On the fluctuation relation

1. **Jarzynski equality**

   (a) The system is initially at thermal equilibrium (temperature T) : $\rho(\Gamma) = \dfrac{e^{-\beta\mathcal{H}_0(\Gamma)}}{Z_0}$

   (b) Inserting $W = \mathcal{H}_1(\Gamma(t_f)) - \mathcal{H}_0(\Gamma(0))$ into the definition of $\overline{e^{-\beta W}}$, we have

$$\overline{e^{-\beta W}} = \frac{1}{Z_0}\int e^{-\beta\mathcal{H}_1(\Gamma(t_f))}\, d\Gamma(0) = \frac{1}{Z_0}\int e^{-\beta\mathcal{H}_1(\Gamma(t_f))}\, d\Gamma(t_f). \tag{1}$$

   The last equality has been obtained by changing variables $\Gamma(0) \equiv \{\mathbf{r}^N(0), \mathbf{p}^N(0)\} \to \Gamma(t) \equiv \{\mathbf{r}^N(t), \mathbf{p}^N(t)\}$. Hamilton's equations imply that the corresponding Jacobian is 1. This is a facet of Liouville theorem, expressing the conservation of volume in phase space : let $\vec{V}$ be the velocity vector associated to phase space displacements,

$$\text{div}\,\vec{V} = \sum_i \left(\frac{\partial}{\partial \mathbf{r}_i}\dot{\mathbf{r}}_i + \frac{\partial}{\partial \mathbf{p}_i}\dot{\mathbf{p}}_i\right) = \sum_i \left(\frac{\partial}{\partial \mathbf{r}_i}\frac{\partial \mathcal{H}}{\partial \mathbf{p}_i} - \frac{\partial}{\partial \mathbf{p}_i}\frac{\partial \mathcal{H}}{\partial \mathbf{r}_i}\right) = 0. \tag{2}$$

   Finally, the rhs of equation (1) reads Z_1/Z_0, and from $F_\lambda = -kT\log Z_\lambda$, we get

$$\boxed{\overline{e^{-\beta W}} = e^{-\beta\Delta F}}, \tag{3}$$

   where $\Delta F = F_1 - F_0 = F(A_1) - F(A_0)$. This is a remarkable results, on two counts : (a) it connects the free energy difference (an equilibrium property), to an average taken over an ensemble of *non equilibrium* measures and (b) it states that the expression of the lhs is protocol independent. Not trivial !

   (c) From a convexity argument $\overline{e^{-\beta W}} \geq e^{-\beta\overline{W}} \implies \overline{W} \geq \Delta F$: second principle of thermodynamics !

2. **Limiting cases**

   (a) In the reversible limit, each realization should yield the same result : $W = \overline{W} = \Delta F$. Hence, the distribution of work should be a Dirac $\delta(W - \overline{W})$. Jarzynski relation then yields $\overline{e^{-\beta W}} = e^{-\beta\overline{W}} = e^{-\beta\Delta F}$, as expected.

   (b) The Hamiltonian depends on time through phase space point $\Gamma(t)$. Such a dependence can be parameterized by λ, and

$$W \equiv \mathcal{H}_1(\Gamma(t_f)) - \mathcal{H}_0(\Gamma(0)) = \int_0^1 \frac{\partial\mathcal{H}_\lambda(\Gamma(\lambda))}{\partial\lambda}\, d\lambda, = \int_0^{t_f} \frac{d\lambda}{dt}\frac{\partial\mathcal{H}_\lambda(\Gamma(t))}{\partial\lambda}\, dt. \tag{4}$$

