[Readings] Topics in Non-Equilibrium Statistical Mechanics

Chaos Tang

September 21, 2021

0 Preface

Here is the reading assignment of the course *Non-Equilibrium Statistical Mechanics*. Apart from the course textbook by Pottier, I have referred to several other books in preparing these notes:

- Pottier, *Nonequilibrium Statistical Mechanics: Linear Irreversible Processes*, Oxford University Press 2010.
- Kardar, *Statistical Physics of Particles*, Cambridge University Press 2007.
- Peliti, *Statistical Mechanics in a Nutshell*, Princeton University Press 2011.

Regarding the division into formalism and application, I choose to lean towards the former. Therefore, among the many examples and applications in the textbook, I have included only the particularly enlightening (and interesting) ones. On the other hand, in a few places where the discussion in the book is not detailed enough, I have tried to borrow examples from or build toy models inspired by the other books.

I find my (re)organization of materials very helpful in clarifying the ideas for myself. I hope it should also help whoever comes across these notes as well.

1 Probability and Random Processes

1.1 Random Variables

A random variable X takes values in the sample space Ω with probability density function (PDF) $p(x)$. The characteristic function $\tilde{p}(k) := \langle e^{ikx} \rangle$ generates the moments. Its logarithm generates the cumulants.

$$\tilde{p}(k) = \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} \langle x^n \rangle, \quad \ln \tilde{p}(k) = \sum_{n=1}^{\infty} \frac{(ik)^n}{n!} \langle x^n \rangle_c. \tag{1.1}$$

The moments and cumulants satisfy a graphical relation:

$$\begin{aligned} \langle x \rangle &= \langle x \rangle_c & \langle x^3 \rangle &= \langle x^3 \rangle_c + 3\langle x^2 \rangle_c \langle x \rangle_c + \langle x \rangle_c^3 \\ &= (.), & &= (...) + 3(..)(.) + (.)(.)(.), \\ \langle x^2 \rangle &= \langle x^2 \rangle_c + \langle x \rangle_c^2 & \langle x^4 \rangle &= \langle x^4 \rangle_c + 4\langle x^3 \rangle_c \langle x \rangle_c + 3\langle x^2 \rangle_c \langle x^2 \rangle_c + 6\langle x^2 \rangle_c \langle x \rangle_c^2 + \langle x \rangle_c^4 \\ &= (..) + (.)(.), & &= (....) + 4(...)(.) + 3(..)(..) + 6(..)(.)(.) + (.)(.)(.)(.), \end{aligned} \tag{1.2}$$

$$\cdots \tag{1.3}$$

$$\begin{aligned} p(k) &= \exp \ln \tilde{p}(k), \\ [\text{diagrams}] &= \exp[\text{connected diagrams}]. \end{aligned} \tag{1.4}$$

In the multi-variate case, each dot is dressed with a label indicating which variable it represents.

If the distribution $p(x)$ has a long tail, some of the moments may not converge.

Example: Cauchy distribution.

$$p(x) = \frac{a}{\pi}\frac{1}{(x-\mu)^2+a^2} \implies \langle x^n\rangle \sim \int^{\infty} x^{n-2}\mathrm{d}x = \infty. \tag{1.5}$$

In fact, with proper regularization to justify the translation $x \mapsto x+\mu$, it's possible to define $\langle (x-\mu)^{2n-1}\rangle = 0$.

The characteristic function is well-defined, but is not differentiable at $k=0$. To see this, close the contour from above if $k>0$ and from below if $k<0$:

$$\tilde{p}(k) = \frac{a}{\pi}\int_{-\infty}^{+\infty}\frac{e^{ikx}}{(x-\mu)^2+a^2}\mathrm{d}x = 2ai\,\mathrm{sgn}(k)\operatorname*{res}_{x=\mu+ia\,\mathrm{sgn}(k)}\frac{e^{ikx}}{(x-\mu)^2+a^2} = e^{-a|k|+ik\mu}. \tag{1.6}$$

