Soft Matter Dynamics

Lecturer: Rob Jack

Lent Term 2020

These notes[1] cover the four lectures delivered by Rob Jack in the 2020 Lent term as part of the Part III course *Theoretical Physics of Soft Condensed Matter.* They present an introduction to non-equilibrium steady states and stochastic thermodynamics. The 'classical' review of the subject, which may be of interest, is that by Udo Seifert (2012 Rep. Prog. Phys. **75** 126001).

Contents

1 Introduction

1.1 Non-equilibrium Steady States

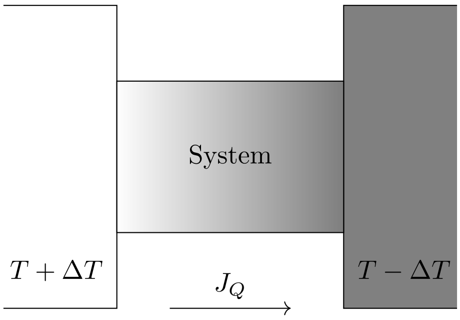
A thermodynamic system is in a steady state if its state variables such as S, N, V are unchanging in time. Equilibrium is a type of steady state for which no external work is being done and there is no dissipation. Crucially, a system in equilibrium has time-reversal symmetry (TRS), meaning evolution of the system is the same whether time is increasing or decreasing.

It is possible for a system to be in a steady state without being in equilibrium, however. Then there is an exchange of energy across the system's boundary, dissipation occurs and external work

[1] Typeset by P. Fowler-Wright (MASt, 2020).

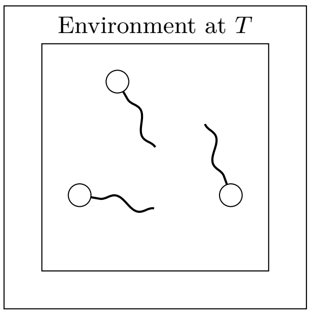
may be done. TRS is necessarily broken. The system is held from equilibrium by a 'driving' source that balances dissipation.

Example: Soft material between heat baths



A heat J_Q flows from the hot reservoir through the cold reservoir, maintaining a non-equilibrium steady state (heat in $\equiv$ dissipation). The lack of TRS is simply the fact that heat flows from hot to cold, never from cold to hot. No work is being done.

Example: Bacterial culture (active material)



Bacteria consume food in the culture and swim. Chemical energy is converted into kinetic energy (mechanical motion) and finally lost to the surroundings as heat. Provided the food source is replenished, this is sustainable. The driving is in the bulk.

Rather than focus on specific examples, we wish to establish the basic principles governing non-equilibrium steady states. We will be guided by a simple stochastic system: a particle with position $\mathrm{x}(t)$ obeying the Langevin equation (refer to the lectures by M. Cates)

$$\dot{\mathrm{x}}(t) = M_1 \mathrm{F}\left(\mathrm{x}(t)\right) + \sqrt{2M_1 T}\eta(t) \qquad (k_B = 1) \tag{1.1}$$

where F is the force on the particle, M_1 a mobility and

$$\langle \eta(t) \rangle = 0, \quad \langle \eta(t)\eta(t') \rangle = \delta(t - t') \tag{1.2}$$

describes unit white noise (Gaussian). This models a particle moving in a high friction environment at temperature T. For simplicity we take $M_1 = 1$. In general x is a vector although we do not use boldface notation. By increasing the number of components of this vector we could encode information for many particles and indeed the conclusions drawn from this simple system will have application to larger number of particles.

2 Laws of Stochastic Thermodynamics

The basic fact we start from is that the heat transfer *into* the environment for our one particle system is

$$Q(\tau) = \int_0^\tau \mathrm{F}\left(\mathrm{x}(t)\right) \cdot \dot{\mathrm{x}}(t) dt \tag{2.1}$$

It is useful to introduce the *instantaneous* heat flux

$$q(t) = \mathrm{F} \cdot \dot{\mathrm{x}} \tag{2.2}$$

The equilibrium case corresponds to F being the negative gradient of a potential (conservative), $\mathrm{F} = -\nabla V$, and

$$\begin{aligned} Q(\tau) &= -\int_0^\tau \nabla V \cdot \dot{\mathrm{x}} && (2.3)\\ &= -\int_0^\tau \frac{d}{dt} V\left(\mathrm{x}(t)\right) && (2.4)\\ &= -\left[V\left(\mathrm{x}(t)\right) - V\left(\mathrm{x}(0)\right)\right] && (2.5) \end{aligned}$$

i.e. the heat transfer into the environment is equal to the decrease in energy $-\Delta V$ of the particle (conservation of energy).

