Non-equilibrium work relations & fluctuation theorems

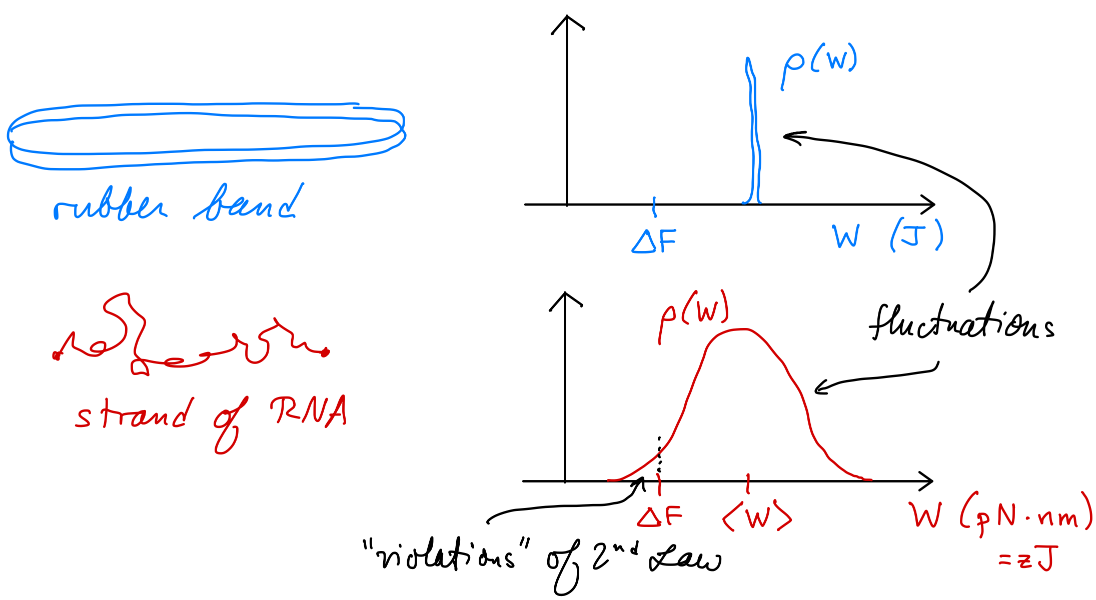
Macroscopic systems:

2nd Law of thermodynamics is expressed in terms of inequalities, e.g. $\Delta S_{tot} \geq 0$, $W \geq \Delta F$, $\int \frac{đQ}{T} \leq \Delta S$, etc.

(or equivalently as prohibitions: "No process is possible whose sole result ...")

Microscopic systems:

These inequalities remain true on average, but now fluctuations are substantial.



It's tempting to view these fluctuations as random, "boring", uninformative noise. However, it has become recognized that these fluctuations satisfy strong, unexpected, & potentially useful laws, which effectively allow us to express the 2^{nd} Law itself in the form of an equality (actually, equalities).

These results provide insights into the thermodynamic "arrow of time" (Why do movies run backward look strange?) the relationship between thermodynamics and information processing (Maxwell's demon), and the probability to observe "violations" of the 2^{nd} Law in very small systems.

Most of these results fall into two classes:

- <u>non-equilibrium work relations</u>

  ~ relationship between work and free energy in a system that is driven away from an intial state of equilibrium

- <u>fluctuation theorems for entropy production</u>

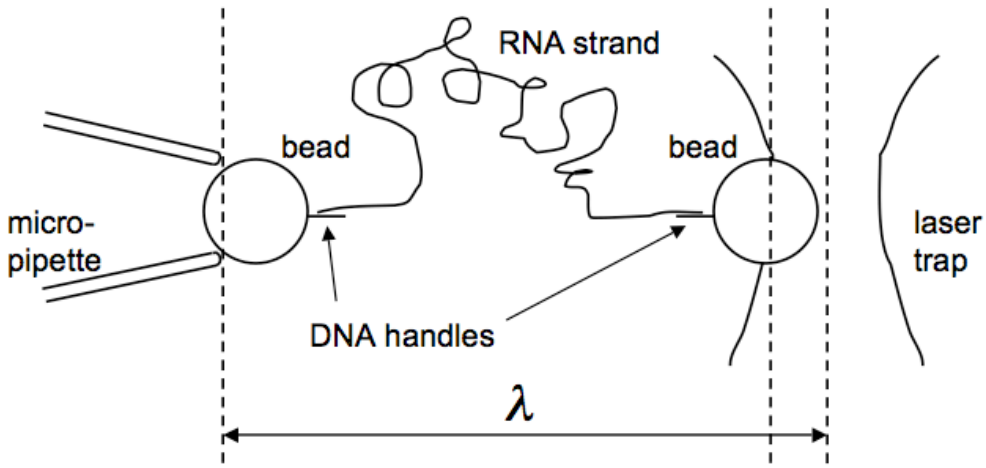
  ~ production of entropy in a system that is either in a nonequilibrium steady state, or else relaxing toward such a state

I'll begin with the 1^{st} class of results …

(see Annu. Rev. Cond. Matt. Phys. 2011 review article posted to course website.)

