

Physics 210B: Non-Equilibrium Statistical Mechanics

Fall 2025

Lecturer: McGreevy

These lecture notes live [here](#). Please email corrections to mcgreevy at physics dot
ucsd dot edu.

Last updated: 2026/02/19, 15:59:37

Contents

0.1	Introductory remarks	4
0.2	Sources and acknowledgement	7
0.3	Conventions	8
I	Stochastic Evolution of Particles	9
1	Random Walks and the Renormalization Group	9
1.1	RG treatment of random walk	9
1.2	Anatomy of an RG scheme	11
1.3	Scaling behavior near a fixed point	13
1.4	Biased gaussian walk	17
1.5	Universality classes of random walks	18
1.6	Anomalous diffusion	24
1.7	Large deviations	28
1.8	Random matrices	31
2	Stochastic Evolution	42
2.1	Brownian motion	43
2.2	Whence the noise and whence the friction?	55
2.3	First-passage/exit problems and BCs for FP	58
2.4	Overcoming barriers with path integrals	63
2.5	A systematic path integral description	68
2.6	Memory, correlation and response	71
2.7	Getting in trouble with the UV	77
3	Markov chains	84
3.1	Examples	90
3.2	Two nice facts about nice Markov chains	103
3.3	Monte Carlo	105
3.4	Langevin and detailed balance	107
3.5	A cautionary tale about demographic noise	109
4	Fluctuation theorems	116
4.1	Quantum Crooks relation	119
4.2	Fluctuation relations for NESS	121
4.3	Non-equilibrium free energy and Maxwell's demon	124

II	Stochastic Evolution of Fields	131
5	Near-equilibrium transport, very briefly	134
5.1	Boltzmann equation	134
5.2	Hydrodynamics	137
5.3	Transport and the approach to equilibrium	145
6	Hydrodynamics of magnets	149
6.1	Coarsening	154
6.2	Hydrodynamic description of generic non-ergodicity	158
6.3	Dynamical HMW theorem and its violation	158
7	Critical phenomena in stochastic dynamics	164
7.1	Transitions to absorbing states	164
7.2	“Self-organized criticality”	185
7.3	Reaction-diffusion problems and spin chains	193
7.4	Second quantization for reaction-diffusion problems	195
7.5	Surface growth and roughening	201

0.1 Introductory remarks

Let me begin by saying that the thing that gets me really excited (about this subject and pretty much in general) is: complex collective phenomena emerging from simple local rules governing the interactions of individual actors or agents. I usually say ‘degrees of freedom’ (dofs) but I think for this class I am allowed to broaden the scope.

For better or worse, unlike other branches of science, like condensed matter physics or particle physics or gravitational physics, statistical physics is defined, not by a set of phenomena or physical questions, but by an approximation. Some words about this approximation: This approximation is based on the idea that when things are so complicated and numerous that we can’t keep track of them all, we can instead benefit from some Law of Large Numbers (is it true? when?) and instead sample from some ensemble. Sometimes this approximation even transcends physics itself, and many of these ideas are useful in ... other parts of the world (e.g. the Monte Carlo method, simulated annealing, various physics-motivated improvements of gradient descent). So there is some temptation when organizing a course on the subject to focus on techniques rather than phenomena. I will do my best to resist this temptation. I would prefer to say instead that this branch of science is based on the *idea* of an ensemble approach rather than on the approximation it leads to.

This course is about *non-equilibrium* statistical mechanics. Our point of view in this course, as in biology, is that¹

$$\boxed{\text{equilibrium} = \text{death}}.$$

This is not an exaggeration; a living organism’s main goal, with all its effortful scurrying around, is to postpone the inevitable return to equilibrium.

Comment on equilibration and ergodicity. Many systems equilibrate, and in a reasonable amount of time reach a state where short-time averages can be computed by ensemble averages. This last condition is called the ergodic hypothesis. One way to think about what ensemble to use, associated with Jaynes, is simply to ask for the distribution over the configuration space that minimizes the amount of unwarranted assumptions. If the system has a conserved energy, it will stay the same and we should include this information in our choice of ensemble. The canonical ensemble (also known as Gibbs distribution) is the distribution with the maximum entropy given that there is a conserved energy, where the temperature acts as a Lagrange multiplier for this fact.

¹with due apology to the gay rights activists of the 1980s

That is, given a system with energies E_j , and known average energy E , consider the function

$$I_E(p_j, \alpha, \beta) \equiv S(p) + \alpha(1 - \sum_j p_j) + \beta(E - \sum_j E_j p_j) \quad (0.1)$$

where $S(p) \equiv -\sum_j p_j \log p_j$ is the Shannon entropy, the essentially unique measure of the lack of bias we introduce by using the distribution p_j . Extremizing this function with respect to α, β, p_j says that the distribution is normalized $\sum_j p_j = 1$, the average energy is correct $\sum_j E_j p_j = E$ and

$$0 = \partial_{p_j} I = -\log p_j - 1 - \alpha - \beta E_j, \quad \text{i.e. } p_j \propto e^{-\beta E_j}, \quad (0.2)$$

the Boltzmann distribution.

This was my review of equilibrium statistical mechanics.

Sometimes the ergodic hypothesis is true, and indeed it can be proven under some assumptions. But there are some real exceptions.

If there are conserved quantities, the approach to equilibrium is parametrically slower (power law rather than exponential in time). Even the conserved energy can be followed around. This is the subject of hydrodynamics.

There are integrable systems, with so many conserved quantities that the late time behavior is labelled not just by the energy, but by macroscopically many quantities. The analog of Jaynes' argument in this case leads to a 'generalized Gibbs ensemble' with Lagrange multipliers not just for energy but for all these other quantities.

There are (non-integrable) systems where the approach to equilibrium is so slow that experimenters can't wait long enough for it to happen. Such systems are called *glassy*.

Non-equilibrium. Just as there are many ways to not be a banana or to [not be on a boat](#), there are very many ways to not be in equilibrium. We can divide them into two categories. The first category is *near*-equilibrium states, systems that have been somehow gently disturbed from their restful equilibrium by a gentle linear poke. There is a lot to say about such states in terms of the equilibrium properties themselves. I believe I am contractually obliged to say something about such things.

The second category is everything else. For example, there are systems for which there is no notion of equilibrium at all, or where the equilibrium state is very degenerate and uninteresting, such as one where all the particles are gone, or where everything is dead.

A nice benefit of departing from equilibrium is that we grow an extra dimension – the one whose coordinate is time.

Are there underlying principles? Some books (like the otherwise good one by Krapivsky et al) will explicitly tell you that there are not! I claim that the basic principle, i.e. the basic meta-physical strategy for developing understanding, is the same as in other areas of physics with many dofs: the renormalization group (RG). For this reason, we'll start with a brief discussion of the RG in the context of random walks. It will give us a chance to remind ourselves about the Central Limit Theorem, and to understand it in its proper context as a statement about a fixed point of the RG.

Narrative arc. I've divided the course into two parts, along the same lines as the two volumes of the wonderful book by Kardar. Part I is about non-equilibrium statistical physics of *particles*, in the sense of a finite number of degrees of freedom. Part II is about non-equilibrium statistical physics of *fields*, meaning that the degrees of freedom are distributed over space, and we are interested in the thermodynamic limit where the volume is large (because this is where magical things like spontaneous symmetry breaking (and ergodicity-breaking more generally) and critical phenomena happen).

0.2 Sources and acknowledgement

Thanks to Dan Arovas, Aparna Baskaran, Nigel Goldenfeld, Tarun Grover, Mike Levy, Leo Radzihovsky, and Slava Rychkov for helpful advice about preparing this class. Thanks to Dominic Holland for helping to decrease the number of errors and to increase the clarity. The material in these notes is collected from many places, among which I should mention in particular the following:

Livi and Politi, *Non-equilibrium statistical physics: a Modern Perspective*. (I will abbreviate this as ‘LP’.)

Mehran Kardar, *Statistical Physics of Particles* and *Statistical Physics of Fields*

Krapivsky, Redner, Ben-Naim, *A Kinetic View of Statistical Physics*

Dan Arovas, [Lecture Notes on Nonequilibrium Statistical Physics \(A Work in Progress\)](#) (210B notes from Fall 2021). (I will abbreviate this as ‘DPA’.)

Zwanzig, *Non-equilibrium Statistical Mechanics*

Reichl, *Modern course in statistical physics*, for a straight-ahead account of standard topics.

Chaikin and Lubensky, *Principles of condensed matter physics*

John Toner, *The Physics of Flocking*

Uwe Tauber, *Critical Dynamics*

David Nelson, Grace Zhang, *The Renormalization Group and Condensed Matter Physics* [This is a brand new book that I just got my hands on at the end of the quarter. It looks good!]

Leticia Cugliandolo, [Introduction to Langevin Stochastic Processes](#), 2504.11994

for KPZ (and Tracy-Widom): Takeuchi, [1708.06060](#)

Avanzini, Esposito et al, [Methods and conversations in \(post\)modern thermodynamics](#)

For fluctuation and work relations: Le Bellac, [Les Houches Lectures on Non-equilibrium Statistical Mechanics](#); Ritort, [2404.04582](#), [0705.0455](#) (the latter includes a section on large deviations).

I have also found Google Gemini helpful for clarifying a number of my confusions (*e.g.* about Ito vs Stratonovich, about the van Kampen expansion, about fluctuation theorems, about the meaning of ‘self-organized criticality’).

0.3 Conventions

$k_B = 1$, that is, since we are not doing thermometry, we measure temperature in units of energy. With all due respect to the great Boltzmann, as Krapivski et al accurately call it, it is a “worthless conversion factor”.

The convention that repeated indices are summed is always in effect unless otherwise indicated. d is the number of space dimensions, D is the number of spacetime dimensions (it’s bigger!).

$\equiv$ means ‘equals by definition’. $A \stackrel{!}{=} B$ means we are demanding that $A = B$. $A \stackrel{?}{=} B$ means A probably doesn’t equal B .

A useful generalization of the shorthand $\hbar \equiv \frac{\hbar}{2\pi}$ is $\mathrm{d}k \equiv \frac{\mathrm{d}k}{2\pi}$. I will also write $\delta^d(q) \equiv (2\pi)^d \delta^{(d)}(q)$. I will try to be consistent about writing Fourier transforms as

$$\int \frac{\mathrm{d}^d k}{(2\pi)^d} e^{ikx} \tilde{f}(k) \equiv \int \mathrm{d}^d k e^{ikx} \tilde{f}(k) \equiv f(x).$$

I will often drop the twiddle on $\tilde{f}(k) \equiv f(k)$ and distinguish it by its argument from $f(x)$.

IFF $\equiv$ if and only if. dofs $\equiv$ degrees of freedom.

RHS $\equiv$ right-hand side. LHS $\equiv$ left-hand side. BHS $\equiv$ both-hand side.

IBP $\equiv$ integration by parts. WLOG $\equiv$ without loss of generality.

$+\mathcal{O}(x^n) \equiv$ plus terms that go like x^n (and higher powers) when x is small.

$+o(x^n) \equiv$ plus terms that are smaller than ax^n for any a when x is small.

$+h.c. \equiv$ plus hermitian conjugate.

$\mathcal{L} \ni \mathcal{O}$ means the object $\mathcal{L}$ contains the term $\mathcal{O}$.

Please tell me if you find typos or errors or violations of the rules above.

Part I

Stochastic Evolution of Particles

1 Random Walks and the Renormalization Group

Consider a *random walk* in Euclidean space of dimension d . Starting from some origin, we take discrete steps, where the displacement $\vec{r}$ of each step is chosen (*e.g.*) independently from some distribution $p(\vec{r})$. For example, we could take our steps to have fixed length a_0 , and uniformly-distributed direction. For now, we assume this distribution and that the walk is otherwise unrestricted.

Of interest is the net displacement after M steps

$$\vec{R}_M \equiv \sum_{i=1}^M \vec{r}_i.$$

This is a random variable with average

$$\langle \vec{R}_M \rangle_M \equiv \int d^d r_1 \cdots d^d r_M p(\vec{r}_1) \cdots p(\vec{r}_M) \sum_{i=1}^M \vec{r}_i$$

which vanishes by our assumption of rotation invariance of the individual distributions. On the other hand, $\langle \vec{r}_i \cdot \vec{r}_j \rangle \propto \delta_{ij}$, so the square-length of the net displacement has

$$\langle |\vec{R}_M|^2 \rangle_M = \sum_i^M \sum_j^M \langle \vec{r}_i \cdot \vec{r}_j \rangle_M = \sum_{j=1}^M \langle |\vec{r}_j|^2 \rangle = M a_0^2.$$

The RMS displacement $R(M) \equiv \sqrt{\langle |\vec{R}_M|^2 \rangle_M} = \sqrt{M} a_0$ goes like the square root of the number of steps, a probably-familiar result on which we are going to get some new perspective now.

1.1 RG treatment of random walk

[Some of this discussion is inspired by *Introduction to Renormalization Group Methods in Physics*, by R. J. Creswick, H. A. Farach, C. P. Poole. Two wonderful books about the RG in general are *Lectures on Phase Transitions and the Renormalization Group*, by Nigel Goldenfeld, and *Scaling and Renormalization in Statistical Physics*, by John Cardy.] Now we'll study the random walk a bit more precisely, and use it to introduce

the RG machinery. To be specific, suppose that each microscopic step is sampled from the (Gaussian) distribution

$$p(\vec{r}) = \mathcal{N} e^{-\frac{|\vec{r}|^2}{2\sigma_0^2}}, \quad \mathcal{N} = (2\pi\sigma_0^2)^{-d/2}.$$

As before, the detailed form of the single-step distribution will be unimportant for the questions of interest to us – the technical term is ‘irrelevant’; this will be an outcome of the RG analysis. In this case, we have $\langle \vec{r} \rangle = 0$, $\langle \vec{r} \cdot \vec{r} \rangle = d\sigma_0^2$.

Let $\vec{r}' \equiv \sum_{i=1}^n \vec{r}_i$. Think of this as a ‘coarse-grained step’ – imagine that the single steps (of RMS size σ_0) are too small to see, but for n big enough, n of them can get somewhere. The distribution for the coarse-grained step is:

$$\begin{aligned} P(\vec{r}') &= \int d^d r_1 \cdots d^d r_n p(\vec{r}_1) \cdots p(\vec{r}_n) \underbrace{\delta\left(\vec{r}' - \sum_{i=1}^n \vec{r}_i\right)}_{=\int d^d k \, e^{i\vec{k} \cdot (\vec{r}' - \sum_i \vec{r}_i)}} \\ &\quad \text{(Do } n \cdot d \text{ Gaussian integrals.)} \\ &\quad \text{Note that } d^d k \equiv \frac{d^d k}{(2\pi)^d} \text{.)} \\ &= \int d^d k \exp\left(-n|\vec{k}|^2 \frac{\sigma_0^2}{2} - i\vec{k} \cdot \vec{r}'\right) \\ &\quad \text{(One more Gaussian integral.)} \\ &= \mathcal{N}' \exp\left(-\frac{|\vec{r}'|^2}{2n\sigma_0^2}\right), \quad \mathcal{N}' \equiv (2\pi n\sigma_0^2)^{-d/2}. \end{aligned} \quad (1.1)$$

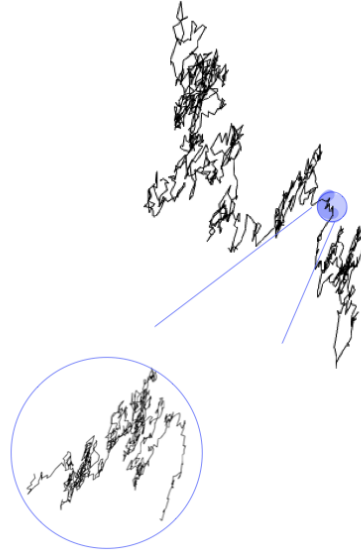
So the variance after n steps grows like $\sqrt{n}$, characteristic of diffusion. We’ll have a lot more to say about diffusion, but I want to emphasize a different point about this result now.

This is the same form of the distribution, with the replacement $\sigma_0 \rightarrow \sigma' \equiv \sqrt{n}\sigma_0$. We can make it *actually* the same distribution if we rescale our units (I will call this the second half of the RG transformation below): rescale $r' \equiv \sqrt{n}r''$, where the zoom factor is chosen to keep the width of the distribution the same after the coarse-graining step. Remembering that distributions transform under change of variables by

$$P(\vec{r}') d^d r' = P(\vec{r}'') d^d r''$$

we have

$$P(\vec{r}'') = \frac{1}{(2\pi\sigma_0)^{d/2}} e^{-\frac{|\vec{r}''|^2}{2\sigma_0^2}}$$



– the same distribution as we had for single step. Therefore, a random walk is (probably) a fractal – it is self-similar on average.

The two steps above – (1) coarse graining and (2) rescaling – constitute a *renormalization group transformation* (more on the general notion next). The ‘coupling constant’ σ_0 transforms under this transformation, in this case as

$$\sigma_0 \mapsto \sigma_{\text{renormalized}} = \sigma_0,$$

i.e. it maps to itself; such a parameter is called *marginal* and is a special case.

Consider the RMS distance covered by a walk in M steps,

$$R(M)^2 \equiv \left\langle \left| \sum_{i=1}^M \vec{r}_i \right|^2 \right\rangle_M.$$

It depends on M and the typical step size, which is σ (since $\sigma^2 = \frac{1}{d} \int d^d r |\vec{r}|^2 p(\vec{r})$). Dimensional analysis tells us that we must have $R(M) \propto \sigma$ and the statistical self-similarity we’ve just found suggests a power law dependence on M :

$$R(M) \sim \sigma M^\nu$$

which scaling relation defines the exponent ν . The coarse-grained walk (no rescaling) takes $M' = M/n$ steps. Demanding the same outcome for the RMS displacement in both the microscopic description and in the coarse-grained description says

$$\sigma M^\nu = \underbrace{\sigma'}_{\sigma'=\sqrt{n}\sigma} (M')^\nu = \sqrt{n}\sigma \left(\frac{M}{n} \right)^\nu = n^{\frac{1}{2}-\nu} \sigma M^\nu. \quad (1.2)$$

(Cultural remark: in the context of quantum field theory, a relation with the same logical content is called a Callan-Symanzik equation.) In order for this to be true for all n , we must have

$$\nu = \frac{1}{2}.$$

$1/\nu$ is a notion of fractal dimension of the random walk.

1.2 Anatomy of an RG scheme

As we saw in (1.2), we are taking a passive point of view on the RG transformations: the observable physics (whatever it may be, $R(M)$ in the example above) is the same, and we are only changing our description of that physics.

An RG transformation has two steps:

1. **Coarse-graining or decimation:** The idea of this step is familiar from the central idea of how thermodynamics emerges from statistical mechanics: we should average over the stuff we can't keep track of (microscopic configurations of the system), holding fixed the stuff we do keep track of (the thermodynamic variables like energy and volume). In the situation mentioned in the previous sentence, we do it all at once.

The key new idea of the RG is to do it a little bit at a time. That is: Integrate out or average over some set of short-distance/fast degrees of freedom, holding fixed a set of long-wavelength/slow degrees of freedom.

Notice that this step is not necessarily reversible: the value of a definite integral (or sum) does not uniquely determine the integrand (or summand). We lose information in this step. This means that a set of transformations defined this way is not in fact a group in the mathematical sense, since there is no inverse element (it is a *semigroup*). So much for that.

The idea is that we are squinting, so that the smallest distance Δx we can resolve gets a little bigger at each step: say before the coarse-graining, we had a resolution $\Delta x = \epsilon$, and afterwards we only keep track of stuff that varies on distances larger than $\Delta x = \lambda \epsilon$ for some scale factor $\lambda > 1$.

2. **Rescaling:** Now we change units to map the coarse-grained system back onto the original one, so that $\lambda \epsilon \mapsto \epsilon$. We do this so that we can compare them.

Now we're going to think about the space on which this transformation is acting. Its coordinates are the parameters of the system, such as the parameters defining the probability distribution such as σ_0 for the random walk, or the couplings in the Hamiltonian if $p = e^{-\beta H}/Z$. Let's call the set of such parameters $\{h_j\}$, where j is an index which runs over as many parameters as we need to consider²³. These parameters get transformed according to

$$\{h_j\} \xrightarrow{\text{steps 1, 2}} \{h'_j \equiv \mathcal{R}_j(\{h\})\}.$$

²For example, in the random walk case, other parameters we could include are $b, c, \dots$ in

$$p(\vec{r}) = \exp - \left(\vec{b} \cdot \vec{r} + \frac{r^2}{2\sigma^2} + cr^4 + \dots \right).$$

³One of the many crucial contributions of Ken Wilson to this subject was (I think) allowing for the possibility of including arbitrarily many parameters. The terror you are feeling at this possibility of an infinite-dimensional space of coupling parameters will be allayed when we discover the correct way to organize them two pages from now.

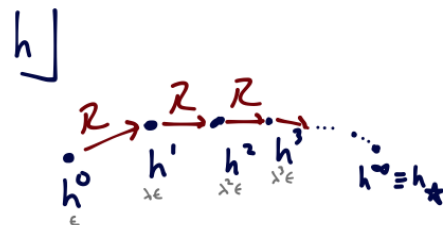
This map is something we can do over and over, coarse-graining (zooming out) by a factor of λ each time, until we get to macroscopic sizes. The repeated application of the map $h'_j \equiv \mathcal{R}_j(h)$ describes a dynamical system on the space of parameters. If we are interested in macroscopic physics, we care about what happens when we do it lots of times:

$$h \mapsto \mathcal{R}(h) \mapsto \mathcal{R}(\mathcal{R}(h)) \equiv \mathcal{R}^2(h) \mapsto \mathcal{R}^3(h) \rightarrow \dots$$

(When studying such a possibly-nonlinear dynamical system more generally, it is a good idea to ask first about the possible late-time behavior.)

What can happen? There are three possibilities:

1. We can reach a *fixed point*, $h_\star = \mathcal{R}(h_\star)$. (We'll include $h \rightarrow \infty$ in some direction in this case. That just means we chose a bad parametrization.)
2. We can go around in circles forever. This is called a *limit cycle*.
3. We can **jump around chaotically forever**, e.g. near some kind of **strange attractor**.



The first case, where there is a fixed point, is the one about which we have a lot to say, and fortunately is what seems to happen usually.

A crucial point: the distribution described by such a fixed point of the RG is self-similar, by the definition of the RG transformation. (If this is true when our zooming size $\lambda \rightarrow 1$, then it is actually scale-invariant.)

1.3 Scaling behavior near a fixed point

Now, suppose we have found a fixed point $h_\star$ of our RG transformation, which is a candidate for describing the macroscopic behavior of our system. It is then a good idea to look at the behavior in the neighborhood of the fixed point (this is also a good piece of advice for general dynamical systems): linearize about the fixed point. We will see that this analysis immediately spits out the phenomenology of scaling behavior near a critical point. If that is not a familiar notion, don't worry, we'll come back to it.

First, define the set of points that flow to the fixed point to be the

$$\text{critical surface of } h_\star \equiv \{h \mid \lim_{n \rightarrow \infty} \mathcal{R}^n(h) = h_\star\} \equiv \mathcal{S}(h_\star)$$

– this is the basin of attraction of the fixed point in question. All these starting points for the couplings produce the same long-wavelength (often called infrared or IR) behavior. They are said to be in the same *universality class*.

Linearizing about the fixed point, let $h_j \equiv h_j^* + \delta_j$, where $|\delta| \ll 1$ will be our small parameter. This maps under the RG step according to

$$h_j \equiv h_j^* + \delta_j \mapsto h'_j = \mathcal{R}_j(h^* + \delta) \stackrel{\text{Taylor}}{=} \underbrace{\mathcal{R}_j(h^*)}_{=h_j^*} + \delta_k \underbrace{\frac{\partial h'_j}{\partial h_k} \Big|_{h^*}}_{\equiv R_{jk}} + \mathcal{O}(\delta^2)$$

where in the last step we assumed that the RG map $\mathcal{R}$ is analytic in the neighborhood of the fixed point, *i.e.* that it has a Taylor expansion. How could it not be? We got it by doing some finite sums of analytic functions. By $+\mathcal{O}(\delta^2)$ I mean plus terms that go like δ^2 and higher powers of delta which are small and we will ignore them. If we ignore them, then the map on the deviation from the fixed point δ is a linear map:

$$\delta_j \mapsto \delta'_j = R_{jk} \delta_k.$$

[\[End of Lecture 1\]](#)

We know what to do with a linear map: find its eigenvalues and eigenvectors:

$$R_{jk} \phi_k^{(n)} = \rho_n \phi_j^{(n)}. \quad (1.3)$$

Notice that nothing we've said guarantees that R_{jk} is a symmetric matrix, so its right and left eigenvectors need not be the same (the eigenvalues are), so we'll also need

$$\tilde{\phi}_j^{(n)} R_{jk} = \rho_n \tilde{\phi}_k^{(n)}. \quad (1.4)$$

Together, these are orthonormal

$$\sum_j \tilde{\phi}_j^{(n)} \phi_j^{(n')} = \delta_{n,n'} \quad (1.5)$$

and complete

$$\sum_n \tilde{\phi}_j^{(n)} \phi_k^{(n)} = \delta_{jk}. \quad (1.6)$$

Real non-symmetric matrices. Later on we're going to think about equations of the following form:

$$p_j^{t+1} = R_{jk} p_k^t \quad (1.7)$$

where p_j^t is the probability distribution for some variable j at time t , and this equation is some update rule for how the probability distribution changes with time. The matrix R is again real but not necessarily symmetric. To solve this equation, we'll again want to understand the eigensystem of the matrix R . So let's spend another minute thinking about such matrices.

The left eigenvector equation can be rewritten as the eigenvalue equation for the transpose of R ($(R^T)_{ij} \equiv R_{ji}$), $R^T \tilde{\phi} = \rho \tilde{\phi}$. The right and left *eigenvalues* ρ_n in (1.3) and (1.4) are the same because they satisfy the same equation

$$0 = F(\rho) = \det(\rho \mathbb{1} - R) = \det(\rho \mathbb{1} - R^T). \quad (1.8)$$

Note also that $F(\rho) = F(\rho^*)^*$ so eigenvalues are real or come in complex-conjugate pairs.

The combination $\tilde{\phi}_j^{(n')} (1.3)_j^{(n)} - (1.4)_k^{(n')} \phi_k^{(n)}$ implies

$$0 = (\rho_n - \rho'_n) \langle n' | n \rangle \quad (1.9)$$

where the inner product is just $\langle n' | n \rangle \equiv \sum_j \tilde{\phi}_j^{(n')} \phi_j^{(n)}$. So either the eigenvalues are different or the L and R eigenvectors are orthogonal. As usual, if there are degenerate eigenvalues we have to use Gram-Schmidt to make an ON basis.

About the eigenvalues, notice the following. We've defined the RG transformation $\mathcal{R} \equiv \mathcal{R}_\lambda$ to accomplish a coarse-graining by a scale factor λ . We can imagine defining such a transformation for any λ , and these operations form a semigroup under composition

$$\mathcal{R}_\lambda \mathcal{R}_{\lambda'} = \mathcal{R}_{\lambda\lambda'}.$$

This is useful because it says that the eigenvalues of the linearized operators

$$R_\lambda \phi^{(n)} = \rho_n(\lambda) \phi^{(n)}$$

must satisfy the same multiplication law⁴

$$\rho_n(\lambda) \rho_n(\lambda') = \rho_n(\lambda\lambda'). \quad (1.10)$$

⁴Why do R_λ for different λ have the same eigenvectors?

It really follows from the semigroup property. The eigenvectors are physical things – an eigenvector determines some operator O with the following property: if I add O to the fixed-point hamiltonian, $H_\star + gO$, an RG transformation does not generate any other operators, i.e. it gives $H = H_\star + \alpha gO$ for some α .

But a function that satisfies this rule must have the form⁵

$$\rho_n(\lambda) = \lambda^{y_n} \quad (1.11)$$

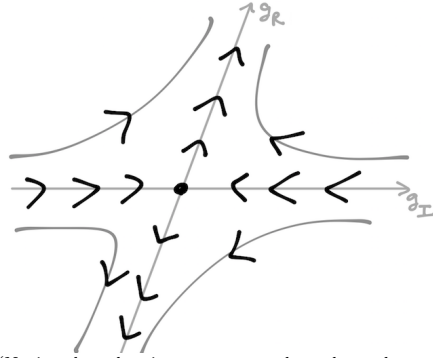
for some y_n independent of λ .

The eigenvectors of R give a preferred coordinate basis near the fixed point:

$$\delta_j = \sum_n g_n \phi_j^{(n)}, \quad g_n \stackrel{(1.5)}{=} \sum_k \tilde{\phi}_k^{(n)} \delta_k,$$

which we will use from now on. y_n in (1.11) is called the *scaling dimension* of the coupling g_n .

Now we can see the crucial RG dichotomy which tames the infinitely many possible couplings: If $|\rho_n| < 1$ ($y_n < 0$) then as we act with $\mathcal{R}$ many times to get to long wavelengths, then $g_n \rightarrow 0$. Such a coupling is called *irrelevant*: it goes away upon repeated RG transformations and its effects on macroscopic physics can be ignored. Notice that since the perturbation is getting smaller, the approximation $|\delta| \ll 1$ becomes better and better in this case.



(Notice that the eigenvectors need not be orthogonal.)

In contrast, if $|\rho_n| > 1$ ($y_n > 0$) then as we act with $\mathcal{R}$ many times to get to long wavelengths, then g_n grows. Such a parameter is called *relevant*, and represents an instability of the fixed point: our linearization breaks down after repeated applications of $\mathcal{R}$ and we leave the neighborhood of the fixed point.

The special case of a coupling with $y_n = 0$ which doesn't change is called *marginal*. Usually apparently marginal couplings can either be removed by a choice of units (as with σ in our example above), or are only marginal in some approximation.

On the other hand, the choice of by how much to zoom out (λ) is an arbitrary one. Doing the RG step by λ twice should give the same result as doing it once by λ^2 . So in particular either one should give the same set of special directions.

⁵The function $y(\lambda) \equiv \log \rho_n(\lambda)$ then satisfies $y(\lambda) + y(\lambda') = y(\lambda\lambda')$. First this implies $y(1) = 0$. If we consider $\lambda' = 1 + \epsilon$, we have

$$\begin{aligned} y(\lambda) + y(1 + \epsilon) &= y(\lambda + \lambda\epsilon) \\ y(\lambda) + y(1) + \epsilon y'(1) &= y(\lambda) + \lambda\epsilon y'(\lambda) + \mathcal{O}(\epsilon^2) \end{aligned}$$

which says that y satisfies the differential equation $y'(\lambda) = \frac{y'(1)}{\lambda}$ which is solved by

$$y(\lambda) = y'(1) \ln \lambda.$$

I'm not sure if the statement (1.11) follows if we only know (1.10) for discrete values of λ . Does it?

In these terms, the critical surface (actually its tangent space near the fixed point) is determined by

$$\mathcal{S}(h_*) = \{g_n = 0 \text{ if } y_n > 0\}.$$

In particular, the *codimension* of the critical surface in the space of couplings is the number of relevant perturbations of the fixed point.

1.4 Biased gaussian walk

Next we generalize our ensemble of random walks to illustrate some features of the RG that were missing from our simple pure Gaussian example above.

First, we can see an example of a relevant operator if we study a *biased* walk, with

$$p(\vec{r}) = (2\pi\sigma^2)^{-d/2} \exp\left(-\frac{|\vec{r} - \vec{r}_0|^2}{2\sigma^2}\right). \quad (1.12)$$

Again define the distribution for the coarse-grained step to be

$$\begin{aligned} P(\vec{r}') &= \int \prod_{i=1}^n (d^d r_i p(\vec{r}_i)) \delta\left(\vec{r}' - \sum_i^n \vec{r}_i\right) \\ &\quad \text{(more Gaussian integrals)} \\ &= (2\pi n\sigma^2)^{-d/2} \exp\left(-\frac{|\vec{r}' - n\vec{r}_0|^2}{2n\sigma^2}\right). \end{aligned} \quad (1.13)$$

So, after the coarse-graining step, we have

$$\begin{cases} \sigma' &= \sqrt{n}\sigma \\ \vec{r}'_0 &= n\vec{r}_0 \end{cases}.$$

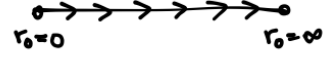
After the rescaling step, to keep the width of the distribution fixed (when expressed in the new units of length), we have

$$\begin{cases} \sigma^{(R)} &= \sigma \\ \vec{r}_0^{(R)} &= \sqrt{n}\vec{r}_0 \end{cases}.$$

So R is diagonal already. This says that the bias of the walk is a *relevant* operator of dimension $y_0 = \frac{1}{2} > 0$. But notice, by the way, that this one relevant operator violates the rotation symmetry of the distribution of steps. This means that as long as we preserve the rotation symmetry, the Gaussian fixed point is stable under the RG (at least to this perturbation).

We have here an explicit example of an RG map $\mathcal{R}$. Let's study its fixed points. There's one at $(\sigma, \vec{r}_0 = 0)$ (for any σ , so actually it is a family of fixed points parametrized by the marginal coupling σ) which is the unbiased walk we studied earlier. This fixed point is unstable because if we turn on a little r_0 it will grow indefinitely.

And there's another fixed point at $(\sigma, \vec{r}_0 = \infty)$. This is where we end up if we perturb the unbiased fixed point. The distribution (1.12) says (by direct calculation) that



$$R(M) = \sqrt{\langle |\vec{R}_M|^2 \rangle_M} = \sqrt{M^2 |\vec{r}_0|^2 + M\sigma^2} \xrightarrow{M \gg 1} M |\vec{r}_0|. \quad (1.14)$$

This means that we'll need $N(a) \sim 1/a$ spheres of diameter a (when a is large compared to σ) to cover the walk – it will be one dimensional.

This means that a system defined by some microscopic distribution of the form (1.12) with some value of $\vec{r}_0$ and σ will look like a Brownian walk of the type described above, with fractal dimension $D = 2$, if you look at it closely, with a resolution $\delta x \ll \sigma$. But from a distance (resolution worse than $\delta x \gg \sigma$), it will look like a one-dimensional path ($D = 1$) in the $\vec{r}_0$ direction. For example, the number of balls of size a needed to cover it behaves as

$$N(a) \sim \begin{cases} a^{-2}, & a \ll \sigma \\ a^{-1}, & a \gg \sigma \end{cases}.$$



1.5 Universality classes of random walks

Now let the distribution from which a single step is sampled be *any* distribution $p(\vec{r})$ with finite moments (if we want to rule out the relevant coupling r_0 , we can choose it to

be rotation invariant, $p(\vec{r}) = p(|\vec{r}|)$, but we won't assume this below). For example, the fixed-step-length distribution $p(\vec{r}) = \frac{1}{4\pi a^2} \delta(|\vec{r}| - a)$ is a good one to keep in mind. (This is still not the most general walk, since we're still assuming the steps are independent. More on that next.) The distribution for the coarse-grained step is

$$\begin{aligned} P(\vec{r}') &= \int \prod_{i=1}^n d^d r_i p(\vec{r}_1 \cdots \vec{r}_n) \delta\left(\vec{r}' - \sum_i \vec{r}_i\right) = \int d^d k e^{-i\vec{k} \cdot \vec{r}'} \left\langle \prod_i e^{i\vec{k} \cdot \vec{r}_i} \right\rangle \\ &\stackrel{\text{iid}}{=} \int d^d k e^{-i\vec{k} \cdot \vec{r}'} \left(\left\langle e^{i\vec{k} \cdot \vec{r}} \right\rangle \right)^n. \end{aligned} \quad (1.15)$$

The quantity

$$\left\langle e^{i\vec{k} \cdot \vec{r}} \right\rangle = \int d^d r p(\vec{r}) e^{i\vec{k} \cdot \vec{r}} \equiv g(k)$$

is called the characteristic function of the distribution $p(\vec{r})$, and is a generating function for its moments:

$$\langle r^m \rangle = (-i\partial_k)^m g(k)|_{k=0}.$$

The Taylor expansion in k of its logarithm is the cumulant expansion:

$$\log g(k) = \sum_m \frac{(i\mathbf{k})^m}{m!} C_m, \quad C_m = (-i\partial_k)^m \log g|_{k=0}.$$

(In particular, C_2 is the variance.) The important point for us is the expansion:

$$\left\langle e^{i\vec{k} \cdot \vec{r}} \right\rangle = \exp \left(i\vec{k} \cdot \langle \vec{r} \rangle - \frac{1}{2} \sum_{\mu, \nu} k_\mu k_\nu \underbrace{(\langle r_\mu r_\nu \rangle - \langle r_\mu \rangle \langle r_\nu \rangle)}_{=\sigma_0^2 \delta_{\mu\nu}} + \frac{(i\mathbf{k})^3}{3!} C_3 + \mathcal{O}(k^4) \right).$$

If we don't truncate the sum in $\sum_m \frac{(i\mathbf{k})^m}{m!} C_m$, then the $\{C_m\}$ are just another set of coordinates on the space of couplings for the walk. Why should we treat the integration variable k in (1.15)

$$P(\vec{r}') = \int d^d k e^{-i\vec{k} \cdot \vec{r}'} \left\langle e^{i\vec{k} \cdot \vec{r}} \right\rangle^n = \int d^d k e^{-i\vec{k} \cdot \vec{r}'} e^{in\vec{k} \cdot \langle \vec{r} \rangle - \frac{n}{2} \sigma_0^2 |\vec{k}|^2 + \mathcal{O}(nk^3)}$$

as small? Because the integrand is suppressed by the Gaussian factor. If the Gaussian bit dominates, then the integrand has support at $k \lesssim \frac{1}{\sqrt{n\sigma_0^2}}$, at which the m th term in the cumulant expansion contributes to the exponent in (1.15) as

$$nk^m C_m \sim n^{1-\frac{m}{2}} \xrightarrow{n \rightarrow \infty} 0 \quad \text{if } m > 2,$$

where the important thing for getting zero is just that C_m is finite and independent of n and k . Since n is the amount by which we are coarse graining, $n \rightarrow \infty$ is the IR

limit, and this is the statement that the couplings C_m for $m > 2$ are irrelevant. Then we can do the remaining Gaussian integral (ignoring the small corrections which are suppressed by $e^{-n^{1-\frac{m}{2}}C_m}$)

$$P(\vec{r}') \stackrel{n \gg 1}{\approx} (2\pi n\sigma_0^2)^{-d/2} e^{-\frac{1}{2} \frac{|\vec{r}' - n\langle\vec{r}\rangle|^2}{n\sigma_0^2}}.$$

What's this? This is the Gaussian we used at the beginning, with $r_0 = n \langle\vec{r}\rangle$.

Now let's set the bias to zero. Notice that the bias term can be forbidden by imposing a symmetry, such as the $\text{SO}(d)$ rotation symmetry, of the distribution. The result we've found in that case, that the distribution for a sum of many random variables independently distributed according to some distribution with finite moments is a Gaussian, is usually called the Central Limit Theorem or the Law of Large Numbers. (For more on the derivation I recommend the discussion in Kardar volume 1.)

In the framework of the RG it is an example of *universality*: all such probability distributions are in the basin of attraction of the gaussian random walk – they are said to be in the same *universality class* meaning that they have the same long-wavelength physics. In particular, their RMS displacement goes like $R_M \sim M^{1/2}$ for large number of steps M , and (for $d \geq 2$) their fractal dimension is $D = 2$.

Notice that we did not prove that the Gaussian fixed point is the only one: we had to assume that we were in its neighborhood in order to use the $k \sim n^{-1/2}$ scaling – this scaling is a property of the neighborhood of the fixed point, just like the exponents y we got by linearizing about the general fixed point in §1.3.

We could try to find other fixed points in the space of d -dimensional walk distributions. For example, we could have chosen the scaling to fix the coefficient C_m for any m . In that case we would find that the $m - 1$ perturbations $C_{l < m}$ are relevant and all the $C_{l > m}$ are irrelevant. The special case where we fix C_1 (i.e. choose $k \sim 1/n$) gives the same fixed-point we reached for the biased walk. The fixed points with $C_{m > 2}$ fixed would have more than one relevant operator (this is called ‘multicritical’), which means reaching them requires tuning several parameters. For better or worse, these fixed point distributions with $m > 2$ don't seem to exist as probability distributions, because they would have to have zero variance⁶.

One crucial lesson, which generalizes well, is that most of the possible perturbations of the fixed point are irrelevant, and there is only a small number of relevant or marginal perturbations. This is why it is possible to do physics. [End of Lecture 2]

Self-avoiding walks have their own universality class. Suppose that the random 1d objects we are studying are actually *polymers* – long chain molecules made

⁶Thanks to Tarun Grover for pointing this out to me. Maybe they do exist as simple analogs of ‘[complex fixed points](#),’ where we drop some positivity assumptions?

of ‘monomers’ which cannot be in the same place, *i.e.* they have some short-ranged repulsion from each other. We can model this as lattice paths without self-intersection, or *self-avoiding walks* (SAWs). Does this microscopic modification of our ensemble change the long-wavelength physics? In fact, yes, this perturbation of our ensemble leads to a flow to a different fixed point, with its own basin of attraction.

Not every form of correlation between the steps of the walk takes us to a new universality class. While it’s true that we used exact independence of the steps in

$$\left\langle e^{\mathbf{i} \frac{k}{\sqrt{n}} \sum_i (x_i - \langle x \rangle)} \right\rangle \stackrel{?}{=} \left\langle e^{\mathbf{i} \frac{k}{\sqrt{n}} (x - \langle x \rangle)} \right\rangle^n, \quad (1.16)$$

in fact such a strong condition is not necessary. Suppose $\langle x_i x_{i+k} \rangle$ depends only on k and falls off as k grows. *Weak* correlations means

$$\sum_{k=1}^L P(k) C(k) \equiv \sum_{k=1}^L P(k) (\langle x_i x_{i+k} \rangle - \langle x_i \rangle \langle x_{i+k} \rangle) \stackrel{L \gg 1}{\sim} L^0, \quad (1.17)$$

where $P(k)$ is some finite-degree polynomial in k , so that this sum of correlations is finite as $L \rightarrow \infty$. If this condition holds, the CLT is still true, and we can again see it by an RG step, *i.e.* by blocking our RW into blocks of L steps, where L is large compared to the range of the correlations. That is, let $y_k \equiv \sum_{i=(k-1)L+1}^{kL} (x_i - \langle x \rangle)$, for $k = 1..N/L$ (take N to be a multiple of L), so that

$$\sigma_y^2 = \langle y_k^2 \rangle - \langle y_k \rangle^2 = \sigma^2 + \frac{2}{L} \sum_{1 \leq i < j \leq L} C(j-i) \approx \sigma^2 + 2 \sum_{k=1}^L C(k) \quad (1.18)$$

and

$$\langle y_k y_{k+1} \rangle - \langle y_k \rangle \langle y_{k+1} \rangle = \frac{1}{L} \sum_{i=1}^L \sum_{j=L+1}^{2L} C(j-i) = \frac{1}{L} \left(\sum_{k=1}^L k C(k) + \sum_{k=L+1}^{2L-1} (2L-k) C(k) \right) \stackrel{L \gg 1}{\rightarrow} 0 \quad (1.19)$$

using (1.17) with these linear polynomials in k . Now we can apply the CLT to y_k , and conclude that

$$\frac{1}{\sqrt{N/L}} \sum_{k=1}^{N/L} y_k = \frac{1}{\sqrt{N}} \sum_i (x_i - \langle x \rangle) = X \quad (1.20)$$

is governed by a Gaussian distribution near its mean, which is the usual CLT for x_i .

This strategy and its conclusion is used in practice all the time when doing Monte Carlo simulations, as I’ll explain below.

Walks with long tails have their own universality class. A SAW is a big departure from what we’ve been talking about so far, in the sense that the process

generating a SAW is no longer *Markov*. A Markov process is one that is determined by the most recent condition imposed on it:

$$P(x_N, t_N; \dots; x_1, t_1 | y_M, s_M; \dots; y_1, s_1) \stackrel{\text{Markov}}{=} P(x_N, t_N; \dots; x_1, t_1 | y_M, s_M) \quad (1.21)$$

where I've assumed $t_i > t_{i-1}, s_i > s_{i-1}$. (Here $P(X|Y)$ denotes the probability distribution on X conditioned on the value of another variable Y . It is related to the joint distribution $P(X, Y)$ by normalization: $P(X|Y)P(Y) = P(X, Y)$. This normalization step is called Bayes' theorem.)

Returning to the Markov case, what if we just choose a single-step distribution that has a long tail, like the Cauchy (aka Lorentzian) distribution, $p(r) = \frac{\sigma/\pi}{r^2 + \sigma^2}$ (e.g. in $d = 1$)?

We can try to take a continuum limit, which helps to see the distinction between the Gaussian case and a case where the single-step distribution has a long tail: think of $t = \tau n$ where n is the step number and τ is some fixed interval, which we can then make very small. Then we can try to take time derivatives and use calculus.

Continuous Markov processes (the resulting limit) come in two kinds: those where

$$\text{Prob}(\text{sample path is continuous}) = 1 \quad (1.22)$$

(this is sometimes called “Lindeberg criterion”), and those where this probability is less than 1. I believe that this condition is equivalent to

$$\lim_{\tau \rightarrow 0} \frac{1}{\tau} \int_{|x-y| > \epsilon} dy P(y, t + \tau | x, t) \stackrel{?}{=} 0 \quad \forall \epsilon > 0. \quad (1.23)$$

For the nice 1d random walk universality class we've been talking about above (“Wiener process”),

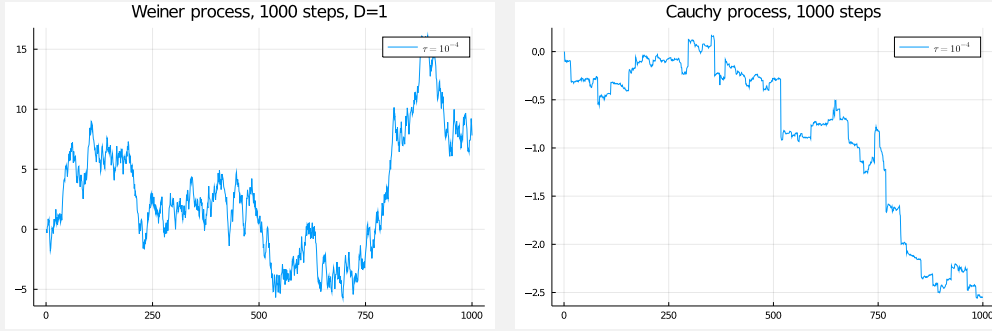
$$P_W(x, t | x', t') = \frac{1}{\sqrt{4\pi D |t - t'|}} e^{-\frac{(x - x')^2}{4D |t - t'|}} \quad (1.24)$$

with $D = \frac{\sigma^2}{2\tau}$; notice that with this distribution, the uncertainty in x grows like $\sqrt{t - t'}$ (the variance of the sum of two steps of the usual random walk is $\sqrt{2} = 2^{1/2}$ times that of each step), characteristic of *diffusion*, and D is called the diffusion constant. For this distribution, you can check that (1.23) is satisfied.

An example where (1.23) is not satisfied is the Cauchy process, where

$$P_C(x, t | x', t') = \frac{|t - t'|}{\pi ((x - x')^2 + |t - t'|^2)}. \quad (1.25)$$

Note that in both cases, $\lim_{t-t' \rightarrow 0} P(x, t | x', t') = \delta(x - x')$.



To simulate either of these processes: given $x_n = x(t = n\tau)$, choose Δx uniformly on $[-1, 1]$, let $y = x_n + \Delta x$. If $P(y, (n+1)\tau | x, n\tau) \geq \text{rand}()$ accept (here $P(y, t' | x, nt)$ is either P_W or P_C , and $\text{rand}()$ is a uniformly random number between $[0, 1]$), *i.e.* set $x_{n+1} = x(t = (n+1)\tau) = y$, otherwise reject, *i.e.* try again.

A stronger sense in which the Cauchy process represents a different universality class is that we can show that it is not in the basin of attraction of the Gaussian distribution⁷. As a slight generalization, consider an iid process $p(x_1 \dots x_N) = \prod_i p_1(x_i)$ where the individual steps are governed by a distribution with (mean zero and) a long tail, specifically:

$$p_1(x) \stackrel{x \rightarrow \infty}{\sim} \frac{1}{|x|^{1+z}}, \quad 0 < z \leq 2. \quad (1.26)$$

$z > 0$ is required in order that $\int_{-\infty}^{\infty} dx p(x) < \infty$ the distribution can be normalized; But if $z \leq 2$ then the variance $\int dx x^2 p_1(x)$ is infinite. Such processes are called *Levy flights*. For the Levy_z process, the variance of the sum of two steps is $2^{1/z}$ times that of each step; this is sometimes called *superdiffusive* (since for $z < 2$ the variance grows faster with time than for diffusion), an example of *anomalous diffusion*.