It is natural to consider deviation from this case by writing $\mathrm{F} = \mathrm{F}^{\mathrm{ext}} - \nabla V$ where $\mathrm{F}^{\mathrm{ext}}$ is some non-conservative force. Then

$$\begin{aligned} Q(\tau) &= \int_0^\tau \mathrm{F}^{\mathrm{ext}}\left(\mathrm{x}(t)\right) \cdot \dot{\mathrm{x}}(t) dt - \Delta V(\tau) && (2.6)\\ &= W(\tau) - \Delta V(\tau) && (2.7) \end{aligned}$$

where $W(\tau)$ defines the work done by the external force *on* the system. This statement of energy conservation (Heat lost − Work received = Decrease in Energy), which applies in the general non-equilibrium case, is the *first law* of stochastic thermodynamics. Note that we are beyond the remit of standard thermodynamics since Q, W and ΔV are random variables, dependent on the path $\mathrm{x}(t)$ (a stochastic process).

For a second law concerning entropy, recall that the probability of $\mathrm{x}(t)$ realising a particular trajectory follows[2]

$$\mathbb{P}_F\left[\mathrm{x}(t)\right] \propto \exp\left(-\int_0^\tau \frac{\left(\dot{\mathrm{x}} - \mathrm{F}\right)^2}{4T} dt\right)` \tag{2.8}$$

Implicit is a conditioning on the initial state:

$$\mathbb{P}_F\left[\mathrm{x}(t)\right] = \mathrm{Prob}\big[\text{path } \mathrm{x}(t) \mid \text{initial condition } \mathrm{x}(0)\big] P_{\mathrm{ic}}\left[\mathrm{x}_0\right] \tag{2.9}$$

where $P_{\mathrm{ic}}\left[\mathrm{x}_0\right]$ denotes the probability of initial condition $x(0)$.

The subscript F distinguishes this forward probability from the *backward*

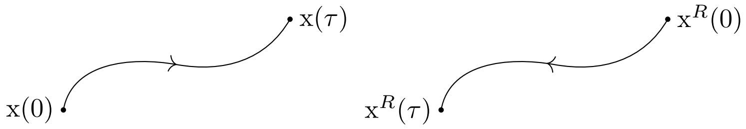
$$\mathbb{P}_B\left[\mathrm{x}(t)\right] \equiv \mathbb{P}_F\left[\mathrm{x}^R(t)\right] \tag{2.10}$$

[2]This is determined by the probability of observing the noise $\eta(t)$ associated with the path.

with $\mathrm{x}^R(t) := \mathrm{x}(\tau - t)$ defining the *time-reversed path.* This probability also follows a Gaussian distribution, but with the rate of change of x negated, and is conditioned on the *final* state (initial state of x^R):

$$\mathbb{P}_B\left[\mathrm{x}(t)\right] \propto \exp\left(-\int_0^\tau \frac{(-\dot{\mathrm{x}} - \mathrm{F})^2}{4T} dt\right) \tag{2.11}$$

$$= \mathrm{Prob}\left[\mathrm{x}^R(t)|\mathrm{x}^R(0)\right] P_{\mathrm{ic}}\left[\mathrm{x}_0^R\right] \tag{2.12}$$



It turns out that the normalisation constants for $\mathbb{P}_F$, $\mathbb{P}_B$ are the same so that

$$\log \frac{\mathbb{P}_F\left[\mathrm{x}(t)\right]}{\mathbb{P}_B\left[\mathrm{x}(t)\right]} = -\int_0^\tau \frac{(\dot{\mathrm{x}} - F)^2 - (\dot{\mathrm{x}} + F)^2}{4T} dt \tag{2.13}$$

$$= \frac{1}{T}\int_0^\tau \mathrm{F} \cdot \dot{\mathrm{x}} dt \tag{2.14}$$

$$= \frac{Q(\tau)}{T} \tag{2.15}$$

So that the ratio of path probabilities is the heat transferred to the environment, scaled by the temperature. This last quantity is identified as the entropy change of the environment ($k_B = 1$):

$$\Delta S_{\mathrm{env}} = \frac{Q}{T} = \log \frac{\mathbb{P}_F\left[\mathrm{x}(t)\right]}{\mathbb{P}_B\left[\mathrm{x}(t)\right]} \tag{2.16}$$

To establish that ΔS_{end} is always non-negative (the second law), firstly use conditioned form of the probabilities (2.9), (2.12):

$$\frac{Q(\tau)}{T} = \log \frac{\mathrm{Prob}\left[\mathrm{x}(t)\right]}{\mathrm{Prob}\left[\mathrm{x}^R(t)\right]} + \log P_{\mathrm{ic}}\left[\mathrm{x}(0)\right] - \log P_{\mathrm{ic}}\left[\mathrm{x}(\tau)\right] \qquad \left(\mathrm{x}^R(0) = \mathrm{x}(\tau)\right) \tag{2.17}$$

Now a steady state is one in which the statistics of $\mathrm{x}(t)$ are unchanging (time-translation invariant), meaning

$$\langle \log P_{\mathrm{ic}}\left(\mathrm{x}(0)\right)\rangle = \langle \log P_{\mathrm{ic}}\left(\mathrm{x}(\tau)\right)\rangle \tag{2.18}$$

Hence the average of (2.17) reduces to

$$\langle Q(\tau)\rangle = \int \mathrm{Prob}\left[\mathrm{x}(t)\right] \log \frac{\mathrm{Prob}\left[\mathrm{x}(t)\right]}{\mathrm{Prob}\left[\mathrm{x}^R(t)\right]} \mathcal{D}\left[\mathrm{x}(t)\right] \tag{2.19}$$

in the steady state.[3] The integral here is to be taken over all possible paths from $t = 0$ to $t = \tau$.