Work and fluctuations in irreversible processes

(useful example to keep in mind)

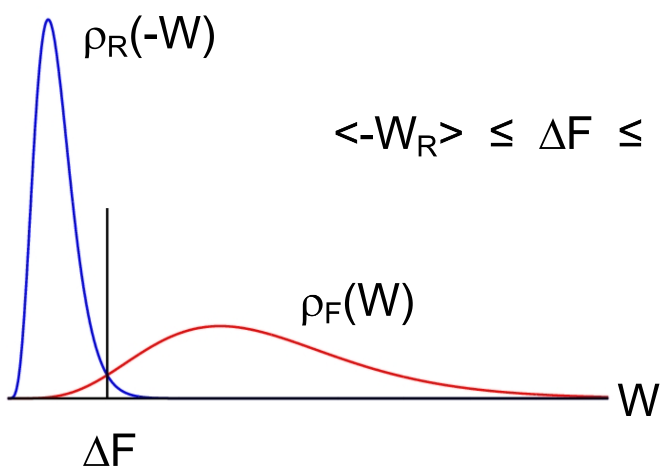


1. Start at λ=A , in equilibrium w/ water

forward
2. Stretch rapidly, λ : A$\rightarrow$B
3. Hold λ fixed and allow to re-equilibrate

reverse
4. "Unstretch" rapidly, λ : A$\leftarrow$B
5. Hold λ fixed and allow to re-equilibrate
6. *Repeat*

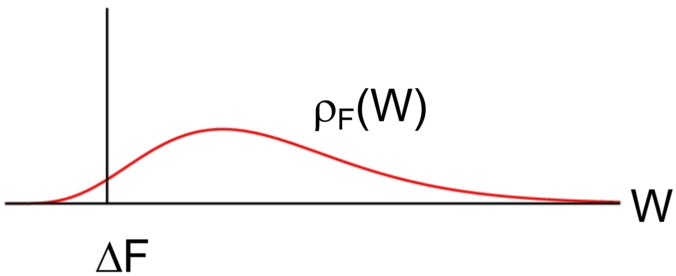
expect the 2nd law to hold *on average*:



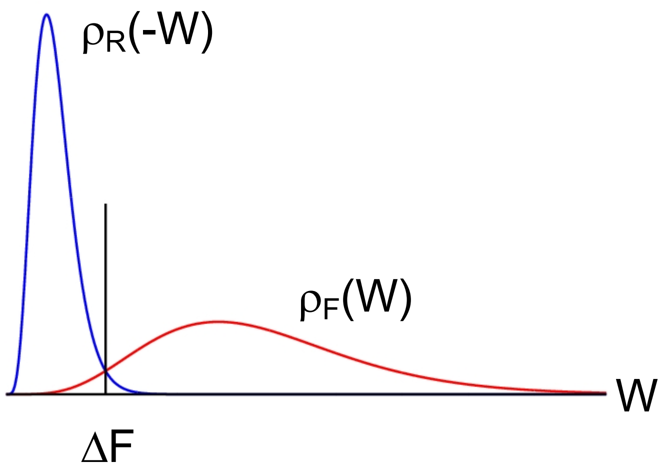
$$\langle -W_R \rangle \leq \Delta F \leq \langle W_F \rangle$$

What about fluctuations around the average?

Three predictions related to work fluctuations in small systems driven away from equilibrium:



$$\left\langle e^{-\beta W} \right\rangle = e^{-\beta \Delta F}$$



$$\frac{\rho_F(+W)}{\rho_R(-W)} = e^{\beta(W-\Delta F)}$$

$$\left\langle \delta(x - x_t) e^{-\beta w_t} \right\rangle = \frac{1}{Z_A} e^{-\beta H(x,\lambda_t)}$$

x = point in phase space (independent variable)
x_t = microscopic trajectory during one realization
w_t = work performed up to time t
λ_t = value of work parameter at time t
<…> = average over ensemble of realizations