Recap. One way to think about the CLT is in terms of the generating function for moments of the N -step displacement $X = \sum_{i=1}^N x_i$, with distribution $P_X(x) = \int d^N x p(x_1 \dots x_N) \delta(x - \sum_i x_i)$, namely $\tilde{P}_X(k) \equiv \langle e^{-\mathbf{i}k \sum_{j=1}^N x_j} \rangle$. Its log is the generating function of the cumulants:

$$\log \tilde{P}_X(k) = -\mathbf{i}k \underbrace{\sum_{i_1=1}^N \langle x_{i_1} \rangle_c}_{=\langle X \rangle_c} + \frac{(-\mathbf{i}k)^2}{2} \underbrace{\sum_{i_1, i_2} \langle x_{i_1} x_{i_2} \rangle_c}_{=\langle X^2 \rangle_c} + \dots \quad (1.27)$$

⁷Whether the Lindeberg criterion is equivalent to this is not clear to me.

The key point is that, if it is analytic at $k = 0$, we can Taylor expand this cumulant generating function in k . If the steps are independent $P(x) = \prod_i p_i(x_i)$ then $\tilde{P}_X(k) = \prod_i \tilde{p}_i(k)$ and $\langle X^n \rangle_c = \sum_{i=1}^N \langle x_i^n \rangle_c$. If further they are iid, $\langle X^n \rangle_c = N \langle x^n \rangle_c$, and $y \equiv (X - N \langle X \rangle) / \sqrt{N}$ has $\langle y \rangle = 0$, $\langle y^n \rangle_c \sim N^{1-n/2}$. More generally, as long as $\sum^N \langle x_{i_1} \cdots x_{i_m} \rangle_c \sim \mathcal{O}(N^{m/2})$, the CLT holds:

$$P(y) \stackrel{N \rightarrow \infty}{\simeq} \frac{1}{\sqrt{2\pi \langle X^2 \rangle_c}} e^{-\frac{y^2}{2\langle X^2 \rangle_c}}. \quad (1.28)$$

For the Cauchy distribution, $p(x) = \frac{a}{\pi(x^2 + a^2)}$, the moment generating function is

$$\tilde{p}_X(k) = \langle e^{-ikx} \rangle = \int dx e^{-ikx} \frac{a}{\pi(x^2 + a^2)} = \frac{1}{2\pi i} \left(\int_{\mathbb{R}} dx \frac{e^{-ikx}}{x - ia} - \int_{\mathbb{R}} dx \frac{e^{-ikx}}{x + ia} \right) = e^{-|ka|}. \quad (1.29)$$

In the last step we use the fact that we can close the contour in the UHP (LHP) if $k > 0$ ($k < 0$). Now $\log \tilde{p}_X(k) = |k|a$ (WLOG assume $a > 0$) is not analytic at $k = 0$, and so the cumulant generating function does not even have a Taylor expansion there.

More generally the large- x behavior $p_1(x) \stackrel{x \rightarrow \infty}{\sim} \frac{1}{|x|^{1+z}}$ implies the small- k behavior $\tilde{p}_1(k) \sim |k|^z + \dots$ of the Fourier transform, which is again non-analytic at $k = 0$. For the iid case,

$$\log \tilde{p}_X(k) = N \log \tilde{p}_1(k) \sim -N|k|^z + \dots \quad (1.30)$$

and as in the Gaussian case we can define a new variable $y = X/N^{1/z}$ to remove the N -dependence from the leading term. The higher order terms (with more powers of k) in the distribution for y are then suppressed by powers of N . For the Cauchy case $z = 1$, the limiting distribution is again the Cauchy distribution $p_y(y) = \frac{a}{\pi(y^2 + a^2)}$. In this sense, it is a fixed point of the RG.

1.6 Anomalous diffusion

[LP §3.5] We've seen above that for many classes of random walks (in particular, when the CLT applies because the cumulants are finite), the mean square deviation after t steps $\langle \delta x^2(t) \rangle \equiv \langle x^2(t) \rangle - \langle x(t) \rangle^2$ goes like

$$\langle \delta x^2(t) \rangle \propto t, \quad (1.31)$$

i.e. the particle diffuses. I also claimed that there are other universality classes, where the behavior is different. Here is a class of models that produces other behaviors, of the form

$$\langle \delta x^2(t) \rangle \propto t^\alpha, \quad (1.32)$$

where the names for the various regimes are:

$$\begin{array}{ccccccc}
 & \text{OU} & & \text{diffusive} & & \text{ballistic} & \\
 & | & & | & & | & \\
 \text{---} & 0 & \text{subdiffusive} & 1 & \text{superdiffusive} & 2 & \text{---} \rightarrow \alpha
 \end{array} \tag{1.33}$$

As appropriate for describing an RG fixed point, the behavior (1.32) has a scale invariance, under which space and time scale differently if $\alpha \neq 2$. The parameter $z \equiv 2/\alpha$ can be called the *dynamical exponent*, which parameterizes the relative scaling of time and space that preserves the equation of motion.

The class of models is called *continuous-time random walk* (CTRW), and it proceeds by instantaneous jumps of length x , chosen from a distribution $p_x(x)$, which occur with a distribution $p_t(t)$ of wait times or residence times between jumps. The summary is: if $\langle x^2 \rangle_{p_x} = \sigma^2 < \infty$ and $\langle t \rangle_{p_t} \equiv \tau < \infty$, we get $\alpha = 1$, diffusion. If $\sigma^2 < \infty$ but $\tau \rightarrow \infty$ because $p_t(t) \stackrel{t \gg 1}{\sim} \frac{1}{t^{1+\alpha}}$, we get subdiffusive behavior labelled by $\alpha \in (0, 1)$. If in contrast $\tau < \infty$, but $\sigma^2 \rightarrow \infty$ because $p_x(x) \stackrel{x \gg 1}{\sim} \frac{1}{|x|^{1+z}}$ ($z > 0$ is required for this to be normalizable), we get $\alpha = \frac{2}{z}$, superdiffusive behavior.

Notice that, especially in the case where $\tau \rightarrow \infty$, the wait times can be quite long. Imagine for a moment trying to simulate that case by stepping through a grid of times. There is a very clever fast way to simulate such a model, that avoids actually waiting for the waiting times, called the Gillespie algorithm. It is also useful for many other problems where we're given reaction rates. You'll get to learn it in a homework problem.

Let's assume $p_x(x) = p_x(-x)$ so $\int_{-\infty}^{\infty} dx p_x(x) = 1$. In contrast it's difficult (but not impossible) to assign meaning to negative waiting times, so $\int_0^{\infty} dt p_t(t) = 1$, $p_t(t < 0) = 0$. Let the *survival probability* be

$$\psi(t) = \text{prob (no step until } t) = \int_t^{\infty} dt p_t(\tau) d\tau = 1 - \int_0^t p_t(\tau) d\tau. \tag{1.34}$$

Let's start our walkers from $x = 0$ at $t = 0$ and let

$$p(x, t) \equiv \text{prob (walker arrives at } x \text{ at time } t) \tag{1.35}$$

$$= \int_{-\infty}^{\infty} dy \int_0^t d\tau p_x(y) p_t(\tau) p(x - y, t - \tau) + \delta(x) \delta(t). \tag{1.36}$$

Here the idea is that x, t is reached by a jump of length y (after waiting time τ) by a walker that arrived at $x - y$ at time $t - \tau$. Notice that I really mean *arrives* at x at time t . In contrast, the probability that the walk is *at* x at time t is

$$\rho(x, t) = \int_0^t d\tau p(x, t - \tau) \psi(\tau) \tag{1.37}$$

– the walker arrived at some earlier time τ and remained there until t .

These equations involve convolutions, so let's solve them by Fourier transformation.

$$p_x(k) \equiv \int_{-\infty}^{\infty} dx p_x(x) e^{-\mathbf{i}kx}, \quad p_t(\omega) = \int_0^{\infty} dt e^{-\omega t} p_t(t). \quad (1.38)$$

Notice that I used imaginary frequency ω here, anticipating that it would decrease the number of factors of $\mathbf{i}$ appearing; sometimes this is called Laplace transform, but it's just the same thing. So (1.36) is

$$p(k, \omega) = p_x(k) p_t(\omega) p(k, \omega) + 1 \quad (1.39)$$

which is solved by

$$p(k, \omega) = \frac{1}{1 - p_x(k) p_t(\omega)}. \quad (1.40)$$

Similarly, the Fourier transform of (1.37) is

$$\rho(k, \omega) = \frac{1}{1 - p_x(k) p_t(\omega)} \psi(\omega) \quad (1.41)$$

and we need to know the FT of the survival probability $\psi(\omega)$. But (1.34) says

$$p_t(\omega) = \int_0^{\infty} dt e^{-\omega t} p_t(t) = - \int_0^{\infty} dt e^{-\omega t} \psi'(t) \stackrel{\text{IBP}}{=} 1 - \omega \psi(\omega) \quad (1.42)$$

so $\psi(\omega) = \frac{1 - p_t(\omega)}{\omega}$.

Now recall that the Fourier transform of a probability distribution has a wonderful interpretation as the generating function for the moments of the distribution:

$$p_x(k) = \int dx e^{-\mathbf{i}kx} p_x(x) = \langle e^{-\mathbf{i}kx} \rangle = \int dx p_x(x) \left(1 - \mathbf{i}kx - \frac{k^2}{2} x^2 + \dots \right) \stackrel{?}{=} 1 - \mathbf{i}k \underbrace{\langle x \rangle}_{=0} - \frac{k^2}{2} \sigma^2 + \dots \quad (1.43)$$

The small k behavior gives the long-distance behavior, *i.e.* the low moments. Similarly, $p_t(\omega) \simeq 1 - \langle t \rangle \omega + \dots = 1 - \tau \omega + \dots$.

In the step with the '?' we switched the order of integration and can run into trouble. But let's assume first that all these moments are finite, and therefore the characteristic functions are actually analytic at $k = 0$ and $\omega = 0$. Putting these results together, the low-momentum behavior is $1 - p_x(k) p_t(\omega) \simeq \sigma^2 k^2 / 2 + \tau \omega + \dots$ and $1 - p_t(\omega) \simeq \tau \omega$, so $\psi(\omega) = \tau$ and

$$\rho(k, \omega) \simeq \frac{1}{\tau \omega + \sigma^2 k^2 / 2} \tau = \frac{1}{\omega + \frac{\sigma^2}{2\tau} k^2} \quad (1.44)$$

which you recognize as the Fourier transform of

$$\partial_t \rho(x, t) = D \partial_x^2 \rho(x, t) + \delta(x) \delta(t), \quad D = \frac{\sigma^2}{2\tau}, \quad (1.45)$$

the diffusion equation.

In contrast, suppose the distribution from which we sample the step lengths is long-tailed,

$$p_x(x) \stackrel{x \gg 1}{\sim} \frac{\sigma^z}{|x|^{1+z}}, \quad 0 < z < 2 \quad (1.46)$$

(where σ is a length scale to make up the units). Then $\langle x^2 \rangle = \lim_{R \rightarrow \infty} R^{2-z}$ diverges if $z < 2$. Such a thing is called a *Levy flight*. With some work you can see that⁸

$$\underbrace{p_x(k=0) - p_x(k)}_{=1} = \int_{-\infty}^{\infty} dx (1 - e^{-ikx}) p_x(x) \simeq (\sigma|k|)^z \quad (1.48)$$

so that $p_x(k) = 1 - (\sigma|k|)^z$ plus terms with larger powers of k , and

$$\rho(k, \omega) \simeq \left(\omega + \frac{\sigma^z}{\tau} |k|^z \right)^{-1} \quad (1.49)$$

and

$$\rho(k, t) \simeq e^{-D_z t |k|^z} \quad (1.50)$$

which is the Fourier transform (actually the definition) of

$$\partial_t \rho = D_z \partial_x^z \rho + \delta(x) \delta(t). \quad (1.51)$$

(The analog of the diffusion constant is $D_z \equiv \frac{\sigma^z}{2\tau}$.) Maybe it is already clear from dimensional analysis, but you can check that⁹

$$\langle |x(t)|^q \rangle = \int_{-\infty}^{\infty} dx \rho(x, t) |x(t)|^q = \int dk \int dx |x|^q e^{ikx} e^{-D_z t |k|^z} \sim t^{q/z}. \quad (1.52)$$

Setting $q = 2$, you can see this is superdiffusive behavior with $\alpha = 2/z$.

You can guess what happens if we take σ^2 finite but choose the waiting times from a distribution with $p_t(t) \stackrel{t \gg 1}{\sim} \frac{t_0^\alpha}{t^{1+\alpha}}$.

⁸The key step is to do the integral by *scaling*:

$$\int_{-\infty}^{\infty} dx e^{-ikx} \frac{\sigma^z}{|x|^{1+z}} \stackrel{y \equiv |k|x}{=} \sigma^z |k|^z \underbrace{\int_{-\infty}^{\infty} dy \frac{e^{-iy}}{|y|^{1+z}}}_{\text{independent of } k}. \quad (1.47)$$

⁹Again, we can do the integral by scaling: $y \equiv Ax, p \equiv k/A, A = (D_z t)^{1/z}$.

An interesting question, I think, is: can we get ballistic behavior from CTRW? It seems to require $\langle t \rangle = 0$, so that we keep the ω^2 term in the expansion of $p_t(\omega)$. That is, it requires negative waiting times. Actually it seems that it is possible to make sense of this by introducing an extra two-valued variable, somewhat analogous to the role of spin in the Dirac equation.

By the way the label ‘OU’ on $\alpha = 0$ in (1.33) stands for Ornstein-Uhlenbeck. We’ll explain this name in §2. [End of Lecture 3]

1.7 Large deviations

[LP §3.3; H. Touchette, *The large deviation approach to statistical mechanics*, Phys. Rep. 478, 1 (2009); Vergassola and Majumdar 0811.2290; Esposito et al in Avanzini et al (eds), *Methods and conversations in (post)modern thermodynamics*; an application to estimating UV sensitivity in cosmic inflation: Cohen, Green, Premkumar.]

The cumulants of a random walk, assuming that they are finite and assuming small-enough correlations between the steps, approach a universal behavior governed by the Gaussian distribution. Is this all the information? In fact, even when the CLT holds, the agreement with $P(X)$ need only hold within a few σ of $\langle X \rangle$.

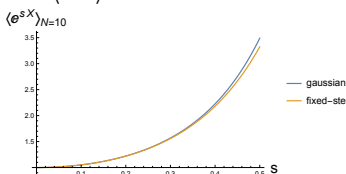
We can also ask questions whose answers are not so universal. For example, we can ask about very unlikely events, events in the tails of the distribution, like what is the probability after N steps of the walker ending up a distance ba_0N away from the starting point, for some number $b > 1$. If the step sizes are all exactly a_0 , this probability is zero, but if each step is governed by a Gaussian distribution with $\sigma = a_0$, there is a very small but nonzero probability. So this is a well-defined quantity that remembers more information than just the universality class.

An observable that probes this quantity is the dangerous object $\langle e^{sX} \rangle$ for some real number $s > 0$ (note the similarity to the moment-generating function), where $X \equiv \sum_{i=1}^N x_i$ as above. With identically-distributed independent steps, this is

$$\langle e^{sX} \rangle = \langle e^{sx} \rangle^N \stackrel{\text{gaussian steps}}{=} e^{N\sigma^2 s^2/2} \quad (1.53)$$

where I used the very useful identity for Gaussian distributions $\langle e^A \rangle = e^{\langle A^2 \rangle/2}$.

This is actually a big number – the factor of e^{sx} can be very large at large x and overcomes the suppression by the distribution. (For the equal-step distribution, we have instead $\langle e^{sX} \rangle^{\text{equal steps}} = \cosh^N(sa_0)$. They agree at small $s \ll 1/a_0$.)



A systematic method to study such dangerous observables that probe the tail of the distribution is called the *method of large deviations*. The idea is just that for such

an observable, we can do the integrals by saddle point – the enormous suppression by the tail of the distribution and the enormous growth of the operator struggle with each other to create a sharply peaked integrand.

There are many reasons we might care about extreme events. For example, people seem to care about record-best achievements in sports. More objectively importantly but harder to quantify, dramatic scientific progress occurs by outlying events. Other examples where we might want to know about tails of the distribution include rogue waves, solar flares, earthquakes.

First, let's rescale the displacement variable and study instead the average of the displacements $S \equiv X/N$. The idea then is to write the distribution for S as

$$P_N(S) \simeq f(S)e^{-NI(S)}; \quad (1.54)$$

such a distribution with positive $I(S)$ (which is called the rate function or Cramer function) is said to obey a large deviation principle (the prefactor $f(S)$ by definition has no N in the exponent). More precisely, it is defined by

$$I(S) = - \lim_{N \rightarrow \infty} \frac{1}{N} \log P_N(S) \quad (1.55)$$

to kill everything but the exponential dependence on N .

If the variables are iid, I claim that the rate function I is

$$I(S) = \sup_q (qS - \lambda(q)) \quad (1.56)$$

where $\lambda(q)$ is the cumulant generating function for each variable x : $\lambda(q) \equiv \log \langle e^{qx} \rangle$.

To see this, write the probability distribution for S as

$$P_N(S) = \int dx_1 \cdots dx_N p(x_1 \cdots x_N) \delta(S - \sum_i x_i/N) \quad (1.57)$$

$$= \int dk e^{ikS} \langle e^{-ik \sum_i x_i/N} \rangle \quad (1.58)$$

$$\stackrel{\text{iid}}{=} N \int dp e^{iNpS} \langle e^{-ipx} \rangle^N = N \int dp e^{-NQ(p,S)} \quad (1.59)$$

where $p \equiv k/N$, $Q(p, S) \equiv -ipS - \log \langle e^{-ipx} \rangle$, and you can see that at large N the p integral wants to be done by saddle point (or stationary phase). The saddle-point condition is $\partial_p Q(p, S) = 0$, which can be rewritten as

$$S = \partial_q \lambda(q) \quad (1.60)$$

with $q \equiv -ip$, the same Legendre transform condition we found above.

The idea is just that if (1.54) holds, we can do the integrals by saddle point:

$$\langle e^{qX} \rangle = \langle e^{NqS} \rangle \simeq \int dS e^{N(qS - I(S))} \simeq e^{N \sup_S (qS - I(S))}. \quad (1.61)$$

But the LHS is $\langle e^{qX} \rangle \stackrel{\text{iid}}{=} \langle e^{qx} \rangle^N \equiv e^{N\lambda(q)}$, so (1.61) says that $I(S)$ and $\lambda(q)$ are related by Legendre transform, which is (1.56).

You can check, for example, that (1.56) gives the correct expressions for the examples of uniform-step and gaussian-step or exponentially-distributed-step random walks, including bias. For the Gaussian case, we already found in (1.53) that $\lambda(q) = \sigma^2 q^2/2$, so we need to extremize the function $I(q, s) = qs - \sigma^2 q^2/2$ wrt q , which happens when $0 = \partial_q I(q, s)|_{q=q_*}$, so $q_* = s/\sigma^2$, and at the extremum we find $I(s) = I(q, s)|_{q=q_*} = \frac{s^2}{2\sigma^2}$. This agrees (exactly in this case, up to normalization) with our result (1.1), $P(s) = \mathcal{N} e^{-N \frac{s^2}{2\sigma^2}}$.

This story that I've told has other names in other areas of physics. In quantum field theory, it is called the background field method. In condensed matter physics $I(S)$ is called the Landau-Ginzburg effective potential. The idea in each case is that we are interested in some macroscopic variable (here it is S , in QFT it is the “classical field”, in condensed matter physics it is some order parameter, like a magnetization field), and want to know the “cost” for that variable to take a particular value (the meaning of cost varies – in QFT it is the effective action, in condensed matter in equilibrium it is the Gibbs free energy, here it is the log of the probability (divided by N)). To do this, we apply an external force to that variable (here, this is the qX term, so q is a generalized force conjugate to the variable X); then we choose the value of the force q so that the variable attains the value we've chosen. This is the content of the saddle point condition in (1.56).

A few properties of $I(S)$: $I(S) \geq 0$ or else the probability P_N would be > 1 . From the fact that it arises by a Legendre transform, $I(S)$ is convex. $I(\langle x \rangle) = 0$ is a minimum so that $P_N(S) \xrightarrow{N \rightarrow \infty} \delta(S - \langle x \rangle)$ is the law of large numbers.

It's interesting to Taylor expand about $S = \langle x \rangle$, letting $z \equiv S - \langle x \rangle$, so

$$P_N(S) \approx e^{-N \left(\frac{z^2}{2\sigma^2} + c_3 z^3 + c_4 z^4 + \mathcal{O}(z^5) \right)} \quad (1.62)$$

$$= P_{\text{CLT}}(S) e^{-N(c_3 z^3 + c_4 z^4 + \mathcal{O}(z^5))} \quad (1.63)$$

From this expression you can see that the CLT answer (which we wrote in terms of $X \equiv \sum_i (x_i - \langle x \rangle)/\sqrt{N}$, so that $z = X/\sqrt{N}$) is good when

$$\begin{cases} Nz^3 \ll 1 \Leftrightarrow |X| \ll N^{1/6}, & \text{for } c_3 \neq 0 \\ Nz^4 \ll 1 \Leftrightarrow |X| \ll N^{1/4}, & \text{for } c_3 = 0 \end{cases}. \quad (1.64)$$

As you can see, whether the CLT gives the whole story depends on what questions we ask.

1.8 Random matrices

The dynamics of a big nucleus (or some other complicated strongly-interacting system) is certainly a non-equilibrium question. To answer this question, we need to know the Hamiltonian, and in particular its spectrum $H|\psi_n\rangle = E_n|\psi_n\rangle$; then the time evolution of an initial state $|a(0)\rangle = \sum_n a_n |\psi_n\rangle$ is

$$|a(t)\rangle = e^{-iHt} |a(0)\rangle = \sum_n a_n e^{-iE_n t} |\psi_n\rangle. \quad (1.65)$$

But you might not think it is a statistical mechanics question. To answer it, we need to answer the question: what is H ? A big nucleus is made of a whole lot of nuclei (hundreds) interacting in a complicated way, and so H is a big ($N \times N$) complicated matrix. Following Wigner, let's apply the Idea of Statistical Mechanics, *i.e.* the strategy that if something is complicated enough, we don't need to know exactly what it is, but rather can sample it from a suitable distribution. Let's suppose H is a *random matrix*, distributed according to

$$P(H) \propto e^{-\text{tr}V(H)} \quad (1.66)$$

where $V(H)$ is a polynomial in the matrix H , where the products are matrix multiplication. Since we are trying to model a Hamiltonian of a closed system, we want it to be hermitian. We can impose time-reversal symmetry so H is a real (and so hermitian implies) symmetric matrix. So this is a distribution on $N(N-1)/2$ variables $H_{ij} = H_{ji}$.

When $V(X) = X^2$, this distribution actually factorizes

$$P(H) \propto e^{-\sum_{ij} H_{ij}^2} \quad (1.67)$$

(though the variances of the off-diagonal entries are half as big as those of the diagonal entries). Let's focus on this Gaussian case (actually there is a version of the CLT here, and as long as the moments are finite the result for the eigenvalue distribution is the same in an appropriate large- N limit). Given that the distribution for steps of a Gaussian random walk was of the form $p(r_i) \propto e^{-r_i^2}$, you can think of this as just a collection of $N(N+1)/2$ Gaussian random walkers, but we are asking a strange collective question about them.

Since the distribution is basis-independent (invariant under $H \rightarrow UHU^{-1}$), it actually only depends on the eigenvalues of H , whose distribution is what we care about.

Changing variables and doing the integral over the basis choice U gives

$$Z = \int dH e^{-\text{tr} H^2} = \int \prod_{i=1}^N d\lambda_i \Delta(\lambda) e^{-\sum_i \lambda_i^2}, \quad \text{where } \Delta(\lambda) \equiv \prod_{i < j}^N (\lambda_i - \lambda_j) \quad (1.68)$$

is called the Vandermonde determinant. Writing $\Delta(\lambda) = e^{\sum_{i < j} \log(\lambda_i - \lambda_j)}$, you can see that it produces a logarithmic repulsion between the eigenvalues, realizing very explicitly the familiar phenomenon of eigenvalue repulsion for a parametrized family of Hamiltonians¹⁰.

Let's think for a moment about the more general case

$$Z = \int dM e^{-N \text{tr} V(M)} = \int \prod_{i=1}^N d\lambda_i \Delta^\beta(\lambda) e^{-N \sum_i V(\lambda_i)}, \quad (1.69)$$

where $\beta = 1$ for GOE. I put a factor of N (redefine $M \equiv \sqrt{N}H$) to emphasize what we should do next. A sign that we are doing something right is that at large N the integral simplifies – we can do it by saddle point. The saddle point condition is

$$\frac{\beta}{N} \sum_{j \neq i} \frac{1}{\lambda_i - \lambda_j} = V'(\lambda_i). \quad (1.70)$$

Let's collect the information about the distribution of eigenvalues into a function of a complex variable z :

$$\omega(z) \equiv \frac{1}{N} \text{tr} \frac{1}{z - M} = \frac{1}{N} \sum_i \frac{1}{z - \lambda_i} \quad (1.71)$$

called the resolvent. Using the fact that $\frac{1}{x \pm i\epsilon} = P \frac{1}{x} - i\pi \delta(x)$ (where P is for 'principal part'), the resolvent has the virtue that it encodes the density of eigenvalues

$$\rho(x) \equiv \frac{1}{N} \sum_i \delta(x - \lambda_i) \quad (1.72)$$

(normalized so that $\int dx \rho(x) = 1$). $\rho(x)$ and the resolvent are related by

$$\omega(z) = \int \frac{\rho(\lambda) d\lambda}{z - \lambda} \quad \text{or} \quad \rho(x) = \frac{1}{2\pi i} (\omega(z - i\epsilon) - \omega(z + i\epsilon)). \quad (1.73)$$

¹⁰Focus on two energy levels that might collide and forget about the rest. The general Hermitian 2×2 matrix is then $h_2 = d_0 \mathbb{1} + \vec{d} \cdot \vec{\sigma}$ with spectrum $d_0 \pm \sqrt{|\vec{d}|^2}$. The two levels only collide if $0 = d_x = d_y = d_z$ – three real conditions on the parameters. This means that in a generic one-parameter family of Hamiltonians, the levels never collide, but rather are seen to repel each other.

That is, the eigenvalue distribution ρ is proportional to the discontinuity of ω across the real axis, so it is only nonzero where ω has branch cuts. So we can rewrite the saddle point equation as

$$\beta P \int \frac{\rho(x)dx}{\lambda - x} = V'(\lambda) \quad \text{or} \quad \omega(z + \mathbf{i}\epsilon) + \omega(z - \mathbf{i}\epsilon) = \frac{2}{\beta} V'(z). \quad (1.74)$$

How to solve such equations? I can offer two relatively simple suggestions, both valid for large N . (Let me set $\beta = 2$ now; as you can see from the preceding equation, this amounts to a rescaling of V .) The first is to multiply (1.70) by $\frac{1}{\lambda_i - z}$ and sum over i , to get

$$\omega^2(z) + \frac{1}{N} \omega'(z) - V'(z) \omega(z) = -\frac{1}{N} \sum_i \frac{V'(z) - V'(\lambda_i)}{z - \lambda_i}. \quad (1.75)$$

This is a nice combination because the second term on the LHS is small when N grows. It becomes

$$\omega^2(z) - V'(z) \omega(z) + R(z) = 0 \quad (1.76)$$

where

$$R(z) \equiv \int d\lambda \rho(\lambda) \frac{V'(z) - V'(\lambda)}{z - \lambda} \xrightarrow{V(x) \rightarrow x^2/2} 1 \quad (1.77)$$

is a polynomial in z of degree $\ell - 2$ when $V(z)$ is a polynomial of degree ℓ . In particular, in the Gaussian case, it is just 1. So then (1.76) is a quadratic equation for the resolvent, whose solution is

$$\omega(z) = \frac{1}{2} (V'(z) - \sigma(z)), \quad \sigma(z) \equiv \sqrt{V'(z)^2 - R(z)}. \quad (1.78)$$

The roots of the polynomial $V'(z)^2 - R(z)$ are branch points of the resolvent function, and are therefore boundaries of the support of the eigenvalue density $\rho(x)$. We choose the negative root because the other one has the wrong behavior at large z .

The second method is a bit less tangible. The idea is to observe that the resolvent is a function with the following properties:

1. $\omega(z)$ is analytic for z not in the support of ρ .
2. $\omega(z) \xrightarrow{z \rightarrow \infty} \frac{c}{z}$ where, using (1.73) the constant $c = \int dx \rho(x) = 1$.
3. The saddle point condition says that $\omega(x + \mathbf{i}\epsilon) + \omega(x - \mathbf{i}\epsilon) = V'(x)$.

So what is a function with a cut running between $-a$ and a on the real axis, analytic otherwise, and going to $1/z$ at large z ? The claim is that these properties specify the function to be of the form

$$\omega(z) = \frac{1}{2} \left(V'(z) - M(z) \sqrt{z^2 - a^2} \right) \quad (1.79)$$

where $M(z)$ is a polynomial. The third condition then says, for the Gaussian case,

$$\omega(z) = \frac{1}{2} (z - \sqrt{z^2 - 4}), \quad (1.80)$$

where $a = 2$ is determined by condition 2.

Either way, the result for the discontinuity across the cut is

$$\rho(x) = \frac{1}{2\pi i} (\omega(x - i\epsilon) - \omega(x + i\epsilon)) = \frac{1}{2\pi} \sqrt{4 - x^2}, \quad x \in [-2, 2]. \quad (1.81)$$

You can check the normalization condition $\int dx \rho(x) = 1$. This is the Wigner semicircle distribution.

As we've discussed above, it is possible to think about the central limit theorem as an example of universality explained by the existence of a fixed point of a renormalization group that coarse-grains a random walk into blocks of steps. Can we provide an analogous RG understanding of the universality of the Wigner semicircle distribution, i.e. that the distribution of eigenvalues of a random matrix in a general potential behaves like $\rho(x) \sim \sqrt{a - x}$ near the edge of the distribution? (Note that we can change the behavior in the interior of the distribution by changing the form of the potential $V(M)$, but the behavior near the edge of the distribution is harder to modify.)

So let me try to offer a third, somewhat incomplete, RG-based perspective on the solution to the one-matrix integral. One possibility is to define a coarse-graining procedure for the matrix by decomposing it into blocks:

$$H_{N+M} = \begin{pmatrix} A & C \\ C^T & B \end{pmatrix} \quad (1.82)$$

and performing the integral over C, B to get an effective action for $H_N = A$:

$$\frac{1}{Z_{N+M}} \int d^{M^2} B d^{NM} C e^{-(N+M)\text{tr}V(H)} = \frac{1}{Z_N} e^{-N\text{tr}\tilde{V}(A)}. \quad (1.83)$$

(Actually, the exact effective potential will not be a polynomial in A , and also it will involve multiple-trace terms, like $\text{tr}A^2\text{tr}A^2$. These effects can be ignored at large N .)

I will say two things about this. First, let's think about the resolvent:

$$G(z) = (H_{N+M} - z\mathbb{1})^{-1} = \begin{pmatrix} G_{AA}(z) & G_{AB}(z) \\ G_{BA}(z) & G_{BB}(z) \end{pmatrix}. \quad (1.84)$$

The Schur complement formula, a fact about matrices, says that

$$G_{AA}(z) = (A - z\mathbb{1}_N - \Sigma(z))^{-1}, \quad \Sigma(z) = C(B - z\mathbb{1}_M)^{-1}C^T \quad (1.85)$$

that is the 'self-energy'

$$\Sigma_{ij}(z) = \sum_{\alpha, \beta=1}^M C_{i\alpha} (B - z\mathbb{1})_{\alpha\beta}^{-1} C_{\beta j}^T, \quad i, j = 1..N \quad (1.86)$$

is a sum of M^2 terms, each of which is a rank-one $N \times N$ matrix. Now we must try to integrate out the off-diagonal stuff. Let's retreat to the Gaussian case for a moment. Then Σ_{ij} is a sum of squares of iid Gaussians, with $\langle C_{i\alpha} C_{j\beta} \rangle_C = \frac{\sigma^2}{N+M} \delta_{ij} \delta_{\alpha\beta}$. Thus we can approximate

$$\langle \Sigma(z)_{ij} \rangle_C = \sum_{\alpha\beta} \langle C_{i\alpha} C_{j\beta} \rangle (B - z)^{-1}_{\alpha\beta} = \frac{\sigma^2 M}{N + M} \omega_B(z) \delta_{ij} \xrightarrow{N \rightarrow \infty, M/N=c} \sigma^2 c \omega_B(z) \quad (1.87)$$

where $\omega_B(z) \equiv \frac{1}{M} \text{tr}(B - z\mathbb{1}_M)^{-1}$ is the trace of the resolvent for the $M \times M$ B block. Now the matrix model is basis-independent, so $\omega_A(z) = \omega_B(z)$, and $\omega(z) = \frac{1}{1+c} (\omega_A(z) + c\omega_B(z)) = \omega_A = \omega_B$, where $c = \frac{M}{M+N}$. Can you complete this argument to conclude that

$$\omega(z) = \frac{1}{z - \omega(z)} \quad (1.88)$$

which is the quadratic equation for $\omega(z)$ we found earlier? I believe the idea is that the mean of A is zero and its fluctuations are parametrically smaller in $1/N$ than the $z - \omega$ term.

Now let's consider a perturbation of the gaussian fixed point, following Brezin and Zinn-Justin, and see if it is relevant or irrelevant. Take $V(X) = \frac{1}{2}X^2 + \frac{g}{4}X^4$. The claim is that the effective potential is $\tilde{V}(X) = \frac{1}{2}X^2 + \frac{\tilde{g}}{4}X^4 + \dots$ where

$$\tilde{g} = g - \frac{\beta(g)}{N}, \quad \beta(g) = -g - 6g^2 + \mathcal{O}(g^3). \quad (1.89)$$

This map has fixed points ($\beta(g) = 0$) at $g_\star = -1/6$ and $g_0 = 0$. The former is an unstable fixed point related to a collision of cuts, while the latter is the stable Gaussian fixed point. 

The result is that the average distribution of eigenvalues (with our previous normalization) is

$$\rho(\lambda) = \frac{1}{\pi N} \sqrt{2N - \lambda^2}, \quad (1.90)$$

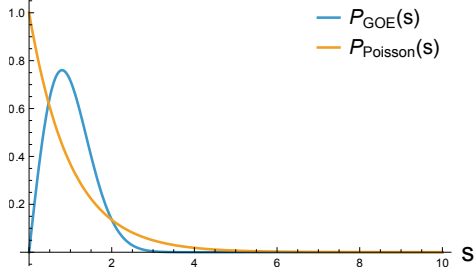
the Wigner semicircle distribution. More generally, for other choices of potential, it is still true that (for generic potentials) the eigenvalues form bands at the edges of which ρ vanishes like the square root of the deviation from the edge. By tuning the potential to cause collisions, one can make it vanish like higher powers.

So I wanted to mention random matrices partly as a perhaps-unfamiliar example of the application of the Idea of Statistical Mechanics (to a non-equilibrium question). The statistics of the energy levels computed from this ensemble actually match very well with statistics of spectra of quantum chaotic systems¹¹ such as large nuclei, once the (dimensionful) average spacing is rescaled to match. The average value of $\min(r, 1/r)$ with $r \equiv \frac{E_i - E_{i-1}}{E_{i+1} - E_i}$, is a universal number that differs between random matrix energy levels (where it is ≈ 0.53 for GOE, the time-reversal-invariant ensemble we studied above; the answer is different for GUE (the ensemble of general Hermitian matrices without any assumption of time-reversal symmetry which is therefore invariant under the unitary group) or GSE (an ensemble of matrices preserved by the symplectic group) and Poisson-distributed levels like those of an integrable system (where the expected value of $\min(r, 1/r)$ is ≈ 0.38 ; you will find the exact number on the homework).

¹¹What is a quantum chaotic system? One definition is a quantum system with a classical limit that is chaotic (in the sense that the dynamics ergodically explores the phase space, rather than being trapped on some level sets of conserved quantities). This version is called the Bohigas-Giannoni-Schmit conjecture, and RMT statistics have been verified in the quantum versions of lots of simple chaotic systems like billiards and things like that.

However, the existence of a classical limit is not really necessary. A better definition of a quantum chaotic system is just a quantum system that is not integrable. Quantum integrability is a special property that means that there exist enough conserved charges to decompose the Hamiltonian into small blocks. Moreover, the conserved charges need to be *simple* operators, like sums over space of local operators. An example of a not simple operator is a projector onto an eigenstate $|E\rangle$; such operators are always conserved quantities (they commute with H), but they do not actually constrain the dynamics. Such an operator is extremely fragile: a small perturbation of the Hamiltonian will completely rearrange it, in contrast with the usual kind of conserved quantity that does constrain the dynamics.

So, basically, if you don't leave out terms in your Hamiltonian, it will have RMT statistics. People will try to tell you that a phenomenon called 'Many Body Localization' is somehow an exception to this statement. The evidence for its existence is only in $d = 1$ and not entirely convincing, and in any case it can be understood as a sort of hidden integrability: despite the fact that you included enough interactions, there is a set of conserved charges that are approximately sums of local operators.



More precisely, the probability distribution for successive events in a Poisson process (with mean spacing 1) is $P_{\text{Poisson}}(s) = e^{-s}$. This is biggest when $s = 0$. In contrast, because of level repulsion, $P_{\text{RMT}}(s)$ is zero when $s = 0$. The level-spacing distribution for random matrix eigenvalues is

$$P_{\beta}(s) \approx A_{\beta} s^{\beta} e^{-B_{\beta} s^2} \quad (1.91)$$

where $\beta = 1, 2, 4$ for the three ensembles GOE, GUE and GSE of real symmetric, hermitian, and quaternionic random matrices.

The approximate formula (1.91) for the level spacing distribution is called the Wigner surmise¹². You can get a sense of it by deriving $P_{\text{RMT}}(s)$ for the case $N = 2$.

The theory of random matrices has many other, often-surprising, applications, including as a toy model low-dimensional string theory and 2d gravity. The idea here is that the calculation of observables can be written in terms of Feynman diagrams that have a locally planar structure (the propagators are double lines associated to the two matrix indices of the field), and the diagrams that dominate are the ones that can be drawn on the smallest-genus Riemann surface. Each diagram can be regarded as a triangulation of a Riemann surface, and so their sum is a sum over random 2d geometries. (Whether this connection to random surfaces is related to the one I'll mention below I do not know, please tell me if you do.)

It also makes a striking appearance in number theory: the level statistics of the imaginary part of the nontrivial zeros of the Riemann zeta function (which famously have real part one half) match GUE. The Riemann zeta function is defined (in some domain of the complex s plane) by

$$\zeta(s) \equiv \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{\text{primes, } p} \sum_{k_p=0}^{\infty} \frac{1}{p^{k_p s}} = \prod_{\text{primes, } p} (1 - p^{-s})^{-1}. \quad (1.93)$$

So you can see that this function knows about the distribution of primes. Its analytic continuation is only zero at negative even integers and when $\text{Re}(s) = \frac{1}{2}$. The imaginary parts of the latter zeros can be regarded as a time series like the times of decay of radioactive atoms or the eigenvalue of a random matrix. It turns out that they repel each other like the latter case. This is a deep mystery.

¹²The constants for GOE, with the mean level spacing set to unity, are $A_{\beta=1} = \frac{\pi}{2}, B_{\beta=1} = \frac{\pi}{4}$. More generally,

$$A_{\beta} = 2 \frac{\Gamma((\beta+2)/2)^{\beta+1}}{\Gamma((\beta+1)/2)^{\beta+2}}, \quad B_{\beta} = \frac{\Gamma((\beta+2)/2)^2}{\Gamma((\beta+1)/2)^2}. \quad (1.92)$$

Here is another application of random matrix theory to a non-equilibrium statistical problem, which I learned about from the Foreword by Freeman Dyson to the *Oxford Handbook of Random Matrix Theory*. I don't think I can do better than to just quote Dyson, who describes

the work of...Milan Krbalek and Petr Seba, who studied the bus system of the city of Cuernavaca in Mexico. Unlike other metropolitan bus systems, this system is totally decentralized, with no central authority and no time tables. Each bus is the property of the driver. At each bus stop, the drivers obtain information from bystanders about the time when the previous bus left. Using this information, the drivers independently adjust their speeds so as to maximize their incomes. Adjustments are made using mental arithmetic without any advanced algorithms. Krbalek and Seba recorded the actual departure times of buses at various stops for a month while the system was in normal operation. They found that the spacings between buses agree accurately with the Gaussian unitary ensemble of random matrix theory. The Gaussian unitary ensemble gives the best approximation to uniform spacing that the bus drivers can achieve, based on the limited information available to them.

The benefit of the bus drivers' self-regulation to the public is measured by R , the ratio between the average waiting-time of a passenger and the average spacing-time between buses. The best possible R is 0.5, which would occur if the bus spacings were exactly equal. If the buses are uncorrelated so that their spacings have a Poisson distribution, then $R = 1$. In Cuernavaca, with the buses correlated according to the Gaussian unitary ensemble, the value of R is $\frac{3\pi}{16} \approx 0.589$, much closer to the equal-spacing value than to the Poisson distribution value. It is remarkable that so large a public benefit results from so simple an optimization process. I am not able to determine whether the application of random matrix theory to financial markets, as described in Chapter 40 of this book by Bouchaud and Potters, yields any comparable benefits. When an expert on markets tells me that some piece of financial wizardry is sure to benefit mankind, I am inclined to believe that a Cuernavaca bus driver might do the job better.

[End of Lecture 4]

A very interesting question is about the distribution of the *groundstate* of H , *i.e.* the

minimum eigenvalue of H . This is governed by the *Tracy-Widom distribution*, which also appears in the statistics of surface growth, a topic that I hope to discuss below. The basic statement is the following. Let a be the upper boundary of the eigenvalue distribution, near which $\rho(x) \sim \sqrt{a-x}$ for $x < a$ (and $\rho(x) = 0$ for $x > a$). ($a = 2$ for the Gaussian case, with the matrix normalized as in (1.69).) One statement is

$$\text{Prob}(\lambda_{\max} > x) \sim e^{-N^2 \Phi(x)} \quad (1.94)$$

where $\Phi(x)$ is the rate function, which as $x \rightarrow a$ goes like $\Phi(a + \delta) = C\delta^{3/2}$.

To see the idea here, consider pulling a single eigenvalue from the saddle point distribution to a value $x > a$. The cost in probability for doing this is determined by the rate function:

$$\Phi(x) = V(x) - 2 \int \rho(y) \log |x - y| dy - c \quad (1.95)$$

where the constant c must be chosen so that $\Phi(a) = 0$.

Where did this come from? Recall that we can write the joint probability distribution for all the eigenvalues as a Coulomb gas (with the matrix normalized as in (1.69)):

$$P(\lambda_i) \propto e^{-N \sum_i V(\lambda_i) - \sum_{i < j} \log |\lambda_i - \lambda_j|} = e^{-N^2 \mathcal{E}[\rho]} \quad (1.96)$$

which at large N it is useful to think about in terms of an energy functional of the eigenvalue distribution $\rho(\lambda)$:

$$\mathcal{E}[\rho] = \int d\lambda V(\lambda) \rho(\lambda) - \frac{1}{2} \int \int d\lambda d\mu \rho(\lambda) \rho(\mu) \log |\lambda - \mu| + \ell \left(\int d\lambda \rho(\lambda) - 1 \right). \quad (1.97)$$

One use of this is that the saddle point equation (1.70) (actually its integral wrt λ) is obtained by varying $\mathcal{E}[\rho]$ with respect to ρ (ℓ is a Lagrange multiplier that keeps $\int \rho = 1$). If we plug in $\rho(\lambda) = \rho_0(\lambda) + \delta(\lambda - x)$, the change in the exponent is proportional to $\Phi(x)$. So $\Phi(x)$ is the effective action for one eigenvalue in the mean field of all the other eigenvalues. In the Coulomb gas picture, it is like an effective potential, including both the external potential V as well as the potential generated by all the other eigenvalues.

To figure out the behavior of $\Phi(x)$ (*i.e.* to see (1.94)), let's take a small detour of first differentiating (1.95) wrt x :

$$\Phi'(x) = V'(x) - 2 \int \frac{\rho(y) dy}{x - y} = V'(x) - 2\omega(x) \quad (1.98)$$

$$\stackrel{V(x)=x^2/2}{=} x - 2 \frac{x - \sqrt{x^2 - 4}}{2} = \sqrt{x^2 - 4}. \quad (1.99)$$

And therefore we can integrate to get

$$\Phi(x) = \int_2^x \sqrt{u^2 - 4} du. \quad (1.100)$$

Setting $x = 2 + \delta$, with $\delta \ll 1$, we get (let $u \equiv 2 + t$)

$$\Phi(2 + \delta) = \int_0^\delta dt 2\sqrt{t} \left(1 + \frac{t}{8} + \mathcal{O}(t^2) \right) = \frac{4}{3} \delta^{3/2} + \mathcal{O}(\delta^{5/2}). \quad (1.101)$$

A similar analysis works for near the edge of the support of ρ for generic potential using (1.79).

The Tracy-Widom law (of which I've only mentioned the asymptotic behavior; the full distribution is known) appears in many places, including random interface growth, like the shape of a coffee stain (see below). It also governs the fluctuations in the length of the largest increasing subsequence of a random permutation (this is the subject of the Baik-Deift-Johansson theorem). For the permutation $\{3, 1, 4, 2, 5\}$, the longest increasing subsequence (LIS) is $\{1, 2, 5\}$ of length 3. For a permutation of length n , the mean LIS length is well-peaked at $2\sqrt{n}$ for large n . The length of the LIS of a random permutation of length N can be mapped to the free energy of a directed polymer in a random 2d medium (*in a random medium* means that there is a potential that makes some locations of monomers more favorable than others, *directed* means the polymer can only stretch up and to the right), which in turn is governed by the KPZ universality class. The map is actually quite understandable: A permutation σ of length n , determines a collection of n points in the plane: $\{(i, \sigma(i))\}_{i=1}^n$. Think of this as determining a random medium, and the LIS is longest directed (up-and-to-the-right) path going through the specified points.

How are these problems related to the largest eigenvalue of a random matrix? The best way I've found to see this connection is via Dyson brownian motion, where we introduce an auxiliary time-dependent non-equilibrium problem whose stationary distribution is the RMT ensemble we studied above. The eigenvalues are random walkers in 1d, with the added feature that they repel each other with a logarithmic potential. The highest eigenvalue can be followed along the trajectory and is special because it's only repelled from one side. This leads to the TW distribution, and to the relation to surface growth. In this way I believe one can find an RG understanding of this universality, too (like we did for CLT), which (I hope) we'll discuss when we get to KPZ.

I can't resist mentioning one more application of RMT, related to the onset of chaos. Suppose we have a physical system, like an ecosystem, that is in some stable state, which by a choice of coordinates we can call $x_i = 0$. In the original work by Robert May, x_i was the deviation from the equilibrium population density of species $i = 1..N$. The assumption of stability says that the unperturbed system has a collection of damping times, which by a choice

of units we set to 1: the linearized equation of motion is $\dot{x}_i(t) = -x_i(t)$. Now imagine perturbing the system somehow. For the question of stability, the only thing that matters is the linearized dynamics near the previously-stable fixed point, and so the information about the perturbation is encoded in the matrix J :

$$\dot{x}_i(t) = -x_i(t) + \alpha J_{ij} x_j(t). \quad (1.102)$$

(Actually I guess we have to shift the origin of the x s to account for a possible constant term in the perturbation.) α is a parameter that says how strong the perturbation is. Since this perturbation is some complicated thing, we may as well sample J_{ij} from the GOE ensemble.

The original fixed point is stable as long as -1 plus the largest eigenvalue of J stays negative. At large N , the largest eigenvalue of J is very probably at $\lambda_{\max} = 2$. The probability that the original fixed point remains stable under the perturbation with parameter α is then

$$P_{\text{stable}}(\alpha, N) = \text{Prob}(\lambda_{\max} \leq 1/\alpha) \xrightarrow{N \rightarrow \infty} \begin{cases} 1, & \alpha < \alpha_c = 2 \\ 0, & \alpha > \alpha_c \end{cases}. \quad (1.103)$$

At finite N , the transition is broadened out in a way determined by the Tracy-Widom distribution. For more on this see [1311.0580](#) (but beware that their random matrices differ from ours by a normalization factor of $\sqrt{2}$).