We now appeal to the fact that, for normalised probability densities p & q,

$$D_{\mathrm{KL}}\left(p//q\right) := \int p(x) \log \frac{p(x)}{q(x)} dx \geq 0 \tag{2.20}$$

[3]In what follows, $\langle\cdot\rangle$ always denotes a steady state average.

with equality if and only if $p = q$ almost everywhere (i.e. up to a set of measure 0). The quantity appearing on the left-hand side is known as the Kullback-Leibler divergence and is a measure of the extent to which the two probability distributions differ. In our case the inequality implies

$$\langle Q(\tau)\rangle \geq 0 \leftrightarrow \langle \Delta S_{\text{env}}\rangle \geq 0 \tag{2.21}$$

Thus, in a steady state heat flows (up to fluctuations) *from* the system *to* the environment, and the entropy of the environment, and so the universe ($\Delta S_{\text{sys}} = 0$ in the steady state) is never decreasing. This is the second law of stochastic thermodynamics. It is worth noting:

1. $\langle Q\rangle = 0$ only if $\text{Prob} = \text{Prob}^*$ where $\text{Prob}^*[\text{x}(t)] := \text{Prob}\left[\text{x}^R(t)\right]$ is a probability distribution we will continue to encounter. This is the time-reversal symmetric case, for which heat cannot flow (assuming no additional complications such as magnetic fields).

2. As

$$Q(\tau) = \int_0^\tau q(t)dt \quad \text{with} \quad q = \text{F}\cdot\dot{\text{x}} \tag{2.22}$$

We have

$$\langle Q(\tau)\rangle = \tau\,\langle \text{F}\cdot\dot{\text{x}}\rangle > 0 \tag{2.23}$$

if TRS is broken.

We now have the basis for a theory of heat and work in a system with random noise. The first law is an *equality* for individual paths, the second an *inequality* for the *average* heat flow.

Next we study the fluctuations $Q(\tau)$ from its average.

3 Fluctuations & Large Deviations of Q

So far we have described the first moment of

$$Q(\tau) = \int_0^\tau q(t)dt \tag{3.1}$$

Now we consider the second moment

$$\begin{aligned}\text{Var}\,(Q(\tau)) &= \int_0^\tau \int_0^\tau \langle q(t)q(t')\rangle - \langle q(t)\rangle\,\langle q(t')\rangle\,dtdt' \quad (3.2)\\ &= \int_0^\tau \int_0^\tau c_q(t-t')dtdt' \quad (3.3)\end{aligned}$$

where we used time translation-invariance in the steady state (the integrand can only depend on the difference $t-t'$). We investigate this quantity at *long times* τ. Our basic working assumption is that the correlations c_q decay as

$$c_q(t) \to 0 \quad \text{for} \quad |t| \to \infty \tag{3.4}$$

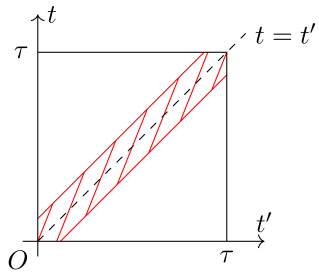
Consequently, the non-negligible contribution to the integral comes from $|t-t'| = O(1)$ and[4]

$$\text{Var}\,(Q(\tau)) \simeq \tau\chi_q \quad \text{with} \quad \chi_q = \int_{-\infty}^{\infty} c_q(t)dt \quad (\tau \gg 1) \tag{3.5}$$

[4]At each t' (vertical line in the plot shown) the inner integral in $\int_0^\tau \int_{-t'}^{\tau-t'} c_q(u)dudt'$ is determined by the same set of values close to the line $t = t'$ and (according to the decay of c_q) evaluates to $\int_{-\infty}^{\infty} c_q(t)dt$.

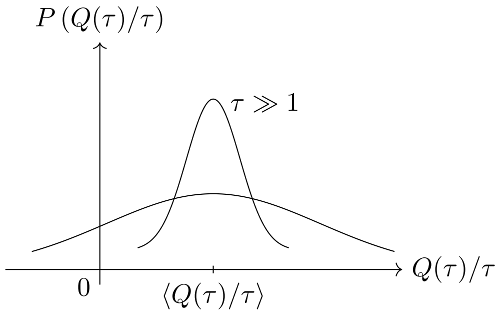
where $\simeq$ indicates a form of asymptotic approximation, as $\tau \to \infty$, and χ_q is a constant (we assume this integral to exist). So while $\langle Q(\tau)/\tau \rangle = O(1)$ as $\tau \to \infty$ (see (2.23)),

$$\mathrm{Var}\,(Q(\tau)/\tau) = \frac{1}{\tau^2}\mathrm{Var}\,(Q(\tau)) \simeq \frac{\chi_q}{\tau} \to 0 \quad \text{as} \quad \tau \to \infty \tag{3.6}$$



It is convenient to work with the scaled $Q(\tau)/\tau$. This is the average heat flux $q(t)$ per unit time over $[0, \tau]$ which coincides with the steady state expectation value of q at long times.[5]

As $\tau \to \infty$ then, the distribution of $Q(\tau)/\tau$ becomes sharply peaked (width $\sim \tau^{-1/2}$) about this mean value:



As illustrated by the above plot, at any τ there is a non-zero probability of observing values $Q(\tau) < 0$ i.e. there exist paths with 'negative dissipation' where heat flows back into the system. Large deviation theory provides tools to address the likelihood of such rare events.

