\right] \tag{35}$$

for $0 < x < \ln 2$, the Jensen–Shannon divergence A satisfies

$$A = \left\langle f_A \circ f_s^{-1}\left(f_s\left(|\Sigma|\right)\right)\right\rangle_s \leqslant f_A \circ f_s^{-1}(\langle\Sigma\rangle_s) \leqslant b_3^{-1} \circ b_1(\langle\Sigma\rangle_s). \tag{36}$$

Here, we use the Jensen's inequality with the fact that the composite function $f_A \circ f_s^{-1}(x)$ is concave and satisfies $f_A \circ f_s^{-1}(x) \leqslant b_3^{-1} \circ b_1(x)$ for all $x \geqslant 0$ (the illustration of the function $b_3^{-1} \circ b_1$ and $f_A \circ f_s^{-1}$ is presented in figure 1(b)). Since b_3 is a monotonic increasing function for $0 \leqslant x \leqslant \ln 2$, we finally get

$$\mathcal{B}_3 = b_3(A) \leqslant b_1(\langle\Sigma\rangle_s) = \mathcal{B}_1. \tag{37}$$

To derive the inequality between $\mathcal{B}_2$ and $\mathcal{B}_3$, we obtain an upper bound of A as:

$$\begin{aligned} A \leqslant \ln\langle \mathrm{e}^{f_A(\Sigma)}\rangle_s &= \ln\langle \mathrm{e}^{f_A(\Sigma)}(1 + \mathrm{e}^{-\Sigma})/2\rangle_F \\ &= \ln\langle \mathrm{e}^{f_A(\Sigma)}(1 + \mathrm{e}^{-\Sigma})/2 + (1 - \mathrm{e}^{-\Sigma})\rangle_F \\ &\equiv \ln\langle f_{\exp A}(\Sigma)\rangle_F \end{aligned} \tag{38}$$

where the fluctuation relation equation (1) is used in the first line and the second line comes from equation (2). The function $f_{\exp A}(x) = \mathrm{e}^{f_A(x)}(1 + \mathrm{e}^{-x})/2 + (1 - \mathrm{e}^{-x})$ is a monotonic increasing function less than 2, and $\langle f_{\exp A}(\Sigma)\rangle_F$ is always greater than or equal to 1 since $A \geqslant 0$. If we define monotonic increasing functions $f_{\ln}(x)$ and $b_2(x)$ as

$$f_{\ln}(x) = \begin{cases} \ln(x) & \text{if } 1 \leqslant x \\ x - 1 & \text{if } x < 1 \end{cases} \tag{39}$$

$$b_2(x) = \begin{cases} 2(e^x - 1) & \text{if } 0 \leqslant x \\ 2x & \text{if } x < 0 \end{cases} \tag{40}$$

and extend the domain of the function $b_3(x)$ to $x < \ln 2$ by setting $b_3(x) = 4x$ for $x < 0$, we get following from Jensen's inequality,

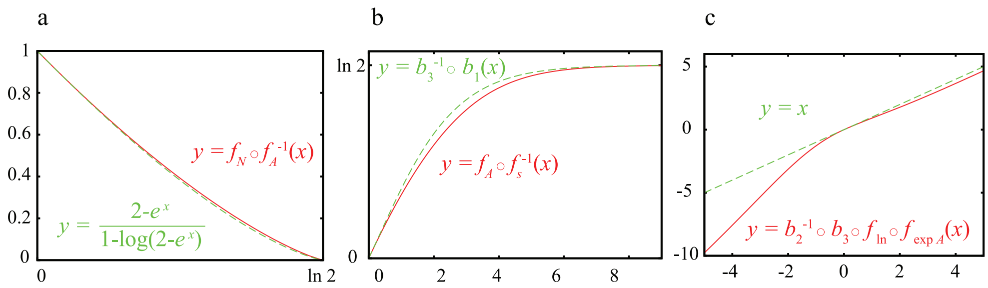

Figure 1. The illustrations of functions used in derivation. (a). The graphs of $y = f_N \circ f_A^{-1}(x)$ (red solid line) and $y = (2 - e^x)/(1 - \log(2 - e^x))$ (greed dotted line). The function $f_N \circ f_A^{-1}(x)$ is convex and greater than or equal to $(2 - e^x)/(1 - \log(2 - e^x))$ in the range of $x \in [0, \log 2]$. (b). The graphs of $y = f_A \circ f_s^{-1}(x)$ (red solid line) and $y = b_3^{-1} \circ b_1(x)$ (greed dotted line). The function $f_A \circ f_s^{-1}(x)$ is concave and less than or equal to $b_3^{-1} \circ b_1(x)$ for $x \geqslant 0$. (c). The graphs of $y = b_2^{-1} \circ b_3 \circ f_{\ln} \circ f_{\exp A}(x)$ (red solid line) and $y = x$ (greed dotted line). The function $b_2^{-1} \circ b_3 \circ f_{\ln} \circ f_{\exp A}(x)$ is less than or equal to x for all x.