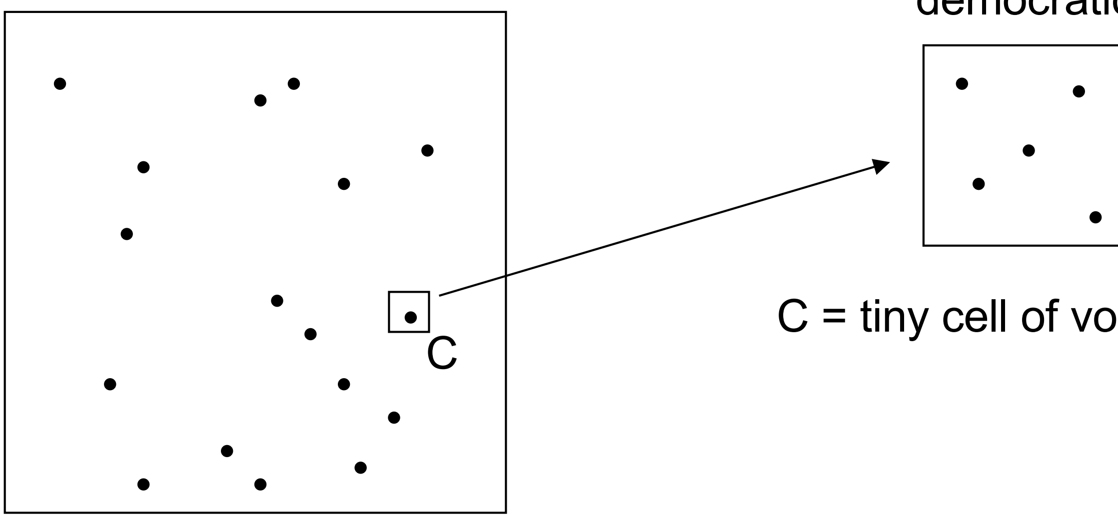
ordinarily we use the time-dependent phase space density
to represent the evolution of an ensemble of trajectories:

$$f(x,t) = \left\langle \delta(x - x_t) \right\rangle$$

be careful w/ ordering of limits …

phase space

“democratic”



C = tiny cell of volume ε

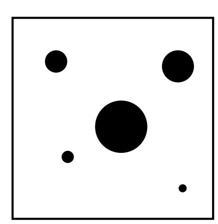
“snapshot” at time t

$$f(x,t) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \lim_{N \to \infty} \frac{1}{N} \sum_{x_t \in C} 1$$

another statistical representation *of the same ensemble* :

$$g(x,t) = \left\langle \delta(x - x_t) \exp(-\beta w_t) \right\rangle$$
$$= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \lim_{N \to \infty} \frac{1}{N} \sum_{x_t \in C} \exp(-\beta w_t)$$

“undemocratic”



claim : $g(x,t) \propto f^{eq}(x,\lambda_t)$

1 $$\left\langle e^{-\beta W} \right\rangle = e^{-\beta \Delta F}$$

2 $$\frac{\rho_F(+W)}{\rho_R(-W)} = e^{\beta(W-\Delta F)}$$

3 $$\left\langle \delta(x - x_t) e^{-\beta w_t} \right\rangle = \frac{1}{Z_A} e^{-\beta H(x,\lambda_t)}$$

general remarks:

- exact results, valid far from equilibrium
- place strong, unexpected constraints on fluctuations
- nonequilibrium fluctuations encode equilibrium information
- closely related to other results -
  - Bochkov & Kuzovlev , <exp(-βW')> = 1
  - entropy fluctuation theorems, p(+σ)/p(-σ) = exp(σ)

Derivations

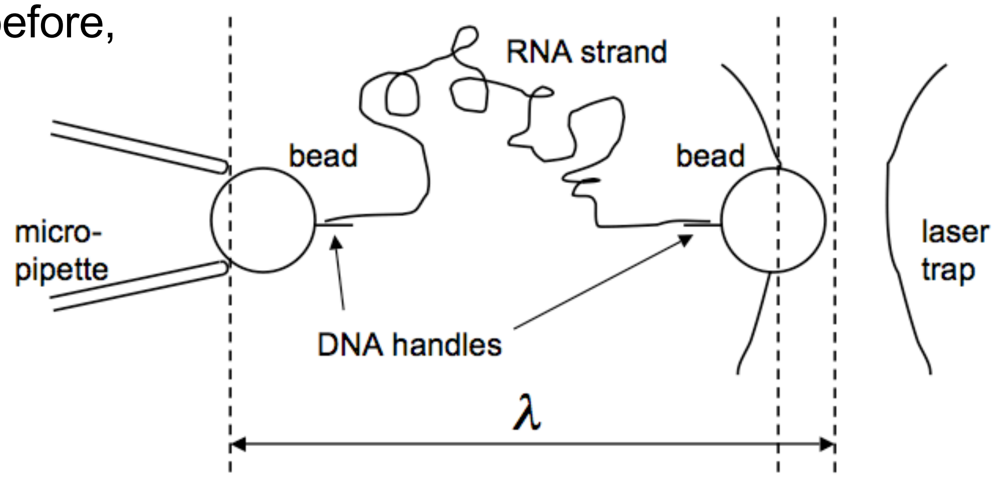
- Hamiltonian dynamics
- Langevin dynamics ("Brownian motion")
- discrete-time Markov processes
- deterministic thermostats (Gaussian, Nosé-Hoover)
- quantum dynamics

  & others

I'll derive #1 using Hamiltonian evolution, #2 using discrete-state dynamics, and #3 using diffusive dynamics with inertia.