1.2 Gaussian Distribution and the Central Limit Theorem

In the most general form with n variables,

$$p(\mathbf{x}) = \frac{1}{\mathcal{N}}\exp\left(-\frac{1}{2}\mathbf{x}^T\mathsf{A}\mathbf{x}+\mathbf{x}^T\mathbf{J}\right), \quad \mathcal{N} = \sqrt{\frac{(2\pi)^n}{\det\mathsf{A}}}\exp\left(\frac{1}{2}\mathbf{J}^T\mathsf{A}^{-1}\mathbf{J}\right). \tag{1.7}$$

The characteristic function is also of Gaussian form:

$$\tilde{p}(\mathbf{k}) = \exp\left(-\frac{1}{2}\mathbf{k}^T\mathsf{A}^{-1}\mathbf{k}+i\mathbf{k}^T\mathsf{A}^{-1}\mathbf{J}\right), \tag{1.8}$$

$$\langle x_i\rangle_c = (A^{-1})_{ij}J_j, \quad \langle x_ix_j\rangle_c = (A^{-1})_{ij}, \quad \langle x_ix_jx_k\cdots\rangle_c = 0. \tag{1.9}$$

For the centered distribution ($\mathbf{J}=0$ so that $\langle x_ix_j\rangle_c = \langle x_ix_j\rangle$), we obtain Wick's theorem by applying the graphical relation, e.g.,

$$\langle x_ix_jx_kx_l\rangle = \langle x_ix_j\rangle\langle x_kx_l\rangle + \langle x_ix_k\rangle\langle x_jx_l\rangle + \langle x_ix_l\rangle\langle x_jx_k\rangle. \tag{1.10}$$

* * * * *

Consider the normalized deviation $Y = \sum_{i=1}^N (X_i-\mu)/\sqrt{N}$ of N independently and identically distributed (IID) variables taken from the distribution $p(x)$, with $\langle x\rangle_c = \mu$ and $\langle x^2\rangle_c = \sigma^2$.

$$\tilde{p}_Y(k) = \prod_{i=1}^N\left\langle\exp\left(i\frac{k}{\sqrt{N}}(X_i-\mu)\right)\right\rangle = \left(e^{-ik\mu}\tilde{p}(k/\sqrt{N})\right)^N \tag{1.11}$$

$$\ln\tilde{p}_Y(k) = N\ln(e^{-ik\mu}\tilde{p}(k/\sqrt{N})) = N\times\left(-\frac{1}{2}\sigma^2\frac{k^2}{N}+O\left(\frac{k^3}{N^{3/2}}\right)\right)\xrightarrow{N\to\infty}-\frac{1}{2}\sigma^2k^2. \tag{1.12}$$

Therefore, $p_Y(y)$ converges to a centered Gaussian distribution of variance σ^2.

- Independence is not necessary for convergence.

A special case with finite cumulants. Suppose all cumulants $\langle x\cdots x\rangle_c$ are finite. Consider the sum $S = \sum_{i=1}^N X_i$ of variables taken from the joint distribution $p(\mathbf{x})$.

$$\tilde{p}_S(k) = \langle e^{ikS}\rangle = \sum_{n=1}^{\infty}\frac{(ik)^n}{n!}\sum_{i_1,\cdots,i_n}\langle x_{i_1}\cdots x_{i_n}\rangle_c \implies \langle s^n\rangle_c = \sum_{i_1,\cdots,i_n}\langle x_{i_1}\cdots x_{i_n}\rangle_c. \tag{1.13}$$

The variable $Y = (S-\langle s\rangle_c)/\sqrt{N}$ satisfies $\langle y\rangle_c = 0$, and $\langle y^n\rangle_c = \langle s^n\rangle_c N^{-n/2}$ for $n\geq 2$. Now, if $\langle s^n\rangle_c = o(N^{n/2})$ for all $n>2$ and $\langle s^2\rangle_c = O(N)$, only $\langle y^2\rangle_c$ survives in the large-N limit.

$$\lim_{N\to\infty}p_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{y^2}{2\sigma^2}\right), \quad \sigma^2 = \lim_{N\to\infty}\frac{\langle s^2\rangle_c}{N}. \tag{1.14}$$

From $\sum_{ij}\langle x_ix_j\rangle_c = O(N)$ we see that there cannot be any long-range correlation, i.e., the number of j such that $\langle x_ix_j\rangle_c = O(1)$ for any fixed i must remain finite as $N\to\infty$.