$$\begin{aligned} b_2^{-1}(\mathcal{B}_3) &\leqslant b_2^{-1} \circ b_3 \circ f_{\ln}(\langle f_{\exp A}(\Sigma)\rangle_F) \\ &\leqslant \langle b_2^{-1} \circ b_3 \circ f_{\ln} \circ f_{\exp A}(\Sigma)\rangle_F, \end{aligned} \tag{41}$$

since the composite function $b_2^{-1} \circ b_3 \circ f_{\ln}(x)$ is convex for $x < \ln 2$. In figure 1(c), we can see that $b_2^{-1} \circ b_3 \circ f_{\ln} \circ f_{\exp A}(\Sigma)$ is always less than or equal to x, so equation (41) gives $b_2^{-1}(\mathcal{B}_3) \leqslant \langle\Sigma\rangle_F$. Since $b_2(x)$ is a monotonic increasing function, we finally get

$$\mathcal{B}_3 \leqslant b_2(\langle\Sigma\rangle_F) = \mathcal{B}_2. \tag{42}$$

3. Implications of equation (5)

Here, we discuss the implication of equation (5) with equations (6) and (7). When cycle F is identical with cycle R, the relation (5) with $\mathcal{B}_\Sigma = \mathcal{B}_1$ is reduced into equation (4) because $\langle W\rangle_F = \langle W\rangle_R$, $\mathrm{var}(W)_F = \mathrm{var}(W)_R$, and $\langle\Sigma\rangle_s = \langle\Sigma\rangle_F = \langle\Sigma\rangle_R$. An engine cycle having the identical reverse cycle is only realizable if the engine protocol satisfies time-symmetric property as $\Lambda(t) = \Lambda(T - t)$ and $\Gamma_{H,C}(t) = \Gamma_{H,C}(T - t)$. Suppose time-symmetric cycles that happen to produce positive work average, $\langle W\rangle_F > 0$, as exemplified in [8] and [9]. Such engine cycles are less effectual: on one hand, their reverse cycles cannot furnish us with a refrigerator because $P_F(W, \Sigma) = P_R(W, \Sigma)$ and hence, $\langle W\rangle_R$ is also positive. On the other hand, according to the relation, they are allowed to yield a finite work average without entropy production, only when paying the cost of infinite work fluctuations.

In sharp contrast, the Carnot cycle are time- *asymmetric* because of $\Gamma_{H,C}(T - t) \neq \Gamma_{H,C}(t)$ (see figure 2), and they act as a refrigerator when running the cycles backward in time. For this case, equation (5) with $\mathcal{B}_\Sigma = \mathcal{B}_1$ shows an intimate relation between the work yield of a heat engine and that of a refrigerator. Especially for the Carnot cycle, we have $\langle\Sigma\rangle_s = 0$ which leads $\langle W\rangle_F = \langle W\rangle_R$.

From equation (5) with $\mathcal{B}_\Sigma = \mathcal{B}_2$, we can see that the entropy production in one engine cycle (say $\langle\Sigma\rangle_F$) restricts not the average work $\langle W\rangle_F$ resulting from cycle F but rather the sum of $\langle W\rangle_F$ and $\langle W\rangle_R$. Provided that the work fluctuation is not infinitely large, the inequality