Our extensive use of the saddle point method in this section has been brought to you by the large letter N .

2 Stochastic Evolution

Interlude. In equilibrium statistical mechanics, the focus of our attention is the partition function Z . It provides the bridge between the microscopic picture $Z = \sum_{\text{states}, s} e^{-\beta H(s)}$ and the macroscopic world of thermodynamics $Z = e^{-\beta F}$. Away from equilibrium there is no perfect generalization of partition function. Zwanzig claims that the objects that should replace the focus of our attention are correlation functions. They have the advantage of often being measurable. Instead of equilibrium correlation functions

$$\langle \phi(x_1)\phi(x_2) \rangle = Z^{-1} \text{tr} e^{-\beta H} \phi(x_1)\phi(x_2) \quad (2.1)$$

now they also depend on *time*.

Given a time series of data $A(t), t \in [0, \tau]$ (such as a simulation of a particle trajectory, or a list of arrival times of buses, or the number of organisms in an ecosystem at a given time, or ...) there are a few basic operations we should do. The mean is

$$\bar{A} \equiv \frac{1}{\tau} \int_0^\tau dt A(t) \quad (2.2)$$

(or it could be a discrete sum, as it will always be in practice). Let $\delta A(t) \equiv A(t) - \bar{A}$ be the instantaneous deviation from the mean. Then

$$C(t) \equiv \frac{1}{\tau} \int_0^\tau ds \delta A(s) \delta A(t+s) \quad (2.3)$$

is the “time correlation function”. $C(0) = \overline{A^2 - \bar{A}^2}$ is the variance of the time series. Often it is interesting to consider the Fourier transform $C_\omega = \int dt e^{-i\omega t} C(t)$ since this (or rather its imaginary part, which is called the spectral density) is what is often measured, for example, in optical absorption experiments (in the case where $A(t)$ is an electric dipole moment). When people get excited about this, they use the name ‘Wiener-Khinchin theorem’ which is the statement that

$$C_\omega = \frac{1}{T} |X_T(\omega)|^2, \quad X_T(\omega) = \int_0^T dt X(t) e^{-i\omega t} \quad (2.4)$$

which you can just check from the definitions.

Here is where the key step of statistical mechanics goes: we approximate the time averages $\frac{1}{\tau} \int_0^\tau ds$ by a suitable ensemble average. This is the same idea as in equilibrium statistical mechanics. The word ‘ergodic’ is used when this step works.

2.1 Brownian motion

[Kardar, *Fields*, §9; [Cardy](#); Zwanzig, chapter 1; Cugliandolo, 2504.11994] In the previous section we studied a class toy models where a particle moved by jumps. Knowing something about Newtonian mechanics, one might wonder about a more accurate microscopic description that produces something like such motion.

As a paradigmatic example of stochastic evolution, let's think about a pollen grain suspended in a quiescent liquid (no turbulent flow for now, please). The model of this situation that we'll discuss (Langevin equation) has several great virtues. First, it played an important role in establishing the atomic nature of matter. A pollen grain is big enough to watch in a microscope and small enough to be buffeted by individual molecules in the liquid. A quantitative understanding of its meandering played an important role in convincing people that stuff is made of atoms.

Second, the same dynamics arises in a great many other contexts. For example, the role of the position of the pollen grain could be replaced by the concentration of some chemical reactant, which also experiences irregular fluctuations in time. We'll see many generalizations, for example, generalizations where instead of the position of a single particle, the degree of freedom in question is a field, like say the magnetization of a magnetic medium.

Unsurprisingly, the evolution of the position of the pollen grain satisfies Newton's law:

$$m\ddot{x} = F_{\text{total}} \quad (2.5)$$

where the RHS accounts for all forces on the particle. They all come from its interactions with the surrounding medium.

Let's first try just friction: $F_{\text{total}} = -\dot{x}/\mu$. This gives $m\dot{v} = -v/\mu$ so $v(t) \equiv \dot{x}(t) = v_0 e^{-t/\tau} \rightarrow 0$ if we wait long enough, meaning $t \gg \tau \equiv \mu/m$. But this disagrees with the idea that the particle should at late times equilibrate with the bath and have $\langle \frac{1}{2}mv^2 \rangle = \frac{d}{2}T$ (by equipartition) for some $T \neq 0$. So friction can't completely dominate.

To do better, we can include a noisy force. We emphasize here the key step of replacing deterministic but intractable dynamics by probabilistic but simple dynamics. (We'll briefly go back to the former in the next subsection.) So consider:

$$m\ddot{x} = \underbrace{-\frac{\partial U}{\partial x}}_{\text{external force}} \underbrace{-\frac{\dot{x}}{\mu}}_{\text{friction force}} + \underbrace{f_{\text{random}}(t)}_{\text{stochastic force}}. \quad (2.6)$$

I didn't write indices or little arrows but if the particle is moving in $d > 1$ dimensions this is a vector equation. The first term is an external (conservative, deterministic)

force, due to *e.g.* gravity and/or a laser trap keeping the particle in some region of space. The second term is a frictional force due to the viscosity of the fluid in which the particle is moving. For example, for a spherical particle of radius R , $\mu^{-1} = 6\pi\tilde{\eta}R$, where $\tilde{\eta}$ is a measure of the viscosity of the fluid. The last term is a stochastic force due to collisions with the fluid particles, which we will specify further below.

Where does the viscous force come from microscopically? Collisions of the tiny fluid particles with our probe particle (and each other). So the last two terms are not unrelated.

To make progress, let's think about the overdamped limit, where we simply ignore the inertia of our particle (the $\ddot{x}$ term):

$$\dot{x} = -\mu\partial_x U + \eta(t) \quad (2.7)$$

where $\eta(t) \equiv \mu f_{\text{random}}(t)$ is a random function of t with $\langle \eta(t) \rangle = 0$ – if the average is not zero, we can absorb it in the external force, *i.e.* add a linear term to the potential in that direction. (Such a random function with zero average is generally called *noise*.) Of course, this information does not yet specify the ensemble from which we are sampling the noise η . We will assume that the noise is *Gaussian*, meaning that the distribution is of the form

$$P[\eta] \propto e^{-\frac{1}{2} \int ds dt \eta(s) \eta(t) F(s,t)} \quad (2.8)$$

for some kernel $F(s,t)$. Gaussian means that Wick's theorem is true: correlation functions of many η s factorize into products of two-point functions, and hence are completely determined by the two-point functions, as a sum of Wick contractions *e.g.*:

$$\langle \eta_1 \eta_2 \eta_3 \eta_4 \rangle \stackrel{\text{Gaussian}}{=} \langle \eta_1 \eta_2 \rangle \langle \eta_3 \eta_4 \rangle + \langle \eta_1 \eta_4 \rangle \langle \eta_2 \eta_3 \rangle + \langle \eta_1 \eta_3 \rangle \langle \eta_2 \eta_4 \rangle \quad (2.9)$$

(where I wrote $\eta_i \equiv \eta(t_i)$ as a shorthand). The motivation for this choice is the central limit theorem discussed above: the force on our big particle is a sum of many approximately-independent forces.

The kernel $F(s,t)$ encodes correlations in time of the noise – in particular I claim that $\langle \eta(t) \eta(s) \rangle = F^{-1}(t,s)$, where F^{-1} is defined by $\int ds F^{-1}(t,s) F(s,r) = \delta(t-r)$, *i.e.* think of its two arguments as matrix indices. At this point I think I should pause and say something about path integrals.

Interlude about functional integrals. Functional integrals have a bad reputation among mathematicians and among students, but they are very useful (and we will use them quite a bit). The right way to think about them is as a convenient notation for a bunch of discrete integrals. That is, think of a function, say $\eta(t)$, on an interval $t \in [0, T]$ as a collection of values $\eta_n \equiv \eta(t_n)$ at a set of

grid points $t_n = \epsilon n$, $\epsilon \equiv T/N$. The integration measure $\int D\eta \cdots$ is defined to be $\int \prod_n d\eta_n \cdots$ where we take $N \gg 1$, $\epsilon \ll 1$ fixing $T = \epsilon N$.

Really the only functional integral that we know how to do is the special class of *Gaussian* functional integrals, where the integrand is of the form (2.8) (times polynomials in η). We can do this because it is just a whole bunch of coupled single-variable gaussian integrals

$$\int D\eta e^{-\frac{1}{2} \int dt ds \eta(t) F(t,s) \eta(s)} = \int \prod_{n=1}^N d\eta_n e^{-\frac{1}{2} \eta_n F_{nm} \eta_m} = \sqrt{\frac{(2\pi)^N}{\det F}}$$

where $F_{nm} \equiv F(t_n, t_m)$. (The last step is accomplished by changing variables to the eigenbasis of the matrix $F_{nm} = F_{mn}$ and using the fact that $\det F = \prod_\alpha f_\alpha$, where f_α are the eigenvalues of F .) The factor of $(2\pi)^{N/2}$ will always drop out of physical quantities and is therefore difficult to remember. The normalization condition (2.8) is then

$$1 = \int D\eta P[\eta] = \int D\eta e^{-\frac{1}{2} \int dt ds \eta(t) F(t,s) \eta(s)} / Z = \sqrt{\frac{(2\pi)^N}{\det F}} / Z. \quad (2.10)$$

More useful than the normalization factor Z is the two-point correlation function of the noise, which can be written as the functional integral

$$\langle \eta(t) \eta(s) \rangle \equiv \int [d\eta] P[\eta] \eta(t) \eta(s). \quad (2.11)$$

On the LHS $\eta(t)$ is a random variable, on the RHS it is just a number, a dummy integration variable. The discrete version of this equation is

$$\langle \eta_n \eta_m \rangle = \int \prod_{k=1}^N d\eta_k P(\eta_1 \cdots \eta_N) \eta_n \eta_m \stackrel{\text{Gaussian}}{=} Z^{-1} \int \prod_{k=1}^N d\eta_k \eta_n \eta_m e^{-\frac{1}{2} \eta_r F_{rs} \eta_s} \quad (2.12)$$

which makes it clear how to do it. We can make a generating function(al) for the correlations by studying

$$Z[J] \equiv \left\langle e^{\int dt J(t) \eta(t)} \right\rangle = \int [d\eta] P[\eta] e^{\int dt J(t) \eta(t)} \quad (2.13)$$

the discrete version of which is

$$\begin{aligned} Z[J] &= \int \prod_{k=1}^N d\eta_k e^{\sum_n J_n \eta_n} P(\eta_1 \cdots \eta_N) \stackrel{\text{Gaussian}}{=} Z^{-1} \int \prod_{k=1}^N d\eta_k e^{-\frac{1}{2} \eta_n F_{nm} \eta_m + \sum_n J_n \eta_n} \\ &= e^{\frac{1}{2} \sum_{nm} J_n (F^{-1})_{nm} J_m} = e^{\frac{1}{2} \int dt \int ds J(t) F^{-1}(t,s) J(s)} \end{aligned} \quad (2.14)$$

by completing the square. Arbitrary correlation functions of the η s are then given by

$$\langle \eta(t_1)\eta(t_2)\cdots\eta(t_k) \rangle = \frac{\delta}{\delta J(t_1)} \cdots \frac{\delta}{\delta J(t_k)} Z[J]|_{J=0}, \quad (2.15)$$

where $\frac{\delta}{\delta J(t)}$ is a functional derivative, defined by the fact that $J(t)$ and $J(s)$ are independent variables: $\frac{\delta J(t)}{\delta J(s)} = \delta(t-s)$. The discrete version of this last relation is just $\frac{\partial J_n}{\partial J_m} = \delta_{nm}$, and the discrete version of the generating function equation is

$$\langle \eta_{n_1} \cdots \eta_{n_k} \rangle = \partial_{J_{n_1}} \cdots \partial_{J_{n_k}} Z(J_1 \cdots J_N)|_{J=0}. \quad (2.16)$$

This relation (2.15) is more general than the Gaussian case, but only in (or near) the Gaussian case do we know $Z[J]$.

Irrelevant comment: *connected* correlation functions of η are given by

$$\langle \eta(t_1)\eta(t_2)\cdots\eta(t_k) \rangle_c = \frac{\delta}{\delta J(t_1)} \cdots \frac{\delta}{\delta J(t_k)} \log Z[J]|_{J=0} \quad (2.17)$$

and in the Gaussian case $\log Z[J] = \frac{1}{2} \int J F^{-1} J$ is quadratic in J so you can see that the only nonzero connected correlation function is the two-point function.

Note that I use square brackets in $P[\eta]$ to emphasize that this probability distribution is a functional, *i.e.* a machine that eats functions and spits out numbers.

[End of Lecture 5]

Thus, $\langle \eta(t)\eta(s) \rangle = F^{-1}(t, s)$. What should we choose for the kernel? If we think that the correlations between the collisions are very short-ranged in time, that is, that the memory of the previous collision is forgotten in a time shorter than we can resolve, it is tempting to approximate the noise as *white*, meaning that $F(s, t) = F(s, t)_{\text{white}} \equiv \delta(s-t)/\Gamma$ for some constant Γ . Why is this called white? It's because the log of $P[\eta]$ is then proportional to

$$\int dt \eta(t)^2 = \int d\omega \eta(\omega) \eta(-\omega) \quad (2.18)$$

meaning that every frequency is equally likely, no color is special in the spectrum of such noise. Thus, we'll assume

$$P[\eta] \propto e^{-\int dt \frac{\eta(t)^2}{2\Gamma}} \quad (2.19)$$

and therefore

$$\langle \eta_i(t) \eta_j(s) \rangle = \Gamma \delta_{ij} \delta(t-t') \quad (2.20)$$

where for the first time I made explicit possible spatial indices on η .

White noise does not really exist. Correlations in η are short-ranged because collisions *between* the fluid particles are rapid (that's why it's a fluid and not a gas), so later collisions with the pollen grain are uncorrelated with earlier ones. If τ is the average collision time of the fluid particles (with each other), we can approximate

$$\langle \eta(t)\eta(s) \rangle = F^{-1}(t, s) \approx e^{-|t-s|/\tau} \stackrel{t-s \gg \tau}{\approx} \Gamma \delta(t-s). \quad (2.21)$$

Example 1: $U = 0$. Let's go back to (2.7) and think about the case where there is no external force, so $\dot{x} = \eta$. This is solved by

$$x(t) = x(0) + \int_0^t ds \eta(s), \quad (2.22)$$

so $\langle x(t) \rangle = x(0)$, no drift. In turn the correlation function is

$$\langle (x(t) - x(0))^2 \rangle = \int_0^t ds_1 \int_0^t ds_2 \langle \eta(s_1)\eta(s_2) \rangle = d \cdot \Gamma \cdot t \quad (2.23)$$

where d is the number of spatial dimensions. So $x(t)$ is a sum of Gaussian-distributed random variables, which is then also Gaussian-distributed, with a variance that grows linearly in the number of terms. If at $t = 0$ we release a swarm of particles at the origin, where will they be at time t ? Let

$$P(\vec{x}, t) = P(x, t | 0, 0) = \langle \delta(x - x_\eta(t)) \rangle_\eta \quad (2.24)$$

be the probability density for the particle to be at $\vec{x}$ *given* that it was at $\vec{0}$ at time $t = 0$. The stochasticity in the last expression comes from entirely the random variable η : $x_\eta(t)$ is the trajectory obtained by solving the Langevin equation with a given η , with initial condition that the particle starts at the origin, $P(\vec{x}, t = 0) = \delta^3(\vec{x})$.

Let's solve for the evolution in steps, by finding $x_n \equiv x(t_n) = x(ndt)$. Integrating the Langevin equation gives

$$x_{n+1} - x_n = \int_{t_n}^{t_n+dt} ds \dot{x} = \int_{t_n}^{t_n+dt} ds \eta(s) \equiv \eta_{dt}. \quad (2.25)$$

(η_{dt} depends on n but I omit that decoration.) That is

$$x_{n+1} = x_n + \eta_{dt} \quad (2.26)$$

Each η_{dt} is a Gaussian random variable with

$$\langle \eta_{dt}^2 \rangle = \int_t^{t+dt} ds \int_t^{t+dt} ds' \langle \eta(s)\eta(s') \rangle = \int_t^{t+dt} ds \Gamma = \Gamma dt. \quad (2.27)$$

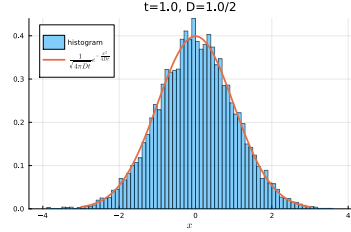
That is, $\eta_{dt} = \sqrt{\Gamma dt} \text{randn}()$ where $\text{randn}()$ denotes a Gaussian random variable with width 1. So the evolution rule is

$$x_{n+1} = x_n + \sqrt{\Gamma dt} \text{randn}() . \quad (2.28)$$

You might be surprised (I was) by the fact that $\sqrt{dt}$ appears here.

We know what distribution we get if we take $N = t/dt$ steps of a random walk with Gaussian steps of variance $\sigma_0 = \sqrt{\Gamma dt}$:

$$P(\vec{x}, t) = e^{-\frac{x^2}{4Dt}} \left(\frac{1}{4\pi Dt} \right)^{3/2}, \quad (2.29)$$



with $D = \Gamma/2$. This function satisfies $\partial_t P = D \vec{\nabla}^2 P$ (for $t > 0$), the diffusion equation.

Example 2: Harmonic trap. Now suppose $U(x) = \frac{1}{2}\kappa x^2$. For example this could come from some laser trap, trying to keep our particle at the origin. This is the mysterious thing called ‘Ornstein-Uhlenbeck process’ I mentioned above, which gives a hint for what the outcome will be. The Langevin equation becomes

$$-\mu \vec{\nabla} U = -\mu \kappa \vec{x} \quad (2.30)$$

and can be solved by integrating factor:

$$\partial_t (e^{\mu \kappa t} \vec{x}(t)) = e^{\mu \kappa t} \vec{\eta}(t). \quad (2.31)$$

Integrating the BHS gives

$$e^{\mu \kappa t} \vec{x}(t) - \vec{x}(0) = \int_0^t ds e^{\mu \kappa s} \vec{\eta}(s) \quad (2.32)$$

and solving for x gives

$$\vec{x}(t) = \vec{x}(0) e^{-\mu \kappa t} + \int_0^t ds e^{-\mu \kappa (t-s)} \vec{\eta}(s). \quad (2.33)$$

(This is the quantity we called $x_\eta(t)$ earlier.) The expectation value $\langle \vec{x}(t) \rangle = \vec{x}(0) e^{-\mu \kappa t}$ decays to the origin with relaxation time $\tau \equiv \frac{1}{\mu \kappa}$ and the variance is

$$\langle (x(t) - \langle x(t) \rangle)^2 \rangle = \int_0^t ds_1 \int_0^t ds_2 e^{-\mu \kappa (2t-s_1-s_2)} \underbrace{\langle \eta(s_1) \cdot \eta(s_2) \rangle}_{\text{Gaussian, white } 2D\delta(s_1-s_2)} \quad (2.34)$$

$$= 2dD \int_0^t ds e^{-2\mu \kappa (t-s)} = \frac{dD}{\mu \kappa} (1 - e^{-2\mu \kappa t}) \quad (2.35)$$

$$\xrightarrow[t \rightarrow \infty]{} \frac{dD}{\mu \kappa}. \quad (2.36)$$

But we already have a strong expectation for what the late-time behavior should be: the dust particle should be in thermal equilibrium with the liquid at the liquid's temperature T , so that the distribution for its position is

$$P(\vec{x}, t) \xrightarrow{t \rightarrow \infty} P_{\text{eq}}(\vec{x}) = e^{-\frac{\kappa \vec{x}^2}{2T}} \left(\frac{\kappa}{2\pi T} \right)^{d/2}, \quad (2.37)$$

which says indeed $\langle x \rangle \xrightarrow{t \rightarrow \infty} 0$ with variance $\langle x^2 \rangle \xrightarrow{t \rightarrow \infty} \frac{dT}{\kappa}$. So we should have

$$\boxed{D = T\mu}, \quad (2.38)$$

a relation due to Einstein. This is a Fluctuation-Dissipation Relation (FDT): The LHS parametrizes the fluctuations of the noise force resulting from collisions of fluid particles with the dust grain; the RHS is proportional to μ , the dissipation resulting from the dust grain's motion through the medium.

By the way, we didn't mention such a relation in the $U = 0$ example because there isn't an interesting equilibrium in that case – the particle just wanders off. Question: we should be able to get back $U = 0$ from the above by taking $\kappa \rightarrow 0$. How? Notice that the limits $\kappa \rightarrow 0$ and $t \rightarrow \infty$ very much do not commute. Above we took $t \rightarrow \infty$ first. If we instead take $\kappa \rightarrow 0$ first in (2.35), we get

$$\frac{dD}{\mu\kappa} (1 - (1 - 2\mu\kappa t + \dots)) \rightarrow 2dDt \quad (2.39)$$

the correct answer for diffusion.

So far this is a necessary condition for consistency. The FDT (in this form) is a signature of equilibrium. Here is an interesting question: Under the overdamped Langevin dynamics (2.7), does the distribution $P(\vec{x}, t)$ approach the corresponding $P_{\text{eq}} \propto e^{-\beta U(\vec{x})}$ for any $U(\vec{x})$? The answer is yes. To see this, let's build some machinery and introduce some useful notation: We can write the probability distribution as

$$P(\vec{x}, t) = \langle \vec{x} | \hat{T}(t) | 0 \rangle \quad (2.40)$$

where the *transition probability* is

$$\langle \vec{x} | \hat{T}(\epsilon) | \vec{x}' \rangle \equiv \text{Prob of being at } \vec{x} \text{ at time } t + \epsilon \text{ given that it was at } \vec{x}' \text{ at time } t. \quad (2.41)$$

In the problem at hand, this quantity is independent of t . The crucial feature is the Markovian nature of the evolution: we don't need to know the entire past history of the particle to predict its later distribution. And to get to x at time $t + \epsilon$ from 0 at $t = 0$ it had to be *somewhere* at the intermediate time t' (let's call that somewhere $\vec{x}'$), so

$$P(\vec{x}, t + \epsilon) = \langle x | \hat{T}(\epsilon) \hat{T}(t) | 0 \rangle = \langle x | \hat{T}(\epsilon) \underbrace{\mathbb{1}}_{=\int d^d x' |\vec{x}'\rangle\langle\vec{x}'|} \hat{T}(t) | 0 \rangle = \int d^d x' \langle \vec{x} | \hat{T}(\epsilon) | \vec{x}' \rangle P(\vec{x}', t). \quad (2.42)$$

For very small ϵ , we know the relation between $\vec{x}$ and $\vec{x}'$ from the Langevin equation:

$$\vec{x} = \vec{x}' - \underbrace{\mu \vec{\nabla}_x U|_{x'} \epsilon + \int_t^{t+\epsilon} ds \vec{\eta}(s)}_{\equiv \vec{\eta}_\epsilon}. \quad (2.43)$$

So we can think of the transition probability as

$$\langle \vec{x} | \hat{T}(\epsilon) | \vec{x}' \rangle = \text{Prob of finding the correct } \eta_\epsilon \quad (2.44)$$

$$\equiv P(\eta_\epsilon) \quad (2.45)$$

$$= \exp \left(-\frac{(x - x' + \epsilon \mu \partial_x U|_{x'})^2}{4D\epsilon} \right) \left(\frac{1}{4\pi D\epsilon} \right)^{d/2} \quad (2.46)$$

$$\stackrel{\epsilon \text{ small}}{=} \exp \left(-\frac{\epsilon (\dot{x} - v(x))^2}{4D} \right) \left(\frac{1}{4\pi D\epsilon} \right)^{d/2}, \quad \vec{v}(x) \equiv -\mu \vec{\nabla}_x U(\vec{x}) \quad (2.47)$$

At the penultimate step we used the fact that we *know* (or at least have assumed) the distribution of the noise $\eta(t)$ in (2.19). We also used the fact that $v(x') = v(x) + \mathcal{O}(\epsilon)$.

Doing this repeatedly gives

$$\begin{aligned} P(\vec{x}, t) &= \langle \vec{x} | \left(\hat{T}(ds) \right)^{t/ds} | \vec{0} \rangle = \langle \vec{x} | \hat{T}(ds) \underbrace{\mathbb{1}}_{= \int d^d x_1 |x_1\rangle\langle x_1|} \hat{T}(ds) \cdots \underbrace{\mathbb{1}}_{= \int d^d x_N |x_N\rangle\langle x_N|} \hat{T}(ds) | \vec{0} \rangle \\ &= \int \prod \frac{dx(s)}{\mathcal{N}} \exp \left(-ds \sum_n \frac{(\dot{x}_n - v(x_n))^2}{4D} \right) \end{aligned} \quad (2.48)$$

$$\equiv \int_{\vec{x}(0)=0, \vec{x}(t)=\vec{x}} Dx \exp \left(- \int_0^t ds \frac{(\dot{x} - v(x))^2}{4D} \right) \quad (2.49)$$

a path integral expression for the probability distribution, weighted by a measure of its distance to the deviation from the ‘classical’ (*i.e.* noise-free) trajectory with velocity $v(x)$.

One way to think about this path integral is that it is a solution of the Fokker-Planck equation:

$$\partial_t P + \vec{\nabla} \cdot \vec{J} = 0, \quad \vec{J} \equiv \vec{v}P - D\vec{\nabla}P \quad (2.50)$$

(for D constant). Notice that it takes the form of a continuity equation. The substance that is conserved is probability.

The path integral as solution to FP equation. Let’s remove the last step

of time evolution:

$$P(\vec{x}, t) = \int d^d x' \left(\frac{1}{4\pi D\epsilon} \right)^{d/2} e^{-\frac{(x-x'-\epsilon v(x'))^2}{4D\epsilon}} P(\vec{x}', t - \epsilon) \quad (2.51)$$

$$\stackrel{\text{def of } y}{=} \left(1 - \epsilon \vec{\nabla} \cdot v(x) \right) \int d^d y \left(\frac{1}{4\pi D\epsilon} \right)^{d/2} e^{-\frac{y^2}{4D\epsilon}} P(x + y - \epsilon v(x), t - \epsilon) \\ \stackrel{\text{Taylor}}{=} \left(1 - \epsilon \vec{\nabla} \cdot v(x) \right) \int d^d y \left(\frac{1}{4\pi D\epsilon} \right)^{d/2} e^{-\frac{y^2}{4D\epsilon}} \left(P(x, t) + (y - \epsilon v(x)) \cdot \vec{\nabla} P \right. \\ \left. + \frac{1}{2} (y_i y_j - 2\epsilon y_i v_j + \epsilon^2 v_i v_j) \vec{\nabla}_i \vec{\nabla}_j P - \epsilon \partial_t P + \mathcal{O}(\epsilon^2) \right) \quad (2.52)$$

$$= \left(1 - \epsilon \vec{\nabla} \cdot v(x) \right) \left(P - \epsilon \vec{v} \cdot \vec{\nabla} P + \epsilon D \vec{\nabla}^2 P - \epsilon \partial_t P + \mathcal{O}(\epsilon^2) \right) \quad (2.53)$$

where $y \equiv x' - x + \epsilon v(x')$ so that $d^d y = d^d x' \left(1 + \epsilon \vec{\nabla} v(x') \right) = d^d x' \left(1 + \epsilon \vec{\nabla} v(x) \right) + \mathcal{O}(\epsilon^2)$. The integration table for the last step is

$$\int d^d y \left(\frac{1}{4\pi D\epsilon} \right)^{d/2} e^{-\frac{y^2}{4D\epsilon}} (1, y_i, y_i y_j) = (1, 0, 2D\epsilon \delta_{ij}). \quad (2.54)$$

Now collect the $\mathcal{O}(\epsilon)$ terms and you'll get the Fokker-Planck equation, $\partial_t P + \vec{\nabla} \cdot \vec{J} = 0$.

Maybe we'll give another derivation of the FP equation below.

[End of Lecture 6]

So let's think about the stationary solution of this equation: $0 = \partial_t P$ requires $\vec{\nabla} \cdot \vec{J} = 0$ but is solved if $\vec{J} = 0$.

$$P(\vec{x}, t) = P_{\text{eq}}(\vec{x}) \propto e^{-U(x)/T} \quad (2.55)$$

has $\vec{\nabla} P_{\text{eq}} = \frac{1}{\mu T} \vec{v} P_{\text{eq}}$ and hence is stationary iff the FDT is true $D = T\mu$, for any $U(x)$. I claim that the evolution of $P(x, t)$ by the FP equation leads to a relaxation to $P_{\text{eq}}(x)$ if you wait long enough. (Actually, we haven't shown that an arbitrary initial P will decay to this one. In principle, based on what we know at the moment, there could be other stationary solutions, or there could be limit cycles.)

Diffusion equation. For $U = 0$ the FP equation reduces to the diffusion equation. Let's spend a little time with the diffusion equation in $d = 1$ and solve it in several different ways using different principles of physics.

$$\partial_t P = D \partial_x^2 P, \text{ with initial condition } P(x, 0) = \delta(x). \quad (2.56)$$

Recall that $P(x, t)dx \equiv \text{Prob}(\text{particle in } (x, x + dx) \text{ at time } t)$.

First let's apply dimensional analysis. There is no bias, so $\langle x \rangle = \int x P(x, t) dx = 0$. $\langle x^2 \rangle = \int dx x^2 P(x, t)$ depends on D, t . These quantities have units

$$[D] = L^2/T, [t] = T, [\langle x^2 \rangle] = L^2. \quad (2.57)$$

Therefore

$$\langle x^2 \rangle = Dt \cdot \text{const.} \quad (2.58)$$

In fact that constant is 2, as you can see by differentiating with respect to time:

$$\partial_t \langle x^2 \rangle = \int_{-\infty}^{\infty} dx (x^2 \partial_t P = x^2 D \partial_x^2 P) \stackrel{\text{IBP}}{=} 2D. \quad (2.59)$$

Those are the first two moments, what about the distribution itself? In $d = 1$, $[P] = L^{-1}$, which means that $\sqrt{Dt} P(x, t)$ is dimensionless, and therefore can only depend on the dimensionless combination $\xi \equiv \frac{x}{\sqrt{Dt}}$, which we can call a 'scaling variable':

$$P(x, t) = \frac{1}{\sqrt{Dt}} P(\xi) \quad (2.60)$$

This reduces the PDE (2.56) to an ODE, $2P'' + \xi P' + P = 0$ whose solution is $P(\xi) = e^{-\xi^2/4}/\sqrt{4\pi}$.

By the same logic as above, the Markov nature of the random walk implies that

$$P(x, 2t) = \int_{-\infty}^{\infty} dy P(y, t) P(x - y, t). \quad (2.61)$$

This is called the Chapman-Kolmogorov equation. It says that the probability later is obtained by the convolution of two intermediate probabilities. Convolution is a bit difficult, but Fourier transform turns it into ordinary multiplication. So let $P(k, t) \equiv \int_{-\infty}^{\infty} dx e^{ikx} P(x, t)$. Then (2.61) is equivalent to $P(k, 2t) = P(k, t)^2$. Since $[k] = 1/L$, we again have a (dimensionless) scaling variable $\kappa \equiv k\sqrt{Dt}$, and $P(k, t) = P(\kappa)$. So (2.61) tells us the functional equation $P(\sqrt{2}\kappa) = P(\kappa)^2$. In terms of $z \equiv \kappa^2$ and $Q(z) \equiv \log P(\kappa)$, it implies the self-similarity relation $Q(2z) = 2Q(z)$, which is solved by $Q = cz$ for some constant c . Thus

$$P(k, t) \propto e^{-k^2 Dt} \stackrel{\text{small } k}{\simeq} 1 - \frac{1}{2} k^2 \langle x^2 \rangle + \dots \quad (2.62)$$

where the last relation implies $c = -2$. Fourier transforming back gives back our answer for $P(x, t)$. This strategy was a bit like coarse-graining in time, like we did for the random walk.

More generally in any dimension, in the absence of a potential, the diffusion equation is linear and translation invariant and so solvable by Fourier transform:

$$\partial_t P(x, t) = D \vec{\nabla}^2 P(x, t) \implies \partial_t P_k(t) = -D k^2 P_k \implies P_k(t) = P_k(0) e^{-D k^2 t}. \quad (2.63)$$

So the time evolution is

$$P(x, t) = \int d^d k e^{i k x} e^{-k^2 D t} P_k(0). \quad (2.64)$$

If the initial condition is $P(x, 0) = \delta^d(x)$, $P_k(0) = 1$ we have

$$P(x, t) = \int d^d k e^{i k x} e^{-k^2 D t} = e^{\frac{-x^2}{4 D t}} \frac{1}{(4 \pi D t)^{d/2}} \quad (2.65)$$

by completing the square.

Suppose we want to know: how long does it take for $P(x, t)$ to reach the equilibrium distribution under its evolution by the FP equation? This last strategy is the one that generalizes best. What did we do above? We wrote the equation in the form

$$\partial_t P = -\hat{L} P \quad (2.66)$$

where $\hat{L}$ is a linear (differential) operator. The general FP equation

$$\partial_t P + \vec{\nabla} \cdot \vec{J} = 0, \quad \vec{J} \equiv \vec{v} P - D \vec{\nabla} P \quad (2.67)$$

is also of this form. (Note that the diffusion coefficient D could also depend on x and t . Based on the interpretation in terms of conservation of probability, which is still true in that case, it is clear that this way of writing it is still the correct one.) To think about this equation, the following analogy is extremely useful:

stochastic evolution	quantum mechanics
Langevin equation	Heisenberg equations of motion
Fokker-Planck equation, (2.66)	Schrödinger equation, $\partial_t \psi = -i \hat{H} \psi$
path integral	path integral
randomness from the noise η	quantum fluctuations
noise-free EoM	classical EoM

The last column is the answer to the question: what trajectory gives the saddle point of the path integral?¹³

From this point of view, the thing to do to solve (2.66) is to find the eigenstates of $\hat{L}$:

$$\hat{L} |n\rangle = E_n |n\rangle \quad (2.68)$$

¹³Actually the most direct analog of the Langevin equation in quantum mechanics is a less familiar formulation of QM called stochastic quantization, which introduces an explicit noise variable to provide the quantum uncertainty. It is designed precisely using this analogy and differs from Langevin by some factors of i .

and expand the initial state in this basis. The only wrinkle is that as written we have no guarantee that $\hat{L}$ is hermitian so that its right and left eigenvectors are the same. One way to proceed is simply to distinguish right and left eigenvectors, as we did in (1.6). We still have a resolution of the identity $\mathbb{1} = \sum_n |n\rangle\langle n|$ or in terms of $\phi_n(x) = \langle x|n\rangle$ (right evec), $\tilde{\phi}_n(x) = \langle n|x\rangle$ (left evec), $\delta(x-y) = \sum_n \phi_n(x)\tilde{\phi}_n(y)$ in terms of orthogonal states $\langle n|n'\rangle = \int dx \phi_n(x)\tilde{\phi}_{n'}(x)$. We used another resolution of the identity here

$$\mathbb{1} = \int dx |x\rangle\langle x| \quad (2.69)$$

so we can define $|P\rangle = \mathbb{1}|P\rangle = \int dx |x\rangle P(x)$, with $P(x) = \langle x|P\rangle$. So

$$\partial_t |P\rangle = -\hat{L} \underbrace{1}_{=\sum_n |n\rangle\langle n|} |P\rangle = -\sum_n E_n |n\rangle P_n(t) \quad (2.70)$$

where $P(x, t) \equiv \langle x|P(t)\rangle$ and $P_n(t) \equiv \langle n|P(t)\rangle$. The LHS is $\partial_t |P\rangle = \sum_n |n\rangle \partial_t P_n(t)$. Acting with $\langle m|$ on the left gives

$$\partial_t P_n(t) = E_n P_n(t) \quad (2.71)$$

whose solution we know well: $P_n(t) = P_n(0)e^{-E_n t}$. So the full solution is just (act with $\langle x|$ from the left on $|P(t)\rangle = \sum_n |n\rangle P_n(t)$)

$$P(x, t) = \sum_n P_n(t)\phi_n(x) = \sum_n P_n(0)e^{-E_n t}\phi_n(x). \quad (2.72)$$

Notice that (2.65) is a special case of this equation where $\hat{L} = -D\partial_x^2$, so the eigenvectors are Fourier modes e^{ikx} .

The problem of finding the eigenvectors, for example for a 1d particle diffusing in a potential $U(x)$ has the same level of difficulty as the analogous Schrödinger problem. And our intuition from such problems is very useful. For example, you can ask: does any initial state evolve to the Boltzmann distribution? We showed above that $\phi_0(x) \equiv P_{\text{eq}}$ is an eigenvector $\hat{L}$ with eigenvalue zero. If there is a gap in the spectrum of $\hat{L}$, the late-time evolution is

$$P(x, t) \xrightarrow{t \rightarrow \infty} P_{\text{eq}}(x)P_0(0) + \phi_1(x)P_1(0)e^{-E_1 t} + \dots \quad (2.73)$$

where $P_n(0) \equiv \int dy P(y, t=0)\tilde{\phi}_n(y)$ is the initial condition. So at a time $t \sim E_1$ we can ignore the deviation from P_{eq} . When is there such a gap? We saw an example above, where $U(x) = 0$ where there is not a gap, and instead $E_k = Dk^2$. This only happens when the particle is able to move on a compact space, *i.e.* the potential falls off at ∞ so that there are scattering states. But, in any such case, as in the case $U = 0$, this means that the probability can leak off to infinity and there isn't really a nonzero

equilibrium distribution. So my answer is: any time there *is* an equilibrium distribution (roughly, $U(x)$ grows at infinity), there is a gap above a unique groundstate of $\hat{L}$. The fact that the groundstate is unique is the familiar fact that even if the potential has multiple minima, as long as there is a finite barrier, tunneling will produce a unique groundstate. We'll come back later to discuss metastable states and their lifetime using the technology we've just developed.

The difficulties of right and left eigenvectors actually can be avoided in any case where there is an equilibrium distribution by studying the evolution of $P(x, t)/\sqrt{P_{\text{eq}}(x)}$, which evolves by the operator $\hat{H} = \sqrt{P_{\text{eq}}}^{-1} \hat{L} \sqrt{P_{\text{eq}}}$ which is Hermitian.

Note that the operator $\hat{L}^\dagger$ whose right eigenvectors are the left eigenvectors of $\hat{L}$ has a physical role to play: it appears in the *backwards* FP equation obtained by varying the initial time:

$$\partial_{t'} P(x, t|y, t') = \hat{L}^\dagger P(x, t|y, t') . \quad (2.74)$$

2.2 Whence the noise and whence the friction?

[Cugliandolo 2504.11994, Zwanzig]

The pollen grain in water is not an isolated system; it is interacting with a bath. Let's study a solvable microscopic classical model of the bath to see where the friction and the dissipation come from (they come from the same thing so it will not be surprising that they are related). We'll model the bath as a collection of simple harmonic oscillators (sometimes this is called Caldeira-Leggett model).

So the system will be a single particle with phase space coordinates x, p , and Hamiltonian $H_{\text{system}} = \frac{p^2}{2m} + U(x)$ (where $U(x)$ is the same external potential we've been discussing) and the bath will be a collection of particles with phase space coordinates q_a, π_a . We'll take the full Hamiltonian to be $H = H_{\text{system}} + H_{\text{env}}$ with

$$H_{\text{env}} = \sum_{a=1}^N \left(\frac{\pi_a^2}{2m_a} + \frac{m_a \omega_a^2}{2} \left(q_a + \frac{c_a}{m_a \omega_a^2} x \right)^2 \right) . \quad (2.75)$$

Because of the $c_a q_a x$ term, the position of the system particle x shifts the equilibrium position of the bath mode q_a by an amount proportional to the coupling c_a . The additive normalization (completing the square) is chosen for convenience. Effectively we are subtracting off the renormalization by the bath of the mass (the coefficient of x^2) of the system mode, because that's not what we're interested in here.

So the equations of motion are

$$\begin{aligned}\dot{x}(t) &= \frac{\partial H}{\partial p(t)} = p(t)/m, & \dot{p}(t) &= -\frac{\partial H}{\partial x(t)} = -U'(x(t)) - \sum_a c_a q_a(t) - \sum_a \frac{c_a^2}{m_a \omega_a^2} x(t), \\ \dot{q}_a &= \pi_a(t)/m_a, & \dot{\pi}_a(t) &= -m_a \omega_a^2 q_a(t) - c_a x(t).\end{aligned}\quad (2.76)$$

Now we will do the step called “integrating out the bath”. In many contexts this actually involves doing integrals, but here it just means solve the equations of motion for the q_a and plug the solutions back in to the equations we care about for x and p :

$$q_a(t) = q_a(0) \cos(\omega_a t) + \frac{\pi_a(0)}{m_a \omega_a} \sin(\omega_a t) - \frac{c_a}{m_a \omega_a} \int_0^t dt' \sin(\omega_a(t-t')) x(t'). \quad (2.77)$$

Now let’s plug (2.77) into the EOM for $\dot{p}$ and do an IBP in the source term to get

$$\dot{p}(t) = \underbrace{-U'(x(t))}_{\text{deterministic}} + \underbrace{\xi(t) - \int_0^t dt' \Gamma(t-t') \dot{x}(t')}_{\text{deterministic result of coupling to bath}} \quad (2.78)$$

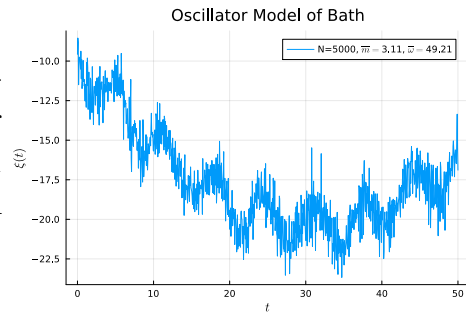
with

$$\xi(t) = - \sum_{a=1}^N c_a \left(\frac{\pi_a(0)}{m_a \omega_a} \sin(\omega_a t) + \left(q_a(0) + \frac{c_a x(0)}{m_a \omega_a^2} \right) \cos(\omega_a t) \right) \quad (2.79)$$

$$\Gamma(t-t') = \sum_{a=1}^N \frac{c_a^2}{m_a \omega_a^2} \cos(\omega_a(t-t')) = \Gamma(t'-t), \quad (2.80)$$

a stationary and symmetric kernel describing a retarded frictional force. Notice that this force depends on the whole history of $x(t')$ for $t' \in [0, t]$ – it is not Markovian. This equation is still completely deterministic – given initial conditions $\{q_a(0), \pi_a(0), x(0), p(0)\}$, the whole trajectory is determined.

But consider the function ξ in (2.79). For incommensurate $\{\omega_a\}$, $\xi(t)$ is a pseudorandom number generator, with *seeds* $\{q_a(0), \pi_a(0)\}$. As $N \gg 1$ grows, it becomes harder and harder to discern structure in plots of this function.



Now, *if* the bath is in thermal equilibrium at temperature T at $t = 0$ (because somehow it interacted before this point) then

$$P(\{q_a(0), \pi_a(0)\}, x(0)) = \frac{1}{Z(x(0))} e^{-\beta H_{\text{env}}(\{q_a(0), \pi_a(0)\}, x(0))} \quad (2.81)$$

where $H_{\text{env}}(\{q_a(0), \pi_a(0)\}, x(0)) = \sum_{a=1}^N \left(\frac{\pi_a^2(0)}{2m_a} + \frac{m_a \omega_a^2}{2} \left(q_a(0) + \frac{c_a x(0)}{m_a \omega_a^2} \right)^2 \right)$. This is a big step, this is the statistical mechanics step.

Claim: now ξ is a Gaussian random variable with $\langle \xi(t) \rangle = 0$, $\langle \xi(t) \xi(t') \rangle = T \Gamma(t-t')$, the same kernel as above, *i.e.*

$$P[\xi] = Z^{-1} e^{-\frac{1}{2T} \int dt \int dt' \xi(t) \Gamma^{-1}(t-t') \xi(t')} \quad (2.82)$$

where Γ^{-1} is defined by $\int dt'' \Gamma^{-1}(t-t'') \Gamma(t''-t') = \delta(t-t')$. It's Gaussian because of our choice to model the bath with SHOs. (To check the claim, use the equipartition theorem for Gaussian variables several times.) The fact that the bath is in equilibrium implies that the statistics of the noise ξ and the friction kernel are the same Γ . This is the origin of the fluctuation-dissipation theorem.

Various choices of bath arise according to the couplings $\{\omega_a, m_a, c_a\}$. To parameterize these choices, set $c_a \equiv \tilde{c}_a / \sqrt{N}$ (to have a nicer $N \rightarrow \infty$ limit), and let

$$S(\omega) \equiv \frac{1}{N} \sum_{a=1}^N \frac{\tilde{c}_a^2}{m_a \omega_a} \delta(\omega - \omega_a) \quad (2.83)$$

be the ‘spectral density’ so that

$$\Gamma(t-t') = \int_0^\infty \frac{d\omega}{\omega} S(\omega) \cos(\omega(t-t')) \quad (2.84)$$

So for example, if

$$S(\omega)/\omega = \frac{2\mu}{\pi} \left(\frac{|\omega|}{\omega_0} \right)^{\alpha-1} e^{-\frac{|\omega|}{\Lambda}} \quad (2.85)$$

with $\alpha = 1$ (this is called an ‘Ohmic bath’) then

$$\Gamma(t) = \frac{2\mu}{\pi} \frac{\Lambda}{1 + (\Lambda t)^2} \xrightarrow{\Lambda \rightarrow \infty} 2\mu \delta(t) \quad (2.86)$$

becomes white noise. So you see white noise requires a bath with a spectrum of oscillators that goes on forever.

What do you get for $\Gamma(t)$ if

$$S(\omega)/\omega = \frac{2\mu}{\pi} \left(\frac{|\omega|}{\omega_0} \right)^{\alpha-1} e^{-\left(\frac{\omega}{\Lambda}\right)^2} \quad (2.87)$$

with $\alpha = 1$? I get $\Gamma(t) = \frac{2\mu}{\sqrt{\pi}} e^{-t^2 \Lambda^2/4} \xrightarrow{\Lambda \rightarrow \infty} 2\mu \delta(t)$ – the limit where we remove the cutoff Λ is the same.

Origin of irreversibility. Famously, Newton’s equations are reversible, in the sense that $\{Q(-t), -P(-t)\}$ at time $-t$ satisfy the same equations as $\{Q(t), P(t)\}$ at

time t . This is true of $\{Q(t), P(t)\} = \{q_a, \pi_a, x, p\}$. It is not true of the Langevin equation. For example in the Markov case, the $\gamma_0 v(t)$ is not invariant under the time-reversal transformation $t \rightarrow -t, x(t')|_{t'=t} \rightarrow x(t')|_{t'=-t}, v(t')|_{t'=t} \rightarrow -v(t')|_{t'=-t}$. In the model of the bath as oscillators, this is because we do not remember $q_a(0), \pi_a(0)$. Try to infer the initial conditions from a plot of (2.79). In fact (at least if our bath were a chaotic system), it would be enough to not remember the 419th digit of $q_1(0)$. The ideal fact of reversibility is completely crushed by the computational complexity of the time evolution, and we should just forget about it. We put initial conditions and time goes forward. [End of Lecture 7]

Dissipation. With $H_{\text{system}} = \frac{mv^2}{2} + V(x)$, the change in the system energy under the Langevin evolution $m\dot{v} - F = -\mu v + \xi$ is

$$\dot{E} \equiv \frac{dH_{\text{system}}}{dt} = mv\dot{v} + V'v = (m\dot{v} - F)v = v(-\mu v + \xi) \neq 0. \quad (2.88)$$

If somehow $\xi = 0$ (for example for $T \rightarrow 0$ with μ fixed), this becomes

$$\dot{E} \stackrel{\xi=0}{=} -\mu v^2 \leq 0, \quad (2.89)$$

gradient descent. Energy would always just flow into the bath.

2.3 First-passage/exit problems and BCs for FP

We showed above that the (overdamped) Langevin equation implies that $P(x, t) \equiv P(x, t|0, 0)$ satisfies the Fokker-Planck (FP) equation:

$$\partial_t P = -\hat{L}P = -\vec{\nabla} \cdot (\vec{v}P - D\vec{\nabla}P) \quad (2.90)$$

where $\vec{v} = \mu \vec{F}_{\text{ext}}$ is the terminal velocity. This equation expresses the indestructibility of probability.

Suppose I have a particle undergoing Brownian motion in a domain with forbidden regions. For example, let's say we are in 1d and $x < a$ is inaccessible to our 1d walker for some reason. What boundary condition should we put on the FP equation for the probability distribution $P(x, t)$?