- Finiteness of variance σ^2 is not necessary for convergence. If $\langle x^2\rangle_c = \infty$, the attraction domain of the Gaussian distribution is characterized by the following condition:

$$\lim_{y\to\infty} y^2 \frac{\int_{|x|>y} p(x)\mathrm{d}x}{\int_{|x|<y} x^2 p(x)\mathrm{d}x} = 0. \tag{1.15}$$

  The above condition is satisfied if $p(x)$ decreases as $|x|^{-3}$ or faster as $|x|\to\infty$, since

$$0 < y^2 \frac{\int_{|x|>y} p(x)\mathrm{d}x}{\int_{|x|<y} x^2 p(x)\mathrm{d}x} < \mathcal{C}y^2 \frac{\int_{|x|>y} |x|^{-3}\mathrm{d}x}{\int_{|x|<y} x^2 p(x)\mathrm{d}x} = \mathcal{C}y^2 \frac{1/y^2}{\int_{|x|<y} x^2 p(x)\mathrm{d}x} \xrightarrow{y\to\infty} 0.$$

- Distributions $p(x)$ that decreases as $1/|x|^{1+\alpha}$ have infinite variance for $0 < \alpha \le 2$. Simple power counting suggests $\tilde{p}(k)$ is singular at $k = 0$ with leading order $|k|^\alpha$. ????? The normalized deviation is now $Y = \sum_{i=1}^N (X_i - \mu)/N^{1/\alpha}$.

$$\ln \tilde{p}_Y(k) = N \ln \tilde{p}(k/N^{1/\alpha}) = N \times \left(-\mathcal{C}\frac{|k|^\alpha}{N} + O\left(\frac{|k|^{1+\alpha}}{N^{1+1/\alpha}}\right)\right) \xrightarrow{N\to\infty} -\mathcal{C}|k|^\alpha. \tag{1.16}$$

  It's been proved that $p_Y(y)$ converges to a stable distribution. A distribution is stable if any linear combination of two independent variables taken from the distribution has the same distribution.

1.3 Random Processes

This is simply the continuous version of the multi-variate random distributions. Instead of $\{X_i\}_{i=1}^N$, we now have a continuum of random variables $\{X(t)\}$. Each realization is a function $x(t)$.

The n-time densities $p_n(x_1, t_1; \cdots; x_n, t_n)$ are hierarchically marginalized:

$$\int p_n(x_1, t_1; \cdots; x_n, t_n)\mathrm{d}x_{s+1}\cdots\mathrm{d}x_n = p_s(x_1, t_1; \cdots; x_s, t_s). \tag{1.17}$$

The n-time averages are the continuous analogs of moments. We are often interested in the autocorrelation function $\kappa(t_1, t_2)$, the continuous analog of variance:

$$\kappa(t_1, t_2) = \langle X(t_1)X(t_2)\rangle - \langle X(t_1)\rangle\langle X(t_2)\rangle = \int x_1 x_2 [p_2(x_1, t_1; x_2, t_2) - p_1(x_1, t_1)p_1(x_2, t_2)]\mathrm{d}x_1\mathrm{d}x_2. \tag{1.18}$$

In the above expression, we have assumed that $X(t)$ is a real-valued signal. If it is complex-valued, $\kappa(t_1, t_2) = \langle X(t_1)^* X(t_2)\rangle - \langle X(t_1)^*\rangle\langle X(t_2)\rangle$.