start w/ derivation of 1 using Hamilton's equations

as before,



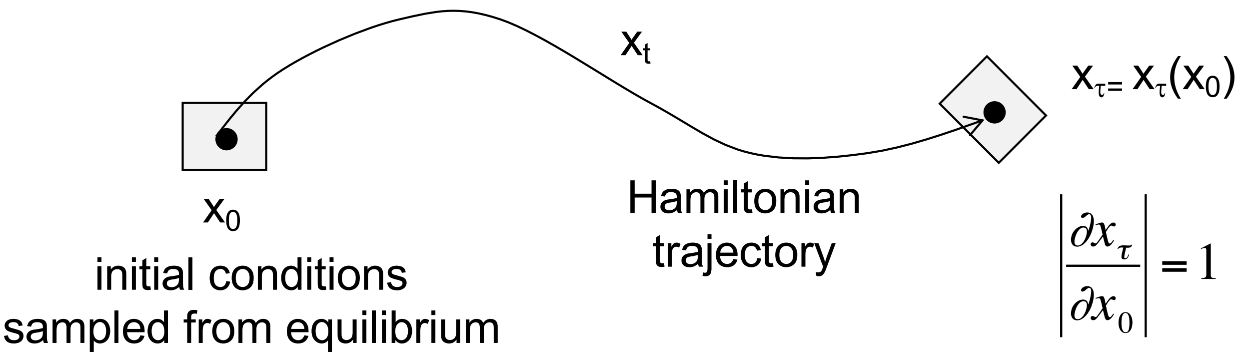
$$x = (q,p) = (\vec{r}_1, \cdots \vec{r}_m, \vec{p}_1, \cdots \vec{p}_m)$$ microstate

$$H(x,\lambda)$$ Hamiltonian

$$\dot{q} = \frac{\partial H}{\partial p} \quad , \quad \dot{p} = -\frac{\partial H}{\partial q} \longrightarrow x(t)\,,\, x_t$$ trajectory

$$\lambda : A \rightarrow B \qquad \lambda(t)\,,\, \lambda_t$$ protocol

For simplicity, consider *special case* of a system that is thermally isolated as the work parameter is varied, $0<t<\tau$ (… not plausible in real single-molecule pulling experiment !)



$$\left|\frac{\partial x_\tau}{\partial x_0}\right| = 1$$

$$W = W\left(x_0\right) = H\left(x_\tau\left(x_0\right), B\right) - H\left(x_0, A\right)$$

$$\left\langle e^{-\beta W}\right\rangle = \int dx_0\, p_A^{eq}\left(x_0\right) e^{-\beta W\left(x_0\right)}$$

$$= \frac{1}{Z_A}\int dx_0\, e^{-\beta H_A\left(x_0\right)}\, e^{-\beta\left[H_B\left(x_\tau\left(x_0\right)\right) - H_A\left(x_0\right)\right]}$$

$$= \frac{1}{Z_A}\int dx_0\; e^{-\beta H_B\left(x_\tau\left(x_0\right)\right)}$$

$$= \frac{1}{Z_A}\int dx_\tau\; e^{-\beta H_B\left(x_\tau\right)} = \frac{Z_B}{Z_A} = e^{-\beta\Delta F}$$

☺

only assumptions:

1. initial conditions sampled from (canonical) equilibrium
2. isolated, classical system (Hamiltonian evolution)

This derivation can be generalized to a system in contact w/ heat reservoir, e.g. biomolecule in solution; no need to assume weak coupling, provided we properly account for solvation free energy.

$$H\left(x, y; \lambda\right) = H_{sys}\left(x; \lambda\right) + H_{env}\left(y\right) + H_{\text{int}}\left(x, y\right)$$

The quantum version of this derivation is simple.

$\hat{H}(\lambda)$ or $\hat{H}_\lambda$ = parameter-dependent Hamiltonian operator

$$\hat{H}(\lambda)\,|n_\lambda\rangle = E_n(\lambda)\,|n_\lambda\rangle$$

$$\hat{\pi}_\lambda = \frac{1}{Z_\lambda}\sum_n |n_\lambda\rangle\langle n_\lambda|\, e^{-\beta E_n(\lambda)}$$

Boltzmann-Gibbs density matrix

(1) prepare in equilibrium ($\lambda = A$), measure energy

$\longrightarrow$ "collapse" to $|n_A\rangle$

not really

(2) $\lambda : A \to B$ as $t : 0 \to \tau$

unitary evolution

(3) measure final energy $\to$ collapse to $|m_B\rangle$

$$W = \underbrace{E_m(B)}_{\text{final energy}} - \underbrace{E_n(A)}_{\text{initial energy}}$$

$$W = E_m(B) - E_n(A)$$

$$\hat{\pi}_\lambda = \frac{1}{Z_\lambda} \sum_n |n_\lambda\rangle\langle n_\lambda| \, e^{-\beta E_n(\lambda)}$$