I find the answer a bit counterintuitive. You might have thought that if the region is forbidden, there should be zero probability for the walker to reach there, so we should just demand that the probability vanishes there. But then the probability flux will not necessarily vanish, and the evolution with the FP equation will lead to probability to be lost into the resulting probability sink (or worse, source). A forbidden region is described by the boundary condition that there is no probability *flux* into that region.

There are questions for which a probability sink is what we want. Suppose that sitting at $x = a$ is a monster that eats our walkers. Then, indeed, the probability $P(x, t)dx$ of finding our walker somewhere with $x > a$ at time t *should* decrease with t , since there is a finite and growing probability that the walker was eaten. This problem is described by the Dirichlet condition at $x = a$. If we want to know instead what is the probability of finding our walker somewhere with $x < a$, *conditioned* on him not having been eaten before that is

$$P(x, t|\text{alive}) = \frac{P(x, t)}{P_{\text{survival}}(t)} \quad (2.91)$$

where $P_{\text{survival}}(t)$ is the probability that the walker has not been eaten by time t .

Here is a less carnivorous example of an interesting question that is answered by this situation. Suppose we would like to know the probability distribution for the time when the walker first reaches $x = a$. Such a question is called a *first-passage problem*. To warm up, let's think about a 1d problem with monsters at the ends of an interval $x \in [a, b]$, so a walker that hits a or b does not return to the interval. Let

$$G(x, t) \equiv \int_a^b dy P(y, t|x, 0) = \text{Prob}(x(t) \in [a, b], \text{ given that } x(0) = x). \quad (2.92)$$

This is the probability that a walker starting at $x \in [a, b]$ has *not* hit either a or b by time t . It is strictly decreasing in time (since walkers are never created), and the events where it decreases are the deaths of the walkers:

$$-\partial_t G(x, t)dt = \text{Prob}(\text{to be eaten in } (t, t + dt) \text{ when starting at } x \text{ at } t = 0). \quad (2.93)$$

So the expected value of a function of the exit time (given that the walker starts at x) is

$$\langle f(t) \rangle_x = \int_0^\infty dt f(t) (-\partial_t G(x, t)). \quad (2.94)$$

For example, the mean exit time is

$$\langle t \rangle_x = \int_0^\infty dt t (-\partial_t G(x, t)) \stackrel{\text{IBP}}{=} \int_0^\infty dt G(x, t). \quad (2.95)$$

If the evolution is time-translation invariant, we can rewrite G by shifting the time arguments, as

$$G(x, t) = \int_a^b dy P(y, 0|x, -t), \quad (2.96)$$

which satisfies the *backwards* FP equation

$$\partial_t G(x, t) = A(x)\partial_x G(x, t) + \frac{1}{2}B(x)\partial_x^2 G(x, t) = \hat{L}^\dagger G(x, t). \quad (2.97)$$

So (2.97) is directly an equation for the probability distribution of exit times.

The backward FP equation

$$\partial_{t_0} P(x, t | x_0, t_0) = A(x_0) \partial_{x_0} P(x, t | x_0, t_0) + \frac{1}{2} B(x_0) \partial_{x_0}^2 P(x, t | x_0, t_0) \quad (2.98)$$

is what happens if we evolve the initial time. It differs from the forwards FP equation in that

- The differential operator acts on x_0 .
- The sign of the drift is different.
- The derivatives don't act on A and B .

Where did the backward FP equation come from? Start from the (obvious) Chapman-Kolmogorov equation

$$P(x, t | x_0, t_0) = \int dy P(x, t | y, t_0 + \Delta t_0) P(y, t_0 + \Delta t_0 | x_0, t_0) \quad (2.99)$$

that says that to get from here to there we have to be somewhere in between. The idea behind the following derivation is just that in a small time Δt_0 the probability distribution doesn't change much.

Trivially rewrite the second factor on the RHS as

$$P(y, t_0 + \Delta t_0 | x_0, t_0) = \int dz \delta(y - z) P(z, t_0 + \Delta t_0 | x_0, t_0) \quad (2.100)$$

so that we can use the following representation of the delta function, a Taylor expansion about $z = x_0$:

$$\delta(y - z) = \sum_{n=0}^{\infty} \frac{(z - x_0)^n}{n!} \partial_{x_0}^n \delta(x_0 - y). \quad (2.101)$$

You can check this by integrating the BHS of this equation against any nice-enough test function. Using this in (2.100) gives

$$P(y, t_0 + \Delta t_0 | x_0, t_0) = \delta(x_0 - y) \cdot 1 + \sum_{n=1}^{\infty} \frac{(z - x_0)^n}{n!} P(z, t_0 + \Delta t_0 | x_0, t_0) \partial_{x_0}^n \delta(x_0 - y) \quad (2.102)$$

where the 1 comes from

$$\int dz P(z, t_0 + \Delta t_0 | x_0, t_0) = 1 \quad (2.103)$$

since $P(z, t_0 + \Delta t_0 | x_0, t_0)$ is a probability distribution in z .

Therefore

$$\begin{aligned} P(x, t | x_0, t_0) - P(x, t | x_0, t_0 + \Delta t_0) &= -\Delta t_0 \partial_{t_0} P(x, t | x_0, t_0) \\ &= \sum_{n=1}^{\infty} \frac{1}{n!} \underbrace{\int dz (z - x_0)^n P(z, t_0 + \Delta t_0 | x_0, t_0)}_{\equiv \Delta t_0 \mu_n(x_0, t_0)} \underbrace{\partial_{x_0}^n \int dy \delta(x_0 - y) P(x, t | y, t_0 + \Delta t_0)}_{= P(x, t | x_0, t_0 + \Delta t_0)} \end{aligned} \quad (2.104)$$

which says that

$$\partial_{t_0} P(x, t | x_0, t_0) = - \sum_{n=1}^{\infty} \frac{\mu_n(x_0, t_0)}{n!} (\partial_{x_0})^n P(x, t | x_0, t_0). \quad (2.105)$$

Now, μ_n here is a moment of the noise distribution over a time interval Δt_0 . But the noise satisfies $\langle \text{noise}^2 \rangle \sim \Delta t_0$, which implies that $\lim_{\Delta t \rightarrow 0} \mu_{n \geq 3} = 0$. So we can truncate the sum over n at $n = 2$, and $A \equiv \mu_1, B \equiv \mu_2$.

We got this by operating on the second factor in (2.99). If we had instead operated on the first factor, we could derive the forwards FP equation.

If you are worried about various mathematical infelicities of interchanging sums and integrals in the above discussion, redo the whole thing while integrating against a nice test function (as in LP §2.5). It might make you feel better but nothing will be different.

Integrating (2.97) over time shows that the mean exit time T satisfies the ODE

$$\frac{1}{2} B(x) \partial_x^2 T + A(x) \partial_x T = -1 \quad (2.106)$$

using $\int_0^\infty dt \partial_t G(x, t) = G(x, \infty) - G(x, 0) = 0 - 1$.

We can solve this equation (2.106) via an orgy of the method of integrating factor. First, you see that if we find a function $\psi_1(x)$ that satisfies

$$\frac{\psi_1'}{\psi_1} = \frac{2A}{B} \quad (2.107)$$

then we can rewrite (2.106) as

$$(\psi_1 T')' = -2\psi_1 / B. \quad (2.108)$$

The solution to (2.107) is

$$\psi_1(x) = e^{\int_a^x dx' \frac{2A(x')}{B(x')}} \quad (2.109)$$

which satisfies $\psi_1(a) = 1$. Now we can integrate (2.108) to get

$$T'(x) = \frac{T'(a)}{\psi_1(x)} - \frac{\psi_2(x)}{\psi_1(x)} \quad (2.110)$$

with $\psi_2(x) \equiv 2 \int_a^x dx' \frac{\psi_1(x')}{B(x')}$. Notice that $\psi_2(a) = 0$. Integrating one more time gives the result

$$T(x) = T(a) + T'(a)\psi_3(x) - \psi_4(x), \quad (2.111)$$

in terms of two constants of integration $T(a), T'(a)$ and two known functions

$$\psi_3(x) \equiv \int_a^x \frac{dx'}{\psi_1(x')}, \quad \psi_4(x) \equiv \int_a^x dx' \frac{\psi_2(x')}{\psi_1(x')}. \quad (2.112)$$

If both boundaries are absorbing, as we've assumed, $T(a) = T(b) = 0$ (since if you start sitting on top of the monster, you get eaten right away)^a. $T(a) = T(b) = 0$ says $T'(a) = \psi_4(b)/\psi_3(b)$ and the solution is

$$T(x) = \frac{\psi_3(x)\psi_4(b) - \psi_3(b)\psi_4(x)}{\psi_3(b)}. \quad (2.113)$$

The result for the case of no drift and constant diffusion constant ($A = 0, B = 2D$) is

$$T(x) = \frac{(x-a)(b-x)}{2D}. \quad (2.114)$$

($\psi_1(x) = 1, \psi_2(x) = \frac{x-a}{D}, \psi_3(x) = x-a, \psi_4(x) = \frac{(x-a)^2}{2D}$.) We'll compare this with an answer on the lattice below.

^aIf instead the boundary at a were a reflecting boundary (because $x \leq a$ is forbidden), then $\partial_x G|_{x=a} = 0$ which from our expression for T in terms of G says $T'(a) = 0$ instead.

We can also ask about the probability to be eaten by a particular monster, *i.e.* to exit via a specific door. For example,

$$\begin{aligned} G_a(x, t) &\equiv \text{Prob}(\text{a particle starting at } x \in [a, b] \text{ exits through } a \text{ at some time } t' > t) \\ &= - \int_t^\infty dt' J(a, t'|x, 0) \end{aligned} \quad (2.115)$$

$$= \int_t^\infty dt' \left(-A(a)P(a, t'|x, 0) + \frac{1}{2} \partial_a (B(a)P(a, t'|x, 0)) \right) \quad (2.116)$$

where J is the current appearing in the forward FP equation. I think the more general

expression in d dimensions is clearer:

$$G_{\Sigma}(x, t) \equiv \text{Prob}(\text{a particle exits through } \Sigma \text{ at some time } t' > t) \quad (2.117)$$

$$= - \int_t^{\infty} dt' \int_{\Sigma} dS' \hat{n}' \cdot \vec{J}(x', t'|x, 0) \quad (2.118)$$

where $\hat{n}'$ is the normal to Σ at x' , and the current is

$$J_i(x', t'|x, 0) = v_i(x')P(x', t'|x, 0) - \partial_j (D_{ij}(x')P(x', t'|x, 0)) \quad (2.119)$$

(The symbols $A_i \equiv v$ and $2B_{ij} = D_{ij}$ are interchangeable.) Applying $\hat{L}^{\dagger} = A_i(x)\partial_{x_i} + \frac{1}{2}B_{ij}\partial_{x_i}\partial_{x_j}$ to (2.118) and using the FP equation in the form $\hat{L}^{\dagger}\vec{J}(x', t|x, 0) = \partial_t\vec{J}(x', t|x, 0)$ gives

$$\partial_t G_{\Sigma}(x, t) = \hat{L}^{\dagger} G_{\Sigma}(x, t) = \int_{\Sigma} dS' \hat{n}' \cdot \vec{J}(x', t|x, 0) \quad (2.120)$$

which is the in-retrospect-obvious statement that the rate at which probability exits through the door Σ is the flux of the probability current through Σ .

If $x \notin \Sigma$, the flux should vanish at $t = 0$, and therefore the total probability of exiting through Σ satisfies the PDE

$$\hat{L}^{\dagger} G(x, 0) = 0 \quad (2.121)$$

with boundary condition $G(x, 0) = \begin{cases} 1 & x \in \Sigma \\ 0 & x \in \Sigma' \end{cases}$ where Σ' is an absorbing boundary.

Next we're going to use these ideas to say something about the case of escape from a metastable minimum.

2.4 Overcoming barriers with path integrals

[LP, §2.5, Bialek, *Biophysics, searching for principles*, chapter 4 and appendix A.5.]

Suppose you want to solve the Fokker-Planck equation, say in $d = 1$,

$$\partial_t P(x, t) = -\partial_x (A(x)P(x, t)) + \frac{1}{2}\partial_x^2 (B(x)P(x, t)) \quad (2.122)$$

for $P(x, t)$ given some initial conditions $P(0, t)$ (and appropriate boundary conditions). I've assumed here for simplicity that $A(x), B(x)$ are independent of time. Perhaps your first impulse when presented with this problem is to take advantage of the similarity with the Schrödinger equation for a particle in a 1d potential

$$\mathbf{i}\partial_t \psi(x, t) = -\frac{1}{2}\partial_x^2 \psi(x, t) + V(x)\psi(x, t). \quad (2.123)$$

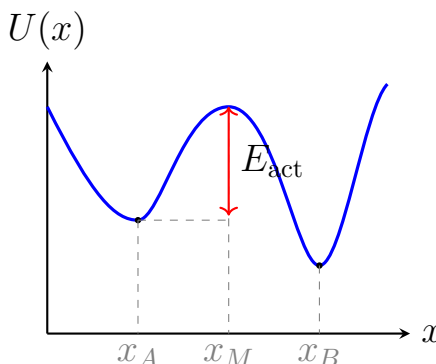
Given an equation of the form (2.122) we can remove the P' term by a change of independent variable. One remaining difference is that pesky factor of $\mathbf{i}$; another is that $\hat{H} = p^2 + V(x)$ is a Hermitian operator, while $\hat{\mathcal{L}}$ in the FP equation $\partial_t p = -\hat{\mathcal{L}}p$ is generally not. The factor of $\mathbf{i}$ is important but easy to put in at the end, and often the hermiticity can be fixed. So the same strategies we use in QM can be used here: namely, if we know the eigenstates of $\hat{H}$, $\hat{H}|\psi_n\rangle = E_n|\psi_n\rangle$, then the solution to (2.123) is

$$\psi(x, t) = \sum_n e^{-\mathbf{i}E_n t} a_n \psi_n(x), \quad a_n = \langle \psi_n | \psi(t=0) \rangle \quad (2.124)$$

(with the wrinkle that $\langle \psi_n |$ is the left eigenvector).

As an important example of this question, we will discuss what is sometimes called ‘the Kramers problem’. Suppose we have a degree of freedom x governed by a potential $U(x)$ that has multiple minima with a barrier in between.

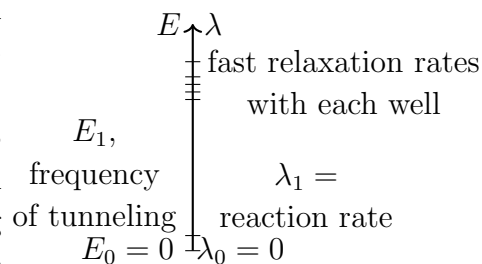
Maybe one of the minima is higher than the other, so the state of a particle sitting in that minimum is only metastable. Metastability means that thermal (or quantum) fluctuations can allow the variable to escape from the local minimum, and the (parametrically slow) rate for that process is what we want to know. For concreteness it’s useful to imagine the variable as a particle rolling around on a hill (with inertia and damping), but the variable could be something more abstract like some reaction coordinate for some chemical reaction between molecules that requires an activation energy E_{act} .



Expectation. This is essentially a 1d Schrödinger problem.

When the potential has the tilted double well form indicated above, we expect that the spectrum of the Schrödinger operator $p^2 + U_{\text{double well}}$ looks like this: $\rightarrow$

and the small nonzero eigenvalue can be obtained by WKB methods, or by doing the path integral by saddle point. In the quantum mechanics case, the latter requires continuing to imaginary time and all kinds of weird stuff. Here, the problem is much more straightforward.



Green’s function method. We can solve this problem using the methods of §2.3. (Afterwards we will solve it again using path integrals.) Recall that we found that the mean escape time satisfied the differential equation

$$-1 = A(x)\partial_x T(x) + D\partial_x^2 T(x) \quad (2.125)$$

where I've specialized to the case where the diffusion constant is indeed constant in time and space. The solution by integrating factor was

$$T(x) = c_1 \int^x \frac{dy}{\psi(y)} - \frac{1}{D} \int^x \frac{dy}{\psi(y)} \int^y dz \psi(z) + c_2 \quad (2.126)$$

where the integrating factor is

$$\psi(x) \equiv e^{\int^x A(y) dy / D} \quad (2.127)$$

and we must determine the constants c_1, c_2 by boundary conditions.

Let's let x_1 be a point far to the left of the metastable minimum at x_A , and let $x_2 \equiv x_B$ be the true minimum. We can set $c_2 = 0$ by setting the lower limits of integration on the y and z integrals at x_1 . So

$$T(x) = c_1 \int_{x_1}^x \frac{dy}{\psi(y)} - \frac{1}{D} \int_{x_1}^x \frac{dy}{\psi(y)} \int_{x_1}^y dz \psi(z) + c_2. \quad (2.128)$$

So let's put a reflecting boundary condition at $x_1 \rightarrow -\infty$. This requires $0 = \partial_x T(x)|_{x=x_1}$ which requires $c_1 = 0$. Then the only absorbing boundary condition is at $x_2 = x_B$, where we'll declare our walkers to have escaped.

Let's see what the answer is. The innocent-looking integrating factor $\psi(x)$ is actually a familiar friend:

$$\psi(x) = e^{-\frac{1}{D\mu} \int_{x_1}^x dx U'(x) dt} \stackrel{\text{FDT}}{=} e^{-U(x)/T}, \quad (2.129)$$

the Boltzmann factor. Thus,

$$T(x) = \frac{1}{D} \int_x^{x_B} dy e^{U(y)/T} \int_{x_1}^y dz e^{-U(z)/T} \stackrel{\text{saddle}}{\approx} \frac{2\pi T/D}{\sqrt{|U_M''| |U_A''|}} e^{(U(x_M) - U(x_A))/T} \quad (2.130)$$

where x_M is the maximum of the barrier.

With path integrals.

$P[x] \equiv \text{Prob}(\text{trajectory is } x(t)) = \text{Prob}(\text{force } \eta/\mu \text{ needed to generate the trajectory } x(t) = x_\eta(t))$

But we know the distribution for the noise force η/μ :

$$P[\eta] \propto e^{-\frac{1}{4\mu T} \int dt \eta^2(t)}. \quad (2.131)$$

Therefore

$$P[x] \propto e^{-\frac{1}{4\mu T} \int dt (\mu \dot{x} + U'(x))^2} \mathcal{J} \equiv \mathcal{J} e^{-S[x]/T}. \quad (2.132)$$

$\mathcal{J}$ here is the Jacobian of the transformation between x and η . Since it is independent of T we will ignore it. Notice that this is the same path integral measure we found above in (2.49). The *action* $S[x]$ is

$$S[x] = \int_{t_i}^{t_f} dt \frac{(\mu \dot{x} + U'(x))^2}{4\mu} \quad (2.133)$$

$$= \int_{t_i}^{t_f} dt \left(\frac{\mu}{4} \dot{x}^2 + \frac{(U'(x))^2}{4\mu} \right) + \frac{1}{2} (U(x_f) - U(x_i)). \quad (2.134)$$

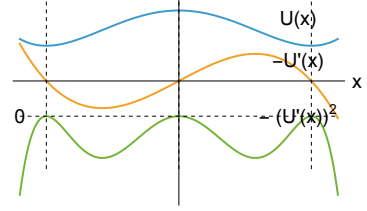
Comparing this with the action in mechanics,

$$S_{\text{cm}}[x] = \int dt \left(\frac{1}{2} m \dot{x}^2 - V(x(t)) \right), \quad (2.135)$$

we can identify the bulk part of S and S_{cm} if we set $m = \mu/2$ and

$$V(x) = -\frac{1}{4\mu} (\partial_x U(x))^2 = -\frac{1}{4\mu} F^2 \quad (2.136)$$

(where $F = -U'$ is the force).



At low T , we can do the path integral by saddle point. The equation to solve is $0 = \frac{\delta S}{\delta x(t)}$. One set of solutions is $x = x_a$, with $x_a \in \{x_A, x_M, x_B\}$; these trajectories where nothing happens don't contribute to the path integral we're asking about.

For purposes of the following discussion let's ignore the imbalance between the two wells and suppose that they actually have the same $U(x_A) = U(x_B)$. The escape from the left well requires climbing from x_A to x_M and then rolling down to x_B . From the point of view of the auxiliary CM problem, each of these steps involves starting at the top of a hill of $V(x)$, rolling down and then up again and stopping at rest in another hill. That is, tunneling means we start at rest at the top of the hill of V , so that the energy is zero:

$$0 = E = \frac{1}{2} m \dot{x}^2 + V(x(t)) \quad (2.137)$$

which says that $V(x(t)) = -\frac{1}{2} m \dot{x}^2$, so

$$\dot{x} = \pm \sqrt{-\frac{2}{m} V(x(t))}. \quad (2.138)$$

The fact that escape occurs from left to right in our setup says we must choose the + root. [\[End of Lecture 8\]](#)

Evaluating on this solution, the bulk part of the action is

$$S_{\text{cm}}[x] = \int_{t_i}^{t_f} dt \left(\frac{1}{2} m \dot{x}^2 - V(x(t)) \right) \Big|_{x=\underline{x}} \quad (2.139)$$

$$= \int dt \left(\frac{1}{2} m \dot{x}^2 + \frac{1}{2} m \dot{x}^2 \right) = m \int dt \dot{x}^2 \quad (\geq 0) \quad (2.140)$$

$$= m \int_{t_i}^{t_f} dt \frac{dx}{dt} \sqrt{-\frac{2}{m} V(x(t))} = \int_{x_i}^{x_f} dx \left(\underbrace{-2mV(x)}_{=\frac{1}{4\mu}(\partial_x U)^2} \right)^{1/2} \quad (2.141)$$

$$= \pm \frac{1}{2} \int_{x_i}^{x_f} dx \partial_x U(x) = \frac{1}{2} |U(x_f) - U(x_i)|. \quad (2.142)$$

Therefore

$$S| = \frac{1}{2} (U(x_f) - U(x_i)) + S_{\text{cm}} = (U_f - U_i) \theta(U_f - U_i) = \begin{cases} E_{\text{act}}, & U_f > U_i \text{ climb up} \\ 0, & U_f < U_i \text{ roll down} \end{cases} \quad (2.143)$$

where $E_{\text{act}} = V(x_M)$ is the activation energy. The log of the probability for escape is then determined by the sum of these two steps, which gives

$$P[x(t)] \propto e^{-E_{\text{act}}/T}, \quad (2.144)$$

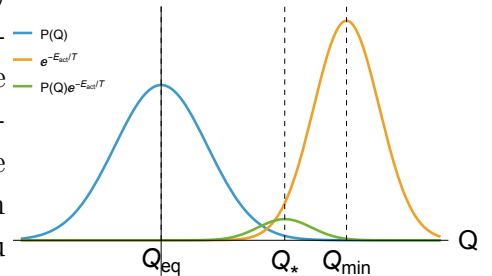
the Arrhenius law.

Here is a nice application of this exponential suppression, from Bialek's book. Suppose that the activation energy depends on some other variable Q , $E_{\text{act}} = E_{\text{act}}(Q)$. In the realization of x as a reaction coordinate, Q could be a variable describing configurations of a protein that acts as a catalyst for the reaction. Then the reaction rate is actually

$$k = c \int dQ P(Q) e^{-E_{\text{act}}(Q)/T} \quad (2.145)$$

where c is some constant, and $P(Q)$ is the distribution of protein structures.

The idea is that the reaction will only happen when $E_{\text{act}}(Q)$ is small. This may not happen for the equilibrium configuration of the protein (the maximum Q_{eq} of $P(Q)$). Rather, the reaction waits for the protein to fluctuate into the right configuration. The configurations of the protein that facilitate the reaction are then actually in the tails of its distribution (around Q_* in the picture at right), and will not be what you see if you sample $P(Q)$.



Bialek explains a nice way to test this picture by freezing the motion of Q so that each molecule has a different E_{act} and hence a different reaction rate. At fixed Q the number of molecules that have not yet reacted will be

$$N(t) = \int dQ P(Q) e^{-r(Q)t}, \quad r(Q) \equiv c e^{-E_{\text{act}}(Q)/T} \quad (2.146)$$

which has a different dependence on t than e^{-kt} with k given by (2.145).

2.5 A systematic path integral description

Above we found a path integral representation of averages of our Brownian motion by observing that the probability of a given trajectory $x_\eta(t)$ is determined by the probability of choosing the noise force η that produces it, and then changing variables in the distribution. Here is a method that involves less thinking, and also offers some other virtues that we'll discuss. It is called the (Martin-Siggia-Rose) MSR path integral method, though I believe the path integral representation itself first appears in papers by Jensen and de Dominicis.

One nice thing about it is that it provides a replacement for the partition function in a non-equilibrium context. The idea is to rewrite $1 = \int D\eta P[\eta]$ in a clever way. Let's call it

$$Z \equiv \int D\eta P[\eta] = \int D\eta P[\eta] 1 \quad (2.147)$$

$$= \int D\eta P[\eta] \int Dx \delta[\dot{x} - v(x) - \eta] \quad (2.148)$$

$$= \int D\eta P[\eta] \int Dx \int D\tilde{x} e^{\mathbf{i} \int dt \tilde{x}(t)(\dot{x}(t) - v(t) - \eta(t))} \quad (2.149)$$

$$= \int Dx D\tilde{x} e^{\mathbf{i} \tilde{x}(\dot{x} - v)} \underbrace{\int D\eta e^{-\int dt \left(\frac{\eta^2}{4D} - \mathbf{i} \tilde{x} \eta \right)}}_{= \mathcal{N} e^{-\int dt D \tilde{x}^2}} \quad (2.150)$$

$$\equiv \int Dx D\tilde{x} e^{-S_{\text{MSR}}[x, \tilde{x}]} \quad (2.151)$$

with

$$S_{\text{MSR}}[x, \tilde{x}] = \int dt (D \tilde{x}^2 - \tilde{x}(\dot{x} - v)). \quad (2.152)$$

I've ignored various normalization factors in the measure. Also, be warned that every possible convention for the factors of $\mathbf{i}$ and signs appears somewhere in the literature. The new variable $\tilde{x}$ is called the *response field* for reasons we'll see in a moment.

For the simple problem we are studying here, we can just do the (Gaussian) integral over $\tilde{x}$ to get

$$P[x] \propto e^{-\frac{(\dot{x}-v)^2}{4D}} \quad (2.153)$$

which is the formula we used above. (The exponent here is sometimes called the Onsager-Machlup action.) Notice that this is the measure we found when we derived the path integral via the Chapman-Kolmogorov equation above (2.49).

So why bother with $\tilde{x}$? I'll give several reasons here. First, Z above is really just the number 1, not so useful. But notice that we can insert sources:

$$Z[J, \tilde{J}] \equiv \int Dx D\tilde{x} e^{-S_{\text{MSR}}[x, \tilde{x}] + \int dt (J(t)x(t) + \tilde{J}(t)\tilde{x}(t))} \quad (2.154)$$

where $J(t), \tilde{J}(t)$ are functions that we get to pick. As usual, this is a generating functional for various two-point functions:

$$C(t, t') \equiv \langle x(t)x(t') \rangle = \frac{\delta}{\delta J(t)} \frac{\delta}{\delta J(t')} Z|_{J=\tilde{J}=0} \quad (2.155)$$

(where the $\langle \rangle$ means average over the noise ensemble) is the correlation function, describing the fluctuations in position our Brownian particle as a result of the noise. Another object we can construct is

$$\mathbf{i}\chi(t, t') \equiv \langle x(t)\tilde{x}(t') \rangle = \frac{\delta}{\delta J(t)} \frac{\delta}{\delta \tilde{J}(t')} Z|_{J=\tilde{J}=0} \quad (2.156)$$

Now who is this χ ? To answer this, consider perturbing the Langevin equation by an external force coupled to x :

$$\dot{x}(t) = v(x) + f(t) + \eta(t). \quad (2.157)$$

Redoing the analysis above amounts to replacing $v \rightarrow v + f$, and we have

$$Z[f] = \int Dx D\tilde{x} e^{-\int dt (D\tilde{x}^2 - \mathbf{i}\tilde{x}(\dot{x} - v(x) - f))} \quad (2.158)$$

and we see that f acts like a source for $\tilde{x}$: $\tilde{J} = \mathbf{i}h$. Thus

$$\chi(t, t') = \frac{\delta \langle x(t) \rangle_f}{\delta f(t')}|_{f=0} \quad (2.159)$$

is the *response function*, describing how the expectation of x changes in response to a (arbitrarily small) force, $\langle x(t) \rangle = \int ds \chi(t, s) f(s)$ (assuming $\langle x(t) \rangle = 0$ in the absence of the force). This is why $\tilde{x}$ is called the response field. Notice that the change in the

expectation value only happens *after* the force acts, so $\chi(t, t')$ has the property that it vanishes if $t' > t$, it is *causal*.

Now let's focus on the equilibrium situation where $D = T/\mu$ and the external force is conservative, meaning

$$v(x) = -\mu \partial_x U(x) \quad (2.160)$$

for some U . This latter condition is required to have an equilibrium state – if the force is not conservative, there will be a current in the steady state. I claim that under this assumption, the following time-reversal operation

$$x(t) \mapsto x(-t) \equiv x'(t), \quad \tilde{x}(t) \mapsto \tilde{x}(-t) + \mathbf{i} \frac{\dot{x}(-t)}{2D} \equiv \tilde{x}'(t) \quad (2.161)$$

is a symmetry, in the sense that

$$S_{\text{MSR}}[x, \tilde{x}] = S_{\text{MSR}}[x', \tilde{x}'] \quad (2.162)$$

up to boundary terms (and the measure of the path integral is also invariant). Notice that if you do this transformation twice, you get back the original configuration, it is a $\mathbb{Z}_2$ symmetry.

This symmetry has the following consequence, called a Ward identity. In a path integral, the integration variables are just dummy variables. The value of the integral doesn't care about what we call our integration variables. Therefore

$$Z[J, \tilde{J}] \equiv \int Dx D\tilde{x} e^{-S_{\text{MSR}}[x, \tilde{x}] + \int dt (Jx + \tilde{J}\tilde{x})} \quad (2.163)$$

$$\stackrel{\text{rename}}{=} \int Dx' D\tilde{x}' e^{-S_{\text{MSR}}[x', \tilde{x}'] + \int dt (Jx' + \tilde{J}\tilde{x}')} \quad (2.164)$$

$$\stackrel{\text{symmetry}}{=} \int Dx D\tilde{x} e^{-S_{\text{MSR}}[x, \tilde{x}] + \int dt (Jx' + \tilde{J}\tilde{x}')} \quad (2.165)$$

$$= \int Dx D\tilde{x} e^{-S_{\text{MSR}}[x, \tilde{x}] + \int dt (J'x + \tilde{J}'\tilde{x})} = Z[J', \tilde{J}'] \quad (2.166)$$

where

$$J'(t) \equiv J(-t) - \frac{1}{D} \partial_t \tilde{J}(-t), \quad \tilde{J}'(t) = \tilde{J}(-t). \quad (2.167)$$

As a result of this identity, we find a relation between fluctuations C and response χ :

$$\frac{\delta}{\delta J(t)} \frac{\delta}{\delta \tilde{J}(t')} Z[J, \tilde{J}] = \frac{\delta}{\delta J(t)} \frac{\delta}{\delta \tilde{J}(t')} Z[J', \tilde{J}'] \quad (2.168)$$

which, setting $J = J' = 0$ says

$$\mathbf{i}\chi(t, t') = \mathbf{i}\chi(-t, -t') - \frac{1}{D} \partial_{t'} C(t, t'). \quad (2.169)$$

Looking at this equation for $t > t'$ (which means $-t < -t'$), the first term on the RHS vanishes, and we have

$$\mathbf{i}\chi(t, t') = -\frac{1}{D}\partial_{t'}C(t, t') \quad (2.170)$$

which is a real-time version of the fluctuation-dissipation relation. We'll see various other appearances of this relation below.

Q: Above we studied $\langle x(t)x(t') \rangle$ and $\langle x(t)\tilde{x}(t') \rangle$. What about $\langle \tilde{x}(t)\tilde{x}(t') \rangle$?

A: Actually, for $t \neq t'$, it is zero in the integral above, and in any MSR integral derived from white noise. It measures the correlations of the noise with itself, it's the propagator for the noise.

2.6 Memory, correlation and response

Memory. Let's not throw away the $\ddot{x}$ term, *i.e.* let's include the inertia of the particle. For illustration, consider the special case where the system is also an SHO, so the EOM are

$$\dot{x} = p/m, \dot{p} = -m\omega^2 x - \frac{\gamma_0}{m}p + \xi. \quad (2.171)$$

Let's integrate out p , like we integrated out the bath:

$$p(t) = \int_0^t ds e^{-\frac{\gamma_0}{m}(t-s)} (-m\omega^2 x(s) + \xi(s)). \quad (2.172)$$

Plugging into the evolution of x gives:

$$\dot{x}(t) = -\int_0^t ds K(t-s)x(s) + \xi_x(t) \quad (2.173)$$

where $K(t) \equiv \omega^2 e^{-\gamma_0 t/m}$ and $\xi_x(t) \equiv \frac{1}{m} \int_0^t ds e^{-\frac{\gamma_0}{m}(t-s)} \xi(s)$ has correlations $\langle \xi_x(t)\xi_x(s) \rangle = \frac{T}{m\omega^2} K(|t-s|)$. This thus has a memory going back a time of order m/γ_0 .

Correlation. Let's take the external force to be zero and study the Langevin equation (with inertia)¹⁴

$$m\dot{v} = -\gamma v + \eta(t) \quad (2.174)$$

in Fourier space¹⁵:

$$\mathbf{i}\omega m v(\omega) = -\gamma v(\omega) + \eta(\omega). \quad (2.175)$$

¹⁴Notice, by the way, that this equation for v is of the same form as that for x in the overdamped Langevin equation with an external harmonic force – they both describe the Ornstein-Uhlenbeck process.

¹⁵with the convention $v(t) = \int d\bar{\omega} e^{\mathbf{i}\omega t} v(\omega)$. If we switch the convention to $e^{-\mathbf{i}\omega t}$ we get a relative sign in the fluctuation dissipation relation (2.183).

The correlation function of the frequency-space velocities is

$$\langle v(\omega)v(\omega') \rangle = \frac{1}{i\omega m + \gamma} \frac{1}{i\omega' m + \gamma} \underbrace{\langle \eta(\omega)\eta(\omega') \rangle}_{=2D\delta(\omega+\omega')} \quad (2.176)$$

$$= \frac{2D}{m^2\omega^2 + \gamma^2} \delta(\omega + \omega') \equiv C(\omega)\delta(\omega + \omega'). \quad (2.177)$$

For those of you who have thought about QFT or many-body perturbation theory, $\frac{1}{i\omega m + \gamma}$ is a propagator for a diffusive mode. Fourier transforming back,

$$\langle v(t)v(t') \rangle = \int_{-\infty}^{\infty} d\omega \frac{2De^{i\omega(t-t')}}{m^2\omega^2 + \gamma^2} \stackrel{\text{FDT}}{=} T e^{-|t-t'|\gamma/m} \xrightarrow{m/\gamma \rightarrow 0} \# \delta(t-t') \quad (2.178)$$

– the velocity fluctuations are correlated on times shorter than the relaxation time m/γ .

Linear response. [LP §4] We can write the damping term as a gradient $-\gamma v = -\gamma \partial_v (H(v) = v^2/2)$ so that the equilibrium distribution is $P(v) \propto e^{-H(v)/T}$. (I'm setting $m = 1$ for a bit.) Now let's perturb this Hamiltonian by a *small*, deterministic, possibly-time-dependent external force

$$H(v) = v^2/2 \longrightarrow H(v) = v^2/2 - v f(t), \quad (2.179)$$

and we'll ask how the system responds at linear order in f to this small perturbation. The Langevin equation becomes

$$m\dot{v} = -\gamma v + \gamma f(t) + \eta(t), \quad \text{or} \quad i m \omega v(\omega) = -\gamma v(\omega) + \gamma f(\omega) + \eta(\omega). \quad (2.180)$$

Since $\langle \eta \rangle = 0$,

$$\langle v(\omega) \rangle = \frac{\gamma}{i m \omega + \gamma} f(\omega) \equiv G(\omega) f(\omega). \quad (2.181)$$

In real time, $\langle v(t) \rangle = \int_{-\infty}^{\infty} ds G(t-s) f(s)$ with the *response function*

$$G(t-s) = \frac{\delta \langle v(t) \rangle}{\delta f(s)} \Big|_{f=0}. \quad (2.182)$$

Correlation $\langle v(\omega)v(\omega') \rangle = C(\omega)\delta(\omega + \omega')$ and response are related by

$$\frac{\omega}{2T} C(\omega) = -\text{Im} G(\omega) \quad (2.183)$$

another expression of the fluctuation-response relation. Notice that $G(\omega)$ is intrinsic to the system – the noise drops out of the RHS, but appears in the LHS quite explicitly. Nevertheless, the condition $D = \gamma T$ makes $C(\omega)$ an equilibrium property.

As a small application, and to clarify why such relations are sometimes called fluctuation-*dissipation* relations, let's suppose that we act with a (weak) external force on our system and ask about the work it does. For example, consider a time-dependent (possibly generalized thermodynamic) force $f(t)$ acting on the variable $v(t)$ (it may be a bit unfamiliar that I'm acting directly on the velocity, sorry). The average work done is

$$W = \int dt f(t) \left\langle \frac{dv(t)}{dt} \right\rangle. \quad (2.184)$$

using (the equation before) (2.182),

$$W = \int dt f(t) \int ds \partial_t G(t-s) f(s) = \int d\omega i\omega G(\omega) f(\omega) f(-\omega). \quad (2.185)$$

Since $f(t)$ is real, $f(-\omega) = f(\omega)^*$, so this is

$$W = \int d\omega i\omega G(\omega) |f(\omega)|^2 \quad (2.186)$$

and we can see that some positivity properties can obtain. Because $\omega |f(\omega)|^2$ is odd under $\omega \rightarrow -\omega$, only the odd part of $G(\omega)$ contributes. But since $G(t)$ is real the odd part of the response function $\frac{G(\omega) - G(-\omega)}{2} = i \text{Im} G(\omega)$ is (proportional to) its imaginary part. So we can use (2.183) to write

$$W = - \int d\omega \omega \text{Im} G(\omega) |f(\omega)|^2 = \frac{1}{2T} \int d\omega \omega^2 C(\omega) |f(\omega)|^2. \quad (2.187)$$

There is a lot more to say about such relations; I am trying to be brief here. See *e.g.* Chaikin and Lubensky chapter 7. Lacking above is a general derivation of (2.183). Let's fill this gap now.

Here is a more general discussion of the relation between correlations and linear response. It is from Landau-Lifshitz vol 5 (Statistical Physics Part 1). It is more general in that we do not make any assumptions about the dynamics of the variable x , and it is fully quantum mechanical. In quantum statistical mechanics there is an extra source of uncertainty beyond the distribution over states – even if you know the state, you can still only predict probabilities for measurements. So we should think about an ensemble of quantum states

$$\rho = \sum_n p_n |n\rangle \langle n|, \quad (2.188)$$

that is, a density matrix. Here, if the $|n\rangle$ are orthonormal, as we'll choose them to be, then p_n is the probability that we are in the state $|n\rangle$. In thermal

equilibrium, $|n\rangle$ are the energy eigenstates $H|n\rangle = E_n|n\rangle$, and $p_n = e^{-\beta E_n}/Z$, i.e. $\rho = e^{-\beta H}/Z$.

In quantum mechanics, the random variables we've been talking about as observables are replaced by operators, which may not commute with each other. So the analog of " $\langle x_\omega x_{\omega'} \rangle$ " requires a choice of the order in which the two operators appear. We follow the perspicacious rule for doing physics and mathematics that when we have a choice to make, we average over all possible choices. So, we define the fluctuation $(x^2)_\omega$ by

$$2\pi\delta(\omega + \omega')(x^2)_\omega = \frac{1}{2} \langle \hat{x}_\omega \hat{x}_{\omega'} + \hat{x}_{\omega'} \hat{x}_\omega \rangle \quad (2.189)$$

where the hats are intended to remind us that $\hat{x}$ is an operator, and the $\langle \rangle$ is with respect to some density matrix that I have not specified yet.

Let's evaluate the fluctuations (2.189) in an energy eigenstate, so the RHS becomes

$$\text{RHS} = \frac{1}{2} \langle \psi_n | (\hat{x}_\omega \hat{x}_{\omega'} + \hat{x}_{\omega'} \hat{x}_\omega) | \psi_n \rangle \equiv \frac{1}{2} (\hat{x}_\omega \hat{x}_{\omega'} + \hat{x}_{\omega'} \hat{x}_\omega)_{nn} \quad (2.190)$$

$$= \frac{1}{2} \sum_m (\hat{x}_\omega)_{nm} (\hat{x}_{\omega'})_{mn} + (\omega \leftrightarrow \omega'). \quad (2.191)$$

The matrix element is

$$(\hat{x}_\omega)_{nm} = \int dt \underbrace{x_{nm}}_{\equiv \langle n | \hat{x} | m \rangle} e^{-i(\omega_{nm} + \omega)t} = 2\pi x_{nm} \delta(\omega_{nm} + \omega) \quad (2.192)$$

where $\omega_{nm} \equiv (E_n - E_m)/\hbar$. (I wrote $(\hat{x})_{nm} = \int dt e^{-i\omega t} \langle n | \hat{x}(t) | m \rangle = \int dt e^{-i\omega t} \langle n | e^{-iHt} \hat{x} e^{iHt} | m \rangle$.) Using this and $x_{nm} = x_{mn}^*$, (2.190) becomes

$$\text{RHS} = 2\pi^2 \sum_m |x_{nm}|^2 (\delta(\omega_{nm} + \omega) \delta(\omega_{mn} + \omega') + (\omega' \leftrightarrow \omega)) \quad (2.193)$$

$$= 2\pi^2 \sum_m |x_{nm}|^2 (\delta(\omega_{nm} + \omega) + \delta(\omega_{mn} + \omega)) \delta(\omega + \omega'). \quad (2.194)$$

Comparing with the definition (2.189) and using $\omega_{mn} = -\omega_{nm}$, we have

$$(x^2)_\omega = \pi \sum_m |x_{nm}|^2 (\delta(\omega_{nm} + \omega) + \delta(\omega_{nm} - \omega)). \quad (2.195)$$

So much for the fluctuations in an energy eigenstate. (The LHS should have some n label on it.)

Now let's think about linear response. Perturb the system by

$$\hat{V} = -f\hat{x} = -\frac{1}{2}\hat{x}(f_0e^{i\omega t} + f_0^*e^{-i\omega t}) , \quad (2.196)$$

a periodic potential. Fermi's golden rule says that the transition rate from state n to state m is

$$w_{mn} = \frac{\pi|f_0|^2}{2\hbar^2}|x_{mn}|^2(\delta(\omega + \omega_{nm}) + \delta(\omega - \omega_{nm})) \quad (2.197)$$

where the two terms in (2.197) come from the two terms in (2.196). The mean energy absorbed per unit time is then

$$Q = \sum_m w_{mn}\omega_{mn} \quad (2.198)$$

$$= \frac{\pi|f_0|^2}{2\hbar^2} \sum_m |x_{mn}|^2 (\delta(\omega + \omega_{mn}) + \delta(\omega + \omega_{nm})) \omega_{mn} \quad (2.199)$$

$$= \frac{\pi|f_0|^2}{2\hbar^2} \omega \sum_m |x_{mn}|^2 (\delta(\omega + \omega_{nm}) - \delta(\omega - \omega_{nm})) \quad (2.200)$$

Let's define the response function or generalized susceptibility $\chi(t)$ by

$$\bar{x}(t) = (\chi f)(t) = \int_0^\infty d\tau \chi(\tau) f(t-\tau) = \frac{1}{2}(\chi(\omega)f_0e^{i\omega t} + \chi(-\omega)f_0^*e^{-i\omega t}) \quad (2.201)$$

(with our choice of monochromatic perturbation). In Fourier space this convolution becomes a product

$$\hat{x}_\omega = \chi(\omega)f_\omega . \quad (2.202)$$

We can write the energy absorbed in terms of χ :

$$Q = -\frac{1}{4}i\omega(\chi^* - \chi)|f_\omega|^2 = -\frac{1}{2}\omega\text{Im}\chi(\omega)|f_0|^2. \quad (2.203)$$

This follows from the definition of work

$$\partial_t E = -\bar{x}\partial_t f \quad (2.204)$$

upon averaging over time. Comparing with our expression for Q in (2.198) we have, in the energy eigenstate n ,

$$\text{Im}\chi(\omega) = -\frac{\pi}{\hbar} \sum_m |x_{nm}|^2 (\delta(\omega + \omega_{nm}) - \delta(\omega - \omega_{nm})) . \quad (2.205)$$

(Again the LHS should have some n label on it.) You can see there are some serious structural similarities between this expression for the imaginary part

of the response function, which determines the absorption of energy, and our expression for the fluctuations of the variable x we are perturbing (2.195). The big difference is just the relative minus sign between the positive and negative frequency terms.

Now let's specify to thermal equilibrium, *i.e.* the Gibbs distribution, $\hat{\rho} = e^{-\beta H}/Z = \sum_n \underbrace{e^{-\beta E_n}/Z}_{\equiv p_n} |n\rangle\langle n|$ and write an equation relating the correlation and response. The fluctuations are

$$(x^2)_\omega|_\beta = \pi \sum_{nm} p_n |x_{nm}|^2 (\delta(\omega + \omega_{nm}) + \delta(\omega - \omega_{nm})) \quad (2.206)$$

$$= \pi \sum_{nm} p_n (1 + e^{\hbar\omega_{nm}/T}) |x_{nm}|^2 \delta(\omega + \omega_{nm}) \quad (2.207)$$

$$= \pi (1 + e^{-\hbar\omega/T}) \sum_{nm} p_n |x_{nm}|^2 \delta(\omega + \omega_{nm}). \quad (2.208)$$

In the second step, we relabelled the summation variables in the second term $n \leftrightarrow m$ and used $p_m = p_n e^{\omega_{nm}/T}$. The response function, by the same logic, is

$$-\text{Im} \chi(\omega)|_\beta = \frac{\pi}{\hbar} (1 - e^{-\hbar\omega/T}) \sum_{nm} p_n |x_{nm}|^2 \delta(\omega + \omega_{nm}). \quad (2.209)$$

Comparing these two expressions, we have

$$(x^2)_\omega|_\beta = -\hbar \text{Im} \chi(\omega)|_\beta \cosh \frac{\hbar\omega}{2T} = -2\hbar \text{Im} \chi(\omega)|_\beta \left(\frac{1}{2} + \frac{1}{e^{\hbar\omega/T} - 1} \right). \quad (2.210)$$

You recognize the last expression in parens as the energy of the harmonic oscillator in thermal equilibrium (divided by $\hbar\omega$). When $T \gg \hbar\omega$, the relation becomes

$$(x^2)_\omega|_\beta = -\frac{2T}{\omega} \text{Im} \chi(\omega)|_\beta \quad T \gg \hbar\omega. \quad (2.211)$$

If we integrate over all frequencies, we have

$$\langle x^2 \rangle_\beta = -2\hbar \int_0^\infty d\omega \text{Im} \chi(\omega) \coth \frac{\hbar\omega}{2T} \xrightarrow{\text{classical}} -4T \int_0^\infty d\omega \frac{\text{Im} \chi(\omega)}{\omega} = -T \chi(0). \quad (2.212)$$

The classical limit holds for a system where the support of $\text{Im} \chi(\omega)$ is in the regime $\hbar\omega \gg T$. In the last step we used the Kramers-Kronig relation for χ , which is true for the same reasons you learned about in the graduate E&M class.

More than one variable. In the above I suppressed all possible indices on our variables. The response function now has two indices: Perturbing by $\hat{V} = -\hat{x}_i f_i(t)$, with f_i again monochromatic,

$$\bar{x}_i(t) = \frac{1}{2} (\chi_{ik}(\omega) f_{0k} e^{i\omega t} + \chi_{ik}^*(\omega) f_{0k}^* e^{-i\omega t}) \quad (2.213)$$

which says

$$\overline{x_{i\omega}} = \chi_{ik}(\omega) f_{k\omega}. \quad (2.214)$$

The definition of work says $\partial_t E = -\dot{f}_i \bar{x}_i$ and so

$$Q = -\frac{1}{4} \mathbf{i} \omega (\chi_{ik}^* - \chi_{ki}) f_{0i} f_{0k}^*. \quad (2.215)$$

Putting these together, the generalization of the fluctuation-dissipation relation (FDT) (2.210) is

$$(x_i x_k)_\omega |_\beta = -\frac{1}{2} \mathbf{i} \hbar (\chi_{ki}^* - \chi_{ik}) \coth \frac{\hbar \omega}{2T}. \quad (2.216)$$

A few comments. The FDT is a signature of equilibrium, which is out of bounds in this class. But we had to talk about it, so at least we can recognize when we're going astray.