- A stationary process is one with time-translation invariance, in which case $\kappa(t_1, t_2) = \kappa(t_2 - t_1)$.
- An ergodic process is one where statistical properties can be deduced from a single sample that's sufficiently long.
- Fact: If a stationary process has absolutely integrable $\kappa(\tau)$, it is ergodic in the mean: $\langle X\rangle = \bar{X}$.

From now on, we focus on stationary and mean-ergodic processes. It's often useful to consider the frequency spectrum of a process. However, any particular realization may not be periodic or square-integrable, so that "IR regularization" is needed.

Given a particular realization $x(t)$, there are two ways to regularize it:

- Fourier transform. Define the cut-off signal $x_T^{\mathrm{cut}}(t)$ that vanishes outside $[0, T]$.

$$\tilde{x}_T(\omega) = \int_0^T x_T^{\mathrm{cut}}(t) e^{i\omega t}\mathrm{d}t, \quad x_T^{\mathrm{cut}}(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \tilde{x}_T(\omega) e^{-i\omega t}\mathrm{d}\omega. \tag{1.19}$$

- Fourier series. Extend the $[0, T]$ segment periodically to obtain $x_T^{\mathrm{per}}(t)$.

$$a_n = \frac{1}{T}\int_0^T x_T^{\mathrm{per}}(t) e^{i\omega_n t}\mathrm{d}t, \quad x_T^{\mathrm{per}}(t) = \sum_{n=-\infty}^{+\infty} a_n e^{-i\omega_n t}. \tag{1.20}$$

Written symbolically,

$$X(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \tilde{X}(\omega)e^{-i\omega t}\mathrm{d}\omega = \sum_{n=-\infty}^{+\infty} A_n e^{-i\omega_n t}, \quad \omega_n = \frac{2\pi n}{T}, \tag{1.21}$$

$$\tilde{X}(\omega) = \int_0^T X(t)e^{i\omega t}\mathrm{d}t, \quad A_n = \frac{1}{T}\int_0^T X(t)e^{i\omega_n t}\mathrm{d}t, \quad A_n = \frac{1}{T}\tilde{X}(\omega_n). \tag{1.22}$$

We would take the $T \to \infty$ limit at the end of a calculation.

The mean of A_n is determined by:

$$\langle A_n\rangle = \frac{1}{T}\int_0^T \langle X\rangle e^{i\omega_n t}\mathrm{d}t = \langle X\rangle \delta_{0n}. \tag{1.23}$$

Notice that $A_0 \sim \tilde{X}(0)$ is not a random variable. Any particular realization a_0 is the temporal average of $x(t)$ over $[0, T]$, which (in the limit $T \to \infty$) is equal to $\langle X\rangle$ because of ergodicity.

In the vicinity of frequency ω_n, each mode carries energy proportional to $|A_n|^2$ and the density of modes are $\frac{T}{2\pi}$. Define the spectral density at that point to be $S(\omega_n) = T|A_n|^2$. In the continuous limit,

$$S(\omega) = \lim_{T\to\infty}\frac{1}{T}\langle|\tilde{X}(\omega)|^2\rangle. \tag{1.24}$$

Since $X(t)$ is stationary, $\langle X(t)\rangle \equiv \langle X\rangle$, hence different frequencies are uncorrelated.

$$\langle \tilde{X}(\omega)\tilde{X}(\omega')^*\rangle = 2\pi\delta(\omega-\omega')S(\omega). \tag{1.25}$$

On the other hand, for a centered process where $\langle X\rangle = 0$,

$$\kappa(\tau) = \langle X^*(t)X(t+\tau)\rangle = \lim_{T\to\infty}\int \langle \tilde{X}^*(\omega_1)\tilde{X}(\omega_2)\rangle e^{i\omega_1 t - i\omega_2(t+\tau)}\frac{\mathrm{d}\omega_1}{2\pi}\frac{\mathrm{d}\omega_2}{2\pi}. \tag{1.26}$$