3.1 Large Deviation Results

Large deviation theory may be used to establish that

$$\mathrm{Prob}\left(\frac{Q(\tau)}{\tau} \simeq q\right) \simeq \exp\left(-\tau I(q)\right) \tag{LDP}$$

where $I(q)$ is a non-negative convex function with a unique zero at

$$q = \langle Q(\tau)/\tau \rangle = \langle q(t) \rangle \tag{3.7}$$

$I(q)$ is known as the rate-function and (LDP) a large deviation principle for $Q(\tau)/\tau$. Two approximate signs appear. Avoiding the technical details, the first indicates that the probability

[5] That is to say the system is *ergodic*. For more on ergodicity and the analysis of time-averaged quantities, see Rob's recent article (2020 Eur. Phys. J. B **93**, 74).

is of $Q(\tau)/\tau$ taking some small but finite range of values centred on q, and the second that the leading order behaviour of this probability is asymptotically equivalent to $\exp(-\tau I(q))$, as $\tau \to \infty$. So that I is non-negative reflects the fact that the probability of a set of non-zero measure cannot diverge and $I(\langle q \rangle) = 0$ that the mean value is observed with probability 1 at long times. In contrast, the probability of any $q \neq \langle q \rangle$ decays exponentially.

Example (Problem Sheet Q1a)
A simple example of a LDP is when $Q(\tau)/\tau$ obeys a central limit theorem:

$$\text{Prob}\left(\frac{Q(\tau)}{\tau} \simeq q\right) = \frac{1}{\sqrt{2\pi\chi_q/\tau}} \exp\left(-\frac{(q-\langle q\rangle)^2}{2(\chi_q/\tau)}\right) \tag{3.8}$$

so that

$$I(q) = \frac{1}{2\chi_q}(q-\langle q\rangle)^2 \tag{3.9}$$

To summarise, large deviation theory tells us that negative dissipation events are exponentially rare in a sense that can be made rigorous.

3.2 Gallavotti-Cohen Fluctuation Theorem

Returning to the entropy change (2.17), note that the second term of

$$\frac{\beta Q(\tau)}{\tau} = \frac{1}{\tau}\log\frac{\text{Prob}[\mathrm{x}(t)]}{\text{Prob}^*[\mathrm{x}(t)]} + \frac{1}{\tau}\left(\log P_{\text{ic}}[\mathrm{x}(\tau)] - \log P_{\text{ic}}[\mathrm{x}(0)]\right) \tag{3.10}$$

$$\left(\beta \equiv \frac{1}{T},\ \text{Prob}^*[\mathrm{x}(t)] \equiv \text{Prob}\left[\mathrm{x}^R(t)\right]\right)$$

becomes negligible at long times ($P_{\text{ic}}[\mathrm{x}(\tau)]$ ought to be bounded). Hence under $\tau \to \infty$,

$$\text{Prob}[\mathrm{x}(t)] \simeq e^{\beta Q(\tau)}\text{Prob}^*[\mathrm{x}(t)] \tag{3.11}$$

Formally then

$$\begin{aligned}
\text{Prob}\left(\frac{Q(\tau)}{\tau} \simeq q\right) &\text{“=”} \int \text{Prob}[\mathrm{x}(t)]\,\overbrace{\delta\left(q - \frac{Q(\tau)}{\tau}\right)}^{\substack{\text{select value(s)}\\ \text{close to } q}}\mathcal{D}[\mathrm{x}(t)] && (3.12)\\
&\simeq \int e^{\beta Q(\tau)}\text{Prob}^*[\mathrm{x}(t)]\,\delta\left(q - \frac{Q(\tau)}{\tau}\right)\mathcal{D}[\mathrm{x}(t)] && (3.13)\\
&= e^{\beta q\tau}\int \text{Prob}^*[\mathrm{x}(t)]\,\delta\left(q - \frac{Q(\tau)}{\tau}\right)\mathcal{D}[\mathrm{x}(t)] && (3.14)\\
&= e^{\beta q\tau}\text{Prob}\left(\frac{Q_B(\tau)}{\tau} \simeq q\right) && (3.15)
\end{aligned}$$

where Q_B is the heat transfer for the reverse path x^R. This is of course the negative of the transfer for the corresponding forward path and so

$$\text{Prob}\left(\frac{Q(\tau)}{\tau} \simeq q\right) = e^{\beta q\tau}\text{Prob}\left(\frac{Q(\tau)}{\tau} \simeq -q\right) \tag{3.16}$$