The word 'dissipation' in FDT comes from the fact that it is the imaginary part of the response function that appears, and this is the part that is related to the dissipation of energy.

You can see from our derivation that there is still a relationship between fluctuations and response even if we don't assume thermal equilibrium, it's just a different relationship.

If we assume time-reversal symmetry and $\mathcal{T}(x) = x$ (no magnetic field, our operators are not momenta) then we can show that $\chi_{ik}(\omega) = \chi_{ki}(\omega)$. See LL v 5 page 391 for the proof. This relation is closely related to the Onsager reciprocity relations about which people get excited. (For example, Onsager won the Nobel prize (in chemistry) for this and not for his solution of the 2d Ising model.) An argument for its importance, due to Dam Son, is that it halves the number of response functions we need to compute.

[End of Lecture 9]

2.7 Getting in trouble with the UV

[DPA, Gardiner, Cugliandolo 2504.11994, page 21, §3.4.7, Discretization of stochastic differential equations]

To see that we can get in trouble if we proceed blindly, let us give here a cautionary tale. There are two alternatives to proceeding blindly: one is to introduce some kind of mathematical rigor and formalism called stochastic calculus; another is to think about the physics we are describing. We will choose the latter option.

To clearly illustrate where the trouble is, consider a stochastic variable that evolves under *multiplicative* noise:

$$\dot{u}(t) = \eta(t)u(t) \quad (2.217)$$

where $\eta(t)$ is a Gaussian random variable with distribution

$$P(\eta) \propto e^{-\frac{1}{2} \int d\omega \eta(\omega) \hat{\phi}(\omega) \eta(-\omega)} . \quad (2.218)$$

Such multiplicative noise happens: for example, suppose $u(t)$ describes a population of some species, and the noise represents births and deaths; clearly the more animals there are, the larger the number of births and deaths. Another equivalent way to describe it would be to say $\dot{u}(t) = \zeta(t)$, but the kernel for $\zeta(t)$ depends on u . We will consider two cases for the kernel $\hat{\phi}$: $\hat{\phi}(\omega) = 1/\Gamma$ (white noise down all the way past the Planck energy) and $\hat{\phi}(\omega) = (1 + \omega^2 \tau^2)/\Gamma$ for some time scale τ that puts a short-time cutoff on how jagged the noise can be.

You can see that there is room for trouble because a sample of white noise $\eta(t)$ is not a nice function – it's not even continuous. So the RHS $\eta(t)u(t)$ is not continuous, and so $\dot{u}(t)$ is not continuous...

Let $t_j = j\Delta t$ to discretize the evolution. The actual solution is

$$u(t_{j+1}) = e^{\int_{t_j}^{t_j+\Delta t} ds \eta(s)} u(t_j). \quad (2.219)$$

Naively, the exponent is of order Δt and we can replace it with:

$$u(t_{j+1}) \stackrel{??}{=} (1 + \eta(t_j)\Delta t) u(t_j) \quad (2.220)$$

(2.220) is solved iteratively by

$$u(t_N) = (1 + \eta(t_{N-1})\Delta t) \cdots (1 + \eta(t_0)\Delta t) u(t_0). \quad (2.221)$$

Now what's $\langle u(t_N) \rangle$? For white noise, $\langle \eta(t_j)\eta(t_k) \rangle = 0$ for all $k \neq j$ – it is uncorrelated down to arbitrarily short times. So there is no drift on average: $\langle u(t_N) \rangle = u(t_0)$. (This answer is named after someone named Ito.)

Alternatively consider the continuum description: $\Delta t \rightarrow 0$, $N\Delta t$ fixed, $t_0 = 0$, $t_N = t$, where the solution is

$$u(t) = u(0) e^{\int_0^t ds \eta(s)} \quad (2.222)$$

and so the average is

$$\langle u(t) \rangle = u_0 e^{\frac{1}{2} \int_0^t ds \int_0^t ds' \langle \eta(s)\eta(s') \rangle} = u(0) e^{\Gamma t/2} \quad (2.223)$$

which gives a nonzero drift named after Stratonovich.

It is tempting to say that this apparent contradiction is caused by a non-uniformity of the limits $\Delta t \rightarrow 0$ and $\tau \rightarrow 0$, that is, these limits do not commute. Ito takes $\tau \rightarrow 0$ first, while Stratonovich takes $\Delta t \rightarrow 0$ first. A defensible point of view is that they are both wrong: τ and Δt are always finite. But this is actually not the explanation for the discrepancy.

If we replace the white noise with our regulated gaussian noise with kernel

$$\langle \eta(\omega)\eta(\omega') \rangle = \delta(\omega + \omega') \frac{\Gamma}{1 + \omega^2 \tau^2} \quad (2.224)$$

what happens? The continuum answer is still

$$\langle u(t) \rangle = u_0 e^{\frac{1}{2} \int_0^t ds \int_0^t ds' \langle \eta(s)\eta(s') \rangle} \quad (2.225)$$

and Fourier transforming gives

$$\langle \eta(s)\eta(s') \rangle = \Gamma \underbrace{\frac{1}{2\tau} e^{-|s-s'|/\tau}}_{\xrightarrow{\tau \rightarrow 0} \delta(s-s')} \quad (2.226)$$

and therefore

$$\langle u(t) \rangle = u_0 \exp \frac{\Gamma}{2} (t + (-1 + e^{-t/\tau})\tau). \quad (2.227)$$

You can see that this has an essential singularity at $\tau = 0$, but returns to the Stratonovich answer as $\tau \rightarrow 0$.

But my statement about the order of limits is not quite correct. The real difference arose in going from (2.219) to (2.220). To understand this, let's see how to get the Stratonovich answer from discrete steps. I claim that it comes from

$$u_{j+1} = (1 + \alpha \Delta t + \Delta W_j) u_j \quad (2.228)$$

where $\Delta W_j \equiv \eta(t_j) \Delta t = \sqrt{\Gamma \Delta t} \text{randn}()$. $\alpha = 0$ is Ito, and I claim that $\alpha = \Gamma/2$ is Stratonovich.

The key ingredient is the observation that $S_N \equiv \sum_{j=0}^{N-1} (\Delta W_j)^2$ is a sum of gaussian random variables, whose distribution at large N , by the CLT, becomes very narrowly peaked around $\sum_{j=0}^{N-1} \Gamma \Delta t = \Gamma t$. This is a statement about the distribution, not just about the mean. So we can make the replacement:

$$\Delta W_j^2 \rightarrow \Gamma \Delta t. \quad (2.229)$$

Let's analyze the log of our random variable

$$\log u_N = \log u_0 \prod_{j=0}^{N-1} (1 + \alpha \Delta t + \Delta W_j) = \log u_0 + \sum_{j=0}^{N-1} \log (1 + \alpha \Delta t + \Delta W_j) . \quad (2.230)$$

Using $\log 1 + x = x - \frac{x^2}{2} + \mathcal{O}(x^3)$, we have

$$\begin{aligned}\log u_N &= \log u_0 + \sum_j \left(\alpha \Delta t + \Delta W_j - \frac{1}{2} (\alpha \Delta t + \Delta W_j)^2 + \dots \right) \\ &= \log u_0 + \sum_j \left(\alpha \Delta t + \Delta W_j - \frac{1}{2} (\alpha^2 \Delta t^2 + 2\alpha \Delta t \Delta W_j + \Delta W_j^2) + \mathcal{O}(\Delta W_j^3) \right).\end{aligned}\quad (2.231)$$

Now let's examine the large- N behavior of each of these terms, where we take $N \rightarrow \infty, \Delta \rightarrow 0$ fixing $t = N\Delta t$, and use the rule (2.229). The first two terms are order 1. The first quadratic term $\sum_j \Delta t^2 \sim t\Delta t \rightarrow 0$ is small, as is the second $\sum_j \Delta t \Delta W \sim \Delta t^{1/2}$. The third term though gives $\sum_j \Delta W_j^2 \sim t$. All the higher order terms go like a positive power of Δ . What's left is:

$$\log u_N = \log u_0 + \sum_j \left(\alpha \Delta t + \Delta W_j - \frac{1}{2} \Delta t + \dots \right) \quad (2.232)$$

and you can see that this stochastic drift $-\frac{1}{2}\Gamma\Delta t$ is cancelled exactly if $\alpha = \Gamma/2$, in which case we get

$$u(t) \equiv u_N = u_0 e^{W(t)}, \quad W(t) \equiv \int_0^t ds \eta(s) \quad (2.233)$$

as an ‘operator equation’, *i.e.* a statement about the whole distribution of the random variables on the BHS. In terms of expectation values, this says $\langle u(t) \rangle = u_0 e^{\Gamma t/2}$, the Stratonovich behavior.

How do we keep both $\tau > 0$ and $\Delta t > 0$? Here is a clever trick.

The noise with kernel $\phi(\omega) = (1 + \omega^2 \tau^2)/\Gamma$, *i.e.* $\langle \eta(\omega) \eta(\omega') \rangle = \delta(\omega + \omega') \frac{\Gamma}{1 + \omega^2 \tau^2}$ has a very physical origin: it is generated by the Ornstein-Uhlenbeck process:

$$d\eta(t) = -\frac{1}{\tau} \eta(t) dt + \frac{\sqrt{\Gamma}}{\tau} dW(t) \quad (2.234)$$

where $dW(t)$ is called ‘Wiener process’, which is a shorthand for

$$\frac{\sqrt{\Gamma}}{\tau} dW(t) \equiv \frac{1}{\tau} \sqrt{\Gamma dt} \text{randn}(). \quad (2.235)$$

(Yes, we are using white noise to generate colored noise.)

Integrating (2.234) from t_n to $t_n + \Delta t = t_{n+1}$ gives

$$\eta_{n+1} = \eta_n e^{-\Delta t/\tau} + \sigma_{\eta \Delta t} \text{randn}() \quad (2.236)$$

with

$$\sigma_{\eta\Delta t}^2 = \langle \eta(\Delta t)^2 \rangle \quad (2.237)$$

$$= \int_{t_n}^{t_{n+1}} ds \left(e^{-(t_{n+1}-s)/\tau} \frac{\sqrt{\Gamma}}{\tau} \right)^2 = \frac{\Gamma}{2\tau} (1 - e^{-2\Delta t/\tau}). \quad (2.238)$$

The noise acts directly only on η . To evolve u , then we just do

$$u_{n+1} = u_n e^{\int_{t_n}^{t_{n+1}} ds \eta(s)} \approx u_n e^{\Delta t \frac{\eta_n + \eta_{n+1}}{2}} \quad (2.239)$$

where we used the trapezoid rule to do the integral.

So $\Delta t \rightarrow 0, \tau$ fixed gives the continuum answer, and if we then take $\tau \rightarrow 0$ we get Stratonovich. If we keep Δt fixed and set $\tau \rightarrow 0$, we actually still get Stratonovich. How do we get Ito from a continuum description? The answer is that we have to study a different continuum equation:

$$\dot{u} = \eta u - \frac{1}{2}\Gamma u. \quad (2.240)$$

The second term is a fine-tuned deterministic drift force designed to precisely cancel out the contributions from ΔW^2 . This seems very unphysical – why might someone want to do this fine tuning? One answer is: if they were doing something other than physics. A process like Ito where there is no drift, where our best guess for the future value of the random variable is its current value, is called a *martingale*. This kind of thing is used in finance, where if there were some drift, people could profit off of it. In that context, the hard work of doing the fine tuning is precisely the job being done by all the finance bros who spend their days looking for any possible arbitrage and greedily gobbling it up, just as the fungus on the floor of the forest spends its days extracting all available free energy from the carcasses of dead animals.

Example. Let's consider the heat dumped into the bath by a 1d Langevin particle in a potential $U(x)$ and verify that the first law of thermodynamics is satisfied. The internal energy of the system is just $U(x)$. The first law says that in each time step, $dU = dW - \mathfrak{d}Q$ where $\mathfrak{d}Q$ is the heat dumped into the bath. If the potential is time-independent, $dW = \partial_t U dt = 0$, no work is done. The heat dumped into the bath $\mathfrak{d}Q = -F_{\text{bath}} dx$ where $F_{\text{bath}} = -\dot{x}/\mu + \xi$ is the force exerted by the bath on the particle. In the overdamped limit,

$$m\ddot{x} \approx 0 = F_{\text{bath}} + F_{\text{ext}} = -\gamma\dot{x} + \xi - \partial_x U. \quad (2.241)$$

Let's solve this for the noise: $\xi = \gamma\dot{x} + \partial_x U$. So the heat dumped into the bath

is

$$-dQ = (-\gamma\dot{x} + \xi)dx = (-\gamma\dot{x} + (\gamma\dot{x} + \partial_x U))dx = \partial_x U dx = dU \quad (2.242)$$

and indeed the first law is satisfied.

This seems quite trivial and of course it had to be the case, but now let's see how it is possible to go wrong by following Ito. Taylor expand the potential about the initial point:

$$U(x + dx) - U(x) = U'(x)dx + \frac{1}{2}U''(x)dx^2. \quad (2.243)$$

But now $dx = -\mu U'(x)dt + \sqrt{2D}dW$ satisfied the Langevin equation, and I've been making a big deal about how $dW^2 \sim dt$ is weirdly large. So it seems we should write

$$dU_{\text{Ito}} = U'(x)dx + DU''(x)dt \quad (2.244)$$

with this weird extra term that violates the first law of thermodynamics.

To avoid this, we should instead use the Stratonovich rule for discretizing integrals of stochastic variables, which is to Taylor expand about the midpoint $x_m \equiv \frac{x_k + x_{k+1}}{2}$:

$$dU = U(x + dx) - U(x) \quad (2.245)$$

$$= \left(U(x_m) + U'(x_m)\frac{1}{2}dx + \frac{1}{2}U''(x_m)\left(\frac{1}{2}dx\right)^2 + \dots \right) \quad (2.246)$$

$$- \left(U(x_m) - U'(x_m)\frac{1}{2}dx + \frac{1}{2}U''(x_m)\left(\frac{1}{2}dx\right)^2 + \dots \right) \quad (2.247)$$

$$= U'(x_m)dx = U'(x + \frac{1}{2}dx)dx \quad (2.248)$$

with no extra term.

So the only weird thing about Stratonovich is that when discretizing an integral of random variables, you should use the midpoint rule:

$$\int f(x)dx \stackrel{\text{yes}}{=} \sum_k f\left(\frac{x_k + x_{k+1}}{2}\right)(x_{k+1} - x_k) \quad (\text{Stratonovich}) \quad (2.249)$$

$$\stackrel{\text{no}}{=} \sum_k f(x_k)(x_{k+1} - x_k) \quad (\text{Ito}) \quad (2.250)$$

rather than the endpoint.

Let me conclude this chapter with the following useful statement generalizing our one-variable results above. If our walkers satisfy the Langevin equation

$$du_a = A_a(u)dt + B_{ab}(u)dW_b \quad (2.251)$$

with $\langle dW_a dW_b \rangle = \delta_{ab}dt$, the probability density $P(u_a, t)$ satisfies the FP equation

$$\partial_t P = -\partial_{u_a} (A_a(u)P(u, t)) + \frac{1}{2} \partial_{u_a} \partial_{u_b} ((BB^T)_{ab}P(u, t)) . \quad (2.252)$$

This can be shown along the same lines we used above, but here is one more point of view. Start from the Chapman-Kolmogorov equation again:

$$P(u, t + dt) = \int du' P(u, t + dt | u', t) P(u', t). \quad (2.253)$$

The transition matrix is

$$P(u, t + dt | u', t) = \langle \delta(u - u' - du(t)) \rangle \quad (2.254)$$

$$\begin{aligned} &= \left(1 + \langle du_a(t) \rangle \partial_{u'_a} + \frac{1}{2} \langle du_a(t) du_b(t) \rangle \partial_{u'_a} \partial_{u'_b} + \dots \right) \delta(u - u') \quad (2.255) \\ &= \delta(u - u') + \left(A_a(u(t)) \partial_{u'_a} + \frac{1}{2} B_{ac}(u) B_{bc}(u) \partial_{u'_a} \partial_{u'_b} \right) \delta(u - u') dt + \dots \end{aligned}$$

where we ignore higher moments of the Brownian step du because they go like higher powers of dt . Now plug this into (2.253), IBP, divide by dt and take $dt \rightarrow 0$.

3 Markov chains

[DPA, LP §2.3] Let's label the states of our system $i = 1..\mathcal{N}$, this can be a very big number. i could even label quantum states if we like. Suppose we do not know exactly what is the state of the system, we would like to keep track of our knowledge of it by a probability distribution P_i . This is a set of $\mathcal{N}$ non-negative numbers satisfying $\sum_i P_i = 1$.

How does this probability distribution evolve in time? A *Markov chain* describes a time evolution of such a distribution that depends only on the distribution at the current time:

$$p_j(t+1) = \sum_i W_{ji} p_i(t). \quad (3.1)$$

The fact that both $p_i(t)$ and $p_j(t+1)$ are probabilities requires that the transition matrix satisfy

$$W_{ij} \geq 0, \forall i, j, \quad \sum_i W_{ij} = 1, \forall j. \quad (3.2)$$

Such a matrix is called a *stochastic matrix*. Notice that the form of the RHS (3.1) is matrix multiplication.

Not all reasonable evolution rules satisfy the Markov property. A good example is a self-avoiding walk. Suppose, for example, that the walker at time step t represents the spatial position of the t th monomer in a long polymer. Clearly it cannot go in exactly the same place as one of the monomers for smaller t .

Notice that (3.1) is *linear* in p , just as the Schrödinger equation is linear in the wavefunction. Why should this always be? I don't know. One perspective on linearity of the time evolution in quantum mechanics is that it prevents polynomial time solutions of NP-hard problems. Maybe I'll comment below about trying to make a system where the time evolution is not linear in p .

Given that (3.1) is linear in p , we can use linear algebra to solve it. A few facts about the eigenvalues λ of W , which satisfy the characteristic equation $\det(W - \lambda \mathbb{1}) = 0$: • $|\lambda| \leq 1$. • There exists at least one eigenvalue $\lambda = 1$. • All the other eigenvectors w_j have $\sum_j w_j = 0$.

Here is an obvious but important statement (suppressing matrix indices): For any t, n ,

$$p(t+n) = W^{t+n} p(0) = W^n W^t p(0). \quad (3.3)$$

This is called the *Chapman-Kolmogorov equation*. It says that a Markov chain gets from here to there by being somewhere in the middle. With the indices maybe it seems more impressive:

$$p_i(t+n) = (W^n)_{ij} (W^t)_{jk} p_k(0) \quad (3.4)$$

– it says that we can cut open the time evolution by summing over the states. It has the same form as inserting $\mathbb{1} = \sum_j |j\rangle\langle j|$ somewhere in a quantum mechanical amplitude, for example as we do a large number of times when deriving the path integral representation.

We can rewrite (3.1) as

$$p_j(t+1) = p_j(t) - \underbrace{1}_{=\sum_i W_{ij}} p_j(t) + \sum_i W_{ji} p_i(t) \quad (3.5)$$

which says

$$p_j(t+1) - p_j(t) = \sum_i (W_{ji} p_i(t) - W_{ij} p_j(t)) \quad (3.6)$$

– the change in p_j is a positive contribution from any state s_i jumping to s_j and a negative contribution of s_j jumping to any state s_i . This rewriting of the equation is designed to make manifest the guarantee that $0 = \sum_i (p_i(t+1) - p_i(t))$. A distribution p_j is stationary if $p_j(t+1) = p_j(t)$, so

$$\sum_i (W_{ji} p_i - p_j W_{ji}) = 0 \quad \forall j. \quad (3.7)$$

If we take the time increment Δt to be infinitesimal¹⁶ and divide (3.6) by Δt , we get a continuous-time master equation:

$$\partial_t p_i = \sum_j (R_{ij} p_j - R_{ji} p_i) \equiv - \sum_j \Gamma_{ij} p_j \quad (3.8)$$

with $R_{ij} \equiv \lim_{\Delta t \rightarrow 0} W_{ij}/\Delta t$. The structure of this equation is designed to guarantee that $0 = \partial_t(1) = \partial_t \sum_i p_i$. The physics of the discrete and continuous-time cases can sometimes differ in detail in important ways (I am thinking of [2508.15726](#)).

Claim: If $R_{ij} \geq 0$ then $p_i(t=0) \geq 0$ implies $p_i(t) \geq 0$ for all later times. Proof: Suppose that at some $t = t_0$, p_i is about to become negative, *i.e.* $p_i = 0, \dot{p}_i < 0$, and is the first p_i to do so. Then $\dot{p}_i = \sum_j W_{ij} p_j(t) \geq 0$, contradiction.

The transition rates R_{ij} (or W_{ij}) could depend on time. If they do not, $\dot{R}_{ij} = 0$, the solution of the continuous-time master equation is simple:

$$P_i(t) = (e^{-\Gamma t})_{ij} P_j(0), \quad (3.9)$$

at least formally. Exponentiating the matrix Γ is best understood in the eigenbasis of Γ . If Γ were a real symmetric matrix, with a spectral decomposition

$$(\Gamma)_{ij} = \sum_{\alpha} \lambda_{\alpha} \psi_i^{\alpha} \tilde{\psi}_j^{\alpha} \quad (3.10)$$

¹⁶More precisely, I should use some other variable k to denote the discrete label above, $p_j(k+1) - p_j(k) = \sum_i (W_{ji} p_i(k) - p_j(k) W_{ji})$, and define $t = k\Delta t$, so that the LHS becomes $\dot{p}_j(t)$ as $\Delta t \rightarrow 0$ fixing $t = \Delta t k$.

(i.e. $\{\lambda_\alpha\}$ are the eigenvalues of Γ , $\{\psi_i^\alpha\}$ are the corresponding eigenvectors), then the exponential is just

$$(e^{-t\Gamma})_{ij} = \sum_{\alpha} e^{-t\lambda_{\alpha}} \psi_i^{\alpha} \psi_j^{\alpha}. \quad (3.11)$$

But actually, while Γ should be real, generally there is no reason to have $\Gamma \stackrel{?}{=} \Gamma^T$. So it will have different left and right eigenvectors:

$$\phi_i^{\alpha} \Gamma_{ij} = \lambda_{\alpha} \phi_j^{\alpha}, \quad \Gamma_{ij} \psi_j^{\beta} = \lambda_{\beta} \psi_i^{\beta}. \quad (3.12)$$

The first equation is the same as $\Gamma^T \phi = \lambda \phi$. We can check that the eigenvalues are the same on both sides: $F(\lambda) \equiv \det(\lambda - \Gamma) = \det(\lambda - \Gamma^T) = F(\lambda^*)^*$. The last step shows that the eigenvalues come in complex conjugate pairs. Combine the two equations in (3.12) to get

$$0 = (\lambda_{\alpha} - \lambda_{\beta}) \langle \phi^{\alpha} | \psi^{\beta} \rangle \quad (3.13)$$

so the right and left eigenvectors for different eigenvalues are orthogonal. The spectral decomposition takes the form

$$\Gamma = \sum_{\alpha} \lambda_{\alpha} |\psi^{\alpha}\rangle \langle \phi^{\alpha}|, \quad i.e. \Gamma_{ij} = \sum_{\alpha} \lambda_{\alpha} \psi_i^{\alpha} \phi_j^{\alpha}. \quad (3.14)$$

So the solution of the (time-independent) master equation is

$$p_i(t) = \sum_{\alpha} (e^{-\lambda_{\alpha} t} c_{\alpha}) \psi_i^{\alpha}, \quad c_{\alpha} \equiv \phi_j^{\alpha} p_j(0). \quad (3.15)$$

Notice that $\sum_i p_i = 1$ says that $\sum_i \Gamma_{ij} = 0$ which means the vector of all 1s: $\phi_i^0 \equiv \frac{1}{\mathcal{N}}(1, \dots, 1)$ is always a left eigenvector of Γ with eigenvalue $\lambda^0 = 0$. The corresponding *right* eigenvector $\psi_i^0 \equiv P_i^{\text{eq}}$ has $\Gamma_{ij} p_j^{\text{eq}} = 0$ and is therefore *stationary*, i.e. $\partial_t P_i^{\text{eq}} = 0$. So there is always at least one stationary distribution of the master equation. But: if $\mathcal{N} = \infty$, this is not the most interesting distribution.

Claim: Under the assumption $R_{ij} \geq 0$, all the eigenvalues of Γ are non-negative.

To summarize, for definiteness in the discrete-time case, $p(t+1) = Wp(t)$, it's a good idea to expand the initial state in the right eigenvectors of W :

$$p(t) = W^t p(0) = W^t \sum_{k=1}^N a_k(0) w^{(k)} = \sum_k a_k(0) (\lambda^{(k)})^t w^{(k)}. \quad (3.16)$$

If the stationary distribution is unique, meaning that all the other eigenvalues of Γ are > 0 (or of W are < 1), then if we wait long enough, the system reaches the

stationary distribution, which can be called (the dreaded word that we are not allowed to talk about this quarter) **equilibrium**,

$$p_i(t) \xrightarrow{t \rightarrow \infty} p_i^{\text{eq}} + a_2(0)e^{-t/\tau_2} \quad (3.17)$$

where the decay time τ_2 can be computed from the next eigenvalue (for the discrete evolution, it is $\tau_2 = -\frac{1}{\log|\lambda_2|}$, where λ_2 is the second largest eigenvalue of W , for the continuous evolution, $\tau_2 = 1/\lambda_2$, where λ_2 is the second smallest eigenvalue of Γ).

A strong condition we can impose on a Markov chain is *detailed balance*:

$$W_{ij}p_j^{\text{eq}} = W_{ji}p_i^{\text{eq}}, \quad \forall i, j \quad (3.18)$$

with $p_j^{\text{eq}} > 0$ for all j . This condition guarantees that there is a steady state of the Gibbs form $p^{\text{eq}}(s) = e^{-\beta H(s)}/Z$, where no currents flow, nothing happens – equilibrium. To see that DB $\iff$ the equilibrium has the Gibbs form, notice that we can solve it for the ratio of transition rates:

$$\frac{W_{ij}}{W_{ji}} = e^{-\beta E(s') - E(s)}. \quad (3.19)$$

We'll understand the implication in the other direction in the argument below. A Markov chain satisfying DB is called *reversible*, because once it reaches equilibrium, there is no arrow of time.

We can contrast equilibrium states that occur in Markov chains satisfying detailed balance with the possibility of a *non-equilibrium steady state* (NESS), a stationary distribution w (eigenvector of W with eigenvalue 1, $Ww = w$) that is not an equilibrium distribution, *i.e.* is not of the Gibbs form. In such a steady state, currents can still flow. A sort of trivial distinction to keep in mind is from the point of view of the FP equation, $\partial_t \rho = -\vec{\nabla} \cdot \vec{J}$: In equilibrium $\vec{J} = 0$, but we can have a steady state with $\vec{J} \neq 0$ as long as $\vec{\nabla} \cdot \vec{J} = 0$.

Failure of DB requires only a pair of states ij such that $W_{ij}w_j^{\text{eq}} \neq W_{ji}w_i^{\text{eq}}$. Q: Given the transition rates $\{W_{ss'}\}$, can we tell whether they lead to an equilibrium steady state? The answer is yes. The idea, called the Kolmogorov condition, comes from the realization that DB is saying that a certain cohomology problem is trivial:

$$\frac{w_{ss'}}{w_{s's}} = \frac{p^{\text{eq}}(s')}{p^{\text{eq}}(s)} \forall s, s' \leftrightarrow \prod_{i=1}^N W_{s_{i+1}s_i} = \prod_{j=1}^N W_{s_{j-1}s_j} \forall \text{ loops of states } s_1 \cdots s_N, s_{N+1} \equiv s_1, s_0 \equiv s_N. \quad (3.20)$$

$\boxed{\Rightarrow}$: DB implies

$$\prod_{i=1}^N \frac{W_{s_{i+1}s_i}}{W_{s_{i-1}s_i}} = \prod_{i=1}^N \frac{p^{\text{eq}}(s_{i+1})}{p^{\text{eq}}(s_i)} = 1. \quad (3.21)$$

[End of Lecture 10]

$\boxed{\Leftarrow}$: How to get $p^{\text{eq}}(s)$ from the transition rates? Pick a reference state s_0 . Then any other state s must have

$$p^{\text{eq}}(s) \equiv p^{\text{eq}}(s_0) \frac{W_{ss_0}}{W_{s_0,s}}, \quad (3.22)$$

which determines $p^{\text{eq}}_{s_0}$ by $\sum_s p^{\text{eq}}(s) = 1$. The idea is we are demanding DB for s_0 and any state it's connected to. Now, does the distribution defined this way satisfy DB for all *other pairs* s, s' ? In general, no. In general, the result for some other s' will depend on the path we take to get from s_0 to s' . The Kolmogorov condition says that in fact the result is path independent, like a condition for no magnetic flux in the state space. Suppose the Kolmogorov criterion holds for the loop s, s_0, s' :

$$W_{s,s_0} W_{s',s} W_{s_0,s'} \stackrel{!}{=} W_{s',s_0} W_{s,s'} W_{s_0,s}. \quad (3.23)$$

Rearranging, this says

$$\left(\frac{W_{s,s_0}}{W_{s_0,s}} \right) W_{s',s} = \left(\frac{W_{s',s}}{W_{s_0,s'}} \right) W_{s,s'} \quad (3.24)$$

which by our construction of $p^{\text{eq}}(s)$ says exactly DB between s and s' :

$$\left(\frac{p^{\text{eq}}(s)}{p^{\text{eq}}(s_0)} \right) W_{s',s} = \left(\frac{p^{\text{eq}}(s')}{p^{\text{eq}}(s_0)} \right) W_{s,s'}. \quad (3.25)$$

■

So, without knowing all the steady states, we can tell just given the update rules whether there can be a NESS or not. The idea is that a failure of the Kolmogorov criterion (3.20) arises when the dynamics produces a flow on the states with a nontrivial curl, so that it can't be written as the gradient of some Hamiltonian, which is what would have to appear in the exponent of the Gibbs equilibrium state.

Note that DB does not imply that the Gibbs state is unique. A stupid counterexample is: stick together two systems,

$$W = W^1 \oplus W^2 = \begin{pmatrix} W^1 & 0 \\ 0 & W^2 \end{pmatrix}. \quad (3.26)$$

So there are no transitions between states in the first system and states in the second. Each W can satisfy DB. If we start with a state in the first (second) system, we end up in the equilibrium state of the first (second) system.

The discussion of equilibration of Markov chains can be confusing. Let me clarify some terminology.

- A Markov chain is *irreducible* if every state can be reached from any other with nonzero probability. That is, $\forall$ states i, j there is some $t < \infty$ such that $(W^t)_{ij} > 0$.

An example of a Markov chain that is not irreducible is $W = W^1 \oplus W^2$ in the example I gave above.

Irreducible implies that if there exists a stationary state, it is unique. I'll give a proof below.

This concept of irreducible is not so innocent as it seems. A family of Markov chains of growing size $\mathcal{N}$ which is irreducible for every finite $\mathcal{N}$ can become reducible in the thermodynamic limit $\mathcal{N} \rightarrow \infty$. This is what's required for spontaneous symmetry-breaking, for example. It's what happens in your brain when you try to remember something – you very much want there to be distinct steady states for each bit of your memory which do not mix with each other under the stochastic dynamics that your brain undergoes.

- How could the late time behavior not result in a stationary state? Here is an example: $W = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Quantum mechanically, we could make a superposition, but classically the state, if it doesn't start at $p = (\frac{1}{2}, \frac{1}{2})$, just hops back and forth forever. A Markov chain is *aperiodic* if this doesn't happen. Let $R_i \equiv \{n \geq 1 \mid (W^n)_{ii} > 0\}$ be the set of return times to state i , and let $d(i) \equiv \text{GCD}(R_i)$. W is aperiodic if $d_i = 1$ for all states i . Note that if W is irreducible, all states have the same d_i . So for an irreducible chain $W_{ii} \neq 0$ for some i is enough.
- A Markov chain is *ergodic* if it is irreducible and aperiodic. For $\mathcal{N} \rightarrow \infty$, we should also demand that all states are *persistent*, meaning that the probability to return after some finite time is 1. Ergodic implies $\exists!$ stationary state.
- A Markov chain is *reversible* if detailed balance holds. In this case, the stationary state is an equilibrium state. The word 'reversible' refers to the fact that in equilibrium there is no arrow of time.

Here is a proof that an irreducible stochastic matrix W has a unique stationary state, *i.e.* right eigenvector of eigenvalue 1. It is part of the Perron-Frobenius theorem.

The idea is to prove that any *left* eigenvector $\sum_i x_i W_{ij} = x_j$ is a constant, x_i is independent of i . Let $M \equiv \max_j (x_j)$ and choose some j that realizes the maximum:

$$M = x_j = \sum_i x_i W_{ij}. \quad (3.27)$$

Since by assumption $x_i \leq x_j = M$, the RHS is bounded above by $\sum_i W_{ij} M$. But conservation of probability says $\sum_i W_{ij} = 1$. But this means that our chain of inequalities must actually all be equalities, so

$$0 = \sum_i \underbrace{(M - x_i)}_{\geq 0} \underbrace{W_{ij}}_{\geq 0}. \quad (3.28)$$

But the RHS is a sum of products of non-negative numbers, which therefore must all separately vanish. We conclude that any state i reachable from j must also have $x_i = M$. Then we can repeat the argument and draw the same conclusion for any state reachable from i , and so on – achieving the maximum is contagious. So the property of attaining the maximum holds for any state reachable by any number of steps from j , which by the irreducible assumption is everybody. So the left eigenspace of eigenvalue 1 is one-dimensional. Therefore so is the right eigenspace of eigenvalue one, since they share a characteristic polynomial.

3.1 Examples

1. Poisson process. Take the configurations to be labelled by non-negative integers $m, n \in \{0, 1, 2, \dots\}$ and take

$$W_{mn} = \begin{cases} \lambda, & m = n + 1, \\ 0, & \text{else} \end{cases}. \quad (3.29)$$

So the master equation is $\partial_t P_n = \lambda(P_{n-1} - P_n)$ (with $P_{-1} \equiv 0$).

This equation looks complicated but is actually the simplest possible thing: it just says that a new event happens at rate λ (independently of all other events), and its result is to increment n by 1. For example: shine a dim light on a photodetector and count how many photons arrive in time intervals T . Their arrival times are independent and the counts will be governed by the above master equation.

It's helpful to organize the information about the probabilities into a generating function

$$P(z, t) \equiv \sum_{n=0}^{\infty} z^n P_n(t) \quad (3.30)$$

which solves $\partial_t P(z, t) = \lambda(z-1)P(z, t)$ and is solved by $P(z, t) = e^{(z-1)\lambda t} P(z, 0)$. If the initial condition is that for sure $n = 0$, *i.e.* $P_n(0) = \delta_{n,0}$, then $P(z, t) = e^{(z-1)\lambda t}$. To extract the probabilities themselves we can do a contour integral

$$P_n(t) = \oint \frac{dz}{z^{n+1}} P(z, t) = e^{-\lambda t} (\lambda t)^n / n!, \quad (3.31)$$

which is the Poisson distribution, $P_n = \alpha^n e^{-\alpha} / n!$ all of whose cumulants are equal to α .

2. Radioactive or otherwise-excited atoms. Let $n \in \{0, 1, 2, \dots\}$ be a number of extant radioactive atoms. Then we can take the transition matrices to be

$$W_{nm} = \begin{cases} n\gamma, & m = n - 1, \\ 0, & \text{else} \end{cases} \quad (3.32)$$

where γ is the decay rate of one atom. The master equation is

$$\partial_t P_n = (n+1)\gamma P_{n+1} - n\gamma P_n. \quad (3.33)$$

Here you can think of $P_n(t) = |\langle \psi(t) | n \rangle|^2$ where $|n\rangle$ is the state of n excited atoms. The average number of excited atoms at time t is $N(t) \equiv \sum_{n=0}^{\infty} n P_n(t)$. Using the master equation, it satisfies

$$\dot{N} = \sum_{n=0}^{\infty} n ((n+1)\gamma P_{n+1}(t) - n\gamma P_n(t)) = -\gamma N(t), \quad (3.34)$$

whose solution is the familiar $N(t) = N(0)e^{-\gamma t}$.

It's interesting to think about this in terms of the generating function: $N(t) = \partial_z P(z, t)|_{z=1}$.

3. Suppose instead of decaying on their own, the particles annihilate each other, *i.e.* in chemistry notation, rather than $A \rightarrow \emptyset$, the dynamical step is $A + A \rightarrow \emptyset$. Then, assuming the events are independent, the master equation is

$$\partial_t P(n, t) = \lambda(n+1)(n+2)P(n+2, t) - \lambda n(n-1)P(n, t) \quad (3.35)$$

where the factors come from the number of ways of choosing a pair of reacting particles.

In the second and third example, the stationary state is just no particles left, quite boring. In the first case, at $t \rightarrow \infty$, all the probability leaks away: $P_n(t) \xrightarrow{t \rightarrow \infty, n \text{ fixed}} 0$. Actually none of these examples satisfy detailed balance, because they are all one-way processes where the number of particles can only decrease.

4. Diffusion (discrete time). Let the configuration space be labelled by $n \in \mathbb{Z}$, and let the time variable be discrete. Suppose n describes the position of a particle on a discretized line, and that at each time step $t = k\Delta t$, with probability r it hops to the left and with probability $1 - r$ hops to the right. Then the (discrete-time) master equation is

$$P_n(k+1) = rP_{n+1}(k) + (1-r)P_{n-1}(k). \quad \text{Diagram: A horizontal line with several circles representing nodes. Two adjacent nodes are highlighted. An arrow labeled '1-r' points from the right node to the left node. Another arrow labeled 'r' points from the left node to the right node. The sequence of nodes ends with an ellipsis '...'.$$

Check that $\sum_n (BHS) = 1$.

This fails detailed balance if $p \neq 1$, but is still ergodic. In that case, with periodic boundary conditions, the steady state has a net current.

Let's replace $n \in \mathbb{Z}$ with a ring of N sites, *i.e.* $n = 1..N$ with periodic boundary conditions. Then the master equation (3.36) can be written as the matrix equation $P(k+1) = WP(k)$ with

$$W = \begin{pmatrix} 0 & 1-r & & r \\ r & 0 & & \\ & & \ddots & \\ 1-r & & & \ddots \end{pmatrix} \quad (3.37)$$

Now there is a steady state independent of t with p_n a constant, independent of n , so $p_n = 1/N$ by normalization.

Two questions: (1) what is the spectrum of decay times to this steady state? (2) How long $t_{\bullet}$ does it take to visit all the sites?

To answer (1), we can recognize that the eigenvalue problem in this case is translation-invariant (and linear) and so solved by Fourier transform, $w_n = \omega^n$, and the PBC implies $w_{N+1} = w_1$, so $\omega^N = 1$. The eigenvectors are then

$$w_k^{(j)} = \frac{1}{\sqrt{N}} e^{2\pi i j k / N} \equiv (\omega_j)^k / \sqrt{N}, j = 0..N-1. \quad (3.38)$$

and the eigenvalues are

$$\lambda_j = (1 - r)\omega_j + r\omega_j^{-1} = \cos \frac{2\pi}{N}j + \mathbf{i}(1 - 2r) \sin \frac{2\pi}{N}j \quad (3.39)$$

which is only real for $r = 1/2$.

To answer (2) for $r \neq 1/2$ note that the particle moves ballistically and the average velocity of the particle is $v = r - (1 - r) = 2r - 1$, so $t_\bullet \sim N/v$. For $r = 1/2$ it moves by diffusion, so $t_\bullet \sim N^2$. (Except that for N even it is always on an odd (even) site for t odd (even).)

Continuum description in space. Suppose we look at the preceding problem at very large N and from very far away. Surely then we can use a continuum description where $x_n \equiv na$, and we can do our accounting in terms of a probability distribution $P(x, t)$. So we can rewrite the master equation (3.36) as

$$P(x, k+1) = rP(x+a, k) + (1-r)P(x-a, k) \quad (3.40)$$

$$\begin{aligned} &= r \left(P(x, k) + a\partial_x P(x, k) + \frac{a^2}{2}\partial_x^2 P(x, k) + \dots \right) \\ &\quad + (1-r) \left(P(x, k) - a\partial_x P(x, k) + \frac{a^2}{2}\partial_x^2 P(x, k) + \dots \right) \\ &= 1 \cdot P(x, k) + (1-2r)a\partial_x P(x, k) + \frac{a^2}{2}\partial_x^2 P(x, k) + \mathcal{O}(a\partial_x)^3 \end{aligned} \quad (3.41)$$

The change in a single time step of $P(x)$ is then, ignoring the higher derivatives,

$$P(x, k+1) - P(x, k) = (1-2r)aP' + \frac{a^2}{2}P''. \quad (3.42)$$

Unsurprisingly, given the conservation of probability, the right-hand side is of the form $-\vec{\nabla} \cdot \vec{J}dt$ with $\vec{J} = \vec{v}P - D\vec{\nabla}P$ where we find $D = \frac{a^2}{2dt}$, $v = \frac{(2r-1)a}{dt}$.

What we've done here is sometimes called Kramers-Moyal expansion. One thing that this is good for is finding a nice expression for the steady-state distribution. In this case, you can see that the steady-state distribution satisfies $\partial_x J_x = 0$, which requires

$$c = (2r-1)aP - DP' \quad (3.43)$$

for a constant c , whose solutions for the steady state distribution are

$$P(x) = Ae^{kx} + B, \quad k = \frac{a(2r-1)}{D}. \quad (3.44)$$

The demand that $P(x+L) = P(x)$ says that A must be zero, and we are left with

$$P_{ss}(x) = \frac{1}{L}, \quad (3.45)$$

a uniform distribution. This uniform distribution is not so innocent, though, because it hosts a nonzero current:

$$J_x^{ss} = vP_{ss} = a\frac{2r-1}{L}. \quad (3.46)$$

Why do we stop at second order in the above expansion? There is an interesting theorem that says that if it doesn't stop at second order, the exact answer never stops (see Tauber page 67 for an argument). I think this is related to the impossibility of constructing sensible fixed points for random walks where the n th cumulant is marginal for $n > 2$.

5. Random walk on an interval with absorbing boundaries. Now let's change the boundary conditions so that if the particle hits site 1 or N it gets *absorbed*, meaning that it stays there forever.

$$W = \begin{pmatrix} 1 & 1-r & 0 & \dots & 0 & 0 \\ 0 & 0 & 1-r & \dots & 0 & 0 \\ 0 & r & 0 & \dots & \vdots & \vdots \\ & & & \ddots & & \\ 0 & & & \dots & r & 1 \end{pmatrix} \quad \begin{array}{c} \text{Diagram of a 1D lattice with } N \text{ sites. Site 1 (left) and Site } N \text{ (right) are red circles representing absorbing boundaries. Transitions between sites } i \text{ and } i+1 \text{ are labeled with } 1-r \text{ (leftward) and } r \text{ (rightward). Ellipses indicate intermediate sites.} \end{array} \quad (3.47)$$

(Beware a bad typo in LP equation 2.80.) Clearly this is not a reversible Markov chain – detailed balance is not satisfied. To see this notice that the loop between sites 1 and 2 $W_{12}p_2^{\text{eq}} = p_1^{\text{eq}} \underbrace{W_{21}}_{=0}$ would require that $p_2^{\text{eq}} = 0$. It's also not irreducible, since no other state is accessible from $n = 1$ or $n = N$.

Here's an interesting question about this case: Suppose at $t = 0$, the particle starts somewhere in the middle, at site j : $p_i(0) = \delta_{ji}$. Eventually it will *definitely* be absorbed by the boundary at either $i = 1$ or $i = N$. What's

$$\wp_j \equiv \text{Prob (walker starting at } j \text{ ends up at 1 and never hits } N) ? \quad (3.48)$$

For example, suppose you have $j - 1$ times M dollars you have that you are willing to gamble. For some reason, at each time step you bet M dollars on some game of chance with probability r of winning and $1 - r$ of losing. For example, I am told that for roulette $r = 18/37$ and $1 - r = 19/37$, which makes it almost sound like maybe it's not an incredibly stupid thing to do. If $j = N$, you end up with $(N - 1)M$ dollars you will declare victory and leave and do something more productive and interesting. If $j = 1$, you end up with 0 dollars, you will also leave and not come back, hopefully. So $\wp_j$ is the *probability of eventual ruin*.

[\[End of Lecture 11\]](#)

Here's a recursive solution. The probability of eventual ruin satisfies, for $j = 2..N - 1$,

$$\wp_j = W_{j+1,j}\wp_{j+1} + W_{j-1,j}\wp_{j-1} = r\wp_{j+1} + (1-r)\wp_{j-1} \quad (3.49)$$

with boundary condition $\wp_1 = 1, \wp_N = 0$. To see (3.49), start with the step where the walker hits $j = 1$ and work backwards. Let's solve it by ansatz (basically by Fourier transform): $\wp_j = \psi^j$, which gives $1 = r\psi + (1-r)\psi^{-1}$ whose solutions for $r \neq \frac{1}{2}$ are $\psi_1 = 1, \psi_2 = \frac{1-r}{r} \equiv s$. For $r = \frac{1}{2}$, the roots collide and the solutions of $\wp_j = \frac{1}{2}(\wp_{j+1} + \wp_{j-1})$ are $\wp_j^{(1)} = 1$ and $\wp_j^{(2)} = j$, a sort of secular solution. Notice that $\wp_j \rightarrow \frac{1}{2}(\wp_{j+1} + \wp_{j-1})$ is the 1d diffusion rule, *i.e.* replace each site with the average of its neighbors, and it's not a coincidence that the two fixed points are the two linearly-independent solutions of $\partial_x^2 \wp = 0$. The equation (3.49) is linear so we can match the boundary conditions by taking a linear combination of the two solutions:

$$\wp_j = As^j + B \implies \wp_j = \frac{s^{j-1} - s^{N-1}}{1 - s^{N-1}}, \quad s \neq 1 \quad (3.50)$$

$$\wp_j = Aj + B \implies \wp_j = \frac{N-j}{N-1}, \quad s \rightarrow 1. \quad (3.51)$$

Notice that the first line approaches the second in the limit $s \rightarrow 1$ by L'Hopital's rule.

Now, what happens at large N , *i.e.* if you have grand ambitions about how much money you want to make at gambling? For $r > \frac{1}{2}, s < 1$, the bias is to the right and $\wp_j \rightarrow s^{j-1}$. For $r < \frac{1}{2}, s > 1$, and $\wp_j \rightarrow 1$. For $r = \frac{1}{2}, s = 1$, $\wp_j \rightarrow 1$ also, so we see that for large N , eventual ruin is guaranteed for any $r \leq \frac{1}{2}$, and approached exponentially rapidly, *i.e.* $\wp_j = 1 - \mathcal{O}(e^{-t/\tau})$. For the roulette odds, starting with $j = 500$ and trying to double your money (betting 1 each round), $\wp_{500} = 1 - 10^{-12}$.

We can compare with the continuum description we gave of such exit problems in §2.3. For example, let's take the case of $r = \frac{1}{2}$ ($s = 1$), in which case there is no drift $A = 0$, only diffusion. We said that the total probability to exit through the left boundary at $x = a$ without ever hitting the right boundary at $x = b$, $G_a(x, 0)$ satisfied

$$L^\dagger G_a(x, 0) = 0, \quad \text{with} \quad G_a(a, 0) = 1, G_a(b, 0) = 0. \quad (3.52)$$

Here $L^\dagger = -A(x)\partial_x + \frac{1}{2}B\partial_x^2 = \frac{1}{2}B\partial_x^2$, so the solution is $G_a(x, 0) = \alpha x + \beta$, and plugging in the boundary conditions gives

$$G_a(x, 0) = \frac{b-x}{b-a} \quad (3.53)$$

which agrees quite well with (3.51).

With drift we can do it, too. Then we found in (3.42) that the FP (Liouville) operator was $L = (1 - 2r)a\partial_x + \frac{1}{2}a^2\partial_x^2$, so the total exit-through-the-left-when-starting-from- x probability $G(x)$ satisfies

$$0 = L^\dagger G(x) = \left(-(1 - 2r)a\partial_x + \frac{1}{2}a^2\partial_x^2 \right) G(x) \quad (3.54)$$

with the same boundary conditions. Integrating twice, we find

$$G(x) = G(a) + \frac{G'(a)}{\alpha} (1 - e^{-\alpha(x-a)}) \quad (3.55)$$

with $\alpha \equiv \frac{2(1-2r)}{a}$. Plugging in the boundary conditions gives

$$G(x) = \frac{e^{\alpha(x-a)} - e^{\alpha(b-a)}}{1 - e^{\alpha(b-a)}} \quad (3.56)$$

which again compares very well with (3.50). (Note that $e^{2(1-2r)}$ and $s = \frac{1-r}{r}$ have the same first three Taylor coefficients about $r = 1/2$.)