Therefore, we obtain the Wiener-Khinchin theorem:

$$\kappa(\tau) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} S(\omega)e^{-i\omega\tau}\mathrm{d}\omega, \quad S(\omega) = \int_{-\infty}^{+\infty}\kappa(\tau)e^{i\omega\tau}\mathrm{d}\tau. \tag{1.27}$$

> **Example: Lorentzian spectral density.**
>
> $$\kappa(\tau) = \frac{1}{2\tau_0}e^{-|\tau|/\tau_0}, \quad S(\omega) = \frac{\omega_0^2}{\omega^2+\omega_0^2}, \quad \omega_0\tau_0 = 1. \tag{1.28}$$
>
> The limit $\tau_0 \to 0$ is the white noise limit, since $\kappa(\tau) \to \delta(\tau)$ and $S(\omega) \to 1$.

If the process is not centered, the spectral density is the sum of $S(\omega)$ for the centered process and $2\pi|\langle X\rangle|^2\delta(\omega)$, and the autocorrelation is the sum of $\kappa(\tau)$ for the centered process and $|\langle X\rangle|^2$.

2 Linear Thermodynamics of Irreversible Processes

2.1 Review of Equilibrium Thermodynamics

[Kardar] builds equilibrium thermodynamics from the bottom up by considering the consequences of the four laws one by one, introducing the concepts of temperature, energy, entropy difference, and absolute entropy. Here, we take the top-down approach by assuming the existence of an extensive state function $S = S(X_i)$, where $\{X_i\}$ are the extensive coordinates (energy, volume, particle numbers, etc.) subject to conservation laws. The intensive variables $\{F_i\}$ are defined such that $\mathrm{d}S = \sum_i F_i \mathrm{d}X_i$ along the equilibrium hypersurface.

> **Case study: Discontinuous phase transition.** Legendre transform vs. convexity?

2.2 Description of Systems in Local Equilibrium

If the system has two characteristic time/length scales, the microscopic degrees of freedom decouple from the macroscopic, thermodynamic desciption:

- If microscopic variables relax much faster than macroscopic variables, we may define instantaneous quantities.
- If the system is composed of a large number of microscopically large pieces, we may define local quantities.

Consider a system satisfying the local equilibrium hypothesis: the local, instantaneous entropy has the same functional form as the global equilibrium entropy. This is equivalent to the statement that the local instantaneous entropy is extensive, ruling out systems where long-range interactions are not negligible, such as the rarefied gas or any strong-coupling system. In mathematical terms, we consider systems where every thermodynamic variable Π satisfies:

$$\frac{\Delta\Pi}{\Pi} \ll \frac{\delta\Pi}{\Pi} \ll 1. \tag{2.1}$$

Here, $\delta\Pi/\Pi$ is the thermodynamic fluctuation over a short time/length, so that $\delta\Pi/\Pi \ll 1$ justifies the local, instantaneous description; $\Delta\Pi/\Pi$ is the variation over the entire time/length scale of the system, so that $\delta\Pi/\Pi \gg \Delta\Pi/\Pi$ justifies the neglection of long-range interactions.

Under the above assumptions, any extensive thermodynamic quantity in a continuum is described by a density field:

$$A(t) = \int_V a(\mathbf{r},t)\mathrm{d}V, \quad \frac{\mathrm{d}A}{\mathrm{d}t} = \int_V \sigma_A \mathrm{d}V - \int_{\partial V} \mathbf{J}_A \cdot \mathrm{d}\mathbf{\Sigma}, \quad \frac{\partial a}{\partial t} = \sigma_A - \nabla\cdot\mathbf{J}_A. \tag{2.2}$$

Here, $a(\mathbf{r},t)$ is the density per unit volume, $\mathbf{J}_A$ is the current density, and σ_A is the source density. On the other hand, if the system is composed of a finite number of subsystems, there is no distinction between a flux and a source (since topologically the subsystems are vertices of a graph). In that case, we define $I_A = \mathrm{d}A/\mathrm{d}t$ for each subsystem.