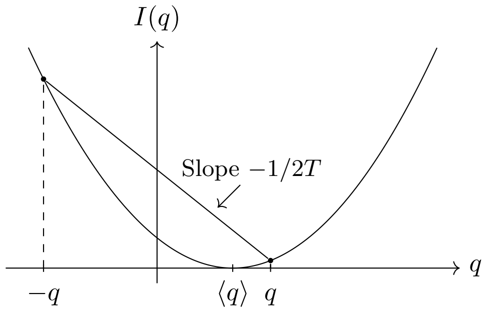


Figure 1: The Gallavotti-Cohen Fluctuation Theorem. That the rate function

$$I(q) = \lim_{\tau\to\infty} \log \mathrm{Prob}\left(Q(\tau)/\tau \simeq q\right)$$

satisfies (3.18) means that the slope of the connecting the graph at q and $-q$ is, at long times, $-1/2T$, with the negative value (heat flow) being exponentially less probable.

This tells us that the rate-function appearing in the (LDP) for $Q(\tau)/\tau$ saitsifes

$$I(q) = -\beta q + I(-q) \tag{3.17}$$

or

$$I(-q) = \frac{q}{T} + I(q) \tag{3.18}$$

This *equality* between rate functions is a version of the *Gallavotti-Cohen Fluctuation Theorem* (the steps leading to (3.18) can be made rigorous). It provides a connection between entropy production (heat flow), time-reversal and path probability. A whole host of similar fluctuation theorems exist for steady state processes.

4 The Adjoint Process

The distribution $\mathrm{Prob}^*[\mathrm{x}(t)]$ for the reversed path has appeared several times so far. We call this the *adjoint process* (in the sense of a Markov process). How to characterise this process? We know that a typical path from $\mathrm{Prob}[\mathrm{x}(t)]$ is a typical trajectory of the system, so a typical path from $\mathrm{Prob}^*[\mathrm{x}(t)]$ is just the time-reverse of such a path (where heat flows from cold to hot, for example).

Being a bit more specific, one may ask whether there is a Langevin equation for sample paths of the adjoint process. It turns out there is, but deriving it is not a simple matter of reversing the 'arrow of time' and will require some additional technology.

4.1 Fokker-Planck & Hamilton-Jacobi Equations

The following are established in the problem sheet.

1. If $P(\mathrm{x}, t)$ is the (instantaneous) probability density for position x at time t then

$$\frac{\partial P}{\partial t} = -\nabla\cdot\left(\mathrm{F}P - T\nabla P\right) \tag{FP}$$

   This is the Fokker-Planck (FP) equation.

2. The stationary (steady state) solution $\partial_t P = 0$ to (FP) is

$$P_\infty(\mathrm{x}) = \frac{e^{-U(\mathrm{x})}}{Z} \tag{4.1}$$

where Z is a normalisation constant and $U(\mathrm{x})$ obeys a Hamilton-Jacobi (HJ) equation

$$\nabla U \cdot (T\nabla U + \mathrm{F}) - \nabla \cdot (\mathrm{F} + T\nabla U) = 0 \tag{HJ}$$

The equilibrium case corresponds to $\mathrm{F} = -\nabla V$, $U = V/T$.

3. A *propagator* $G(\mathrm{x}, \mathrm{y}, t)$ describes the probability density of being at (y, t) given position x at $t = 0$. This obeys the FP equation

$$\frac{\partial G(\mathrm{x}, \mathrm{y}, t)}{\partial t} = -\nabla_\mathrm{y} \cdot (\mathrm{F}(\mathrm{y})G(\mathrm{x}, \mathrm{y}, t) - T\nabla_\mathrm{y} G(\mathrm{x}, \mathrm{y}, t)) \tag{GFP}$$

but also the *backward* FP equation

$$\frac{\partial G(\mathrm{x}, \mathrm{y}, t)}{\partial t} = \mathrm{F}(\mathrm{x}) \cdot \nabla_\mathrm{x} G(\mathrm{x}, \mathrm{y}, t) + T\nabla^2_\mathrm{x} G(\mathrm{x}, \mathrm{y}, t) \tag{GFP2}$$

To characterise the adjoint process, we derive its propagator $G^*(\mathrm{x}, \mathrm{y}, t)$, the probability of being at (y, t) on a reverse path given x at $t = 0$.

4.2 Adjoint Propagator

In the steady state, the probability of going from x to y in a time t in the adjoint process equals the probability going from y to x in same time for the original process (the paths are simply reversed). Hence the two propagators are related via the stationary solution as

$$G^*(\mathrm{x}, \mathrm{y}, t)\, P_\infty(\mathrm{x}) = P_\infty(\mathrm{y}) G(\mathrm{y}, \mathrm{x}, t) \tag{4.2}$$