6. Suppose we have two sites, labelled 1 and 2. And let $n_{1,2}$ be the number of particles at each site. Suppose that a particle can hop from $1 \rightarrow 2$ with rate D , but cannot hop from $2 \rightarrow 1$. Then

$$\partial_t P(n_1, n_2, t) = D(n_1 + 1)P(n_1 + 1, n_2 - 1, t) - Dn_1 P(n_1, n_2, t). \quad (3.57)$$

This example also violates detailed balance.

7. Page rank. [Falkovich]
8. Population dynamics, part 1. [I got this example from the book by Tauber.] We saw in the first chapter that the gaussian random walk is a fixed point of the RG with a single relevant parameter. That means it could govern the universal behavior at a critical point between two phases. What two phases? This example gives one answer to that question. It turns out that it governs a transition to an absorbing state in 0+1 dimensions (we will see later that this means that it is a 0+1d analog of directed percolation).

Consider a system whose dof is a non-negative integer n , think of it as the population of some species of animal on an island. We don't keep track of the position information here, that's for the second part of the course. The possible events that can happen to our organisms are: asexual reproduction $A \xrightarrow{\sigma} A + A$, and death $A \xrightarrow{\kappa} \emptyset$, and these happen at the rates indicated. Not the most exciting life,

but nevertheless, they may be capable of experiencing joy. This information is packaged in the transition matrices: $W^b = W_{n \rightarrow n+1} = \sigma n$, $W^d = W_{n \rightarrow n-1} = \kappa n$. So the state $n = 0$ is *absorbing*, meaning that the dynamics cannot escape from it; once this extinction state is reached, nothing further ever happens.

A question we would like to answer is: what is the late-time behavior of the population? What is the probability that our organism goes extinct? Using this as a sort of order parameter, what is the phase diagram as a function of the parameters κ, σ ? (Since we can absorb one of them into the definition of t , this is a 1d phase diagram as a function of κ/σ .)

You might think that because this system has a finite number of degrees of freedom, it doesn't really make sense to speak about phases and phase transitions. How can anything non-analytic happen? Here it is useful to remember the following slogan of non-equilibrium physics: the time dimension really counts as another dimension. That is, if we ask questions about the late-time limit, $t \rightarrow \infty$, there really can be singularities in the behavior as a function of the couplings. That is what will happen in this example.

A nice way to write the master equation is

$$\partial_t P(n, t) = \sum_{\text{reactions}, I} (\mathcal{E}^{-S_I} - 1) (W^I P(n, t)) \quad (3.58)$$

where the sum runs over the possible reactions, $\mathcal{E}^s f(n) = f(n + s)$ is a sort of creation or annihilation operator, and S_I is the 'stoichiometry vector' of reaction I , *i.e.* by how much does the reaction change the numbers of the reactants. In this case, this is just a fancy (but sometimes useful) way of writing

$$\partial_t P(n, t) = \sigma(n-1)P(n-1, t) + \kappa(n+1)P(n+1, t) - (\sigma + \kappa)nP(n, t). \quad (3.59)$$

From here I can suggest two routes to answering our questions. The first is to introduce a generating function, as we did in the simpler examples above,

$$g(z, t) = \sum_{n=0}^{\infty} z^n P(n, t) \quad (3.60)$$

which satisfies $g(1, t) = \sum_{n=0}^{\infty} P(n, t) = 1$; if the initial condition is $P(n, t=0) = \delta_{n, n_0}$, then $g(z, 0) = z^{n_0}$. Moments of the distribution are given by $\langle n(t) \rangle = \sum_n n P(n, t) = \partial_z g(z, t)|_{z=1}$, and more generally $\langle n(t)^k \rangle = (z \partial_z)^k g(z, t)|_{z=1}$. The probability of extinction at time t is $P_{\text{extinction}}(t) = g(0, t)$, and so the survival probability is $P_{\text{survival}}(t) = 1 - g(0, t)$. In terms of the generating function, the

master equation (which is infinitely many ODEs) becomes a single PDE:

$$\partial_t g(z, t) = \sum_n (\sigma z^{n+1} + \kappa z^{n-1} - (\sigma + \kappa) z^n) n P(n, t) \quad (3.61)$$

$$= \kappa(1 - z) \left(1 - \frac{\sigma}{\kappa} z\right) \partial_z g(z, t) \quad (3.62)$$

with the initial condition $g(z, 0) = z^{n_0}$.

Is this progress? Yes, because this equation is of the form $0 = \partial_t g(z, t) + v(z) \partial_z g(z, t)$ and so can be solved by the method of characteristics, as follows.

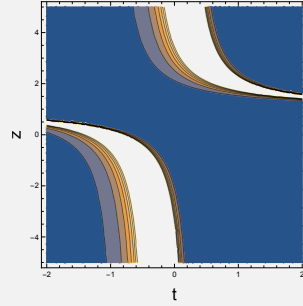
Method of characteristics. The idea is to think of the PDE

$$0 = \partial_t g(z, t) + v(z) \partial_z g(z, t) \stackrel{!}{=} \frac{dg}{ds} \quad (3.63)$$

as the condition that g is constant along a “characteristic” curve parameterized by an auxiliary variable s . That is, we want to identify the two terms in the equation with

$$0 = \frac{dg(z(s), t(s))}{ds} = \frac{dt}{ds} \partial_t g + \frac{dz}{ds} \partial_z g. \quad (3.64)$$

We are trying to directly find the contour lines $(t(s), z(s))$ of our function, as in the figure at right (for $\sigma = \kappa, n_0 = 4$). So to find $g(z, t)$ we find a curve that hits (z, t) and starts somewhere on the initial condition, and the value is just the value at that initial condition.



The first identification in this case is easy: $\frac{\partial t}{\partial s} = 1$, so we can just identify s with t . The second identification is

$$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial t} = v(z) = -\kappa(1 - z) \left(1 - \frac{\sigma}{\kappa} z\right). \quad (3.65)$$

The solution to this ODE is

$$\frac{1 - \frac{\sigma}{\kappa} z}{1 - z} = C_0 e^{-(\kappa - \sigma)t}. \quad (3.66)$$

The constant C_0 labels which characteristic curve we are on. Let $z = z_0$ be the value of $z(t = 0)$ that reaches (z, t) . So we can eliminate C_0 by

$$C_0 = \frac{1 - \frac{\sigma}{\kappa} z_0}{1 - z_0}, \quad (3.67)$$

and we have the relation

$$\frac{1 - \frac{\sigma}{\kappa}z}{1 - z} = \frac{1 - \frac{\sigma}{\kappa}z_0}{1 - z_0} e^{-(\kappa - \sigma)t}. \quad (3.68)$$

The solution of the PDE is then

$$g(z, t) = z_0^{n_0} \quad (3.69)$$

where z_0 is determined from z, t by (3.68). The result of all the algebra to eliminate z_0 is

$$g(z, t) = \begin{cases} \left(\frac{(\kappa - \sigma z) e^{(\kappa - \sigma)t} - \kappa(1 - z)}{(\kappa - \sigma z) e^{(\kappa - \sigma)t} - \sigma(1 - z)} \right)^{n_0}, & \sigma \neq \kappa \\ \left(\frac{1 + (1 - z)(\kappa t - 1)}{1 + (1 - z)\kappa t} \right)^{n_0}, & \sigma = \kappa \end{cases}. \quad (3.70)$$

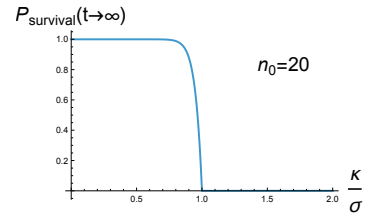
You can check that $g(1, t) = 1$ is satisfied (since the dynamics preserves probability).

The extinction probability is

$$P_{\text{extinction}}(t) = g(0, t) = \begin{cases} \left(\frac{e^{(\kappa - \sigma)t} - 1}{e^{(\kappa - \sigma)t} - \frac{\sigma}{\kappa}} \right)^{n_0} & \sigma \neq \kappa \\ \left(\frac{\kappa t}{1 + \kappa t} \right)^{n_0}, & \sigma = \kappa \end{cases} \quad (3.71)$$

which really has three regimes: if $\kappa < \sigma$ (death is less likely than birth), the extinction probability decays exponentially, if $\kappa = \sigma$ it is a power law, and if $\kappa > \sigma$ it grows exponentially. More interesting is the $t \rightarrow \infty$ behavior:

$$P_{\text{survival}}(t) = 1 - P_{\text{extinction}}(t) \xrightarrow{t \rightarrow \infty} \begin{cases} 1 - \left(\frac{\kappa}{\sigma} \right)^{n_0}, & \sigma > \kappa \\ 0, & \sigma \leq \kappa \end{cases} \quad (3.72)$$



which behaves like a wonderful order parameter for the two phases.

To answer the question of: what is the universality class of this critical point, let's talk about the second method to analyze the master equation. The other method to learn from (3.59) is to what's called van Kampen size expansion, by writing $n = Nc + \sqrt{N}x$, and expanding in powers of $1/\sqrt{N}$. In this example, N is a formal book-keeping parameter, to be regarded as large, but on which no physical answers should depend. c is the mean-field concentration, which will satisfy an ordinary differential equation that we can derive as the leading

term of the expansion. Then we can introduce a distribution for the continuous variable x as $\Pi(x, t) \equiv \sqrt{N}P(n, t)$ and derive, as the next term in the expansion, a Fokker-Planck equation for Π . I will use this method more explicitly in a more complicated example below. Here let me just report the answers: In terms of the intensive rates $w^I(c) = W^I(Nc)/N$, the mean field equation is

$$\dot{c} = \sum_I S_I w_I(c) = (\sigma - \kappa) c \quad (3.73)$$

whose solution is $c(t) = c_0 e^{(\sigma - \kappa)t}$, which either blows up or decays to extinction. In the special (critical) case $\sigma = \kappa$, $c(t) = c_0$ is constant.

The expansion is about $c(t)$: $n(t) = Nc(t) + \sqrt{N}x(t)$. The FP equation for $\Pi(x, t)$ takes the form

$$\partial_t \Pi(x, t) = \partial_x (-A(x)\Pi + D\partial_x \Pi) \quad (3.74)$$

with drift term $A(x) = Jx = (\sigma - \kappa)x$ (where $J = \frac{dc}{dc}$ is the linearization of the mean field equation about its solution), and diffusion coefficient $D = \sum_I S_I^2 w_I(c) = (\sigma + \kappa) c(t)$. The equivalent Langevin equation is just

$$\dot{x} = (\sigma - \kappa) x + \xi(t) \quad (3.75)$$

where

$$\langle \xi(t)\xi(t') \rangle = D(t)\delta(t - t') = (\sigma + \kappa) c(t)\delta(t - t'). \quad (3.76)$$

The new ingredient is that the magnitude of the noise depends on the variable $c(t)$. In particular, when the population goes extinct, $c(t) \rightarrow 0$, the fluctuations completely die too. This is the characteristic property of an absorbing state.

At the critical point, $c(t) = c_0$ is just a constant, and this just becomes ordinary Brownian motion. It's the critical theory of this phase transition. And it is indeed true that the deformation away from the critical point is adding a drift term to the Brownian motion, which was the relevant perturbation. (It is also doing something funny to the variance for which I don't have a simple statement.)

The next two examples are getting ahead of ourselves a bit, in the sense that they are about extensive dofs. But I can't resist.

9. Glauber dynamics. At each site of a graph (let's think about the square lattice) place a spin $s_i \in \{+1, -1\}$. Pick an inverse temperature β . Define the energy of a configuration according to the nearest-neighbor Ising Hamiltonian

$$E(s) = \sum_{\langle ij \rangle} s_i s_j. \quad (3.77)$$

At each time step, we do the following update: Uniformly randomly choose a site i , and compute the change in energy ΔE that would result from flipping the spin at i . Notice that this depends only on the state at i and its neighbors, so the rule is local in space. Then pick a random number $r \in [0, 1]$ from the uniform distribution; if $r < e^{-\beta \Delta E}$, flip the spin. That is,

$$W_{s \leftarrow s'} = \max(1, e^{-\beta(E_s - E_{s'})}). \quad (3.78)$$

Claim: the spin configuration converges to samples of the Boltzmann distribution $P^{\text{eq}}(s) = e^{-\beta E(s)}$. The evolution rule satisfies DB with respect to this distribution, as you can check:

$$\frac{W_{s \leftarrow s'}}{W_{s' \leftarrow s}} = e^{-\beta(E_s - E_{s'})}. \quad (3.79)$$

Notice that this Markov chain is also ergodic.

The equilibrium we are trying to approach does not determine the dynamics. A different dynamics, which is more appropriate if we want to interpret the Ising model as a lattice gas (*e.g.*, $s_i = \downarrow$ means a particle is present at site i , $s_i = \uparrow$ means there is no particle there) is the Kawasaki dynamics: at each time step, pick a link and try to exchange the state of the sites at the ends of it. With the above interpretation, this dynamics conserves the number of particles. But it also converges to the Ising equilibrium distribution (with a sector of fixed magnetization).

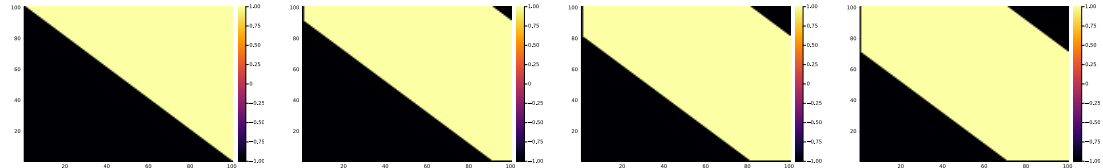
10. Toom rule.

A probabilistic cellular automaton (PCA) is a spatially-extended system with a local update rule, *i.e.* the state of each site at the next time step is determined by the current state of its neighbors. There is in general no Hamiltonian or detailed balance, but there can be very interesting non-equilibrium steady states.

The Glauber dynamics is a simple example of such a PCA. This is an example of a PCA with an equilibrium state which is therefore necessarily of the form $p(s) = e^{-\beta H(s)}/Z$ of a Gibbs state for some Hamiltonian $H(s)$. The phase diagram of such states is completely determined by the free energy $F(g) = -T \log Z$, which is some function of the parameters g of the model. A very simple fact, called the Gibbs phase rule, says that *generically* $F(g)$ will have a unique absolute minimum value. Multiple degenerate minima is what's required to have absolutely stable memory of at least one bit. But this requires tuning g to a codimension-one locus in the space of parameters. This is consistent with our experience in the Ising model: when we impose the $\mathbb{Z}_2$ symmetry flipping all the spins, the Ising model has, at $T < T_c$ and in the thermodynamic limit, two degenerate states related

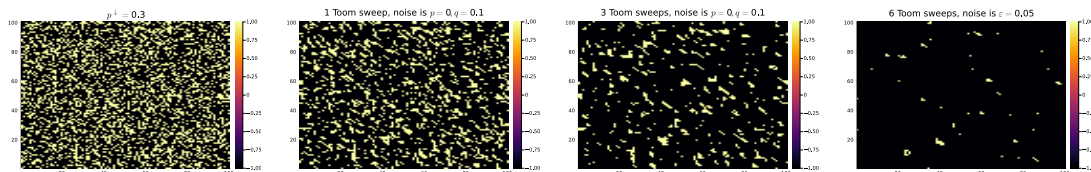
by the $\mathbb{Z}_2$ symmetry; the equilibrium distribution is of the form $p\rho_+ + (1-p)\rho_-$. Imposing the $\mathbb{Z}_2$ symmetry is constraining the couplings to a codimension one (at least) subset by forbidding the magnetic field term $\Delta H = \sum_i h s_i$; adding this term, there is indeed no longer any notion of spontaneous symmetry-breaking and no sharp phase transition between low and high temperature. So, based on these very reasonable facts from equilibrium statistical mechanics, you might have thought that having a stable memory requires imposing some kind of symmetry.

Now let's consider a PCA studied by Toom. The degrees of freedom are the same as the Ising model, but the update rule is now: update s_i to the majority of s_i and its north and east neighbors ($s_i(t+1) = s_i(t)s_{i+\hat{x}}(t)s_{i+\hat{y}}(t)$) with probability $1-p-q$, and to $+1$ with probability p and -1 with probability q . Notice that if $p \neq q$, this model has no Ising spin-flip symmetry exchanging $s = \pm 1$. It also breaks detailed balance. To see this, we just need to find a violation of the Kolmogorov criterion (3.20), that is, a loop of states around which the dynamics is the same in reverse. Here is a series of states evolved according to the zero-noise Toom rule where you can see the directionality of the evolution before your eyes:



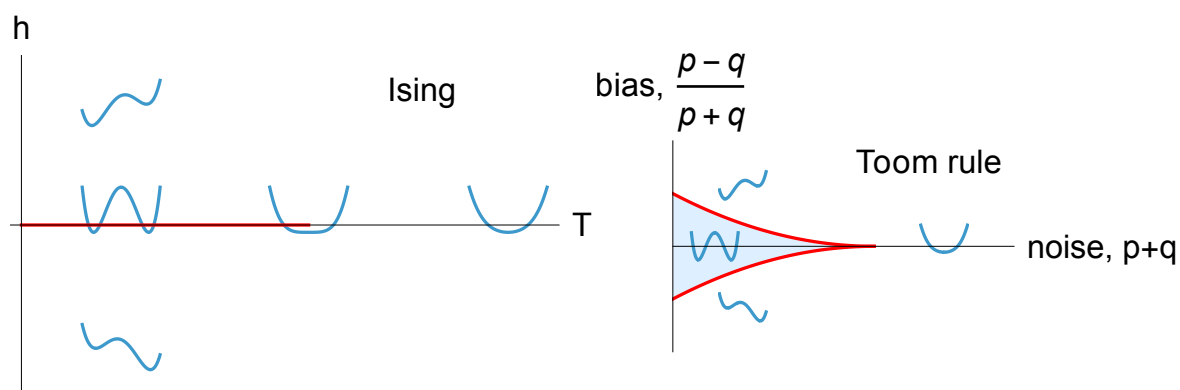
That is, a NW-SE-oriented domain wall simply moves SW. (In contrast, a domain wall oriented NE-SW stays where it is. This can be used to make a [movie that looks like snowfall](#).) If you watch the movie in reverse, it goes NE. So its steady states need not be of the Gibbs form anymore and the Gibbs phase rule is out the window.

So a right triangle of up spins with its hypotenuse running NW to SE in a sea of down spins shrinks away. The same is true of such a triangle of down spins in a sea of up spins. This rule *erodes* finite islands of minority spins, even in the presence of errors – the boundary of a region of up spins in a sea of down spins will drift to the southeast and shrink away. Even though there need be no $\mathbb{Z}_2$ symmetry, this PCA has multiple steady states (roughly all-up and all-down) in the thermodynamic limit. This is something that can happen in a system without detailed balance (and hence no notion of equilibrium) that simply cannot happen in equilibrium.



Toom proved rigorously that this is the case. Claim: For sufficiently small error rate ϵ , the Toom PCA (in the thermodynamic limit) is not irreducible and can therefore have more than one steady state. That is, it has multiple invariant measures (stable ordered phases, i.e. ergodicity-breaking), and this happens even when $p \neq q$.

Sketch: In the noisy system, errors can create new islands. To sustain a macroscopic minority domain, a connected graph of errors must occur. The probability of such an error graph decays exponentially in its size, if ϵ is small enough. Thus large domains are exponentially unlikely, and global order is stable.



The best account I've found of the physics of the Toom rule is [in this lovely essay by Geoffrey Grinstein](#).

We'll continue to discuss more examples of Markov chains and lessons we can extract from them.

[End of Lecture 12]

3.2 Two nice facts about nice Markov chains

1. For an ergodic Markov chain, every state is persistent, meaning that it has unit probability to be visited again in a finite time. This allows us to ask the question: for each state j , what is the *return time* $\langle T_j \rangle$? The answer is called the *Kac lemma*:

$$\langle T_j \rangle = \frac{1}{w_j^{(1)}} \quad (3.80)$$

where $w^{(1)}$ is the equilibrium eigenvector.

To see this, let $T_j^{(n)}$ be the time when the walk hits j for n th time. The total time to return to j M times is then $\mathcal{T}_M = \sum_{n=1}^M T_j^{(n)}$. The fraction of time spent at j is then $\phi_j(\mathcal{T}_M) = M/\mathcal{T}_M$. But as $M \rightarrow \infty$, the fraction of time spent at j approaches the equilibrium occupation probability $w_j^{(1)}$. Therefore

$$\langle T_j \rangle = \lim_{M \rightarrow \infty} \mathcal{T}_M/M = \lim_{M \rightarrow \infty} \frac{1}{\phi_j(\mathcal{T}_M)} = \frac{1}{w_j^{(1)}}. \quad (3.81)$$

So ergodicity implies equivalence of ensemble and time averages.

2. Define a notion of distance to the equilibrium distribution p_j^* ,

$$D(t) \equiv D(p(t), p_*) \equiv \sum_j \frac{1}{p_j^*} (p_j(t) - p_j^*)^2 = \sum_j \frac{p_j^2(t)}{p_j^*} - 1. \quad (3.82)$$

This is well-defined, *i.e.* p^* exists and we can divide by p_j^* either by the ergodic assumption¹⁷ or if we assume DB. This distance satisfies $D(t) \geq 0$ with equality only when $p_j(t) = p_j^*$.

Now let's ask how $D(t)$ changes in time:

$$\Delta D \equiv D(t+1) - D(t) = \sum_j \frac{p_j^2(t+1) - p_j^2(t)}{p_j^*}. \quad (3.83)$$

For *reversible* Markov chains (*i.e.* assume DB), we can show that $\Delta D \leq 0$ and $\Delta D = 0$ iff $p_j = p_j^*$ is the only stationary state (the equilibrium state). Ergodic Markov chains do equilibrate but they need not do so monotonically. We've seen that in the case of reversible chains, we can conjugate the evolution operator to a symmetric (hence hermitian) one, which therefore has real eigenvalues. For ergodic but not reversible chains (an example is the biased walk on a ring), the eigenvalues need not be real, which means that there can be oscillations on the way the inevitable boring conclusion.

Here is the calculation; it's a little complicated, but involves only three pieces of input: the evolution equation, the probability-conservation condition, and

¹⁷Ergodic implies there exists a unique steady state p_i^* . Since $\sum_i p_i^* = 1$ at least one entry must be nonzero, call it p_j^* . But, by irreducibility, any other i is connected by a series of t steps, $(W^t)_{ij} \neq 0$. The steady state condition is $p_i^* = \sum_k (W^t)_{ik} p_k^*$ but the sum of positive terms on the RHS includes $k = i$, so $p_i^* \geq (W^t)_{ij} p_j^* > 0$.

detailed balance. First, repeatedly using $p_i(t+1) = \sum_j W_{ij} p_j(t)$,

$$\Delta D = \sum_i \frac{1}{p_i^*} \left(\sum_{j,k} W_{ij} W_{ik} p_j(t) p_k(t) - \underbrace{1}_{=\sum_j W_{ji}} p_i(t)^2 \right) \quad (3.84)$$

Then using DB, we can replace $W_{ij} W_{ik} = W_{ji} W_{ki} \frac{(p_i^*)^2}{p_j^* p_k^*}$ to get

$$\Delta D = \sum_{ijk} W_{ji} W_{ki} \frac{p_i^* p_j p_k}{p_j^* p_k^*} - \sum_{ij} \underbrace{W_{ij}}_{\stackrel{\text{DB}}{=} W_{ji} p_i^* / p_j^*} \frac{p_j^2}{p_j^*} \quad (3.85)$$

$$\stackrel{1=\sum_k W_{ki}}{=} \sum_{ijk} W_{ji} W_{ki} p_i^* \frac{p_j p_k}{p_j^* p_k^*} - \sum_{ijk} W_{ji} W_{ki} p_i^* \left(\frac{p_j}{p_j^*} \right)^2 \quad (3.86)$$

$$= -\frac{1}{2} \sum_{ijk} W_{ji} W_{ki} p_i^* \left(\frac{p_j}{p_j^*} - \frac{p_k}{p_k^*} \right)^2 \leq 0 \quad (3.87)$$

and $\Delta D = 0$ iff $p_j = p_j^*$ is the equilibrium state.

3.3 Monte Carlo

[For a much more complete account of this subject, the book by Werner Krauth, *Statistical Mechanics: Algorithms and Computations*, is both good and comprehensible.] Suppose our goal were to estimate equilibrium averages of some observable $\mathcal{O}(s)$ with respect to a Gibbs distribution $p^*(s) = e^{-\beta E(s)} / Z$:

$$\langle \mathcal{O} \rangle_{\text{eq}} \equiv \sum_{j=1}^N \mathcal{O}(s_j) p^*(s_j) \equiv \sum_j \mathcal{O}_j p_j^*. \quad (3.88)$$

I know that we're not supposed to talk about equilibrium in this class. But especially when the dofs are spread over space, so that the number of states is $N = e^{L_x L_y L_z} \gg 1$, the number of terms in this sum quickly becomes prohibitively large. Furthermore, because of the exponential in the Boltzmann factor, $p^*(s) \propto e^{-\beta E(s)}$ can be quite negligible for many states, so maybe we don't need to sum over all of them.

We can evaluate this purely equilibrium quantity using an auxiliary ergodic Markov chain¹⁸. The input to the calculation is the ratios $\frac{p_i}{p_j} = e^{-\beta(E_i - E_j)}$. Notice that we do *not*

¹⁸Apparently this brilliant idea occurred to Stanislaw Ulam while he was in hospital recovering from a head injury, unable to work, and forced to play solitaire to entertain himself. I recommend his

need the normalization factor $Z = \sum_j e^{-\beta E_j}$; this is important because that requires summing over all the states, which is what we are trying to avoid. Notice that this means we don't get $F = -T \log Z$ or the entropy $S = -\partial_T F$ from this calculation; I'll say something later about how we might get around this.

The idea is just to find a time evolution rule $s_1 \rightarrow s_2 \rightarrow s_3 \cdots \rightarrow s_n$ so that

$$\langle \mathcal{O} \rangle_{\text{eq}} = \sum_{j=1}^N \mathcal{O}_j p_j^* \approx \frac{1}{n} \sum_{t=1}^n \mathcal{O}(t) \quad (3.89)$$

where $\mathcal{O}(t) \equiv \mathcal{O}(s_t)$. The virtue is that the approximation here can be quite good even when $n \ll N$, and furthermore it can be controlled. The key idea is just the point we made above that for an ergodic Markov chain, the fraction of time spent in state s_i approaches p_i^* . Notice that we are using the idea of statistical mechanics backwards here: usually we want to predict time averages, and we do it by sampling from an ensemble; here we want to sample from an ensemble and we do it by measuring time averages.

Two questions to answer: (1) How do we choose the update rule? (2) What is the error in the approximation (3.89) as a function of n ?

There are many possible answers to (1). The simplest possibility is a rule that preserves locality, where we try to update one degree of freedom at a time. One thing we can do is Glauber dynamics: pick a site. Consider flipping it; that is, compute the change in energy ΔE that would result if you flipped it. Flip it if `rand()` $< e^{-\beta \Delta E}$, where `rand()` is a random number chosen from $[0, 1]$ with the uniform distribution. We argued earlier that this satisfies DB for the Gibbs state $p_i \propto e^{-\beta E_i}$. More generally, if we parametrize our transition rates as $W_{ji} = \omega(\beta(E_j - E_i))$ for some function $\omega(x)$, then detailed balance $\frac{W_{ji}}{W_{ij}} = e^{-\beta(E_j - E_i)}$ holds iff $\frac{\omega(x)}{\omega(-x)} = e^{-x}$ ¹⁹ The Metropolis rule is the special case $\omega(x) = \min(1, e^{-x})$.

To answer (2) we need to estimate the correlation time τ of our Markov process, defined by

$$\langle \mathcal{O}(t) \mathcal{O}(0) \rangle - \langle \mathcal{O}(t) \rangle^2 \simeq e^{-t/\tau}. \quad (3.90)$$

This is something we can measure during the evolution of our simulation. The CLT for sums of weakly correlated random variables then says that since the number of *independent* samples in n steps is n/τ , the error is proportional to $\sqrt{\tau/n}$. This is great because it means that by a use of a polynomial-in-system-size number of resources, we can get a good estimate.

memoirs, *Adventures of a Mathematician*. There are precedents for this idea, such as Buffon's needle method of computing π .

¹⁹This equation has the same form as the equation we'll derive later for the distribution of entropy produced by a NESS (4.35). Is that a coincidence?

So, for example, if you want to compute expectation values in the uniform distribution on N sites, one way to do it (probably not the best) is to sample from the hopping process we described above with $r = 1/2$, say with PBC (actually it's probably a good idea to have some nonzero probability for the walker to stay so there is no even/odd effect). One more question: we have to choose some initial condition, and generally we cannot choose it from the equilibrium distribution we are trying to sample (because if we could we wouldn't need to go through all this trouble). So clearly we cannot start measuring our observable (3.89) right away. How long do we need to wait before we start taking measurements? In the silly example I just mentioned, suppose we start with $p(i, t = 0) = \delta_{i,1}$. But we know how to find the equilibration time: if we expand the initial condition in the eigenvectors of the transition matrix W , the contribution from the second largest eigenvalue of W decays like $\lambda_2^t = e^{-t|\log \lambda_2|}$ – so the equilibration time is $\tau = 1/|\log \lambda_2|$. In a more complicated dynamics, it is not so easy to compute the spectrum of W and one has to find ways to estimate τ from the results themselves.

There is a lot more to say about this method, which I should emphasize is very widely used in both classical and quantum many body physics. It is used essentially any time we want to estimate a quantity averaged with respect to a *positive* distribution. Sometimes in quantum mechanics we want to compute sums of numbers with nontrivial phases, for example if there is a complex number in the action, or minus signs. The use of the MC method in such cases is obstructed by a *sign problem*, which makes the problem exponentially hard again.

As we approach a continuous phase transition, τ grows. This means that we need to run the MC longer to get data of the same quality. This is called *critical slowing down*. A way around this problem exists for certain systems, by designing a Markov chain whose update rule is not local. That is, it tries to update a whole macroscopic amount of sites at the same time. When such a *cluster update* method exists, as it does for the Ising model, we can still get excellent information this way about the critical behavior.

3.4 Langevin and detailed balance

To make contact with the discussion of the previous chapter, let's consider the question: does the Langevin evolution satisfy detailed balance?

Let's consider first the overdamped case, and discretize the evolution as

$$x(t + dt) \equiv x' = x(t) - U'(x(t))/\gamma dt + d\xi/\gamma \equiv F(x, d\xi) \quad (3.91)$$

where $d\xi = \int_t^{t+dt} ds \eta(s)$ is a Gaussian RV with

$$\langle d\xi \rangle = 0, \langle d\xi(t)^2 \rangle = \int_0^{dt} ds \int_0^{dt} ds' \langle \eta(s) \eta(s') \rangle = \Gamma dt, \quad (3.92)$$

that is

$$P(d\xi) = e^{-\frac{d\xi^2}{2\Gamma dt}} / \sqrt{2\Gamma dt}. \quad (3.93)$$

Notice that $d\xi$ is therefore of order $\sqrt{dt}$. So we want to know whether $W_{xx'} p_{x'} \stackrel{?}{=} W_{x'x} p_x$ for the equilibrium distribution $p_x \propto e^{-\beta U(x)}$.

We can solve (3.91) for the Gaussian RV $d\xi = \gamma(x' - x) + U'(x)dt$ to find the probability of a transition from x to x' in the time dt :

$$W_{x'x} = \langle \delta(x' - F(x, d\xi)) \rangle_{d\xi} = \gamma P(d\xi | (3.91)) = \frac{\gamma}{\sqrt{2\pi\Gamma dt}} e^{-\gamma^2 \frac{(x' - x + U'(x)dt/\gamma)^2}{2\Gamma dt}}. \quad (3.94)$$

Similarly,

$$W_{xx'} = \frac{\gamma}{\sqrt{2\pi\Gamma dt}} e^{-\gamma^2 \frac{(x - x' + U'(x')dt/\gamma)^2}{2\Gamma dt}}, \quad (3.95)$$

and therefore

$$\frac{W_{x'x}}{W_{xx'}} = \exp \left(-\frac{\gamma^2}{2\Gamma dt} (2(x' - x)U'(x)dt/\gamma - 2(x - x')U'(x')dt/\gamma) \right) \quad (3.96)$$

$$= e^{-\frac{2\gamma}{\Gamma}(x' - x)U'(x) + \mathcal{O}(dt^2)} = e^{\frac{2\gamma}{\Gamma}(E(x) - E(x')) + \mathcal{O}(dt^2)} \quad (3.97)$$

where in the last step we used $(x' - x)U'(x) \simeq U(x') - U(x) + \mathcal{O}(dt^2)$. So detailed balance holds iff $\Gamma = 2\gamma T$, the familiar fluctuation-dissipation relation.

With inertia there are two wrinkles. First, if we write the Langevin equation as a second order equation in x , it is not Markov – there is a nontrivial memory. This is a bit of a silly legalistic statement, because we can integrate in a momentum variable to make the equation first order again:

$$x' \equiv x(t + dt) = x(t) + \frac{p(t)}{m} dt \equiv F_x(x, p, d\xi) \quad (3.98)$$

$$p' \equiv p(t + dt) = p(t) - U'(x(t))dt - \gamma \frac{p(t)}{m} dt + d\xi \equiv F_p(x, p, d\xi). \quad (3.99)$$

The second wrinkle is that the momentum variable is odd under time-reversal. The actual general detailed balance condition is

$$W_{\beta\alpha} p_\alpha = W_{\alpha^* \beta^*} p_{\beta^*} \quad (3.100)$$

where α^* is the time-reversed state of α . Here, if $\beta = (x, p)$, then $\beta^* = (x, -p)$. By the same logic as above,

$$\begin{aligned} W_{\beta\alpha} &= \langle \delta(x' - F_x(x, p, d\xi)) \delta(p' - F_p(x, p, d\xi)) \rangle_{d\xi} = \delta(x' - x - p dt/m) \frac{1}{\sqrt{2\pi\Gamma dt}} e^{-\frac{(p' - p + U'(x)dt - \gamma p dt/m)^2}{2\Gamma dt}} \\ W_{\alpha^* \beta^*} &= \delta(x - x' + p dt/m) \frac{1}{\sqrt{2\pi\Gamma dt}} e^{-\frac{(p' - p + U'(x)dt - \gamma p dt/m)^2}{2\Gamma dt}}. \end{aligned} \quad (3.101)$$

Using the same approximations as above to ignore terms of higher order in dt ,

$$\frac{W_{\beta\alpha}}{W_{\alpha^*\beta^*}} = e^{\frac{2\gamma}{T}(E(\alpha)-E(\beta))+\mathcal{O}(dt^2)} \quad (3.102)$$

with $E(\alpha) = \frac{p^2}{2m} + U(x) = E(\alpha^*)$. Again DB holds iff $\Gamma = 2\gamma T$.

3.5 A cautionary tale about demographic noise

[McKane and Newman. I learned this story from Nigel Goldenfeld.] Above we studied a number of examples of Markov chains acting on systems whose degrees of freedom were integers. On a previous homework problem, we saw that we could sometimes get away with using a coarse-grained description where we ignored the discreteness of these variables, and instead solve a differential equation. Here is an example where this causes serious trouble.

We have a patch of land populated by foxes and rabbits. For now we ignore spatial variation and just keep track of the total number of each, A and B . It has been observed by ecologists that the populations of such predator and prey species exhibit oscillations in time (see, for example, the book by Murray called *Mathematical Biology*). For a long time, ecologists believed that the origin of these oscillations is explained by the Lotka-Volterra equations:

$$\partial_t A = -\sigma A + \lambda AB \quad (3.103)$$

$$\partial_t B = +\mu B - \lambda AB. \quad (3.104)$$

The story of these equations is: left alone, predators (A) die off at a rate σ (per predator); similarly, in the absence of predation, the net number of bunnies (B) would increase at a rate μ (per bunny); predators eat prey and multiply at a rate λ (per predator per prey).

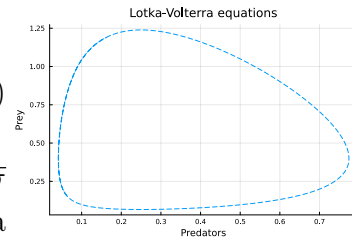
The solution of these equations does give oscillations, but it gives completely deterministic oscillations that depend on the initial conditions. This is because these equations have a conserved quantity – the evolution can be rewritten as

$$\partial_\tau Q_\alpha = \epsilon_{\alpha\beta} \partial_{Q_\beta} H \quad (3.105)$$

with

$$H = \frac{1}{\epsilon} (e^{\epsilon Q_1} - 1) - Q_1 + \epsilon (e^{Q_2/\epsilon} - 1) - Q_2 \quad (3.106)$$

where $A = \frac{\mu}{\lambda} e^{\epsilon Q_1}$, $B = \frac{\sigma}{\lambda} e^{Q_2/\epsilon}$ and $\epsilon \equiv \sqrt{\sigma/\mu}$, $\tau \equiv t\sqrt{\mu\sigma}$ are de-dimensionalized quantities. So H is conserved, and a given population is forced to oscillate around on the fixed- H locus in phase space. This is not the right explanation.

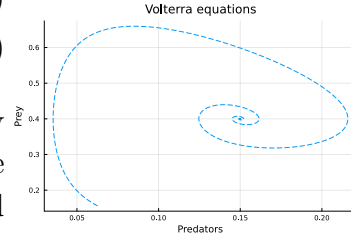


Something missing from (3.104) is the fact that there can't be too many rabbits because there is only a finite amount of grass in the area. Including such a carrying capacity K for the prey changes the equations to the Volterra equations,

$$\partial_t A = -\sigma A + \lambda_1 AB \quad (3.107)$$

$$\partial_t B = +\mu B(1 - B/K) - \lambda_2 AB. \quad (3.108)$$

Now H is no longer conserved, but there are also not any oscillations anymore! With finite K , these equations have a single attractive fixed point with nonzero values of A and B . Naturally the ecologists proceeded to try to change the equations in many ways to get back their oscillations.



The above oscillation-free equations (3.108) arise as a mean-field approximation to the following Markov chain (sometimes called ‘individual-level model’): The state of the system is specified by $n \geq 0$ predators (A), $m \geq 0$ prey (B) and $N - n - m \geq 0$ empty sites (E), though we don't keep track of spatial structure here – I hope to come back to that. The reactions are $BE \xrightarrow{b} BB$ (prey reproduction), death $A \xrightarrow{d_A} E$, $B \xrightarrow{d_B} E$, and predator-prey interactions $AB \xrightarrow{p_1} AA$, $AB \xrightarrow{p_2} AE$, with the rates indicated. The best way to simulate this process is using the Gillespie algorithm that we discussed on an earlier homework.

Let's write a master equation. It takes the form

$$\begin{aligned} \partial_t P(n, m, t) = & (\mathcal{E}_A - 1) (T(n-1, m|n, m)P(n, m, t)) + (\mathcal{E}_B^{-1} - 1) (T(n, m+1|n, m)P(n, m, t)) \\ & + (\mathcal{E}_B - 1) (T(n, m-1|n, m)P(n, m, t)) + (\mathcal{E}_A^{-1}\mathcal{E}_B - 1) (T(n+1, m-1|n, m)P(n, m, t)). \end{aligned} \quad (3.109)$$

Here $\mathcal{E}_A^{\pm 1} f(n, m) = f(n \pm 1, m)$, $\mathcal{E}_B^{\pm 1} f(n, m) = f(n, m \pm 1)$ are raising and lowering operators, and the transition matrices for the four processes are determined by the rates described above²⁰

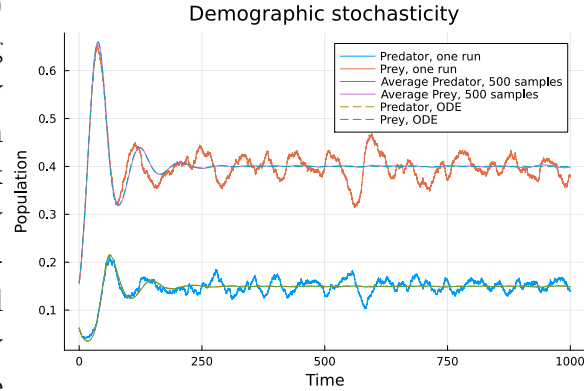
$$T(n-1, m|n, m) = d_A n, \quad T(n, m+1|n, m) = b \frac{m}{N} (N - n - m), \quad (3.110)$$

$$T(n, m-1|n, m) = p_2 \frac{nm}{N} + d_B m, \quad T(n+1, m-1|n, m) = p_1 \frac{nm}{N}. \quad (3.111)$$

To get the two Volterra equations (3.108) from this, multiply by n and m respectively, then sum over m, n , and everywhere replace $\langle m^2 \rangle \rightarrow \langle m \rangle^2$, $\langle n^2 \rangle \rightarrow \langle n \rangle^2$, $\langle mn \rangle = \langle m \rangle \langle n \rangle$, ignore anything with a $1/N$, and set $A \equiv \langle n \rangle / N$, $B \equiv \langle m \rangle / N$. The constants are related by $\sigma = d_A$, $\mu = 2b - d_B$, $K = 1 - \frac{d_B}{b}$, $\lambda_1 = p_1$, $\lambda_2 = p_1 + p_2 + b$.

²⁰I'm not sure why M&N have the factors of 2 in front of their rates b, p_1, p_2 in their equation 1. To match their plots with their stated parameter values, these factors should be removed.

You might think that this step is reasonable as long as N is big. As you can check by doing the Gillespie algorithm, (3.108) does in fact do a good job of predicting the ensemble-averaged population density – it shows transient oscillations ending in a constant steady-state. But if you look at individual trajectories of the Markov chain, they show large oscillations that persist over long times! Like in the ecological observations! What happened to the law of large numbers, which predicts that the fluctuations in a sum of random variables should go like $1/\sqrt{N}$?



[End of Lecture 13]

McKane and Newman show that the law of large numbers is correct: the oscillations do go like $1/\sqrt{N}$, but they are multiplied by a big number because of a resonance that is driven by the demographic stochasticity. Introduce the deviations from the mean-field behavior in the master equation: $n/N = A + x_A/\sqrt{N}$, $m/N = B + x_B/\sqrt{N}$, and expand in powers of N^{-1} . The leading-order result is the Volterra equations above. At the next order is a linear Fokker-Planck equation which is equivalent to the following Langevin equation:

$$\dot{x} = Jx + \eta \quad (3.112)$$

where $x = \begin{pmatrix} x_A \\ x_B \end{pmatrix}$, η is a pair of white-noise Gaussian random variables, and J is precisely the matrix in the linearization of the Volterra equations about the stable fixed point. So we know that the spectrum of this 2×2 matrix is a pair of eigenvalues both with (the same) negative real part, and equal-magnitude imaginary parts describing spiraling in to the stable fixed point. And (3.112) describes the driving of this oscillator by *white* noise. The whiteness of the noise means that *any* resonance of the oscillator, no matter its frequency, will be excited by the noise.

Van Kampen expansion. Let's spend a few minutes on the expansion step that McKane and Newman don't explain. Suppose we have a master equation in the form

$$\partial_t P(\mathbf{n}, t) = \sum_I \left(\prod_j \mathcal{E}_j^{-S_{Ij}} - 1 \right) [W_I(\mathbf{n}) P(\mathbf{n}, t)] \quad (3.113)$$

where the sum over I runs over the possible reactions, and S_{Ij} are called the ‘stoichiometry vectors’ for the reactions – by how much does the population j change during reaction I . (3.109) is of this form with $\mathbf{n} = (n, m)$.

Let’s do the expansion described in the previous paragraph. Let $\mathbf{n} = N\mathbf{c} + \sqrt{N}\mathbf{x}$, where $\mathbf{c} = (A, B)$ are the fixed-point densities and $\mathbf{x} = (x_A, x_B)$ are the fluctuation variables. Notice that a change of $\Delta n_j = 1$ corresponds to a change of $\Delta x_j = 1/\sqrt{N}$. So we’re going to Taylor expand the operators:

$$\mathcal{E}_j^{-S_{Ij}} = \exp\left(-S_{Ij} \frac{\partial}{\partial n_j}\right) = \exp\left(-\frac{S_{Ij}}{\sqrt{N}} \frac{\partial}{\partial x_j}\right). \quad (3.114)$$

Combining these, the full factor for reaction I is

$$\prod_j \mathcal{E}_j^{-S_{Ij}} - 1 = \exp\left(-\frac{1}{\sqrt{N}}(\mathbf{S}_I \cdot \nabla_x)\right) \quad (3.115)$$

$$\approx \left[-\frac{1}{\sqrt{N}}(\mathbf{S}_I \cdot \nabla_x) + \frac{1}{2N}(\mathbf{S}_I \cdot \nabla_x)^2 + O(N^{-3/2})\right]. \quad (3.116)$$

expanding in $1/\sqrt{N}$. So for example, $\mathcal{E}_A - 1$ (meaning $S_{IA} = -1, S_{IB} = 0$) gives

$$\mathcal{E}_A - 1 \approx \frac{1}{\sqrt{N}} \frac{\partial}{\partial x_A} + \frac{1}{2N} \frac{\partial^2}{\partial x_A^2}. \quad (3.117)$$

Next we Taylor expand the reaction rates around the mean-field value $\mathbf{n} = N\mathbf{c}$:

$$W_I(\mathbf{n}) = W_I(N\mathbf{c} + \sqrt{N}\mathbf{x}) \approx W_I(N\mathbf{c}) + \sqrt{N}(\mathbf{x} \cdot \nabla_c W_I)\Big|_{\mathbf{n}=N\mathbf{c}}. \quad (3.118)$$

Let’s define the intensive rate $w_I(\mathbf{c}) = W_I(N\mathbf{c})/N$, which is $\mathcal{O}(1)$. The expansion becomes:

$$W_I(\mathbf{n}) \approx Nw_I(\mathbf{c}) + \sqrt{N} \sum_j x_j \frac{\partial(Nw_I)}{\partial n_j} = Nw_I(\mathbf{c}) + \sqrt{N} \sum_j x_j \frac{\partial w_I}{\partial c_j} \quad (3.119)$$

$$\equiv W_I^{(0)} + W_I^{(1)}, \quad (3.120)$$

where $W_I^{(0)} = Nw_I(\mathbf{c})$ is the $\mathcal{O}(N)$ term, and $W_I^{(1)} = \sqrt{N}(\mathbf{x} \cdot \nabla_c w_I)$ is the $\mathcal{O}(\sqrt{N})$ term.

Plugging these expansions in to the master equation gives

$$\partial_t P = \sum_I \left[-\frac{1}{\sqrt{N}}(\mathbf{S}_I \cdot \nabla_x) + \frac{1}{2N}(\mathbf{S}_I \cdot \nabla_x)^2\right] \left[(W_I^{(0)} + W_I^{(1)})P\right]. \quad (3.121)$$

The leading term on the RHS is

$$-\sqrt{N} \sum_I \mathbf{S}_I \cdot \nabla_x (w_I(\mathbf{c})P) + \mathcal{O}(N^0). \quad (3.122)$$

On the LHS we have $\partial_t P(n, t)$, with the partial derivative taken at fixed $\mathbf{n} = N\mathbf{c}(t) + \sqrt{N}x$, which means that x is not fixed, rather

$$\partial_t x^j = -\sqrt{N} \frac{dc^j}{dt} \quad (3.123)$$

is quite large, of order $\sqrt{N}$. This means

$$\partial_t P(\mathbf{n}, t)|_n = \partial_t P|_x + \sum_j \frac{\partial}{\partial x^j} P \partial_t x^j \quad (3.124)$$

$$= \underbrace{\partial_t P|_x}_{\mathcal{O}(1)} - \underbrace{\sqrt{N} \frac{d\mathbf{c}}{dt} \cdot \vec{\nabla}_x P}_{\mathcal{O}(\sqrt{N})}. \quad (3.125)$$

equating the $\mathcal{O}(\sqrt{N})$ terms gives

$$-\sqrt{N} \frac{d\mathbf{c}}{dt} \cdot \vec{\nabla}_x P = -\sqrt{N} \sum_I (\mathbf{S}_I w_I(\mathbf{c})) \cdot \nabla_x P \quad (3.126)$$

where I used the fact that $w_I(\mathbf{c})$ is independent of x . This can only be true if

$$\frac{d\mathbf{c}}{dt} = \sum_I \mathbf{S}_I w_I(\mathbf{c}) \quad (3.127)$$

– we got a deterministic equation from the leading order of our master equation! You can check this is the mean-field equations (3.108) in our case.