The claim is that local entropy production is always a bilinear form of affinities and fluxes, as long as the system is embracive enough that all extensive quantities are "transferred" across the system instead of being "generated" by external sources. We illustrate this with an example.

> **Example: Rod between two temperatures.** Choose the system to include the rod and the two heat reservoirs. It is described by entropy (density) functions $S_H(t)$, $S_C(t)$, $s(\mathbf{r},t)$.
>
> During the time interval $[t, t+\mathrm{d}t]$, a particular cell located at $\mathbf{r}$ changes from one equilibrium to another: $\partial s/\partial t = \sum_i F_i \partial x_i/\partial t$. However, nothing is said about the relation of equilibrium states between different cells, so that $\nabla\cdot\mathbf{J}_S = \nabla\cdot\sum_i F_i\mathbf{J}_i$.
>
> $$\sigma_S = \frac{\partial s}{\partial t} + \nabla\cdot\mathbf{J}_S = \sum_i \left(F_i \frac{\partial x_i}{\partial t} + F_i \nabla\cdot\mathbf{J}_i + (\nabla F_i)\cdot\mathbf{J}_i \right) = \sum_i \left(F_i\sigma_{X_i} + (\nabla F_i)\cdot\mathbf{J}_i \right). \tag{2.3}$$
>
> In the middle, the extensive quantities are locally conserved ($\sigma_i = 0$) so that $\sigma_S = \sum_i (\nabla F_i)\cdot\mathbf{J}_i$.
>
> At one end of the rod (say, the hot one), one might be tempted to say that $\sigma_E \neq 0$ due to the energy source that is the reservoir. However, since our system includes the reservoir, the heat is not "generated" by an external source, but is "transferred" from the reservoir. Therefore, the correct interpretation is that $\sigma_E = 0$ but we should include another term $-I_E^{(H)}\delta(\mathbf{r}-\mathbf{r}_H)$ in the equation to obtain $\sigma_S = -F_E^{(H)} I_E^{(H)} \delta(\mathbf{r}-\mathbf{r}_H) + \sum_i (\nabla F_i)\cdot\mathbf{J}_i$. For the reservoir itself, $\mathrm{d}S_H/\mathrm{d}t = F_E^{(H)} I_E^{(H)}$. The total energy is conserved:
>
> $$0 = \frac{\mathrm{d}}{\mathrm{d}t}\left(E^{(H)} + E^{(C)} + \int e(\mathbf{r})\mathrm{d}V \right) = I_E^{(H)} + I_E^{(C)} - \int (\nabla\cdot\mathbf{J}_E)\mathrm{d}V \tag{2.4}$$
>
> All heat emitted from the rod, $\int(\nabla\cdot\mathbf{J}_E)\mathrm{d}V$, is absorbed by the reservoirs.

2.3 Interlude: Thermal Fluctuations at Equilibrium

Consider the thermal fluctuations δX_i in the extensive coordinates. For now, we don't assume local equilibrium, but simply require the decoupling of microscopic degrees of freedom so that the extensive coordinates $\{X_i\}$ fully characterize the (out-of-equilibrium) macrostate.

What is the probability distribution of fluctuations? The Einstein formula answers this question ingeniously by inverting the Boltzmann entropy formula. For an isolated system, since $S = k_B \ln \Omega$, the number of microstates corresponding to the fluctuated macrostate is $\exp((S + \delta S)/k_B)$. The probability is proportional to the number of microstates, as all of them are assumed to be equally likely: $w \propto \exp(\delta S/k_B)$.