Taking the time derivative and using (4.1) & (GFP) (note $\mathrm{x} \leftrightarrow \mathrm{y}$),

$$\begin{aligned}
\frac{\partial G^*(\mathrm{x}, \mathrm{y}, t)}{\partial t} &= \frac{P_\infty(\mathrm{y})}{P_\infty(\mathrm{x})} \frac{\partial G(\mathrm{y}, \mathrm{x}, t)}{\partial t} && (4.3)\\
&= \exp(U(\mathrm{x}) - U(\mathrm{y}))\nabla_\mathrm{x} \cdot (-\mathrm{F}(\mathrm{x}) + T\nabla_\mathrm{x})\, G(\mathrm{y}, \mathrm{x}, t) && (4.4)\\
&= \exp\big(U(\mathrm{x}) - \cancel{U(\mathrm{y})}\big)\nabla_\mathrm{x} \cdot (-\mathrm{F}(\mathrm{x}) + T\nabla_\mathrm{x}) \exp(\cancel{U(\mathrm{y})} - U(\mathrm{x}))G^*(\mathrm{x}, \mathrm{y}, t) && (4.5)
\end{aligned}$$

where we used (4.2) (again) and cancelled $U(\mathrm{y})$ in the exponents since the gradients are in x. Denoting $\nabla_\mathrm{x} U(\mathrm{x}) = U'(\mathrm{x})$, we can bring the second exponential through both gradients:

$$\begin{aligned}
\frac{\partial G^*(\mathrm{x}, \mathrm{y}, t)}{\partial t} &= e^{U(\mathrm{x})}\nabla_\mathrm{x} \cdot \left[e^{-U(\mathrm{x})} (-\mathrm{F}(\mathrm{x}) - TU'(\mathrm{x}) + T\nabla_\mathrm{x})\, G^*(\mathrm{x}, \mathrm{y}, t)\right] && (4.6)\\
&= [\nabla_\mathrm{x} - U'(\mathrm{x})]\,[-\mathrm{F}(\mathrm{x}) - TU'(\mathrm{x}) + T\nabla_\mathrm{x}]\, G^*(\mathrm{x}, \mathrm{y}, t) && (4.7)\\
&= T\nabla^2_\mathrm{x} G^*(\mathrm{x}, \mathrm{y}, t) - (\mathrm{F}(\mathrm{x}) + 2TU'(\mathrm{x})) \cdot \nabla_\mathrm{x} G^*(\mathrm{x}, \mathrm{y}, t)\\
&\quad + G^*(\mathrm{x}, \mathrm{y}, t) \underbrace{\left[-\nabla_\mathrm{x} \cdot \mathrm{F}(\mathrm{x}) - T\nabla^2_\mathrm{x} U(\mathrm{x}) + \mathrm{F}(\mathrm{x}) \cdot \nabla_\mathrm{x} U(\mathrm{x}) + T\left(\nabla_\mathrm{x} U(\mathrm{x})\right)^2\right]}_{=0 \text{ by (HJ)}} && (4.8)
\end{aligned}$$

Therefore G^* obeys the backwards FPE

$$\frac{\partial G^*(\mathrm{x},\mathrm{y},t)}{\partial t} = T\nabla_{\mathrm{x}}^2 G^*(\mathrm{x},\mathrm{y},t) - \overbrace{\left(\mathrm{F}(\mathrm{x}) + 2TU'(\mathrm{x})\right)}^{\text{“F(x)”}} \cdot \nabla_{\mathrm{x}} G^*(\mathrm{x},\mathrm{y},t) \tag{4.9}$$

associated with the Langevin

$$\dot{\mathrm{x}}(t) = -\mathrm{F}(\mathrm{x}) - 2T\nabla U(\mathrm{x}) + \sqrt{2T}\eta(t) \tag{4.10}$$

This characterises the adjoint process. On comparison with the Langevin (2.19) of the original process, we see that, in addition to changing sign, the forcing term has acquired a new part $2T\nabla U(\mathrm{x})$ dependent on the stationary solution.

The equilibrium case $\mathrm{F} = -\nabla V$, $U = V/T$, gives

$$\dot{\mathrm{x}}(t) = -\nabla V(\mathrm{x}) + \sqrt{2T}\eta(t) \tag{4.11}$$

for both processes, consistent with TRS.

Note that, in (4.10), $U(\mathrm{x})$ is an unknown that must be obtained by other means – there is no explicit form for this function for a general non-equilibrium steady state.

5 Probability & Particle Currents

So far we have characterised dissipation and TRS for steady state systems described by particle coordinates $\mathrm{x}(t)$. We now discuss the generalisation to fields and at the same time consider systems with arbitrary initial states (not necessarily steady). We develop the formalism for many non-interacting particles, with a glimpse at the interacting case towards the end.

5.1 Average Density and Current

The empirical density for a system of non-interacting particles is defined by assigning a delta peak to each particle,

$$\hat{\rho}(\mathrm{x},t) = \sum_{i=1}^{N} \delta\left(\mathrm{x} - \mathrm{x}_i(t)\right) \tag{5.1}$$

and the empirical current by weighting each peak according to the particle's velocity,

$$\hat{J}(\mathrm{x},t) = \sum_{i=1}^{N} \dot{\mathrm{x}}_i(t)\delta\left(\mathrm{x} - \mathrm{x}_i(t)\right) \tag{5.2}$$