$$\langle e^{-\beta W}\rangle = \sum_n \pi_A(n) \sum_m P(m_B|n_A)\, e^{-\beta W}$$

$$= \sum_{n,m} \frac{1}{Z_A} e^{-\beta E_n(A)} \left|U_{mn}\right|^2 e^{-\beta[E_m(B) - E_n(A)]}$$

$$\hookrightarrow \langle m_B | \hat{U} | n_A \rangle$$

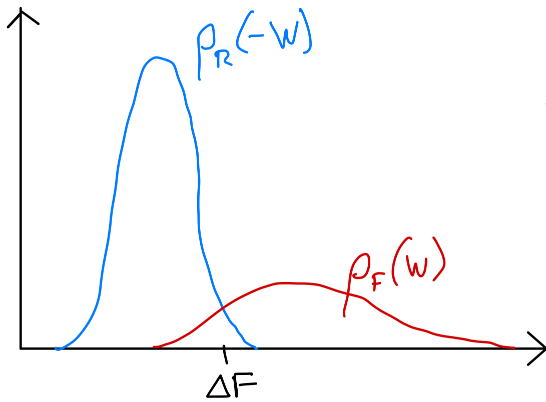
time-evolution operator

$$= \frac{1}{Z_A} \sum_m e^{-\beta E_m(B)} \sum_n \left|U_{mn}\right|^2 = 1 \quad \text{(unitarity)}$$

$$= \frac{Z_B}{Z_A} = e^{-\beta \Delta F} \qquad \checkmark$$

Not so easy to extend this derivation to the case of a system coupled to a heat bath.
(Open problem ...)

Crooks Fluctuation Theorem



$$\frac{\rho_F(W)}{\rho_R(-W)} = e^{\beta(W-\Delta F)}$$

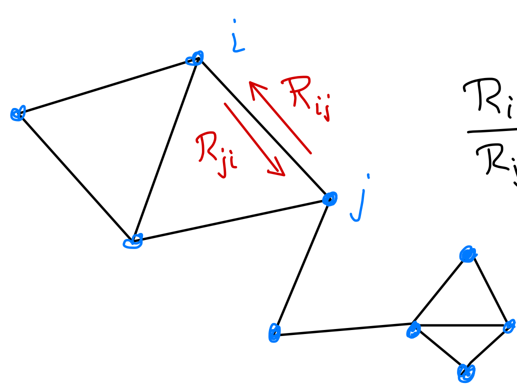
$$\Delta F = F_B - F_A$$

Note: $\rho_F(W)\, e^{-\beta W} = \rho_R(-W)\, e^{-\beta \Delta F}$

$$\langle e^{-\beta W} \rangle_F = e^{-\beta \Delta F}$$ after integrating both sides

(<u>Exercise</u>: derive using Hamiltonian dynamics)

I'll derive this result for a discrete-state process.



$$\frac{R_{ij}}{R_{ji}} = e^{-\beta(E_i - E_j)}$$

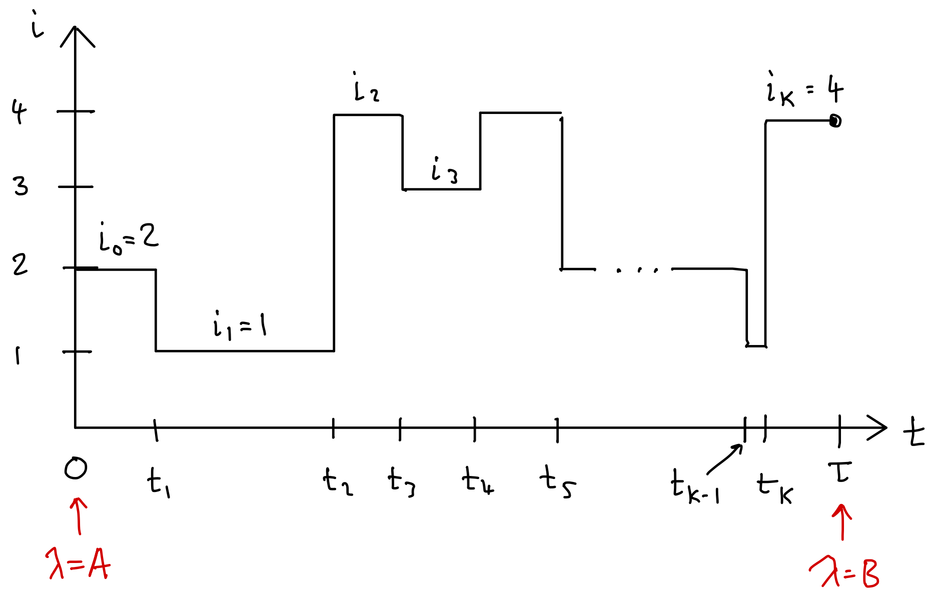
detailed balance

$R = R(\lambda)$
$E_i = E_i(\lambda)$
$\lambda = \lambda(t)$

forward process : $\lambda_F(t)$: $A \to B$
reverse process : $\lambda_R(t)$: $B \to A$ as $t : 0 \to \tau$