If the system is not isolated, the environment states are marginalized, and the probability is proportional to

$$\Omega(\{X\}) := \Omega_{\text{sys}}(\{X\})\Omega_{\text{env}}(\{X^{\text{tot}} - X\}) = \exp\left(\frac{S_{\text{sys}}(\{X\}) + S_{\text{env}}(\{X^{\text{tot}} - X\})}{k_B}\right). \tag{2.5}$$

Since the environment is huge, $X \ll X^{\text{tot}}$. To the lowest order,

$$\begin{aligned}
\Omega(\{X\}) &\approx \exp\left(\frac{S_{\text{sys}}(\{X\}) + S_{\text{env}}(\{X^{\text{tot}}\}) - \sum_i F_{i,\text{env}}(\{X^{\text{tot}}\})X_i}{k_B}\right) \\
&\propto \exp\left(\frac{S_{\text{sys}}(\{X\}) - \sum_i F_{i,\text{env}}(\{X^{\text{tot}}\})X_i}{k_B}\right) \\
&\approx \exp\left(\frac{S_{\text{sys}}(\{X\}) - \sum_i F_{i,\text{env}}(\{X^{\text{tot}} - X\})X_i}{k_B}\right) \\
&= \exp\left(\frac{S_{\text{sys}} - \sum_i F_{i,\text{env}}X_i}{k_B}\right).
\end{aligned} \tag{2.6}$$

The environment value $F_{i,\text{env}}$ does not fluctuate, so the probability of fluctuation

$$w \propto \exp\left(\frac{\delta S - \sum_i F_{i,\text{env}}\delta X_i}{k_B}\right) = \left(\frac{\sum_i (F_{i,\text{sys}} - F_{i,\text{env}})\delta X_i}{k_B}\right) = \left(\frac{\sum_i \delta F_i \delta X_i}{k_B}\right). \tag{2.7}$$

We may interpret the second term as the entropy flow from the environment to the system.

From this, we prove the claim $\langle \delta X_i \delta F_j \rangle = -k_B \delta_{ij}$.

$$\langle \delta X_i \delta F_j \rangle = k_B \int \delta X_i \frac{\partial w}{\partial \delta X_j} \prod_k \mathrm{d}(\delta X_k) = -k_B \int w \frac{\partial \delta X_i}{\partial \delta X_j} \prod_k \mathrm{d}(\delta X_k) = -k_B \delta_{ij}. \tag{2.8}$$

The first term vanishes because $\langle \delta F_j \rangle = 0$ by definition.

What about $\langle (\delta X_i)^2 \rangle$ and $\langle (\delta F_i)^2 \rangle$?

2.4 Kinetic Coefficients

The affinities $\mathcal{F}_i$ indicate the deviation/fluctuation from equilibrium, while the fluxes $\mathcal{J}_i$ record the flow/relaxation as a result. Here, we have flattened all indices such as the vector indices of ∇F and $\mathbf{J}$. As long as the local equilibrium hypothesis is valid, $\mathcal{J}_i(\mathbf{r}, t)$ only depends on the affinities at that place and time. To the lowest order,

$$\mathcal{J}_i(\mathbf{r}, t) = \mathcal{J}_i[\mathcal{F}] = 0 + \sum_k L_{ik}\mathcal{F}_k + O(\mathcal{F}^2), \quad L_{ik} := \frac{\partial \mathcal{J}_i}{\partial \mathcal{F}_k}. \tag{2.9}$$

Since $\sigma_S = \sum_{ik} L_{ik}\mathcal{F}_i\mathcal{F}_k = \sum_{ik} L_{(ik)}\mathcal{F}_i\mathcal{F}_k$ is non-negative-definite, the symmetric part satisfies:

$$L_{ii} \geq 0, \quad L_{ii}L_{kk} \geq L_{(ik)}L_{(ki)} = \frac{1}{4}(L_{ik} + L_{ik})^2. \tag{2.10}$$

Examples.

- Electrical conductivity (isotropic case): $\mathbf{J} = \sigma\mathbf{E}$.

$$\mathbf{J} = q\mathbf{J}_N = qL_{NN}\nabla\left(-\frac{\bar{\mu}}{T}\right) \overset{(*)}{=\!=} \frac{q^2}{T}L_{NN}\mathbf{E}. \tag{2.11}$$

  Here, the step $(*)$ is valid as long as the temperature and the carrier density (and hence μ) is uniform, so that $\nabla\bar{\mu} = \nabla(\mu + q\varphi) = -q\mathbf{E}$.