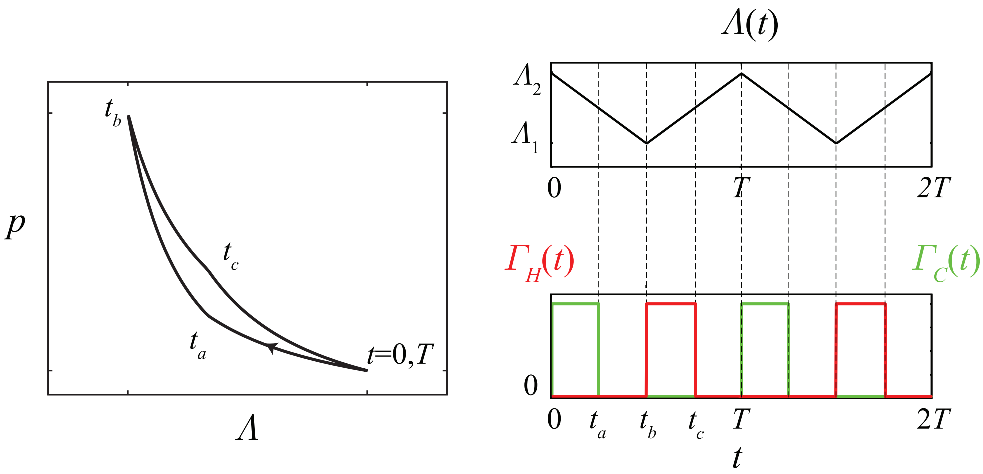

Figure 2. Schematics of the control parameter $\Lambda(t)$ and the thermal connectivity $\Gamma_{H,C}(t)$ during two periods of the Carnot cycle. Here one can regard $\Lambda(t)$ as the volume of gas as a working substance. The Carnot cycle consists of four processes: isothermal compression $(0 \leqslant t < t_a)$ at T_C, adiabatic compression $(t_a \leqslant t < t_b)$, isothermal expansion $(t_b \leqslant t < t_c)$ at T_H, and adiabatic expansion $(t_c \leqslant t \leqslant T)$. For simplicity, we present $\Lambda(t)$ as the linear ramp. Although each process is a quasi-static process and takes a long time, the trajectories of $\Gamma_{C,H}(t)$ are clearly not time-symmetric.

(5) with $\mathcal{B}_\Sigma = \mathcal{B}_2$ for $\langle\Sigma\rangle_F = 0$ indicates that $\langle W\rangle_F = -\langle W\rangle_R$. Because of $\langle W\rangle_F \neq \langle W\rangle_R$ for time-asymmetric engine cycles, $\langle W\rangle_F$ can be non-zero even when entropy production vanishes ($\langle\Sigma\rangle_F = 0$).

On the other hand, we write equation (5) with $\mathcal{B}_\Sigma = \mathcal{B}_2$ as

$$\mathcal{B}_2 \mathrm{var}(W)_F \geqslant (\langle W\rangle_F + \langle W\rangle_R)^2/(1+\eta) \tag{43}$$

with $\eta = \mathrm{var}(W)_R/\mathrm{var}(W)_F$ and may view the left-hand side of equation (43) as the *uncertainty* of engine cycle F, which is probably the reason that equations of such form are called the thermodynamic uncertainty relation. Notice here that $\mathcal{B}_2$ can be written as

$$\mathcal{B}_2 = 2\langle \mathrm{e}^{-(\Sigma-\langle\Sigma\rangle_F)} - 1\rangle_F$$

thanks to $\langle \mathrm{e}^{-\Sigma}\rangle_F = 1$, as mentioned in equation (2), and that it contains the fluctuation of Σ around its average $\langle\Sigma\rangle_F$. The relation (43) shows that the uncertainty of cycle F can have a minimum bound either when $\langle W\rangle_F + \langle W\rangle_R$ vanishes or when η goes to infinite faster than $\langle W\rangle_F + \langle W\rangle_R$. The former happens in the Carnot cycle. What kind of engine cycles accomplishes the latter condition remains as a question, while thermodynamic rules known so far do not exclude the possibility of such engine cycles.

4. Arrow of time

Although the external parameter control and the coupling strength of cycle F and the cycle R are the time-reversal of each other, the statistics of W and Σ obtained in cycle F is not identical to that measured in the reversed time direction in cycle R. This fundamental asymmetry is often called the thermodynamic arrow of time. The Jensen–Shannon divergence A, defined in equation (28), is often called the length of time's arrow because it is a standard measure of this asymmetry [10]. The value of A lies between 0 and $\ln 2$. When $P_F(W, \Sigma) = P_R(-W, -\Sigma)$, meaning that the statistics of W and Σ obtained in cycle F is indistinguishable from that of $-W$ and $-\Sigma$ obtained in the cycle R, A reaches its minimum value 0. On the other hand, A approaches $\ln 2$ when $P_F(W, \Sigma)$ has negligible overlap with $P_R(-W, -\Sigma)$, or in other words when the statistics of W and Σ obtained in running cycle F is perfectly distinguishable from that of $-W$ and $-\Sigma$ obtained in the reversed cycle R.