   In the quasi-static limit, this quantity no longer fluctuates, and

$$\overline{W} = \int_0^1 \left\langle \frac{\partial\mathcal{H}_\lambda}{\partial\lambda}\right\rangle_\lambda d\lambda. \tag{5}$$

   where $\langle\ldots\rangle_\lambda$ denotes canonical averaging with weight $\exp(-\beta\mathcal{H}_\lambda)/Z_\lambda$. Then,

$$\frac{\partial F_\lambda}{\partial\lambda} = -kT\frac{\partial\log Z_\lambda}{\partial\lambda} = -kT\frac{1}{Z_\lambda}\int\frac{\partial}{\partial\lambda}e^{-\beta\mathcal{H}_\lambda(\Gamma)}\, d\Gamma = \left\langle\frac{\partial\mathcal{H}_\lambda}{\partial\lambda}\right\rangle_\lambda. \tag{6}$$

   Going back to equation (5) yields $\overline{W} = \displaystyle\int_0^1 \frac{\partial F_\lambda}{\partial\lambda}\, d\lambda = \Delta F$.

(c) With $W = \mathcal{H}_1(\Gamma(0)) - \mathcal{H}_0(\Gamma(0))$,

$$\overline{e^{-\beta W}} = \left\langle e^{-\beta W} \right\rangle_0 = \frac{1}{Z_0} \int e^{-\beta \mathcal{H}_1} \, d\Gamma = \frac{Z_1}{Z_0} = e^{-\beta \Delta F}. \tag{7}$$

3. **Crooks relation**

(a) For the reverse process, the system evolves from $\Gamma(t_f)$ to $\Gamma(0)$, and the work received is $W = \mathcal{H}_0(\Gamma(0)) - \mathcal{H}_1(\Gamma(t_f))$, so that

$$p_B(W) = \frac{1}{Z_1} \int e^{-\beta \mathcal{H}_1(\Gamma(t_f))} \, \delta\left\{W - \mathcal{H}_0(\Gamma(0)) + \mathcal{H}_1(\Gamma(t_f))\right\} \, d\Gamma(t_f). \tag{8}$$

(b) Back to the definition of p_F :

$$\begin{aligned} p_F(W) \equiv & \; \frac{1}{Z_0} \int e^{-\beta \mathcal{H}_0(\Gamma(0))} \, \delta\left\{W - \mathcal{H}_1(\Gamma(t_f)) + \mathcal{H}_0(\Gamma(0))\right\} \, d\Gamma(0) \\ = & \; \frac{\exp(\beta W)}{Z_0} \int e^{-\beta \mathcal{H}_1(\Gamma(t_f))} \, \delta\left\{W - \mathcal{H}_1(\Gamma(t_f)) + \mathcal{H}_0(\Gamma(0))\right\} \, d\Gamma(0) \\ = & \; \exp(\beta W) \, \frac{Z_1}{Z_0} \, \frac{1}{Z_1} \int e^{-\beta \mathcal{H}_1(\Gamma(t_f))} \, \delta\left\{W - \mathcal{H}_1(\Gamma(t_f)) + \mathcal{H}_0(\Gamma(0))\right\} \, d\Gamma(t_f) \\ = & \; \exp(\beta W) \, \frac{Z_1}{Z_0} \, p_B(-W), \end{aligned}$$

where once more, a change of variables with unit Jacobian has been performed. Therefore :

$$\boxed{p_F(W) \, e^{-\beta W} = e^{-\beta \Delta F} \, p_B(-W)}. \tag{9}$$

(c) Integrating (9) over all W :

$$\int_{-\infty}^{\infty} p_F(W) \, e^{-\beta W} \, dW = e^{-\beta \Delta F} \int_{-\infty}^{\infty} p_B(-W) \, dW = e^{-\beta \Delta F} \int_{-\infty}^{\infty} p_B(W') \, dW' = e^{-\beta \Delta F}, \tag{10}$$

and we retrieve Jarzynski equality.

(d) In the reversible limit,

$$p_F(W) = \delta(W - \Delta F), \quad p_B(W) = \delta(W + \Delta F). \tag{11}$$

Crooks relation is obeyed.