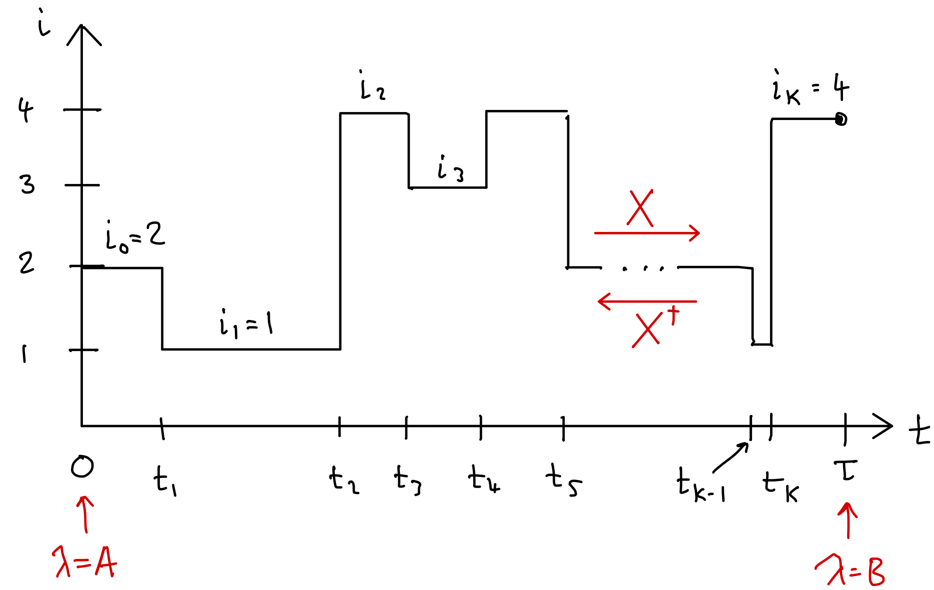
$$\underline{\lambda_F(t) = \lambda_R(\tau - t)}$$

Look @ 1 realization of the forward process:



K = # of transitions
(differs from one realization to the next)

t_k = time of k^{th} transition , $1 \le k \le K$



convenient notation for 1 trajectory:

forward process

$$X = i_0 \xrightarrow{t_1} i_1 \xrightarrow{t_2} i_2 \cdots \xrightarrow{t_K} i_K$$

"conjugate twin":

reverse process

$$X^\dagger = i_0 \xleftarrow{t_K^\dagger} i_1 \xleftarrow{t_{K-1}^\dagger} i_2 \cdots \xleftarrow{t_1^\dagger} i_K$$

$$t_\ell^\dagger = \tau - t_{K+1-\ell}$$

For convenience, define $t_0=0$ & $t_{K+1}=\tau$. Then:

$$\Delta t_\ell = t_{\ell+1} - t_\ell \qquad 0 \le \ell \le K$$

$$= \underline{\text{dwell time}} \text{ in state } i_\ell$$

Note: same set of dwell times in X & $X^\dagger$.

Now let $S_F(i,t',t'')$ denote survival probability:

If the system is in state i @ time t' during the forward process, then $S_F(i,t',t'')$ is the probability that it makes no transitions out of state i during the interval $t' < t < t''$.

$$S_F(i,t',t''+\delta t'') = \underbrace{S_F(i,t',t'')}_{\substack{\text{probability}\\ \text{to survive in}\\ \text{state } i \text{ until}\\ \text{time } t''}} \cdot \Big[1 - \underbrace{\delta t'' \sum_{j \neq i} R_{ji}\big(\lambda^F(t'')\big)}_{\substack{\text{probability}\\ \text{to survive in } i\\ \text{from } t'' \text{ to}\\ t''+\delta t''}}\Big]$$

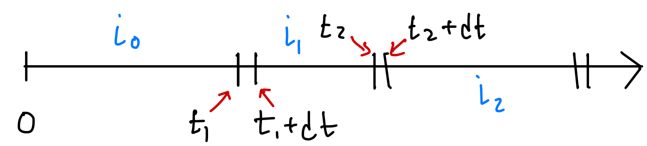
$$\longrightarrow \quad \frac{\partial S_F}{\partial t''} = \Big(-\sum_{j\neq i} R_{ji}\Big) S_F = R_{ii}\, S_F$$

$$S_F(i,t',t'') = \exp\left[\int_{t'}^{t''} R_{ii}\big(\lambda_F(t)\big)\,dt\right]$$

$$= S_R(i,\tau-t'',\tau-t')$$

since $\lambda_F(t) = \lambda_R(\tau - t)$

$P_F[X]$ = probability density to observe X when performing forward process



$$P_F[X](dt)^K = \pi_A(i_0)\cdot S_F(i_0, 0, t_1)\cdot R_{i_1 i_0}\,dt$$
$$\cdot\, S_F(i_1, t_1, t_2)\cdot R_{i_2 i_1}\,dt$$
$$\cdots\, S_F(i_K, t_K, \tau)$$

$$P_R[X^\dagger](dt)^K = \pi_B(i_K)\cdot S_R(i_K, 0, t_1^\dagger)\cdot R_{i_{K-1} i_K}\,dt$$
$$\cdot\, S_R(i_{K-1}, t_1^\dagger, t_2^\dagger)\cdot R_{i_{K-2} i_K}\,dt$$
$$\cdots\, S_R(i_0, t_K^\dagger, \tau)$$