Equation (5) with $\mathcal{B}_3$ allows a lower bound of A to be determined by measuring work in the forward and reversed engine cycle. The value of A approaches to $\ln 2$ when $(\langle W\rangle_F + \langle W\rangle_R)^2$ is much greater than $\mathrm{var}(W)_F + \mathrm{var}(W)_R$, which is consistent with the fact that the two probabilities $P_F(W, \Sigma)$ and $P_R(-W, -\Sigma)$ are completely distinguishable in this limit. On the other hand, equation (37) shows that A is bounded from above by a monotonic increasing function of $\langle\Sigma\rangle_s$ or $\langle\Sigma\rangle_F$. Consequently, from equations (37), (42) and (5) with $\mathcal{B}_3$, we can determine the range of A as

$$b_3^{-1}\left(\frac{(\langle W\rangle_F + \langle W\rangle_R)^2}{\mathrm{var}(W)_F + \mathrm{var}(W)_R}\right) \leqslant A \leqslant b_3^{-1}(\mathcal{B}_{1\text{ or }2}) \tag{44}$$

with $\mathcal{B}_1$, $\mathcal{B}_2$, and b_3 defined in equations (6)–(7) and equation (35). In the case of Carnot cycle, $\langle\Sigma\rangle_s = \langle\Sigma\rangle_F = \mathcal{B}_1 = \mathcal{B}_2 = 0$, which leads $A = 0$.

Here, we discuss the physical meaning of the relation (5) with $\mathcal{B}_\Sigma = \mathcal{B}_1$ or $\mathcal{B}_2$ in terms of time's arrow. The ratio between $(\langle W\rangle_F + \langle W\rangle_R)^2$ and $\mathrm{var}_F(W) + \mathrm{var}_B(W)$ measures how the work measured in cycle F is distinguishable from that measured from the time-reversal of cycle R. If this ratio is finite, the length of time's arrow A is finite by relation (5) with $\mathcal{B}_\Sigma = \mathcal{B}_3$. Since the finite length of the time's arrow indicates that there is a dissipation in cycle F and R as equations (37) and (42) show, the entropy production and the symmetrized entropy production should also be positive.

To sum up, equation (5) shows that entropy production is always positive if cycle F and the time-reversal of cycle R can be distinguished by measuring work. In the case of time-symmetric engine, they can always be distinguished by the sign of the average work if the average work yield is non-negligible compared to its fluctuation. In contrast, in the Carnot engine, even if the harvested work is statistically significant, cycle F cannot be distinguished from the time-reversal of cycle R.

5. Summary

We have derived relation (5) that dictates work, entropy production, and the arrow of time, and a detailed proof have been presented based on the fluctuation theorem (1). Work yield in engine cycle F correlates with work yield in cycle R, the reverse of cycle F, and they enter into relation (5) only through the sums. We have discussed a few aspects that equation (5) explains. A time-symmetric engine cycle can yield a finite work average only when accompanied either by a finite entropy production or by infinite work fluctuations. For time-asymmetric engine cycles such as the Carnot cycle, such restriction disappears; they can deliver a finite work-average in the absence of entropy production. The concept of time's arrow provides a physical understanding of equation (5). The Carnot cycle is, indeed, a very special time-asymmetric engine cycle. It has the minimum length of the arrow of time quantified by equation (28) and the minimum of the uncertainty defined in equation (43).

Acknowledgments

This research was supported by a Basic Science Research Program through the National Rational Research Foundation of Korea(NRF) funded by the Ministry of Education (2017R1D1A1B03029903).

ORCID iDs

Juyeon Yi https://orcid.org/0000-0002-2637-791X

References

[1] Sinitsyn N 2011 *J. Phys. A: Math. Theor.* **44** 405001
[2] Campisi M 2014 *J. Phys. A: Math. Theor.* **47** 245001
[3] Campisi M, Hänggi P and Talkner P 2011 *Rev. Mod. Phys.* **83** 771
[4] Holubec V 2014 *J. Stat. Mech.* P05022
[5] Proesmans K and Van den Broeck C 2017 *Europhys. Lett.* **119** 20001
[6] Hasegawa Y and Vu T V 2019 *Phys. Rev. Lett.* **123** 110602
[7] Holubec V and Ryabov A 2018 *Phys. Rev. Lett.* **121** 120601
[8] Allahverdyan A E, Hovhannisyan K V, Melkikh A V and Gevorkian S G 2013 *Phys. Rev. Lett.* **111** 050601
[9] Campisi M and Fazio R 2016 *Nat. Commun.* **7** 11895
[10] Feng E H and Crooks G E 2008 *Phys. Rev. Lett.* **101** 090602