These are random (path dependent) quantities. Their spatial *averages* $\bar{\rho}(\mathrm{x},t)$ and $\bar{J}(\mathrm{x},t)$ are *not.* The first is obtained from the probability density for a single particle in the obvious way:

$$\bar{\rho}(\mathrm{x},t) = NP(\mathrm{x},t) \tag{5.3}$$

That $P(\mathrm{x},t)$ obeys the (FP) equation means that $\bar{\rho}$ satisfies

$$\frac{\partial \bar{\rho}}{\partial t} = -\nabla \cdot \left(-T\nabla\bar{\rho} + \bar{\rho}\mathrm{F}\right) \tag{5.4}$$

which we recognise as a continuity equation

$$\frac{\partial \bar{\rho}}{\partial t} = -\nabla \cdot \bar{J} \tag{5.5}$$

with average current

$$\bar{J} = -T\nabla\bar{\rho} + \bar{\rho}\,\mathrm{F} \tag{5.6}$$

It follows that the current for the adjoint process (the time-reversed current) is

$$\bar{J}^{*} = -T\nabla\bar{\rho} - (\mathrm{F} + 2T\nabla U)\,\bar{\rho} \qquad (\mathrm{F} \to -\mathrm{F} - T\nabla U \text{ cf. } (4.10)) \tag{5.7}$$

It is natural to decompose these currents as

$$\bar{J} = J^{S} + J^{A}, \;\; \bar{J}^{*} = J^{S} - J^{A} \tag{5.8}$$

where

$$J^{S} = -\left(T\nabla + T\nabla U\right)\bar{\rho} \quad \text{and} \quad J^{A} = \left(\mathrm{F} + T\nabla U\right)\bar{\rho} \tag{5.9}$$

The physical significance is that J^S and J^A are respectively symmetric and antisymmetric under time reversal so that, in equilibrium, $J^A = 0$ and $\bar{J} = \bar{J}^* = J^S$. Several important results are derived from these currents and so they are useful to define even if they cannot be written explicitly (recall U is not generally known). We demonstrate three of these results.

1. Orthogonality relation for J^S and J^A.

$$\int \frac{J^{S} \cdot J^{A}}{\bar{\rho}} d\mathrm{x} = 0 \quad \text{for all } \bar{\rho} \tag{5.10}$$

2. A free energy functional and characterisation of the steady state.

$$J^{S} = -\bar{\rho}\,T\nabla\frac{\delta\mathcal{F}}{\delta\bar{\rho}} \tag{5.11}$$

where

$$\mathcal{F}\left[\bar{\rho}\right] = \overbrace{\int \bar{\rho}\log\frac{\bar{\rho}}{NP_{\infty}} d\mathrm{x}}^{= ND_{\mathrm{KL}}\left(\frac{\bar{\rho}}{N}//P_{\infty}\right)} \tag{5.12}$$

By the properties of the KL divergence (2.20),

$$\mathcal{F}\left[\bar{\rho}\right] \geq 0 \tag{5.13}$$

with equality in the steady state ($\bar{\rho} = NP(\mathrm{x}, t) = NP_{\infty}$). Thus we have a free-energy like quantity for a non-equilibrium system and is in fact *the* free energy for an equilibrium system ($\bar{J} = J^S$) for which the condition

$$J^{S} = 0 \tag{5.14}$$

in the steady state (no currents) implies

$$\frac{\delta\mathcal{F}}{\delta\bar{\rho}} = 0 \tag{5.15}$$

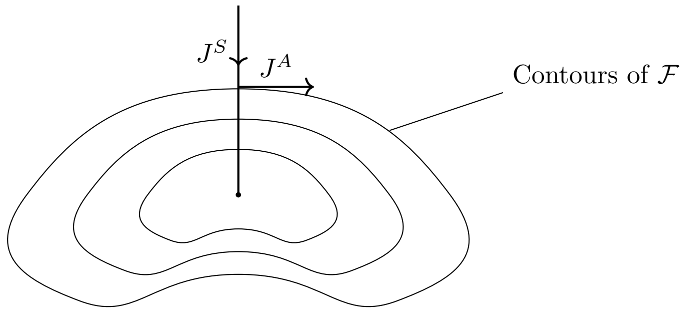
3. Monotonic decrease of free energy as the steady state is approached:

$$\frac{d}{dt}\mathcal{F}[\bar{\rho}] \leq 0 \tag{5.16}$$

Moreover, the rate of decrease is determined by J^S:

$$\frac{d}{dt}\mathcal{F}[\bar{\rho}] = -\int \frac{\left|J^S\right|^2}{\bar{\rho}T} dx \tag{5.17}$$

So, in some abstract sense, J^S defines the direction to the steady state and J^A the contours of $\mathcal{F}$.