The same survival probabilities appear in the expressions for $P_F[X]$ & $P_R[X^\dagger]$.

e.g. $S_R(i_0, t_K^\dagger, \tau) = S_F(i_0, \tau-\tau, \tau - t_K^\dagger)$
$= S_F(i_0, 0, t_1)$

$$P_F[X] = \pi_A(i_0) \cdot R_{i_1 i_0} \cdot R_{i_2 i_1} \cdots R_{i_k i_{k-1}} \cdot (\text{all the } S_F\text{'s})$$

same

$$P_R[X^+] = \pi_B(i_k) \cdot R_{i_{k-1} i_k} \cdot R_{i_{k-2} i_{k-1}} \cdots R_{i_0 i_1} \cdot (\text{all the } S_R\text{'s})$$

$$\frac{P_F[X]}{P_R[X^+]} = \frac{\pi_A(i_0)}{\pi_B(i_k)} \cdot \frac{R_{i_1 i_0}}{R_{i_0 i_1}} \cdots \frac{R_{i_k i_{k-1}}}{R_{i_{k-1} i_k}}$$

$$= \frac{e^{-\beta E(i_0, A)}/Z_A}{e^{-\beta E(i_k, B)}/Z_B} \cdot e^{-\beta[E(i_1,\lambda_1) - E(i_0,\lambda_1)]}$$

$\lambda_\ell = \lambda(t_\ell)$

$$\cdots \cdot e^{-\beta[E(i_k,\lambda_k) - E(i_{k-1},\lambda_k)]}$$

$$= e^{-\beta \Delta F} \cdot e^{+\beta \Delta E} \cdot e^{-\beta \sum_{\ell=1}^{k} \delta E_\ell}$$

$\delta E_\ell = E(i_\ell, \lambda_\ell) - E(i_{\ell-1}, \lambda_\ell)$

= energy change due to ℓ th transition

$\Delta E = E(i_k, B) - E(i_0, A)$

= net change in system's energy

$$\Delta E - Q = W$$

$$\frac{P_F[X]}{P_R[X^+]} = e^{-\beta\Delta F} \cdot e^{+\beta\Delta E} \cdot e^{-\beta \underbrace{\sum_{\ell=1}^{k} \delta E_\ell}_{Q}}$$

$$\boxed{\frac{P_F[X]}{P_R[X^+]} = e^{\beta(W-\Delta F)}}$$

$$W = W_F[X] = -W_R[X^+]$$

$$\rho_F(W) = \int dX \; P_F[X] \, \delta\left(W - W_F[X]\right)$$

$$\hookrightarrow \int dX = \sum_{K=0}^{\infty} \underbrace{\int_0^\tau dt_1 \cdots \int_0^\tau dt_K}_{t_1 < \cdots < t_K} \sum_{\{i_0 \cdots i_K\}}$$

$$\rho_R(W) = \int dX^+ \, P_R[X^+] \, \delta\left(W - W_R[X^+]\right)$$

$$\hookrightarrow \int dX^+ = \sum_{K=0}^{\infty} \underbrace{\int_0^\tau dt_1^+ \cdots \int_0^\tau dt_K^+}_{t_1^+ < \cdots < t_K^+} \sum_{\{i_0 \cdots i_K\}}$$

same!

$$\begin{aligned}
\rho_F(W) &= \int dX \, P_F[X] \, \delta\left(W - W_F[X]\right) \\
&= \int dX \; P_R[X^+] \, e^{\beta(W_F[X] - \Delta F)} \, \delta\left(W - W_F[X]\right) \\
&= e^{\beta(W-\Delta F)} \int dX^+ \, P_R[X^+] \; \delta\left(W + W_R[X^+]\right) \\
&= e^{\beta(W-\Delta F)} \, \rho_R(-W)
\end{aligned}$$

Finally, derive $\langle \delta(x-x_t) e^{-\beta W_t} \rangle = \frac{1}{Z_A} e^{-\beta \mathcal{H}(x,\lambda_t)}$ for diffusive process w/ inertia.