- Diffusion (isotropic case): $\mathbf{J}_N = -D\nabla n$.

$$\mathbf{J}_N = L_{NN}\nabla\left(-\frac{\bar{\mu}}{T}\right) \overset{(*)}{=\!=} -\frac{1}{T}\left.\frac{\partial\mu}{\partial n}\right|_T \nabla n. \tag{2.12}$$

  Here, the step $(*)$ is valid as long as the temperature is uniform.

- Thermal conductivity (isotropic case): $\mathbf{J}_E = -\kappa\nabla T$.

$$\mathbf{J}_E = L_{EE}\nabla\left(-\frac{1}{T}\right) = \frac{1}{T^2}L_{EE}\nabla T. \tag{2.13}$$

Due to time-reversal invariance at the microscopic level, the kinetic coefficients satisfy the Onsager reciprocity relations:

$$L_{ik} = (-)^{\mathsf{T}_i+\mathsf{T}_k}L_{ki}, \quad \text{where } \mathsf{T}_{i,k} \text{ are time-reversal parities of } \mathcal{J}_{i,k}. \tag{2.14}$$

Since they reflect the microscopic aspects of the system, the Onsager relations cannot be proved within the framework of thermodynamics. Here, we merely attempt a justification of the statement.

For simplicity, consider a time-reversal parity-even system without magnetic fields or angular momenta. By time-reversal invariance and stationarity,

$$\langle\delta X_i(0)\delta X_k(t)\rangle = \langle\delta X_i(t)\delta X_k\rangle \implies \langle\delta X_i\delta\dot{X}_k\rangle = \langle\delta\dot{X}_i\delta X_k\rangle. \tag{2.15}$$

By Onsager's regression hypothesis (a consequence of the fluctuation-dissipation theorem unprovable in thermodynamics), fluctuations relax by following the same lows as fluxes:

$$\delta\dot{X}_i = \sum_j L_{ij}\delta F_j \implies \sum_j L_{kj}\langle\delta X_i\delta F_j\rangle = \sum_j L_{ij}\langle\delta F_j\delta X_k\rangle \implies L_{ki} = L_{ik}. \tag{2.16}$$

An important consequence is the minimum entropy production theorem, which holds for systems that are time-reversal parity-even ($L_{ik} = L_{ki}$) in the linear regime. Consider the time derivative of the total entropy production $P_S = \int\sigma_S\mathrm{d}V$ of a system kept out of equilibrium by time-independent constraints:

$$\begin{aligned}
\frac{\mathrm{d}P_S}{\mathrm{d}t} &= \int\frac{\partial\sigma_S}{\partial t}\mathrm{d}V \\
&= 2\int\sum_{ik}L_{ik}(\nabla F_i)\cdot\nabla\left(\frac{\partial F_k}{\partial t}\right)\mathrm{d}V \\
&= 2\int\sum_k\mathbf{J}_k\cdot\nabla\left(\frac{\delta\dot{S}}{\delta x_k}\right)\mathrm{d}V \\
&= -2\int\sum_k(\nabla\cdot\mathbf{J}_k)\frac{\delta\dot{S}}{\delta x_k}\mathrm{d}V \\
&= 2\int\sum_{jk}\frac{\partial x_k}{\partial t}\frac{\delta^2 S}{\delta x_k\delta x_j}\frac{\partial x_j}{\partial t}\mathrm{d}V,
\end{aligned} \tag{2.17}$$

where on the second line, the symmetry $L_{ik} = L_{ik}$ is crucial.

From thermodynamics, we know that the entropy function is convex, i.e., $\delta^2 S/\delta x_k\delta x_j$ is non-positive definite. Hence, $\mathrm{d}P_S/\mathrm{d}t \leq 0$. On the other hand, P_S itself is non-negative. The time-evolution of P_S is therefore towards a stationary limit, where the entropy production P_S reaches a minimum.