Derivations:

1. Using the (HJ) equation,

$$\int \bar{\rho}\overbrace{\left[\nabla U \cdot (\mathrm{F} + T\nabla U) - \nabla \cdot (\mathrm{F} + T\nabla U)\right]}^{=0 \text{ by (HJ)}} dx = 0 \tag{5.18}$$

Integrating the second term by parts,

$$\int \underbrace{(\nabla\bar{\rho} + \bar{\rho}\nabla U)}_{-J^S/T} \cdot \underbrace{(\mathrm{F} + T\nabla U)}_{J^A/\bar{\rho}} dx = 0 \tag{5.19}$$

2. From the definition of $\mathcal{F}$ (5.12) and the form of the stationary distribution $P_\infty = e^{-U}/Z$,

$$\frac{\delta\mathcal{F}}{\delta\bar{\rho}} = \log\frac{\bar{\rho}}{NP_\infty} + 1 = \log\frac{\bar{\rho}}{e^{-U}} + \text{ constant} \tag{5.20}$$

$$\Rightarrow \nabla\frac{\delta\mathcal{F}}{\delta\bar{\rho}} = \frac{\nabla\bar{\rho}}{\bar{\rho}} + \nabla U \equiv -\frac{J^S}{\bar{\rho}T} \tag{5.21}$$

3. Using the generalisation of the chain rule for functional derivatives,

$$\frac{d}{dt}\mathcal{F}[\bar{\rho}] = \int \frac{\partial\bar{\rho}}{\partial t}\frac{\delta\mathcal{F}}{\delta\bar{\rho}} dx \tag{5.22}$$

$$= -\int \left(\nabla\cdot\bar{J}\right)\frac{\delta\mathcal{F}}{\delta\bar{\rho}} dx \qquad \text{(from (5.5))} \tag{5.23}$$

$$= -\int \left(J^A + J^S\right)\cdot\frac{J^S}{\bar{\rho}T} dx \tag{5.24}$$

where we integrated by parts and used (5.11). It remains to use the orthogonality relation from 1.:

$$\frac{d}{dt}\mathcal{F}\left[\bar{\rho}\right] = -\int \frac{\left|J^S\right|^2}{\bar{\rho}T} dx \leq 0 \tag{5.25}$$

Having characterised the average $\bar{\rho}$, a natural next step is to consider fluctuations in the density. Given the limited time, we instead change tack and review the structure considered so far but for systems of *interacting* particles.

5.2 Interacting Systems

For the general case (particle interactions) it is often more productive to consider joint probabilities for density and current. Provided $\partial_t \hat{\rho} = -\nabla \cdot \hat{J}$, path probabilities have Gaussian form

$$\mathbb{P}_F\left[\hat{\rho}(\mathrm{x},t), \hat{J}(\mathrm{x},t)\right] \propto \exp\left(-\ell_0^d \int_0^\tau \int \frac{\left(\hat{J}-\bar{J}\right)^2}{4\sigma\left(\bar{\rho}\right)T} \mathrm{d}\mathrm{x}\mathrm{d}t\right) \tag{5.26}$$

with the variables in the integrand obeying the stochastic PDE

$$\hat{J} = \bar{J} + \sqrt{\frac{2\sigma\left(\hat{\rho}\right)T}{\ell_0^d}}\eta\left(\mathrm{x},t\right) \tag{5.27}$$

Here ℓ_0 is a length scale introduced when coarse-graining the system and $\sigma(\hat{\rho})$ is the noise strength which takes the place of a mobility and reduces to $\sigma = \hat{\rho}$ if interactions are absent.

The structure of the non-interacting case is by and large preserved. In particular, $\bar{J} = J^A + J^S$ can still be decomposed according to time reversal and J^S is still the gradient of a free-energy-like quantity

$$J^S = -\sigma\left(\hat{\rho}\right)T\nabla\frac{\delta\mathcal{F}}{\delta\hat{\rho}} \tag{5.28}$$

with $\mathcal{F}$ such that

$$\frac{d}{dt}\mathcal{F} = -\int \frac{\left|J^S\right|^2}{\sigma\left(\hat{\rho}\right)T} \mathrm{d}\mathrm{x} \leq 0 \tag{5.29}$$

$\mathcal{F}$ in fact has a dual purpose, also describing the probability to observed a density $\hat{\rho}$ in the steady state

$$\mathrm{Prob}\left[\hat{\rho}(\mathrm{x})\right] \simeq \exp\left(-\ell_0^d \mathcal{F}\left[\hat{\rho}\right]\right) \tag{5.30}$$

which is a statement about fluctuations rather than averages.

Finally, the steady state dissipation is

$$\left\langle \int \frac{\hat{J}\cdot J^A}{\sigma} \mathrm{d}\mathrm{x} \right\rangle \tag{5.31}$$

and this continues to be non-negative (second law).

In summary, while there isn't an explicit formula for J^A, J^S in general (constructing the adjoint process is challenging), it is always true that

- J^S characterises relaxation towards the steady state, determines the rate of change of $\mathcal{F}$ and is given by the gradient of $\mathcal{F}$
- J^A vanishes in equilibrium systems (TRS) whilst determining the steady state ($J^S = 0$) dissipation in non-equilibrium systems via $\nabla \cdot J^A = 0$ ($J^A \neq 0$).

6 Conclusion

In this short lecture course we have developed aspects of stochastic thermodynamics with a focus on non-equilibrium steady states. Key concepts included the breaking of time-reversal symmetry, dissipation characterised by fluctuation theorems such as Gallavotti-Cohen, the adjoint process and the role of symmetric and antisymmetric probability currents in the description of stochastic systems.