[If we set $\gamma, D_p = 0$ in the following calculation, we'll get the derivation of this result for Hamiltonian dynamics.]

$$x = (q,p) \in \mathbb{R}^{2N} \quad , \quad \mathcal{H} = \frac{p^2}{2m} + V(q,\lambda)$$

$$\dot{q} = \frac{p}{m} \quad , \quad \dot{p} = -\frac{\partial V}{\partial q} - \frac{\gamma}{m} p + \xi(t)$$

$$f(x,t) = \langle \delta(x-x_t) \rangle$$

$$g(x,t) = \langle \delta(x-x_t) e^{-\beta W_t} \rangle$$

$$W_t = \int_0^t dt' \, \dot{\lambda}_{t'} \, \frac{\partial \mathcal{H}}{\partial \lambda}(x_{t'}, \lambda_{t'})$$

$$\begin{cases} \dfrac{\partial f}{\partial t} = \hat{\mathcal{L}}_\lambda f = \{\mathcal{H}, f\} + D_p \dfrac{\partial}{\partial p}\left[e^{-\beta \mathcal{H}} \dfrac{\partial}{\partial p} \left(e^{\beta \mathcal{H}} f \right) \right] \\ \dfrac{\partial g}{\partial t} = \, ? \end{cases}$$

Note: $\hat{\mathcal{L}}_\lambda \pi_\lambda = \frac{1}{Z_\lambda} \hat{\mathcal{L}}_\lambda e^{-\beta \mathcal{H}(x,\lambda)}$

$= 0$

$$h(x,w,t) \equiv \langle \delta(x - x_t)\, \delta(w - w_t) \rangle$$

= joint probability distribution to get $x_t = x$ & $w_t = w$.

$$\begin{cases} \dot{q} = \frac{p}{m} \\ \dot{p} = -\frac{\partial V}{\partial q} - \gamma \frac{p}{m} + \xi(t) \\ \dot{w} = \dot{\lambda} \frac{\partial \mathcal{H}}{\partial \lambda} \end{cases}$$

only stochastic (random) term

$$\frac{\partial h}{\partial t} = -\frac{\partial}{\partial q}\left(\frac{p}{m} h\right) - \frac{\partial}{\partial p}\left[\left(-\frac{\partial V}{\partial q} - \gamma \frac{p}{m}\right) h\right] - \frac{\partial}{\partial w}\left[\dot{\lambda} \frac{\partial \mathcal{H}}{\partial \lambda} h\right] + D_p \frac{\partial^2 h}{\partial p^2}$$

$$= \hat{\mathcal{L}}_\lambda h - \dot{\lambda} \frac{\partial \mathcal{H}}{\partial \lambda} \frac{\partial h}{\partial w}$$

$$g(x,t) = \langle \delta(x - x_t)\, e^{-\beta w_t} \rangle = \int dw\, e^{-\beta w}\, h(x,w,t)$$

$$\frac{\partial g}{\partial t} = \int dw\, e^{-\beta w} \left(\hat{\mathcal{L}}_\lambda h - \dot{\lambda} \frac{\partial \mathcal{H}}{\partial \lambda} \frac{\partial h}{\partial w}\right)$$

$$= \hat{\mathcal{L}}_\lambda g - \dot{\lambda} \frac{\partial \mathcal{H}}{\partial \lambda} \int dw\, e^{-\beta w} \frac{\partial h}{\partial w}$$

$$= \hat{\mathcal{L}}_\lambda g - \beta \dot{\lambda} \frac{\partial \mathcal{H}}{\partial \lambda} \underbrace{\int dw\, e^{-\beta w} h}_{g}$$

$$\therefore \quad \frac{\partial g}{\partial t} = \left(\hat{\mathcal{L}}_\lambda - \beta \dot{\lambda} \frac{\partial \mathcal{H}}{\partial \lambda}\right) g$$

Now solve this eqn for the init'l. conditions

$$g(x,0) = \langle \delta(x - x_0) \underbrace{e^{-\beta w_0}}_{1} \rangle = f(x,0) = \pi_A(x)$$

$$= \frac{1}{Z_A} e^{-\beta \mathcal{H}(x,A)}$$

Solution: $g(x,t) = \frac{1}{Z_A} e^{-\beta \mathcal{H}(x,\lambda_t)}$

Check: $\frac{\partial g}{\partial t} = -\beta \dot{\lambda} \frac{\partial \mathcal{H}}{\partial \lambda} g$

$$\left(\hat{\mathcal{L}}_\lambda - \beta \dot{\lambda} \frac{\partial \mathcal{H}}{\partial \lambda}\right) g = \cancelto{0}{\frac{1}{Z_A} \hat{\mathcal{L}}_\lambda e^{-\beta \mathcal{H}(x,\lambda)}} - \beta \dot{\lambda} \frac{\partial \mathcal{H}}{\partial \lambda} g$$

Thus, $g(x,t) \equiv \langle \delta(x - x_t) e^{-\beta w_t} \rangle = \frac{1}{Z_A} e^{-\beta \mathcal{H}(x,\lambda_t)}$

☺