The next order gives the Fokker-Planck equation for $\Pi(x, t) = \sqrt{N}P(n = Nc + \sqrt{N}x, t)$:

$$\partial_t \Pi = -\nabla_x \cdot (\mathbf{A}(\mathbf{x})\Pi) + \frac{1}{2} \nabla_x \nabla_x : (\mathbf{D}\Pi). \quad (3.128)$$

The drift vector is

$$A_j(\mathbf{x}) = -\sum_I S_{Ij} \left(\sum_k x_k \frac{\partial w_I}{\partial c_k} \right) = -\sum_k x_k \left(\sum_I S_{Ij} \frac{\partial w_I}{\partial c_k} \right) = -J_{jk} x_k \quad (3.129)$$

where $J_{jk} \equiv \frac{\partial \dot{c}_j}{\partial c_k}$ is the linearization of the mean-field dynamics. The diffusion matrix is

$$D_{jk} = \sum_I S_{Ij} S_{Ik} w_I(\mathbf{c}). \quad (3.130)$$

Plugging in our reaction rates for the McKane-Newman problem, where the reactions are

1. Predator death: $\mathbf{S}_1 = (-1, 0)$, $w_1 = d_A n / N = d_A A$

2. Prey birth: $\mathbf{S}_2 = (0, +1)$, $w_2 = bB(1 - A - B)$

3. Prey death: $\mathbf{S}_3 = (0, -1)$, $w_3 = p_2AB + d_BB$

4. Predation: $\mathbf{S}_4 = (+1, -1)$, $w_4 = p_1AB$

the Jacobian is

$$J = \begin{pmatrix} 0 & p_1A \\ -B(b + p_1 + p_2) & -bB \end{pmatrix} \quad (3.131)$$

and the diffusion matrix is

$$D = \begin{pmatrix} 2d_AA & -d_AA \\ -d_AA & 2(d_AA + d_BB + p_2AB) \end{pmatrix} \quad (3.132)$$

where $\mathbf{c} \equiv (A, B)$ and the steady-state values are $A = \frac{bp_1 - bd_A - d_2p_1}{p_1(p_1 + p_2 + b)}$, $B = d_1/p_1$.

We can see the oscillations in individual trajectories by studying the power spectrum of the Langevin variables x :

$$P(\omega)_{\alpha\beta}\delta(\omega + \omega') \equiv \langle \tilde{x}_\alpha(\omega)\tilde{x}_\beta(\omega') \rangle_\eta = \left\langle \left((i\omega\mathbb{1} - J)^{-1} \eta \right)_\alpha \left((i\omega'\mathbb{1} - J)^{-1} \eta \right)_\beta \right\rangle_\eta \quad (3.133)$$

where $\alpha, \beta \in \{A, B\}$, and the Fourier transform of (3.112) is $\tilde{x} = (i\omega\mathbb{1} - J)^{-1} \tilde{\eta} \equiv G(\omega)\tilde{\eta}$, and

$$D_{\alpha\beta}\delta(\omega + \omega') \equiv \langle \tilde{\eta}_\alpha(\omega)\tilde{\eta}_\beta(\omega') \rangle \quad (3.134)$$

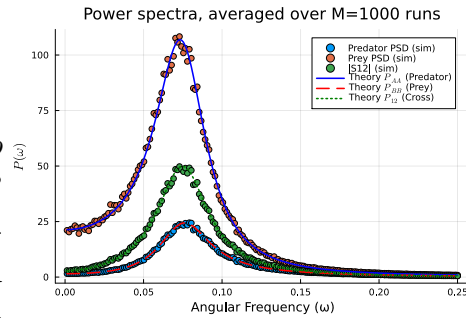
so

$$P(\omega)_{\alpha\beta} = (G(\omega)DG(-\omega))_{\alpha\beta} . \quad (3.135)$$

The result for each component is of the form

$$P_{\alpha\beta}(\omega) = \frac{a + b\omega^2}{(\omega^2 - \Omega_0^2)^2 + \Gamma^2\omega^2} \quad (3.136)$$

with $\Gamma = -\text{tr}J = -J_{BB}$, $\Omega_0^2 = -J_{AB}J_{BA}$ and a, b depend on α, β . (Q: Why aren't there two peaks? It's because the eigenvalues of J are a complex-conjugate pair with the same real part.) Ω_0 is indeed the frequency of the observed oscillations in the individual time series.



There are many other examples where individual trajectories of a stochastic process have quite different behavior from the mean-field deterministic behavior. One example in the recent literature is described by [Xu and Swingle](#) (and reviewed in Appendix C

[here](#)). They are studying the propagation of quantum information under unitary time evolution by chaotic local Hamiltonians. The effect of quantum fluctuations (the analog of demographic stochasticity) is to give the information front a nonzero thickness.

4 Fluctuation theorems

[Ritort, Udo Seifert, ... [Jarzynski, Jarzynski's original papers](#); Tauber §8.4]

Jarzynski's equality is a strengthening of the 2d Law of Thermodynamics. It says that if a system is in contact with a heat bath in equilibrium at temperature T and is subjected to some forces that drive it out of equilibrium, then

$$\langle e^{-W/T} \rangle = e^{-\Delta F/T} \quad (4.1)$$

where the average is taken over system trajectories γ . $W[\gamma]$ is the work done on the system when it follows trajectory γ , and ΔF is the change in free energy between initial and final configurations. Jensen's inequality $\langle e^X \rangle \geq e^{\langle X \rangle}$ then implies the usual 2d Law in the form $\langle W \rangle \geq \Delta F$.

The Jarzynski equality actually follows from a stronger statement called the Crooks fluctuation theorem, which says

$$\frac{P_F(W)}{P_R(-W)} = e^{\beta(W-\Delta F)}. \quad (4.2)$$

Here $P_F(W)$ is the probability distribution for the work done when following a particular driving protocol, while $P_R(W)$ is the corresponding distribution for the time-reversed driving protocol.

To see that (4.2) implies (4.1), start with the definition of the LHS of (4.1):

$$\langle e^{-W/T} \rangle = \int dW P_F(W) e^{-\beta W}. \quad (4.3)$$

Substituting for $P_F(W)$ using (4.2) gives

$$\langle e^{-W/T} \rangle = \int dW P_R(-W) e^{-\beta \Delta F} e^{\beta(W-\Delta F)} = e^{-\beta \Delta F} \int dW P_R(-W) = e^{-\beta \Delta F}. \quad (4.4)$$

So we would like to understand why and when (4.2) is true. We are going to assume a discrete Markov evolution of the probability distribution. A configuration of the system is z , and a trajectory is $\gamma \equiv \{z_0, z_1, \dots, z_M\}$, so z_k is the configuration at time $t_k = k\Delta t$. Imagine if you like that we take $M \rightarrow \infty, \Delta t \rightarrow 0$ holding $t = M\Delta t$ fixed. The master equation is

$$P_{k+1}(z) = \sum_{z'} W_k(z' \rightarrow z) P_k(z'). \quad (4.5)$$

Note that the transition rates W_k are allowed to depend on time. Probability conservation $\sum_z P_k(z) = 1$ for all time k requires

$$\sum_{z'} W_k(z \rightarrow z') \stackrel{!}{=} 1, \quad \forall k, z. \quad (4.6)$$

Now we make a big restriction, related to the words ‘in contact with a heat bath’ above. We assume that this dynamics admits an equilibrium.

- First, we let W depend on time only through some parameter λ .
- Second, we assume that for each λ , (global) detailed balance holds

$$\frac{W_\lambda(z \rightarrow z')}{W_\lambda(z' \rightarrow z)} = \frac{P_\lambda^{\text{eq}}(z')}{P_\lambda^{\text{eq}}(z)}. \quad (4.7)$$

Basically we are assuming that the system is coupled to a bath, and $P_\lambda^{\text{eq}}(z) = e^{-\beta E_\lambda(z)} / Z(\lambda)$ for some β, E_λ .

Now, for each trajectory $\gamma = \{z_0 \rightarrow z_1 \rightarrow \cdots \rightarrow z_M\}$ happening while we change the parameter according to $\lambda_0 \rightarrow \lambda_1 \rightarrow \cdots \rightarrow \lambda_M$, we define a corresponding *reverse* trajectory $\gamma^* = \{z_M \rightarrow z_{M-1} \rightarrow \cdots \rightarrow z_0\}$ happening while we change the parameter according to $\lambda_M \rightarrow \lambda_{M-1} \rightarrow \cdots \rightarrow \lambda_0$, that is, $\lambda_k^R \equiv \lambda_{M-k-1}$ happens in the opposite order. We also assume that the system along the forward trajectory starts in equilibrium according to H_{λ_0} , while the system along the reverse trajectory starts in equilibrium according to H_{λ_M} . Then the probability of a particular forward trajectory is

$$P_F[\gamma] = P_{\lambda_0}^{\text{eq}}(z_0) \prod_{k=0}^{M-1} W_{\lambda_k}(z_k \rightarrow z_{k+1}). \quad (4.8)$$

The probability for the corresponding reverse trajectory is

$$P_R[\gamma^*] = P_{\lambda_M}^{\text{eq}}(z_M) \prod_{k=0}^{M-1} W_{\lambda_k^R}(z_{M-k} \rightarrow z_{M-k-1}) = P_{\lambda_M}^{\text{eq}}(z_M) \prod_{k=0}^{M-1} W_{\lambda_k}(z_{k+1} \rightarrow z_k). \quad (4.9)$$

You can see that in the ratio we can use the detailed balance condition (4.7):

$$\frac{P_F[\gamma]}{P_R[\gamma^*]} = \frac{P_{\lambda_0}^{\text{eq}}(z_0)}{P_{\lambda_M}^{\text{eq}}(z_M)} \prod_{k=0}^{M-1} \frac{W_{\lambda_k}(z_k \rightarrow z_{k+1})}{W_{\lambda_k}(z_{k+1} \rightarrow z_k)} = \frac{P_{\lambda_0}^{\text{eq}}(z_0)}{P_{\lambda_M}^{\text{eq}}(z_M)} \prod_{k=0}^{M-1} \frac{P_{\lambda_k}^{\text{eq}}(z_{k+1})}{P_{\lambda_k}^{\text{eq}}(z_k)}. \quad (4.10)$$

Now what's left is a ratio of Boltzmann factors:

$$\frac{P_F[\gamma]}{P_R[\gamma^*]} = \frac{e^{-\beta E_{\lambda_0}(z_0)} / Z(\lambda_0)}{e^{-\beta E_{\lambda_M}(z_M)} / Z(\lambda_M)} \prod_{k=0}^{M-1} \exp \left(-\beta \underbrace{(E_{\lambda_k}(z_{k+1}) - E_{\lambda_k}(z_k))}_{\equiv -Q_k} \right) \quad (4.11)$$

where Q_k is the heat dissipated into the bath during the step from k to $k+1$ along the forward trajectory. Recall that $Z(\lambda) = e^{-\beta F(\lambda)}$, so

$$\frac{Z(\lambda_M)}{Z(\lambda_0)} = e^{-\beta \Delta F}, \quad \Delta F \equiv F(\lambda_M) - F(\lambda_0). \quad (4.12)$$

There are two ways to organize the remaining factors. One is to observe that the heat dissipated into the bath during the forward trajectory is

$$Q[\gamma] = \sum_{k=0}^{M-1} Q_k \quad (4.13)$$

and the first law says

$$W[\gamma] = \Delta E + Q[\gamma] = E_{\lambda_M}(z_M) - E_{\lambda_0}(z_0) + Q[\gamma]. \quad (4.14)$$

This is exactly what remains in the $e^{-\beta(\dots)}$.

Alternatively (and this is basically just a proof of the first law), we can regroup the terms as

$$\begin{aligned} \frac{P_F[\gamma]}{P_R[\gamma^*]} &= e^{-\beta\Delta F} e^{-\beta E_{\lambda_0}(z_0)} e^{\beta(E_{\lambda_0}(z_0) - E_{\lambda_0}(z_1) + E_{\lambda_1}(z_1) - E_{\lambda_1}(z_2) + \dots + E_{\lambda_{M-1}}(z_{M-1}) - E_{\lambda_{M-1}}(z_M))} e^{\beta(E_{\lambda_M}(z_M))} \\ &= e^{-\beta\Delta F} \exp \left(\underbrace{\beta \sum_{k=1}^M (E_{\lambda_k}(z_k) - E_{\lambda_{k-1}}(z_k))}_{=W[\gamma]} \right) = e^{\beta(W[\gamma] - \Delta F)} \end{aligned} \quad (4.15)$$

since $E_{\lambda_k}(z_k) - E_{\lambda_{k-1}}(z_k)$ is the work done **on** the system in going from λ_{k-1} to λ_k while in state z_k .

The previous equation (4.15) was a statement about the ratios of probabilities of individual trajectories. Crooks' statement is about probabilities of the observable W . To get this, we group all trajectories with the same work. The probability, $P_F(W)$, of finding a particular outcome for the work in the forward trajectory is the sum over trajectories γ_W where the work done is W of $P_F[\gamma_W]$:

$$P_F(W) = \sum_{\gamma} P_F[\gamma] \delta(W[\gamma] - W). \quad (4.16)$$

$P_R(-W)$ is the sum over the corresponding reverse trajectories of $P(\gamma_W^*)$ which therefore have work $W_R = -W$. So

$$P_F(W) = \sum_{\gamma} P_F[\gamma] \delta(W[\gamma] - W) \quad (4.17)$$

$$\stackrel{(4.15)}{=} \sum_{\gamma} P_R[\gamma^*] e^{\beta(W[\gamma] - \Delta F)} \delta(W[\gamma] - W) \quad (4.18)$$

$$= \sum_{\gamma} P_R[\gamma^*] e^{\beta(W - \Delta F)} \delta(W[\gamma] - W) \quad (4.19)$$

$$= e^{\beta(W - \Delta F)} \sum_{\gamma} P_R[\gamma^*] \delta(W[\gamma] - W) = e^{\beta(W - \Delta F)} P_R(-W). \quad (4.20)$$

■

Notice that the combination $\beta(\Delta F - W) = \Delta S$ is the entropy produced by the process. We could rewrite the relation as a relation between probabilities over S instead:

$$\frac{P_F(S)}{P_R(-S)} = e^S. \quad (4.21)$$

So the Crooks relation, like a good generalization of the Second Law, says that trajectories that produce entropy are much more likely than those that consume it.

These relations have been experimentally verified in systems of polymers where experimentalists can pull on the ends of them, and by using a harmonic optical trap to drag a particle through water. A simple solvable model of the latter is the OU process where we can move the minimum of the harmonic well: so the potential is $U(x - \lambda(t))$, with $U(x) = \frac{1}{2}\kappa x^2$. [End of Lecture 14]

Another application is to Monte Carlo simulations, where it is generally difficult to measure free energies (since it is not the average of some local observable). These relations give a way to measure differences of free energies.

4.1 Quantum Crooks relation

The Crooks relation is still true, at least sometimes, even without coupling the system to a bath the whole time. And actually the appearance of thermodynamic quantities is really just coming from the initial conditions. Here is a more general fluctuation theorem for an *isolated* system. It is also manifestly quantum mechanical. Here is the protocol to which it applies:

- Suppose at $t = 0$ we have some initial distribution of energy eigenstates, and we do a projective measurement of the energy. So we get outcome $E_n(0) \equiv E_n(\lambda_0)$ with probability $p_n(0)$. (We could take this to be $p_n(0) = e^{-\beta E_n(\lambda_0)} / Z(\lambda_0)$ but this is not essential.)
- Now we evolve the system unitarily, *i.e.* it is a closed quantum system, with the time-dependent Hamiltonian $H(t) = H_{\lambda_t}$, until time $t = \tau$. So the final state is $|\psi(\tau)\rangle = U |n(0)\rangle$, where U is the time-ordered exponential of $-\mathbf{i} \int dt H(t)$.
- Then at time τ we do a projective measurement of the H_{λ_τ} . We get the outcome $E_m(\lambda_\tau) \equiv E_m(\tau)$ with probability

$$p_{m \leftarrow n} = |\langle m(\tau) | U | n(0) \rangle|^2. \quad (4.22)$$

The work done is $W_{nm} = E_m(\tau) - E_n(0)$, and the total probability for this result of the forwards process is

$$P_F(n \rightarrow m) = p_n(0)p_{m \leftarrow n} = p_n(0) |\langle m(\tau) | U | n(0) \rangle|^2. \quad (4.23)$$

For the reverse process, we

- Start at $t = \tau$ and measure $H(\lambda_1)$ with distribution $p_m(\tau)$ (which could be $p_m(\tau) = e^{-\beta E_m(\tau)}/Z(\tau)$ but need not be).
- Evolve with $U^\dagger$ to go back in time to $t = 0$.
- Measure $H(\lambda_0)$.

The probability for obtaining outcome $|n(0)\rangle$ is

$$P_R(n \rightarrow m) = p_m(\tau)p_{n \leftarrow m}. \quad (4.24)$$

But here comes the key point:

$$p_{n \leftarrow m} = |\langle n(0) | U^\dagger | m(\tau) \rangle|^2 = |\langle m(\tau) | U | n(0) \rangle|^2 = p_{m \leftarrow n}. \quad (4.25)$$

This property is called microscopic reversibility.

Thus in the ratio, all the stuff about the dynamics completely cancels out:

$$\frac{P_F(n \rightarrow m)}{P_R(m \rightarrow n)} = \frac{p_n(0)}{p_m(\tau)}. \quad (4.26)$$

Sticking this into the forward work distributions as above gives

$$P_F(W) = \sum_{n,m} \delta(W - W_{nm}) P_F(n \rightarrow m) \quad (4.27)$$

$$= \sum_{n,m} \delta(W - W_{nm}) P_R(n \rightarrow m) \frac{p_n(0)}{p_m(\tau)}. \quad (4.28)$$

Now, if we had taken the thermal initial conditions, the ratio on the RHS of (4.26) would be

$$\frac{p_n(0)}{p_m(\tau)} = \frac{e^{-\beta E_n(0)}/Z(0)}{e^{-\beta E_m(\tau)}/Z(\tau)} = e^{-\beta \Delta F} e^{\beta W_{nm}} \quad (4.29)$$

in which case

$$P_F(W) = \sum_{n,m} \delta(W - W_{nm}) P_R(n \rightarrow m) e^{-\beta \Delta F} e^{\beta W_{nm}} \quad (4.30)$$

$$= e^{-\beta \Delta F} \sum_{n,m} \delta(W - W_{nm}) P_R(n \rightarrow m) e^{\beta W} \quad (4.31)$$

$$= e^{\beta(W - \Delta F)} \sum_{n,m} \delta(W - W_{nm}) P_R(n \rightarrow m) \quad (4.32)$$

$$= e^{\beta(W - \Delta F)} P_R(-W) \quad (4.33)$$

which means that $P_F(W)/P_R(-W)$ takes precisely the same form as the Crooks relation (4.2). The important thing about the initial condition was that it depended only on the energy difference W_{nm} and not any other property of the initial and final states. I'm not sure how uniquely this restricts us to the thermal distribution.

4.2 Fluctuation relations for NESS

Suppose we have an evolution rule with a non-equilibrium steady state. This means that it is stationary, but does not satisfy (global) detailed balance $W_{ij}P_j^{\text{eq}} = W_{ji}P_i^{\text{eq}}$ for all i, j . Global detailed balance would mean no current flow between any pair of states $J_{ij} = W_{ij}P_j^{\text{eq}} - W_{ji}P_i^{\text{eq}} = 0$, and no entropy production. The resulting state is a Gibbs state. This is not the subject of this course.

Let us suppose instead that the transition rates satisfy a condition called *local detailed balance* (LDB), which says

$$\frac{W_{ij}}{W_{ji}} = e^{\Delta S_{ij}^{\text{bath}}} \quad (4.34)$$

where $\Delta S_{ij}^{\text{bath}}$ is the entropy dumped into the bath when the system transitions from i to j , $\Delta S_{ij}^{\text{bath}} = \beta \Delta Q_{ij}$. A good situation to have in mind is a system coupled to *two* baths which are held at different temperatures (maybe they are at two ends of the system, so it will generate a heat current). Then some of the transitions W_{ij}^1 are mediated by the first bath at temperature T_1 , and others W_{kl}^2 are mediated by the second bath at temperature T_2 . Although each of these rates satisfies DB for some T , there is no global T for all the transitions²¹.

Under this assumption, the same logic as above leads to the following detailed fluctuation theorem, analogous to the Crooks relation. Suppose we start in the NESS. Then:

$$\frac{P(\Sigma_t)}{P(-\Sigma_t)} = e^{\Sigma_t} \quad (4.35)$$

where $P(\Sigma_t)$ is the distribution of entropy production over a time t by the evolution rule. Notice that there is no longer a distinction between forwards and backwards because everything is happening in the NESS. The Crooks relation is now a functional equation for the entropy distribution.

²¹But if you ask me detailed questions about the meaning of this condition I will have a hard time. For example, for any arbitrary set of transition rates W_{ij} , why can't we just *define* a matrix of numbers ΔS_{ij} by this equation (4.34)? Here there is some mysterious physics input about the nature of the bath that's causing the transitions. One possible answer is that for a general choice of W_{ij} the set of differences ΔS_{ij} will not define a state function. Thanks to Kshitij Gupta for this suggestion.

The proof is the same as above. For a given trajectory $\gamma = (z_0, z_1 \cdots z_M)$,

$$P[\gamma] = P_{\text{NESS}}(z_0)W_{z_1, z_0} \cdots W_{z_M, z_{M-1}} \quad (4.36)$$

and its reverse satisfies

$$P[\gamma^*] = P_{\text{NESS}}(z_M)W_{z_{M-1}, z_M} \cdots W_{z_0, z_1} \quad (4.37)$$

The ratios are then

$$\frac{P[\gamma]}{P[\gamma^*]} = \frac{P_{\text{NESS}}(z_0)}{P_{\text{NESS}}(z_M)} \prod_{k=0}^{M-1} \frac{W_{z_{k+1}, z_k}}{W_{z_k, z_{k+1}}} \quad (4.38)$$

$$= e^{\sum_{k=0}^{M-1} \log \frac{W_{z_{k+1}, z_k}}{W_{z_k, z_{k+1}}}} e^{\log P_{\text{NESS}}(z_0) - \log P_{\text{NESS}}(z_M)} \quad (4.39)$$

$$\stackrel{\text{LDB}}{=} e^{\sum_{k=0}^{M-1} \Delta S_{z_{k+1}, z_k}^{\text{bath}}} e^{\Delta S^{\text{system}}} \quad (4.40)$$

$$= e^{\Sigma[\gamma]} \quad (4.41)$$

Here we identify $S(z) = -\log P(z)$ as the microcanonical entropy of a given state, and $\Sigma[\gamma] = \Delta S^{\text{bath}} + \Delta S^{\text{system}}$. The same manipulation as above then implies (4.35). Integrating (4.35) over Σ_t gives the analog of Jarzynski:

$$\langle e^{-\Sigma_t} \rangle = \int d\Sigma P(\Sigma) e^{-\Sigma} = \int d\Sigma P(-\Sigma) = 1. \quad (4.42)$$

Fluctuation theorems for deterministic systems. The previous simple result (4.35) is a stochastic version of a more difficult statement called the Cohen-Gallivotti fluctuation theorem. This is the claim that the same statement is true about *deterministic* systems, where there is actually no stochasticity at all, and the entropy arises by some coarse-graining on phase space.

How is entropy generated in a deterministic system? Here's an example that makes it very vivid, related to the deterministic chaotic system called Baker's map. Our dynamical variable is a real number between 0 and 1. Represent the number in base b , say $b = 10$. At each time step, we multiply by b and remove the largest digit. I emphasize that this is completely deterministic. If you had a precise representation of the state at a given time, the state at the next time step is completely determined.

But now suppose you want to actually simulate this system. That means the number has to be stored in a computer. The default way to do this is as a floating point number, which will keep 18 digits of the real number. This is some inevitable coarse-graining in our description of the system. (Of course we could use higher precision numbers, but eventually it will have to be truncated.) That means that if the initial number is

$$0.314159265358979324 \quad (4.43)$$

then at the second time step, the number is

$$0.14159265358979324? . \quad (4.44)$$

The new last digit gets dredged out of the muck of some adjacent memory location in the computer. So we create $\Delta S = \log 10$ worth of entropy. Notice that in this example, this ΔS is independent of how precise is our representation, of many digits we manage to keep.

The idea is that any sufficiently interesting nonlinear dynamics is like this. The dynamics itself pulls information out of the muck of the shortest distances, so that any uncertainty about initial conditions eventually is amplified into effective macroscopic stochasticity. For more discussion of this beautiful subject, I recommend the book by David Ruelle called *Chance and Chaos*.

4.3 Non-equilibrium free energy and Maxwell's demon

[*e.g.* [Stopnitzky thesis](#); [Takara, Hasegawa, Driebe](#)] Let's briefly return to thinking about systems coupled to a heat bath. Consider the following functional of a probability distribution ρ (it works equally well as a functional of a density matrix, and indeed everything I say here is completely true quantum mechanically), specified by a choice of Hamiltonian H and the temperature T of the equilibrium heat bath to which the system is connected:

$$F[\rho] \equiv \langle H \rangle_\rho - TS(\rho). \quad (4.45)$$

Here $\langle H \rangle_\rho = \text{tr} \rho H \equiv \sum_n p_n E(n)$ is the energy expectation in the state ρ , and $S(\rho) = -\sum_n \rho_n \log \rho_n$ is the Shannon entropy (in the quantum case, it is the von Neumann entropy, *i.e.* replace ρ_n with the eigenvalues of the density matrix). Notice that if we plug in $\rho = \rho_{\text{eq}} \equiv e^{-\beta H}/Z(\beta)$ we get back the usual relation $F_{\text{eq}} (\equiv -T \log Z) = E - TS$, with $E = \langle H \rangle_\rho$ and $S = S(\rho_{\text{eq}})$ is the thermal entropy. I want to discuss two claims about this functional.

1. F is (uniquely) minimized by the Gibbs state. This is the canonical-ensemble version of the statement that the Boltzmann distribution maximizes the entropy at fixed energy (see our review of equilibrium thermodynamics on page 5 of these notes).
2. The usual second law of thermodynamics says that for equilibrium states $W \geq \Delta F_{\text{eq}}$, where $\Delta F_{\text{eq}} = F_{\text{eq}}(1) - F_{\text{eq}}(0)$ is the change of *equilibrium* free energies, and W is the work required to make this change happen.

Here is another generalization of this statement: The functional F bounds the work cost of transitions between non-equilibrium states. That is, suppose we have a system coupled to a thermal bath at temperature T . Initially it is in a state ρ_0 (either density matrix or probability distribution) and uncorrelated with the bath. Then the system evolves according to some time-dependent dynamics to the state ρ_1 . The claim is that

$$\langle W \rangle \geq F_T[\rho_1] - F_T[\rho_0] \quad (4.46)$$

where $\langle W \rangle$ is the average work done on the system during the process. (Actually we will replace this inequality with a more useful equality in the course of the proof.)

A nice way to understand the first statement is to observe that

$$F_T[\rho] = F_{\text{eq}}(T) + TD(\rho||\rho_{\text{eq}}) \quad (4.47)$$

where $D(\rho||\sigma) \equiv \text{tr}\rho \log \rho - \text{tr}\rho \log \sigma$ is the relative entropy (or KL divergence), which has many lovely virtues. In particular, it is non-negative, and only vanishes if $\rho = \sigma$, which immediately gives statement 1.

To see statement 2, let me make it a little more precise. We assume that the system plus bath evolve as a closed system, *i.e.* under unitary dynamics of a (time-dependent) Hamiltonian

$$H_{\text{SB}} = H_t + H_B + H_I(t) \quad (4.48)$$

where the bath Hamiltonian H_B is fixed the whole time and H_t acts only on the system. We assume that either (a) the interactions H_I are weak and we can ignore them, or (b) H_I is turned off at $t = 0$ and turned off again at $t = 1$. Either of these assumptions mean that in the first law of thermodynamics (*i.e.* energy conservation), we can ignore H_I :

$$\Delta E = \langle W \rangle + \langle Q \rangle \quad (4.49)$$

where $\Delta E = \text{tr}\rho_1 H_1 - \text{tr}\rho_0 H_0$.

We also assume that the initial state is a product state $\rho_{SB}(0) = \rho_0 \otimes \rho_B^{\text{eq}}$, with the bath in thermal equilibrium at its temperature T , and no correlations between system and bath. We'll talk below about what happens if we relax this important assumption.

Then here's a Lemma (Clausius' inequality):

$$\Delta S \geq \langle Q \rangle / T \quad (4.50)$$

where Q is the heat flow *into* the system from the bath during the process, and $\Delta S = S[\rho_1] - S[\rho_0]$ is the change in system entropy. Let's show that (4.50) implies (4.46), then we'll prove (4.50). Clausius says $-\langle Q \rangle \geq -T\Delta S$. Using this in (4.49) gives

$$\langle W \rangle \geq \Delta E - T\Delta S = \Delta F_{\text{neq}}. \quad (4.51)$$

Now let's see why (4.50) is true. The initial state is $\rho_{SB}(0) = \rho_0 \otimes \rho_B^{\text{eq}}$. If we wait, it becomes $\rho_{SB}(1) = U\rho_{SB}(0)U^\dagger$ under our unitary, probability conserving dynamics. The final state of the system is the marginal (or reduced density matrix) $\rho_1 = \text{tr}_B \rho_{SB}(1)$ and the final state of the bath is the marginal $\rho_B(1) \equiv \text{tr}_S \rho_{SB}(1)$.

Subadditivity of the entropy (either Shannon or von Neuman) says

$$S(\rho_1) + S(\rho_B(1)) \stackrel{\text{subadditivity}}{\geq} S(\rho_{SB}(1)) \quad (4.52)$$

$$\stackrel{\text{unitarity}}{=} S(\rho_{SB}(0)) = S(\rho_0 \otimes \rho_B^{\text{eq}}) \quad (4.53)$$

$$\stackrel{\text{product}}{=} S(\rho_0) + S(\rho_B^{\text{eq}}). \quad (4.54)$$

In the second step we used the fact that unitary dynamics does not change the entropy of the whole system. This says

$$\underbrace{S(\rho_1) - S(\rho_0)}_{=\Delta S_{\text{system}}} + \underbrace{S(\rho_B(1)) - S(\rho_B^{\text{eq}})}_{=\Delta S_{\text{bath}}^{\text{actual}}} \geq 0 \quad (4.55)$$

– the total entropy production of the parts is non-negative.

Now let's relate the actual entropy change of the bath $\Delta S_{\text{bath}}^{\text{actual}}$ to the heat exchanged $\langle Q \rangle$. Using claim 1,

$$F_T[\rho_B(1)] \geq F_T[\rho_B^{\text{eq}}] \quad (4.56)$$

no matter what the final state of the bath, $\rho_B(1)$, is. Thus,

$$T(S(\rho_B(1)) - S(\rho_B^{\text{eq}})) \leq \text{tr}(H_B \rho_B(1)) - \text{tr}(H_B \rho_B^{\text{eq}}) \quad (4.57)$$

$$\Delta S_{\text{bath}}^{\text{actual}} \leq \frac{1}{T}(\text{tr}(H_B \rho_B(1)) - \text{tr}(H_B \rho_B^{\text{eq}})) = -\langle Q \rangle / T. \quad (4.58)$$

Putting these together, we have

$$\Delta S_{\text{system}} \stackrel{(4.55)}{\geq} -\Delta S_{\text{bath}}^{\text{actual}} \stackrel{(4.58)}{\geq} \frac{\langle Q \rangle}{T}. \quad (4.59)$$

which implies the Clausius inequality, (4.50). ■

Our proof here actually gives us more information – it says what we gave up in each of the inequality steps of (4.59). In fact, it tells us the amount by which this generalized second law is not saturated. The statement above is that

$$W_{\text{diss}} \equiv \langle W \rangle - \Delta F_{\text{neq}} \geq 0. \quad (4.60)$$

But in fact,

$$W_{\text{diss}} = T D(\rho_B(1) || \rho_B^{\text{eq}}) + T I(S : B)_1 \quad (4.61)$$

where we've already met the relative entropy, and the second term is the final mutual information,

$$I(A : B)_\rho \equiv D(\rho_{AB} || \rho_A \otimes \rho_B) = (S_A + S_B - S_{AB})_\rho \stackrel{\text{subadditivity}}{\geq} 0 \quad (4.62)$$

between the system and the bath, $I(S : B)_1 \equiv I(S : B)_{\rho_{SB}(1)}$. The proof of this statement is really just the work we did above without using any of the inequalities (or just plug in the definition of the RHS and witness all the cancellations). The mutual information also has many lovely virtues. $I(A : B)_\rho$ bounds the correlations between A and B in the state ρ . And it is positive (that's subadditivity of the entropy). So (4.61) shows that W_{diss} is a sum of two independently positive terms. They have the

following interpretations. The work done on the system goes into two things: disturbing the equilibrium state of the bath, or creating correlations between the system and the bath. [\[End of Lecture 15\]](#)

Now: what if the initial state is not a product, i.e., what if there are initially correlations between the system and the bath? The last point above is a hint: if there are such correlations, *e.g.*, if the bath knows something about the initial state of the system, they can be used as a *resource* to avoid doing quite as much work. Such correlations are, for example, precisely the result of some part of the bath performing a measurement on the system. The way to quantify the deviation of $\rho_{SB}(0)$ from a product state is precisely the mutual information, as you can see from its definition (4.62). And in this more general case, the second law becomes

$$\langle W \rangle \geq \Delta F_{\text{neq}} - TI(S : B)_0 . \quad (4.63)$$

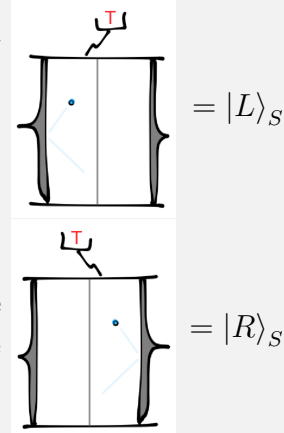
The proof is the same except in (4.54) we have to add in the mutual information in the initial state. To see that this bound, which says that less work is required, is tight, here is an example.

Szilard engine. Take S to be a 1-particle ideal gas in a box to which we may add a partition separating left (L) and right (R) parts of the box. Let B be a 1-(qu)bit memory plus a thermal bath. And take $H_S = 0$.

Assume that the initial state is the state immediately after B has *measured* whether the particle is in the left or the right half of the box, so

$$\rho_{SB}(0) = \frac{1}{2} (|L\rangle\langle L|_S \otimes |L\rangle\langle L|_B + |R\rangle\langle R|_S \otimes |R\rangle\langle R|_B) \quad (4.64)$$

where $|R/L\rangle_S$ are depicted at right, and $|R/L\rangle_B$ are the states of the bath *knowing* that the particle is in the right/left half of the box etc.



The initial entropies are: the particle is equally likely to be in either half, so $S(\rho_0) = \log 2$, and similarly $S(\rho_B) = \log 2$. But there is nontrivial mutual information

$$I(S : B)_0 = S(\rho_0) + S(\rho_B) - S(\rho_{SB}(0)) = \log 2 + \log 2 - \log 2 = \log 2. \quad (4.65)$$

This is an information resource that the bath (Maxwell's demon) can use to extract work. The evolution is the following: if the state is L , the demon attaches a piston to the right side of the partition and lets the gas do work by expanding

against it. If the state is R , the demon attaches the piston to the left. The work done by the gas is

$$W_{\text{out}} = \int F dx = \int P A dx = \int P dV \stackrel{\text{ideal}}{=} \int_{V/2}^V \frac{dV}{V} NT|_{N=1} = T \log 2, \quad (4.66)$$

so the work done *on* the gas (which is what we've been talking about) is $\langle W \rangle = -T \log 2$. At the end of this process, we (and the bath) know nothing about the state of the particle, it could be anywhere, so $\rho_1 = \mathbb{1}/2$, and $F_T[\rho_1] = E[\rho_1] - TS[\rho_1] = 0 - T \log 2$ (which is just the same as for the initial state). Thus

$$-T \log 2 = \langle W \rangle \geq \Delta F_{\text{neq}} - TI(S : B)_0 = 0 - T \log 2 \quad (4.67)$$

and our improved inequality is saturated.

Notice that our path from state 0 to state 1 is not a cycle: at the end of this story, the demon no longer knows the state of the particle. That information has been burnt as fuel.

I want to record the general statement in the form of an equality, with no assumptions about the initial correlations between system and bath, nor about whether the bath is initially in equilibrium. It is

$$W_{\text{diss}} \equiv \langle W \rangle - \Delta F_{\text{neq}} = TD(\rho_{SB}(t) || \rho_S(t) \otimes \rho_B^{\text{eq}})|_{t=0}^1. \quad (4.68)$$

To see this, first write the LHS as

$$W_{\text{diss}} \equiv \langle W \rangle - \Delta F_{\text{neq}} = \langle W \rangle - (\Delta E - T \Delta S^{\text{system}}) = -\langle Q \rangle + T \Delta S^{\text{system}} \quad (4.69)$$

where we used the first law in the form $\Delta E = \langle W \rangle + \langle Q \rangle$. Now start from the fact that

$$F_T[\rho_B(1)] - F_T[\rho_B^{\text{eq}}] = TD(\rho_B(1) || \rho_B^{\text{eq}}) \quad (4.70)$$

and write the LHS as

$$\text{tr} H_B \rho_B(1) - \text{tr} H_B \rho_B^{\text{eq}} - T(S_B(1) - S_B(\text{eq})) = -\langle Q \rangle + \text{tr} H_B(\rho_B(0) - \rho_B^{\text{eq}}) - T(S_B(1) - S_B(\text{eq}))$$

where we used $-\langle Q \rangle = \text{tr} H_B \rho_B(1) - \text{tr} H_B \rho_B(0)$. So we can eliminate

$$-\langle Q \rangle = T(S_B(1) - S_B(\text{eq}) + D(\rho_B(1) || \rho_B^{\text{eq}})) - (E_B(0) - E_B^{\text{eq}}). \quad (4.71)$$

Plugging this into W_{diss} in (4.69) gives

$$\begin{aligned}
W_{\text{diss}}/T &= S(\rho_1) + S_B(1) - (S(\rho_0) + S_B(\text{eq})) + D(\rho_B(1)||\rho_B^{\text{eq}}) - (E_B(0) - E_B^{\text{eq}})/T \\
&= S(\rho_{SB}(1)) + I(S : B)_1 - \left(\underbrace{S(\rho_0) + S_B(0)}_{I(S:B)_0 + S(\rho_{SB}(0))} - S_B(0) + S_B(\text{eq}) \right) + D(\rho_B(1)||\rho_B^{\text{eq}}) - (E_B(0) - E_B^{\text{eq}})/T \\
&= \underbrace{S(\rho_{SB}(1)) - S(\rho_{SB}(0))}_{=0} + I(S : B)|_0^1 + D(\rho_B(1)||\rho_B^{\text{eq}}) - (S_B(\text{eq}) - S_B(0)) - (E_B(0) - E_B^{\text{eq}})/T
\end{aligned}$$

Now we observe from the definition of relative entropy that

$$D(\rho_B(0)||\rho_B^{\text{eq}}) = \text{tr} \rho_B(0) \log \rho_B(0) - \text{tr} \rho_B(0) \log \rho_B^{\text{eq}} \quad (4.72)$$

$$= -S_B(0) + E_B(0)/T + \log Z(T) \quad (4.73)$$

$$= -S_B(0) + E_B(0)/T + S_B(\text{eq}) - \beta E_B^{\text{eq}} \quad (4.74)$$

which is precisely the last four terms in our expression. Therefore

$$W_{\text{diss}}/T = (I(S : B) + D(\rho_B(t)||\rho_B^{\text{eq}}))|_{t=0}^1. \quad (4.75)$$

Furthermore, for each time we can rewrite

$$I(S : B) + D(\rho_B(t)||\rho_B^{\text{eq}}) = D(\rho_{SB}(t)||\rho_S(t) \otimes \rho_B^{\text{eq}}). \quad (4.76)$$

To see this just use the definitions and the fact that

$$\log \rho_S \otimes \rho_B = \log \rho_S \otimes \mathbb{1}_B + \mathbb{1}_S \otimes \log \rho_B. \quad (4.77)$$

Finally, I want to mention that our proof of this generalization of the second law only used some relatively weak properties of the entropy, namely non-negativity of the relative entropy, and subadditivity. There is a much stronger statement (at least quantum mechanically) called *strong* subadditivity, which is always true. One form of strong subadditivity is the monotonicity of the relative entropy under arbitrary channels:

$$D(\rho||\sigma) \geq D(\mathcal{E}(\rho)||\mathcal{E}(\sigma)). \quad (4.78)$$

Here $\mathcal{E}$ is essentially anything we can do to the system. For example, suppose $\mathcal{E}$ describes some dynamics that results in the equilibrium ρ_{eq} . Then $\mathcal{E}(\rho_{\text{eq}}) = \rho_{\text{eq}}$, and so

$$D(\rho||\rho_{\text{eq}}) \geq D(\mathcal{E}(\rho)||\mathcal{E}(\rho_{\text{eq}})) = D(\mathcal{E}(\rho)||\rho_{\text{eq}}) \quad (4.79)$$

is monotonic under the evolution. This is a version of the Boltzmann H -theorem.

I had hoped that we could use strong subadditivity to prove a stronger or more general version of our work inequality. One version is: consider the case where the system Hamiltonian $H(t) = H$ is fixed, so no work is done, $W = 0$. Then our channel takes ρ_0 to $\mathcal{E}(\rho_0) = \rho_1$ and the inequality

$$D(\rho_0 || \rho_{\text{eq}}) \geq D(\rho_1 || \rho_{\text{eq}}) \tag{4.80}$$

says, using (4.47),

$$F_T[\rho_0] - F_{\text{eq}} \geq F_T[\rho_1] - F_{\text{eq}} \implies F_T[\rho_0] \geq F_T[\rho_1] \tag{4.81}$$

or $0 \geq \Delta F_{\text{neq}}$, which is this special case of the result we proved before, proved using a stronger tool.

Part II

Stochastic Evolution of Fields

[Kardar *Fields*, §9.2, Cugliandolo §6, Toner, *Flocking*.]

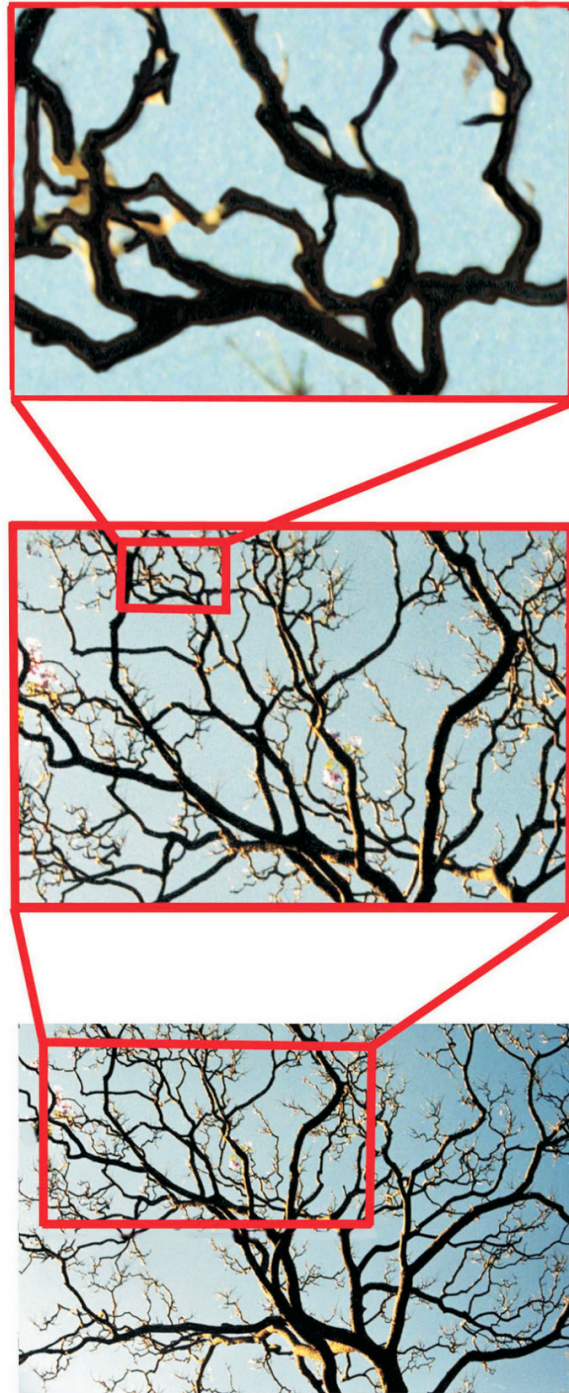
For the rest of the course I want to talk about field theories – systems made of many degrees of freedom spread over space and interacting with each other by local rules.

We began the course with a discussion of the RG idea, in the context of random walks of a single particle. This idea becomes only more important when we think about field theories. The key to understanding such systems (to the extent we do) is to identify the right degrees of freedom.

Which degrees of freedom must we keep in the long-wavelength description? The key ingredient is that at long wavelengths, derivatives are small. We want to leave out everything we can, but some degrees of freedom have slow dynamics for various reasons. These include:

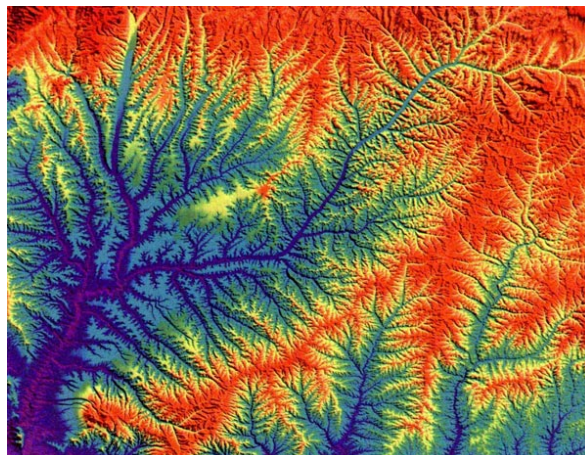
1. Conserved quantities. They can't disappear, they can only move around, which generally happens by diffusion. More quantitatively, $\partial_t \rho = -\vec{\nabla} \cdot \vec{J}$, you see that the RHS necessarily involves derivatives which are small.
2. Order parameters for spontaneous breaking of continuous symmetries (Goldstone modes). Such a degree of freedom shifts under the symmetry, like $h(x) \rightarrow h(x) + \epsilon$ and this means that it only appears in the dynamics via derivatives, $\partial_t h \sim \vec{\nabla}^n h + \dots$, which are small by assumption.
3. Order parameters that might become critical. We could be near a phase transition where some external hand is tuning a knob to make the correlation length or correlation time infinite. The idea here is that $\partial_t h \sim rh + \dots$, but someone has come and tuned r to zero so that only the terms involving derivatives or higher powers of h appear.
4. Other slow modes that might be present for some reason I haven't thought of. I'll discuss some such reasons I haven't thought of below.

Slow modes are generally associated with power laws – the absence of some intrinsic length or time scale, and hence some scale invariance. Such power laws actually occur in Nature: the distribution of magnitudes of earthquakes (Gutenberg-Richter), of river drainage basis of a given area (Horton), of clouds of a given cross-sectional area, of traffic jams of given duration, of branches of a tree of a given diameter (Leonardo’s area-preserving rule for transport of water and nutrients), of sizes of avalanches of neuronal activity in the brain, of frequencies of word usages in natural languages (Zipf’s law), all of these empirically have some self-similarity. Some of these power laws have striking visual consequences, and Mandelbrot wrote a whole book about it, called *The Fractal Geometry of Nature*. The fact that there seems to be some kind of critical phenomena in these examples is something to be understood because there doesn’t seem to be any giant hand tuning the parameters (maybe you could argue otherwise in the biological examples).



(from Taylor, R., Newell, B., Spehar, B., Clifford, C. (2005). *Fractals: A Resonance between Art and Nature*. In: Emmer, M. (eds) *Mathematics and Culture II*. Springer, Berlin, Heidelberg.)

Note that I am not claiming that all of the above examples must be understood in the same way. We will see below a number of possible explanations, from all four of the categories above.