4. **Gaussian fluctuations**

(a) Since W is Gaussian,

$$\log\langle e^{kW} \rangle = \langle kW \rangle + \frac{k^2}{2} \, \sigma, \tag{12}$$

where k is arbitrary. Take then $k = -\beta$. Alternatively, one can calculate explicitly... Jarzynski then allows us to write

$$\beta \Delta F = \beta \overline{W} - \frac{\beta^2 \sigma^2}{2} \qquad ; \qquad \boxed{\overline{W} = \Delta F + \frac{\beta \sigma^2}{2}}. \tag{13}$$

We get $\overline{W} \geq \Delta F$, as it should.

(b) One can view $W_{\text{diss}} \equiv \overline{W} - \Delta F$ as the dissipated work, that quantifies irreversibility of the transformation. Since $W_{\text{diss}} = \beta\sigma^2/2$, we show here that the dissipated work is related to the fluctuations of W from a realization to the next, through the standard deviation σ. This qualifies our connection as of fluctuation-dissipation type. In the reversible limit, $\sigma = 0$ and we recover a Dirac peak for p_F (and p_B).

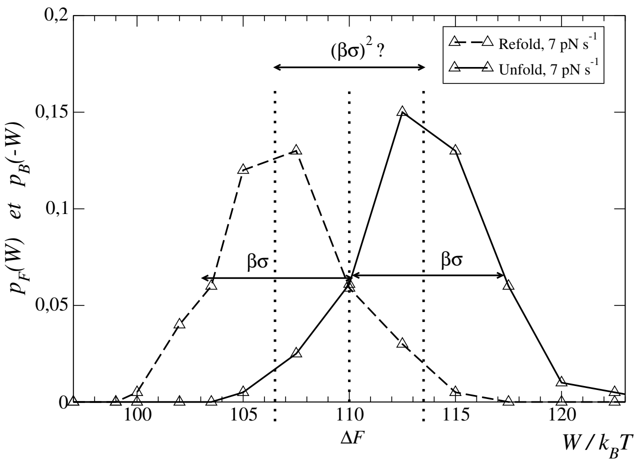

FIGURE C1 – Plots of the probability density distributions of work $p_F(W)$ (continuous curve) and $p_B(-W)$ (dashed line) for the experiment performed at 7 $pN\,s^{-1}$. The three horizontal arrows have length 7. The question mark for the upper arrow is to question the validity of a Gaussian ansatz.

5. **. . . where one measures the length of time's arrow . . .**

   The Kullback-Leibler distance reads, from Crooks relation

$$\int_{-\infty}^{\infty} p_F(W) \log \left[\frac{p_F(W)}{p_B(-W)} \right] = \int_{-\infty}^{\infty} p_F(W)\, \beta[W - \Delta F] \;=\; \beta(\overline{W} - \Delta F) \;=\; \beta W_{\text{diss}}, \tag{14}$$

   which establishes a noteworthy link between irreversibility, and the possibility to distinguish the forward and reverse protocols from one another.

6. **Application to single molecule experiments**

   (a) Crooks relation (9) indicates that the graphs of $p_F(W)$ and $p_B(-W)$ have to cross precisely at $W = \Delta F$, and cannot afford more than one crossing. This yields a strong constraint, quite well obeyed with the data shown. The corresponding estimate is $\Delta F \simeq 110\, kT$.

   (b) The offset follows from the second principle : $\overline{W}_F \geq \Delta F$, $\overline{W}_B \geq -\Delta F$, whence $-\overline{W}_B < \overline{W}_F$. This shift can be considered as another measure of the length of time's arrow.

   The data shown make sense. We can note that the Gaussian assumption is not valid here. Consider indeed the measures at 7 $pN\,s^{-1}$. We read on the graph something like $\beta\sigma$ not too far from 5 or 6, which from the Gaussian ansatz should also be that of p_B. Then, the maxima of p_F and p_B should be separated by $(\beta\sigma)^2 > 25$; this is not the case (see Fig. C1). One can also note that the spread σ is larger for the "fast" protocol, than for the slower one (as expected).