(from Bruce D Malamud, [Tails of natural hazards](#) 2004 Phys. World 17 (8) 25.) This is a false color image of a drainage network in the Loess Plateau, ShānXī Province, China. Color is elevation. The actual size is 300km across, but you really cannot tell from the image itself.

One strategy for trying to answer this kind of question is associated in my mind with the names Landau and Ginzburg, what Toner calls the “hydrodynamic approach”, or what is called effective field theory (EFT). The idea is that we guess the right dofs and symmetries, and then add all terms consistent with the symmetries and see what happens.

Sample physical applications I hope to talk about include:

- diffusion of particle number and other conserved charges
- actual hydrodynamics of fluids!
- spin relaxation dynamics
- coarsening
- absorbing-state (birth/extinction) transitions
- surface or interface growth
- flocking

5 Near-equilibrium transport, very briefly

5.1 Boltzmann equation

[LP §1] Consider a fluid made of N particles, with their phase space locations at time t labelled by $\{\vec{r}_i(t), \vec{v}_i(t)\}_{i=1}^N$, with $N \sim 10^{25}$. Let us briefly pretend that the world is classical and that it is a good idea to keep track of the positions and velocities of so many particles. It evolves according to

$$\dot{r}_i(t) = v_i(t), \quad \dot{v}_i(t) = F_i(t)/m \quad (5.1)$$

where F_i is the sum of external forces on particle i , the force from the walls of the container, and the forces exerted by the other particles. A good step towards coarse-graining this information is to introduce the distribution function

$$f(r, v, t) \equiv \sum_{i=1}^N \delta^d(r - r_i(t)) \delta^d(v - v_i(t)). \quad (5.2)$$

So

$$f(r, v, t) d^d r d^d v = \begin{array}{l} \text{the number of particles in the phase space box} \\ \text{with a corner at } r, v \text{ and a corner at } r + dr, v + dv \end{array}. \quad (5.3)$$

So $n(r, t) = \int d^d v f(r, v, t)$ is the particle density in space, and $N = \int d^d r d^d v f(r, v, t)$ is the total number, which we already fixed. Given some observable $A = A(r, v)$, let

$$\bar{A}(t) = \frac{1}{N} \int d^d r \int d^d v A(r, v) f(r, v, t). \quad (5.4)$$

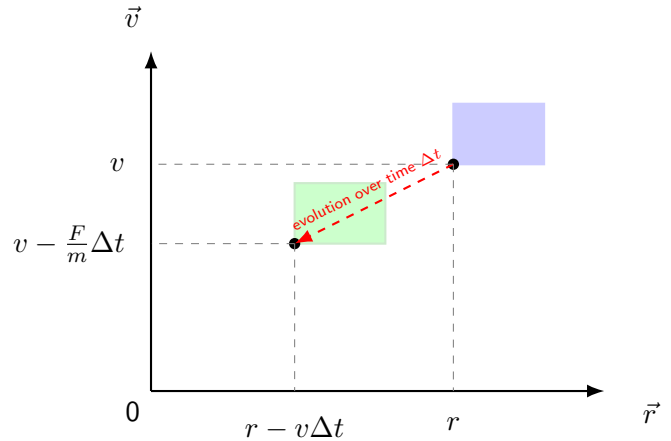
So for example, the energy per particle is

$$e(t) = \frac{1}{N} \int d^d r \int d^d v \frac{1}{2} m v^2 f(r, v, t). \quad (5.5)$$

Now we want to think about the time evolution of the distribution function:

$$\partial_t f(r, v, t) = \partial_t f|_s + \partial_t f|_c \quad (5.6)$$

where the two terms respectively represent the ‘streaming’ or advection of the particles, including any external forces, and the collisions between particles and particles or between particles and walls.



Let's think about the evolution without collisions first. As depicted above, all particles

in the box at $(r - vdt, v - F/mdt)$ at time $t - dt$ end up in the box at (r, v) at time t :

$$f(r, v, t + dt) = f(r - vdt, v - Fdt/m, t) \quad (5.7)$$

$$= f(r, v, t) - dt \left(\vec{v} \cdot \vec{\nabla}_r f(r, v, t) + \frac{\vec{F}}{m} \cdot \vec{\nabla}_v \cdot f(r, v, t) \right) + \mathcal{O}(dt^2). \quad (5.8)$$

We can summarize this by saying that the *total* derivative of the distribution function

$$\frac{df}{dt} = \partial_t f + \dot{x} \cdot \vec{\nabla}_r f + \dot{v} \cdot \vec{\nabla}_v f = \partial_t f + \vec{v} \cdot \vec{\nabla}_r f + \vec{F}/m \cdot \vec{\nabla}_v f \quad (5.9)$$

would be zero in the absence of collisions.

With collisions, some of our friends in the initial box $(r - vdt, v - F/mdt)$ at time $t - dt$ will collide with something and fail to reach their destination (r, v) at time t . Furthermore, some particles that were not headed for (r, v) at time t will indeed end up there. To represent these two processes, we can write

$$\partial_t f|_c d^d r d^d v dt = (\bar{R} - R) d^d r d^d v dt \quad (5.10)$$

where $\bar{R}$ is the gain and R is the loss. This is the number of collisions in the time interval dt resulting in one of these two events.

Finding $R, \bar{R}$ is not so easy, and we will make some simplifying assumptions. First, we will assume collisions are rare (either because our fluid is dilute enough or because we took small enough dr, dv, dt) and therefore neglect multiple collisions. We'll also assume local (short-ranged) interactions. A binary collision is described by a process $(v, u) \rightarrow (v', u')$. Think of the second particle as the target on which the first is scattering. If the process is elastic, which we'll assume, then $|v' - u'| = |v - u|$, and it is parameterized by a scattering angle Ω , let's say this is the angle between v and v' . Then in a time dt *one* test particle of velocity v , in the (Galilean) frame of its target with velocity u sweeps out a cylinder of length $|u - v|dt$ and with cross-sectional area $d\sigma(\Omega)$. Here $d\sigma(\Omega)$ is the scattering cross section for these particles to collide and produce particles with velocities v', u' with scattering angle Ω . It is the effective area of the beam that produces this particular collision. The number of target particles in this cylinder is

$$f(u) d^d u \cdot |u - v| dt d\sigma d^d r. \quad (5.11)$$

This was for *one* test particle, and so the total number of collisions in time dt (into the angle Ω and for the given target velocity) is

$$f(v) d^d v \cdot f(u) d^d u \cdot |u - v| dt d\sigma d^d r. \quad (5.12)$$

The total number of collisions is obtained by integrating this over Ω and over the target velocity u . Dividing this by $d^d v d^d r dt$ gives

$$R = f(r, v, t) \int d\Omega \int d^d u \frac{d\sigma}{d\Omega} (|v - u|) f(r, u, t) |v - u|. \quad (5.13)$$

Similarly, using microscopic reversibility ($d\sigma(u, v \rightarrow u', v') = d\sigma(u', v' \rightarrow u, v)$), and assuming conservative central forces, so that $d^d u d^d v = d^d u' d^d v'$,

$$\bar{R} = \int d\Omega d^d u \frac{d\sigma}{d\Omega} (|v - u|) f(r, v', t) f(r, u', t) |v - u| \quad (5.14)$$

where here (v', u') are determined from (u, v) and the scattering angle Ω .

We're going to need to distinguish the velocity labels on a bunch of $f(r, v, t)$ at the same r, t , so let's write $f(r, v, t) \equiv f_v$. Putting this together,

$$\partial_t f_v = - \left(\vec{v} \cdot \vec{\nabla}_r + \frac{\vec{F}(r)}{m} \cdot \vec{\nabla}_v \right) f_v + \int d\Omega \int d^d u \frac{d\sigma}{d\Omega} (|v - u|) (f_{v'} f_{u'} - f_v f_u) |v - u|. \quad (5.15)$$

This is the Boltzmann equation. I should emphasize that this equation is not always true. In particular, it involved the very important assumption that the collision rate is proportional to the *product* of the single-particle distribution functions, that is, that the 2-particle distribution can be approximated as factorized.

Claim: the entropy $S[f] \equiv - \int d^d r d^d v f_v \log f_v$

1. is non-decreasing under the evolution by the Boltzmann equation, $\partial_t S \geq 0$,
2. is maximized by the Maxwell-Boltzmann distribution

$$\underline{f}_v \propto e^{-\frac{1}{2} m \beta (v - v_0)^2}, \quad (5.16)$$

for some β, v_0 , in the absence of external forces.

Note that I call $S[f]$ an entropy because it takes the form of the Shannon entropy of a probability distribution, though f is not quite a probability distribution. The difference is just some normalization of f , which produces an additive constant in S .

Here's the proof of these statements (which is informatively called the Boltzmann H -theorem, because he called the entropy H). To see the monotonicity of the entropy, take ∂_t of $S[f]$ to get

$$\partial_t S[f] = -\frac{1}{4} \int d\Omega \int d^d v \int d^d u \frac{d\sigma}{d\Omega} (|v - u|) |v - u| \log \frac{f_u f_v}{f_{u'} f_{v'}} (f_{u'} f_{v'} - f_u f_v) \quad (5.17)$$

and use

$$\log(x/y)(y - x) \leq 0 \quad (5.18)$$

(with equality iff $y = x$), and the fact that $f_v \geq 0$. So you see that $\partial_t S = 0$ iff $f_{v'} f_{u'} = f_v f_u$ for all collisions.

To see the second statement, first observe that $n = \int d^d v f_v$, $e = \int d^d v \frac{1}{2} m v^2 f_v$, and $\vec{p} = \int d^d v m \vec{v} f_v$ are conserved by the Boltzmann equation. Thus, in looking for the maximum, we can include Lagrange multipliers for each of them and look for solutions of (where $\delta \equiv \frac{\delta}{\delta f_v}$)

$$0 \stackrel{!}{=} -\delta S + \alpha \delta n + \beta \delta e + \vec{\gamma} \cdot \delta \vec{p} \quad (5.19)$$

$$= (1 + \log f_v) + \alpha + \beta \frac{1}{2} m v^2 + m \vec{\gamma} \cdot \vec{v}. \quad (5.20)$$

The solution to this equation is (5.16) with $\vec{v}_0 = -\vec{\gamma}/\beta$ (with the normalization fixed). So the MB distribution is an extremum; let's see that it is a maximum: $S[f_v] \leq S[\underline{f}_v]$ for all f_v . To see this, note that

$$S[\underline{f}_v] = \int d^d v \underline{f}_v \left(1 + \alpha + \frac{1}{2} m \beta v^2 + m \vec{v} \cdot \vec{\gamma} \right) = (1 + \alpha) n + \beta e + \vec{\gamma} \cdot \vec{p} = \int d^d v f_v \left(1 + \alpha + \frac{1}{2} m \beta v^2 + m \vec{v} \cdot \vec{\gamma} \right) \quad (5.21)$$

where on the RHS we can use *any* f_v . Then

$$S[\underline{f}_v] - S[f_v] = + \int d^d v \left(f_v \log f_v - f_v \log \underline{f}_v \right) \quad (5.22)$$

$$= \int d^d v \left(f_v \log f_v - f_v \log \underline{f}_v - f_v + \underline{f}_v \right) \quad (5.23)$$

$$= D(p, \underline{p}) \geq 0 \quad (5.24)$$

is the relative entropy or KL divergence of two probability distributions, where $p \equiv f/n$. This is always positive by the inequality $x \log x - x \log y - x + y \geq 0$ for $x \geq 0, y > 0$. Boltzmann didn't call it by either of these terms in his paper from 1872.

Actually, with external forces, the result of maximizing S can depend on r, t . It takes the form

$$\underline{f}(r, v, t) = n(r, t) \left(\frac{m \beta(r, t)}{2\pi} \right)^{d/2} e^{-\frac{m \beta}{2} (v - v_0(r, t))^2} \quad (5.25)$$

which is locally the Maxwell-Boltzmann distribution.

5.2 Hydrodynamics

(5.25) is *local* thermodynamic equilibrium, or *hydrodynamics*. Notice the key role played by the conserved quantities. Another way to think about hydrodynamics is: we are trying to coarse-grain away all the details and only keep track of the variables that are non-trivial at late times and on long-wavelengths. There are several reasons why a variable might survive this process, and the one we've discovered so far is because it represents a conserved quantity. A conserved quantity can't simply disappear, it can

only go somewhere else. So if we start in a state where there is an overdensity of a conserved quantity, necessarily it will take a time long compared to the microscopic time scales (like the relaxation time τ) for it to go away – it has to diffuse. The job of hydrodynamics is to follow around such overdensities.

There is an approach to hydrodynamics that makes the answer inevitable. It is the analog of Landau-Ginzburg theory, but historically it came first. The strategy is: (1) guess the degrees of freedom (2) specify the symmetries (3) write down all possible terms involving the dofs consistent with the symmetries.

What we've just shown from the Boltzmann equation tells us that the dofs are: the local fluid velocity $\vec{v}_0(r, t)$, the local temperature $T(r, t)$ and the local density $n(r, t)$. The symmetries are translation symmetry, and (Galilean) boosts, which are broken by the equilibrium configuration, which picks a rest frame.

If instead we had been studying relativistic particles, like a relativistic QFT at finite temperature, the resulting hydrodynamics is called relativistic hydrodynamics because of the underlying Lorentz invariance, and we can do all of this a bit more slickly. (A good reference for this subject is [these lectures by Pavel Kovtun](#).) Then the conservation of energy and momentum and conservation of particle number are compactly written as

$$\partial_\mu T^{\mu\nu} = 0, \quad \partial_\mu J^\mu = 0. \quad (5.26)$$

The remaining question is how to write the stress tensor $T_{\mu\nu}$ and particle current J_μ in terms of the dofs, u_μ, T, μ (now the velocity is a 4-vector, and I'm writing things in terms of a chemical potential rather than a density). These are called 'constitutive relations', and the idea is to write them in a derivative expansion. A general parametrization of the stress tensor and current in terms of the 4-velocity u^μ is

$$T^{\mu\nu} = \mathcal{E}u^\mu u^\nu + \mathcal{P}\Delta^{\mu\nu} + q^\mu u^\nu + q^\nu u^\mu + t^{\mu\nu}, \quad J^\mu = \mathcal{N}u^\mu + j^\mu \quad (5.27)$$

where $\Delta^{\mu\nu} \equiv \eta^{\mu\nu} + u^\mu u^\nu$ annihilates vectors parallel to the timelike vector u^μ ($u^\mu u_\mu = -1$), and q^μ, j^μ are transverse to u^μ and $t^{\mu\nu}$ is transverse and traceless. The coefficients $\mathcal{E}, \mathcal{P}, q^\mu, t^{\mu\nu}, \mathcal{N}, j^\mu$ have to be written in terms of the hydrodynamic variables u_μ, T, μ in a derivative expansion. The zeroth order terms are just:

$$T_{(0)}^{\mu\nu} = \epsilon u^\mu u^\nu + p\Delta^{\mu\nu}, \quad J^\mu = n u^\mu. \quad (5.28)$$

Here $\mathcal{E} = \epsilon$ is the local energy density, $\mathcal{P} = p$ is the local pressure, $\mathcal{N} = n$ is the local charge density. The equilibrium equation of state tells us $p = p(T, \mu)$ from which ordinary thermodynamic relations tell us the entropy density $s = \partial_T p$, the particle density $n = \partial_\mu p$ and energy density, $\epsilon = -p + Ts + \mu n$. Using the conservation laws (5.26) and $dp = sdT + nd\mu$, you can show that these relations imply that

$$\partial_\mu (s u^\mu) = 0 \quad (5.29)$$

– there is no entropy production in this zeroth order ‘ideal’ hydrodynamics. (Note that the idealness of ideal hydrodynamics is distinct from that of the ideal gas – the ideal gas has no hydrodynamics. Ideal hydrodynamics requires the particles to collide with each other and produce local thermodynamic equilibrium. The thing that’s not allowed to happen in ideal hydrodynamics is dissipation.) [\[End of Lecture 16\]](#)

First-order hydrodynamics means that we keep terms with up to one derivative of the hydrodynamic variables in the constitutive relations. Here we encounter an important annoyance – the variables $T(x), u^\mu(x), \mu(x)$ are well-defined in equilibrium, but there is no operator whose expectation value in some non-equilibrium state defines each of these quantities (unlike $T^{\mu\nu}$ and J^μ themselves). This means that we can redefine

$$T(x) \rightarrow T'(x) = T(x) + \delta T(x) \quad (5.30)$$

$$\mu(x) \rightarrow \mu'(x) = \mu(x) + \delta\mu(x) \quad (5.31)$$

$$u^\mu(x) \rightarrow u'^\mu(x) = u^\mu(x) + \delta u^\mu(x) \quad (5.32)$$

where the δ s are first-order in derivatives. $u^2 = -1$ requires $u_\mu \delta u^\mu = 0$. The resulting change in the constitutive relations is

$$\delta\mathcal{E} = 0, \delta P = 0, \delta N = 0, \delta q_\mu = -(\mathcal{E} + \mathcal{P})\delta u_\mu, \delta j_\mu = -\mathcal{N}\delta u_\mu, \delta t_{\mu\nu} = 0. \quad (5.33)$$

So we can choose δu_μ to set $j_\mu = 0$ – no charge flow in the local rest frame of the fluid (this is called Eckhart frame). Or we can choose δu_μ so that $q_\mu = 0$ – no energy flow in the local rest frame of the fluid (this is called Landau frame). So far this specifies $\delta\mu$; the freedom in $\delta T, \delta\mu$ can be used to define T and μ so that $\mathcal{E} = \epsilon, \mathcal{N} = n$.

Let’s choose Landau frame. This means we still have to express $\mathcal{P}, t^{\mu\nu}, j^\mu$ in terms of the hydro variables u, T, μ . This becomes a game of building Lorentz scalars, vectors and tensors out of the ingredients u^μ, T, μ and their derivatives ∂^μ . The conservation laws $u_\mu \partial_\nu T^{\mu\nu} = 0, \partial_\mu J^\mu = 0$ get rid of two of the three possible scalars, so

$$\mathcal{P} = p = -\zeta \partial_\lambda u^\lambda + \mathcal{O}(\partial^2) \quad (5.34)$$

where ζ is a new coefficient called the bulk viscosity. Similarly,

$$j^\mu = -\sigma T \Delta^{\mu\nu} \partial_\nu (\mu/T) + \chi_T \Delta^{\mu\nu} \partial_\nu T + \mathcal{O}(\partial^2) \quad (5.35)$$

where σ is the charge conductivity and χ_T is actually zero. Finally

$$t^{\mu\nu} = -\eta \sigma^{\mu\nu} + \mathcal{O}(\partial^2) \quad (5.36)$$

where $\sigma^{\mu\nu} \equiv \Delta^{\mu\alpha} \Delta^{\nu\beta} (\partial_\alpha u_\beta + \partial_\beta u_\alpha - \frac{2}{d} \eta_{\alpha\beta} \partial_\mu u^\mu)$ is the strain tensor, and η is called the shear viscosity. Writing $\partial_\mu J^\mu$ in terms of u^μ, T, μ gives charge diffusion, while

$\partial_\mu T^{\mu\nu}$ in terms of u^μ, T, μ gives the Navier-Stokes equations. It predicts waves and their dispersion and damping rates in terms of just these four numbers.

The shortcoming of this approach is that it doesn't tell us the values of the four transport coefficients $\eta, \sigma, \zeta, \chi_T$ (in other context, these coefficients in the derivative expansion unspecified by the low-energy physics are often called Wilson coefficients). There are several ways to think about determining them. One is to derive them using the Boltzmann equation, as I'll outline below. Another is to relate them to linear response theory of the currents $T_{\mu\nu}, j_\mu$, which we'll do next.

Unlike ideal hydrodynamics, first-order hydrodynamics produces entropy. There exists a current

$$S^\mu = su^\mu + (\text{gradient terms})^\mu \quad (5.37)$$

(where the gradient terms are linear in derivatives and vanish in equilibrium) with $\partial_\mu S^\mu \geq 0$ expressing the second law of thermodynamics. The covariant version of the Gibbs-Duhem relation (the usual version is $Ts = p + \epsilon - \mu$ and is obtained by setting $u^\mu = (1, \vec{0})^\mu$) says

$$S^\mu = \frac{1}{T} (pu^\mu - T^{\mu\nu}u_\nu - \mu J^\mu) = \left(s + \frac{\mathcal{E} - \epsilon}{T} - \frac{\mu}{T} (\mathcal{N} - n) \right) u^\mu + q^\mu/T - \mu j^\mu/T. \quad (5.38)$$

The positivity of entropy production leads to the constraints:

$$\eta \geq 0, \zeta \geq 0, \sigma \geq 0, \chi_T = 0. \quad (5.39)$$

This is an interesting example of using a macroscopic property of an EFT to constrain microphysics in the way people would also like to do for *e.g.* theories of physics beyond the Standard Model coupled to gravity.

Linear response. [The classic reference for this subject is the book by Dieter Forster, *Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions*, or the paper by [Kadanoff and Martin](#). The paper by Kovtun on relativistic hydrodynamics linked above has a nice summary which I am following.] Let's talk about extracting transport coefficients from microscopic information. I'm not going to do this for the full set of hydrodynamic transport coefficients, but for the simple case of diffusion.

Let's play the Landau game. Suppose the only degree of freedom is a density $n(x)$. And the symmetries are: rotation invariance, plus a $U(1)$ symmetry meaning that the density is conserved:

$$\partial_t n + \vec{\nabla} \cdot \vec{J} = 0. \quad (5.40)$$

Then we need to find a constitutive relation expressing the current $\vec{J}$ in terms of our degree of freedom n , as a derivative expansion. Since n is a scalar (under rotations), in order to make a vector out of it we need to take a derivative. Therefore we have the

constitutive relation $\vec{J} = -D\vec{\nabla}n$ (“Fick’s law”) in terms of the ‘transport coefficient’ D , the diffusion constant (plus terms with more derivatives that we will ignore). Plugging into the continuity equation, this leads to the diffusion equation $\partial_t n - D\vec{\nabla}^2 n = 0$.

In Fourier space, the solution of the diffusion equation is

$$n(z, k) = \frac{n_0(k)}{-\mathbf{i}z + Dk^2} \quad (5.41)$$

with the small disclaimer that we should keep the imaginary part of z positive so that $n(z, k) = \int_0^\infty dt e^{\mathbf{i}zt} n(t, k)$ converges. It will be useful to write the solution in terms of an initial chemical potential using the static susceptibility $\chi \equiv \partial_\mu n|_{\mu=0}$ and assuming that our n is small:

$$n(z, k) = \frac{\chi\mu_0(k)}{-\mathbf{i}z + Dk^2}. \quad (5.42)$$

Now we make contact with some microscopic question. Imagine we adiabatically turn on a gentle source for the density, and then switch it off at $t = 0$:

$$\mu(t, x) = e^{\epsilon t} \mu(x) \theta(-t). \quad (5.43)$$

We’ll treat the question fully quantum-mechanically. What we are doing is adding a time-dependent perturbation to our Hamiltonian $\delta H(t)$, and asking about the linear response of some observable A :

$$\delta \langle A(t, x) \rangle = -\mathbf{i} \int_{-\infty}^t dt' \langle [A(t, x), \delta H(t')] \rangle \quad (5.44)$$

ignoring terms of order δH^2 . In the problem specified above, $\delta H = -\int d^d x \mu(t, x) n(t, x)$ and we are asking about the change in the density $A = n$:

$$\langle n(t, x) \rangle = -\int_{-\infty}^0 dt' e^{\epsilon t'} \int d^d x' \mu(x') G_{nn}^R(t - t', x - x') \quad (5.45)$$

where

$$G_{AB}^R(t - t', x - x') \equiv -\theta(t - t') \langle [A(t, x), B(t', x')] \rangle \quad (5.46)$$

is the retarded Green’s function.

We want to compare this expression with the Fourier transform (5.42). I’m going to put this calculation in a box. First Fourier transform in space to get

$$\langle n(t, k) \rangle = -\int_{-\infty}^0 dt' e^{\epsilon t'} \mu_0(k) G_{nn}^R(t - t', k). \quad (5.47)$$

Notice that at $t = 0$ this equation says

$$\langle n(t=0, k) \rangle = -\mu_0(k) \int_0^\infty dt' e^{-\epsilon t'} G_{nn}^R(t', k) = -\mu_0(k) G_{nn}^R(z=0, k) \quad (5.48)$$

so we can identify $\chi = -G_{nn}^R(z=0, k)$. Now

$$G_{nn}^R(t-t', k) = \int_{-\infty}^\infty d\omega G_{nn}^R(\omega, k) e^{-i\omega(t-t')} \quad (5.49)$$

where $G^R(\omega, k)$ is analytic in the UHP of ω because $G^R(t, k) = 0$ for $t < 0$. Therefore

$$\langle n(t, k) \rangle = -\mu_0(k) \int_{-\infty}^0 dt' e^{\epsilon t'} \int_{-\infty}^\infty d\bar{\omega} G^R(\omega, k) e^{-i\omega(t-t')} \quad (5.50)$$

$$= -\mu_0(k) \int d\omega e^{-i\omega t} \frac{G_{nn}^R(\omega, k)}{i\omega + \epsilon}. \quad (5.51)$$

Integrating the BHS times $e^{(iz-\epsilon)t}$ over t from 0 to ∞ gives

$$\langle n(z, k) \rangle = -\mu_0(k) \int_{-\infty}^\infty d\omega \frac{G_{nn}^R(\omega, k)}{(i\omega + \epsilon)(i(\omega - z) + \epsilon)} \quad (5.52)$$

$$= -\mu_0(k) \frac{G_{nn}^R(z, k) - G_{nn}^R(z=0, k)}{iz} \quad (5.53)$$

where we did the integral by closing the contour in the UHP where $G(\omega, k)$ is analytic, and taking the residues at the two poles $\omega = i\epsilon, \omega = z + i\epsilon$. Comparing with (5.42) we have

$$G_{nn}^R(z, k) - \underbrace{G_{nn}^R(z=0, k)}_{=-\chi} = \frac{-iz\chi}{-iz + Dk^2}. \quad (5.54)$$

The result is

$$G_{nn}^R(z, k) = \frac{\chi Dk^2}{iz - Dk^2} \quad (5.55)$$

which as you can see has the appropriate diffusive pole in the LHP. This implies the Kubo formula

$$D\chi = -\lim_{\omega \rightarrow 0} \lim_{k \rightarrow 0} \frac{\omega}{k^2} \text{Im} G_{nn}^R(\omega, k). \quad (5.56)$$

The order of limits between k and ω matters here as you can see by taking the wrong limit of (5.55). So if you have a way of determining the retarded two-point function of n from some microscopic description, you know what to use for D in the hydrodynamic (diffusion) equation.

You can imagine that we can do the same kinds of manipulations for other transport coefficients. The logic is the same. They each have their own Kubo formula.

What effective field theory produces this kind of response function (5.55)? We can introduce a Langevin description of the current:

$$J_j = -D\partial_j n + r_j \quad (5.57)$$

where r_j is a noise vector, which we can assume to be Gaussian and white. The current should still be conserved:

$$\partial_t n = -\vec{\nabla} \cdot J = D\vec{\nabla}^2 n + \theta \quad (5.58)$$

where $\theta \equiv -\partial_j r_j$ has

$$\langle \theta(x, t) \theta(x', t') \rangle = -2T\chi D\vec{\nabla}^2 \delta^d(x - x') \delta(t - t') \quad (5.59)$$

with the kernel chosen to guarantee equilibrium. The extra $\vec{\nabla}^2$ is a consequence of the conservation law. That is

$$\langle \dots \rangle = \int D\theta e^{-\frac{1}{2} \int dt d^d x d^d x' \theta(x, t) \Gamma(x, x') \theta(x', t)} \dots, \quad -C\vec{\nabla}^2 \Gamma(x, x') = \delta^d(x - x'). \quad (5.60)$$

To compute correlations of n , $\langle n(x, t) n(x', t') \dots \rangle$, one way to do it is to find $n_\theta(x, t)$, the solution for a fixed noise instance, and then average the product $n_\theta(x, t) n_\theta(x', t') \dots$ over the distribution (5.60). A slicker way is to write

$$\langle n(x, t) n(x', t') \dots \rangle = \int Dn \langle \delta[\text{EOM}] \rangle_\theta \mathcal{J} n(x, t) n(x', t') \dots \quad (5.61)$$

where $\mathcal{J} = \det \frac{\delta \text{EOM}}{\delta n} = \text{constant}$ will not matter, and EOM is (5.58). Now we can represent the delta function imposing the equations of motion by

$$\delta[\text{EOM}] = \int D\rho e^{i \int \rho(\text{EOM})}, \quad (5.62)$$

where ρ will play the role of the response field, as in our analogous discussion of diffusion. Using the fact that the θ integral is Gaussian, we have

$$\langle n(x, t) n(x', t') \dots \rangle = \int Dn D\rho e^{-S_{\text{eff}}[\rho, n]} n(x, t) n(x', t') \dots \quad (5.63)$$

with

$$S_{\text{eff}}[\rho, n] = \int dt d^d x \left(\frac{C}{2} (\partial_i \rho)^2 - i\rho (\partial_t n - D\vec{\nabla}^2 n) \right). \quad (5.64)$$

This description is called MSR. If you wanted, you could do the (also Gaussian) integral over ρ to find a description purely in terms of n , with

$$S_{\text{eff}}[n] = \int dt \int d^d x d^d x' \frac{1}{2} B(x, t) \Gamma(x, x') B(x', t), \quad B \equiv \partial_t n - D \vec{\nabla}^2 n. \quad (5.65)$$

This is like the Onsager-Machlup action we discussed in the context of single particle motion; in the context of field theory, it has the shortcoming that it is nonlocal in space (in this example, because of the form of the noise kernel). So this is another virtue of the MSR description involving the response field ρ .

Notice that

$$\int Dn D\rho e^{-S_{\text{eff}} n_1 n_2} = \langle n_1 n_2 \rangle = \langle n_2 n_1 \rangle = \int Dn D\rho e^{-S_{\text{eff}} n_2 n_1} \quad (5.66)$$

since obviously the objects in the path integrand are just numbers and can be written in any order. So correlations of n are actually symmetric objects. With more degrees of freedom $\phi_a = (n, \dots)_a$, this computes the symmetrized correlation function

$$C_{ab} = \frac{1}{2} \langle \{ \phi_a, \phi_b \} \rangle. \quad (5.67)$$

In the present case n is real so the Fourier transform satisfies $C_{nn}(\omega, k)^* = C_{nn}(-\omega, -k)$, and further there is time-reversal symmetry and rotation invariance which implies that $C_{nn}(-\omega, -k) = C_{nn}(\omega, k)$ so $C_{nn}(\omega, k)$ is real. Writing the (Gaussian) action as

$$S_{\text{eff}}[\phi^a] = \frac{1}{2} \int \bar{d}\omega \bar{d}^d k (\phi_{\omega, k}^a)^* P_{ab}(\omega, k) \phi_{\omega, k}^b \quad (5.68)$$

we find $C_{ab}(\omega, k) = (P^{-1})_{ab}$ which in this case gives

$$C_{nn}(\omega, k) = \frac{2T\chi Dk^2}{\omega^2 + (Dk^2)^2}. \quad (5.69)$$

This indeed satisfies the FDT

$$C_{ab}(\omega, k) = -\frac{2T}{\omega} \text{Im} G_{ab}^R(\omega, k) \quad (5.70)$$

(valid for $\beta\omega \ll 1$). The response function is

$$C_{\rho n}(\omega, k) = \frac{1}{\omega - \mathbf{i} Dk^2} = -C_{n\rho}(\omega, k) \quad (5.71)$$

and the two are related by

$$-\mathbf{i} D\chi k^2 C_{n\rho}(\omega, k) = G_{nn}^R(\omega, k). \quad (5.72)$$

A reason you might want to develop such a field theory description further is to understand hydrodynamic fluctuations. Examples where they are important are the power laws that appear in hydrodynamic correlation functions (called long-time tails), and at the onset of fluid instabilities like Rayleigh-Bernard convection (or maybe the transition to turbulence).

5.3 Transport and the approach to equilibrium

Finally, we can return to the Boltzmann equation. With some work, we can show that it indeed produces the Navier-Stokes equation. Thus, one way to think about the Boltzmann equation is as a UV (short-time, short-distance) completion of the fluid dynamics equations. Why do we want such a thing? One, very common, purpose in life for it is to compute, given some microscopic description of the fluid particles and their scattering cross sections, the actual numbers for the transport coefficients (such as the viscosities) appearing in the NS equations. Secondly, the NS equations very likely predict their own demise and require some such UV completion.

You should regard this as a big success of coarse-graining. Not only do we get thermodynamics, but also transport.

Write

$$f(r, v, t) = \underline{f}(r, v) + g(r, v, t) \quad (5.73)$$

where we'll assume g is small and expand in it. We're going to plug this into the Boltzmann equation. This procedure is called Chapman-Enskog expansion. The (log of the) local equilibrium distribution about which we are expanding can be written as

$$\log \underline{f}(v, r, t) = \log n(r, t) - \frac{d}{2} \log T(r, t) - \frac{mw(r, t)^2}{2T(r, t)} + \text{constant} \quad (5.74)$$

where $w_i \equiv v_i - u_i(r, t)$ is the deviation from the background flow $u(r, t)$.

We will make a rather crude approximation now. The idea is that the collision term $(\frac{\partial f}{\partial t})_{\text{collisions}} = \mathcal{C}[\underline{f} + g] = \hat{L}g$ is a linear operator acting on the perturbation g . We will make the approximation that all of its nonzero eigenvalues are the same, and equal to $-1/\tau$. It's a bit like the Einstein model of phonons. The *relaxation time approximation* is

$$\left(\frac{\partial f}{\partial t}\right)_{\text{collisions}} \approx \underbrace{\frac{-f}{\tau}}_{\text{loss}} + \underbrace{\frac{\underline{f}}{\tau}}_{\text{gain}}. \quad (5.75)$$

Furthermore, we make an assumption about the eigenvectors: in each time interval dt , a fraction dt/τ of the particles scatter elsewhere (this produces the loss term), and this same fraction of particles arrive from the equilibrium distribution.

You can see that it has the manifest virtue that it pushes the distribution toward the equilibrium state. Solving for the perturbation gives

$$g_v \approx -\tau \left(\partial_t + \vec{v} \cdot \vec{\nabla}_r + \frac{\vec{F}(r)}{m} \cdot \vec{\nabla}_v \right) \underline{f}_v \quad (5.76)$$

where τ is a relaxation time. Plugging this back in to the Boltzmann equation,

$$\partial_t f_v \approx -\frac{f_v - \underline{f}_v}{\tau} \implies f_v - \underline{f}_v \sim e^{-t/\tau}, \quad (5.77)$$

indeed we decay back to equilibrium in time of order τ .

It is not a perfect approximation. Since it declares all the relaxation rates to be the same, various transport coefficients come out to be the same in this approximation which in reality can be different. For example, the Prandtl number measures the ratio of the viscosity relaxation time to the heat conduction relaxation time, $Pr \equiv C_p \eta / \kappa$, and this approximation predicts it to be 1, while in real fluids it can be 2/3. So if you want quantitatively correct values of transport coefficients, you'll have to find a way to do better. My goal here is just to convey some idea of how the fluid equations arise from the Boltzmann description.

In analyzing the consequences of this approximation, unsurprisingly by now, it's useful to follow the conserved quantities, ϕ . If a quantity ϕ is preserved by the collisions, then

$$0 = \int dv \phi(r, v) \left(\partial_t + \vec{v} \cdot \vec{\nabla}_r + \frac{\vec{F}(r)}{m} \cdot \vec{\nabla}_v \right) f(r, v, t). \quad (5.78)$$

We can apply this to

1. Particle conservation, $\phi = m$. From this we learn, at leading order,

$$0 = \int d^3v m (\partial_t + v_i \partial_{r_i} + F_i / m \partial_{v_i}) \underline{f} \quad (5.79)$$

$$= \partial_t \left(\underbrace{m \int d^d v \underline{f}}_{=mn \equiv \rho} \right) + \partial_{r_i} \underbrace{\int d^d v m v_i \underline{f}}_{=\rho u_i} \quad (5.80)$$

where the term involving F vanishes because $\int d^d v \partial_{v_i} \underline{f} = 0$. So the continuity equation is

$$\partial_t \rho + \vec{\nabla}_r \cdot (\rho \vec{u}) = 0. \quad (5.81)$$

2. Momentum conservation, $\phi_i = mv_i$. From this, using the leading-order solution, we learn

$$0 = \partial_t \int mv_i \underline{f} + \partial_{r_j} \int mv_i v_j \underline{f} + \int d^d v m v_i F_j \partial_{v_j} \underline{f} \quad (5.82)$$

$$\stackrel{\text{IBP}}{=} \partial_t (\rho u_i) + \partial_{r_i} (\rho u_i u_j + P_{ij}^{(0)}) - \frac{\rho}{m} F_i \quad (5.83)$$

In the second term we wrote $v = u + w$ and noticed that the cross terms average to zero. Here the pressure tensor is defined as

$$P_{ij} \equiv \int d^d v m (v_i - u_i) (v_j - u_j) \underline{f} = \int d^d w w_i w_j \underline{f}. \quad (5.84)$$

The leading approximation ($g = 0$) is

$$P_{ij}^{(0)} = \int d^d w \rho w_i w_j \underline{f} = \rho \left(\frac{m}{2\pi T} \right)^{d/2} \int d^d w w_i w_j e^{-\frac{m}{2T} w^2} = \delta_{ij} \underline{P} \quad (5.85)$$

with

$$\underline{P} = \frac{\rho}{d} \left(\frac{m}{2\pi T} \right)^{d/2} \int d^d w w^2 e^{-\beta \frac{m w^2}{2}} = nT, \quad (5.86)$$

the ideal gas law. So the ideal hydro equation is

$$\underbrace{\left(\partial_t + \vec{u} \cdot \vec{\nabla} \right)}_{\equiv D_t} u_\beta = F_\beta / m - \frac{1}{\rho} \partial_{r_\beta} \underline{P}. \quad (5.87)$$

The first-order approximation to the pressure is instead

$$P_{ij}^{(1)} = m \int d^d w w_i w_j \left(\underline{f}_w + g_w \right) = \delta_{ij} \underbrace{\underline{P}}_{=nT} + \Pi_{ij} \quad (5.88)$$

with

$$\Pi_{ij} = -\frac{\tau m^2}{T} \sum_{kl} \frac{\partial u_k}{\partial r_l} \int d^d w w_i w_j \left(w_k w_l - \frac{1}{d} \delta_{kl} w^2 \right) \underline{f}_w. \quad (5.89)$$

This fourth moment of the Maxwell-Boltzmann distribution is determined by Wick contractions. It is a not-entirely trivial step to show that the first-order consequence of momentum conservation is of the same form as the zeroth-order term but with the pressure replaced by the first-order approximation to the pressure; it requires the use of all three zeroth order conservation equations. The first-order version of (5.87) is

$$D_t u_i = \frac{F_i}{m} - \frac{1}{\rho} \partial_{r_i} \left(P - \frac{\eta}{3} \vec{\nabla}_r \cdot \vec{u} \right) + \frac{\eta}{\rho} \vec{\nabla}_r^2 u_i, \quad (5.90)$$

the Navier-Stokes equation. Here η is the shear viscosity, and the result from the relaxation time approximation gives $\eta = \tau nT = \tau \underline{P}$.

3. $\phi = \frac{1}{2}mv^2$. From this, at leading order, we learn

$$D_t T = -\frac{2}{d} T \vec{\nabla} \cdot \vec{u} \quad (5.91)$$

(as long as $\vec{F}$ is independent of velocity and so doesn't lead to dissipation). Arriving at this expression involves some boiling using the leading order momentum-conservation equation²². At first order, it gives

$$D_t T = -\frac{1}{C_V} \vec{\nabla}_r \cdot \vec{u} T + \frac{\kappa}{\rho C_V} \vec{\nabla}_r^2 T \quad (5.98)$$

the heat equation. The thermal conductivity κ and heat capacity C_V are determined from the microscopic information. In the relaxation-time approximation κ is also proportional to τ and this is why the Prandtl number comes out to be 1.

[End of Lecture 17]

²²I record the boiling here: the leading order result for energy conservation is

$$0 = \partial_t \left(\frac{d}{2} u T + \frac{1}{2} \rho u^2 \right) + \vec{\nabla} \cdot \left(\vec{u} \left(\frac{d}{2} n T + n T + \frac{1}{2} \rho u^2 \right) \right). \quad (5.92)$$

This combination $\frac{d}{2} n T + n T = \mathcal{E}_{\text{int}} + P$ is the enthalpy density (of the ideal gas). We can rewrite this as

$$D_t \left(\frac{d}{2} n T + \frac{1}{2} \rho u^2 \right) + \left(\frac{d}{2} n T + \frac{1}{2} \rho u^2 \right) \vec{\nabla} \cdot \vec{u} + \vec{\nabla} \cdot (\vec{u} P). \quad (5.93)$$

The leading-order momentum conservation relation is $\rho D_t u = -\vec{\nabla} P$. Dot this with $\vec{u}$ to get

$$D_t \left(\frac{1}{2} \rho u^2 \right) + \frac{1}{2} \rho u^2 \vec{\nabla} \cdot \vec{u} = -\vec{u} \cdot \vec{\nabla} P. \quad (5.94)$$

Now subtract (5.94) from (5.93):

$$D_t \left(\frac{d}{2} n T \right) + \frac{d}{2} n T \vec{\nabla} \cdot \vec{u} + \underbrace{P}_{=nT} \vec{\nabla} \cdot \vec{u}. \quad (5.95)$$

We still have the derivatives acting on n here. We can get rid of them using the continuity equation

$$D_t n = -n \vec{\nabla} \cdot \vec{u} \quad (5.96)$$

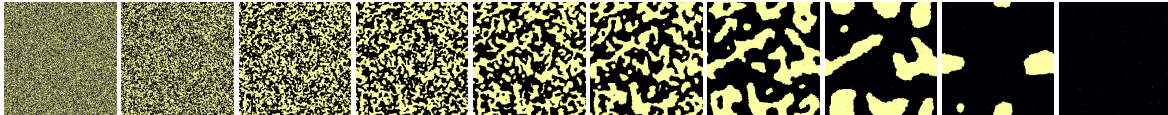
and using $D_t \left(\frac{d}{2} n T \right) = n D_t T + T D_t n$ (that is, D_t satisfies the product rule (“is a derivation”)), we find

$$0 = \frac{d}{2} n D_t T + n T \vec{\nabla} \cdot \vec{u} \quad (5.97)$$

which is the equation I wrote above.

6 Hydrodynamics of magnets

[Cardy’s book, Kardar §9.2, LP §8, Bray, *Theory of Phase-Ordering Kinetics*] Take a $T = \infty$ spin configuration, that is, choose the value of the spin at each site independently uniformly randomly. Now evolve using the Glauber dynamics at $T < T_c$. This is sometimes called a *quench*, meaning that we are evolving an equilibrium state at one value of parameters with the evolution rule for another value. What happens? ([Here](#) is a movie.)



(In the above, the dynamics is sampled at exponentially-spaced time steps.) The idea is that the system forms domains of up spins and down spins, and the domain walls separating them move around. They have a tension which makes them want to minimize their surface area, and they diffuse because of the noise. And they can annihilate each other. In 1d, a domain wall is just a funny kind of particle (area is fixed) and the dynamics is diffusion-limited annihilation. It makes a big difference whether the dynamics conserves the magnetization.

Here is a hydrodynamic or Landau description of the dynamics. At this point I think I’m supposed to say some words of apology for using such coarse-grained descriptions. Even Cardy says some words about “hiding our ignorance of the origins of” irreversibility using phenomenological equations. This is easiest to explain if we are studying the neighborhood of a phase transition where some order parameter field ϕ goes from being ordered (having a nonzero expectation value) to being zero. In that case, ϕ itself is small, and the correlation length is large so it explains why we should be able to use a continuum description and focus on the terms with the fewest derivatives.

The coarse-grained variable $\phi(x)$ that we will follow around is a magnetization averaged over a patch of space near x ,

$$\phi(x) \equiv \frac{\sum_{i \in R_x} s_i}{\text{volume}(R_x)} \quad (6.1)$$

where R_x is a region centered at the site x . The size of the region is a short-distance scale which plays the role of a short-distance cutoff on the wavenumber of the field $\phi(x)$.

Recall that this Langevin equation

$$\partial_t v_I(t) = -\Gamma_{IJ} v_J(t) + \zeta_I(t) = -\Gamma_{IJ} \frac{\partial}{\partial v_J} \mathcal{H} + \zeta_I(t) \quad (6.2)$$

with $\mathcal{H} = \frac{1}{2}v_I^2$ will relax to equilibrium $p(v) \propto e^{-\beta v_I^2/2}$ (the Maxwell-Boltzmann distribution), as long as

$$\langle \zeta(t)_I \zeta_J(t') \rangle = 2D_{IJ} \delta(t - t') \text{ with } D_{IJ} = \Gamma_{IJ} T. \quad (6.3)$$

Now imagine that the indices J on our degrees of freedom include labels x indicating where they are in space.

So let $\phi(x, t)$ be the local magnetization. A dynamics for ϕ that will relax to equilibrium is

$$\partial_t \phi(x, t) = - \int d^d x' \Gamma(x' - x) \frac{\delta \mathcal{H}}{\delta \phi(x', t)} + \zeta(x, t) \quad (6.4)$$

with

$$\langle \zeta(x, t) \zeta(x', t') \rangle = 2T \Gamma(x - x') \delta(t - t'). \quad (6.5)$$

The equilibrium distribution is $p[\phi] \propto e^{-\beta \mathcal{H}[\phi]}$. So we have two choices to make to specify our dynamics: who is $\mathcal{H}[\phi]$ and who is the noise kernel $\Gamma(x - x')$?

$\mathcal{H}[\phi]$ is the equilibrium free energy at fixed local magnetization. This is the subject to which Landau and Ginzburg first applied their famous philosophy: it should be a sum of all possible extensive terms consistent with the symmetries and organized as an expansion in derivatives and powers of the field. So, for example, for the case of a magnet with Ising ($\mathbb{Z}_2$) symmetry: $\phi \rightarrow -\phi$, this looks like

$$\mathcal{H}[\phi] = \int d^d x \left(\frac{1}{2} r \phi^2 + u \phi^4 + \frac{1}{2} \left(\vec{\nabla} \phi \right)^2 \right). \quad (6.6)$$

I've rescaled the field ϕ to set the coefficient of the spatial kinetic term $\frac{1}{2} \left(\vec{\nabla} \phi \right)^2$ to one. Terms with odd powers of ϕ are forbidden by the Ising symmetry.

Now let's talk about the noise kernel, Γ . First consider the generic situation where we impose no further symmetry constraints on the dynamics, except that it is local, so that $\Gamma(x - x')$ is short-ranged in its argument, then its Fourier transform $\Gamma_k = \Gamma_0 + \Gamma_1 k^2 + \dots$. Following Hohenberg and Halperin, this is called model A dynamics. This is the situation, for example, for Glauber dynamics. Then (6.4) becomes, at long wavelengths so we just keep the leading terms in the derivative expansion,

$$\partial_t \phi = -\Gamma \left(r \phi - \vec{\nabla}^2 \phi + 4u \phi^3 \right) + \zeta \quad (6.7)$$

which is creatively called the time-dependent Landau-Ginzburg (TDLG) equation.

Notice that on a trajectory of the dynamics of (6.4),

$$\partial_t \mathcal{H} = \int d^d x \frac{\delta \mathcal{H}}{\delta \phi(x, t)} \partial_t \phi = -\Gamma_0 \int d^d x \left(\frac{\delta \mathcal{H}}{\delta \phi} \right)^2 + \Gamma_0 \int d^d x \frac{\delta \mathcal{H}}{\delta \phi} \zeta. \quad (6.8)$$

The noise average of this is

$$\partial_t \langle \mathcal{H} \rangle = -\Gamma_0 \int d^d x \left(\frac{\delta \mathcal{H}}{\delta \phi} \right)^2 \leq 0 \quad (6.9)$$

– in this sense, the dynamics monotonically approaches equilibrium, it is said to be ‘purely relaxational’. (I assumed $\Gamma_0 > 0$.)

Another possibility is that the dynamics conserve the total spin, as is appropriate for a spin description of a lattice gas (Kawasaki dynamics) or a Heisenberg ferromagnet. This is called model B. In that case $\partial_t \phi_{k=0} = 0$ which requires $\Gamma_{k=0} = 0$, so $\Gamma_k = \Gamma_1 k^2$.

Here’s another way to think about model B: Since the order parameter is conserved, it should satisfy a continuity equation:

$$\partial_t \phi = -\vec{\nabla} \cdot \vec{J}. \quad (6.10)$$

A constitutive relation for the current is

$$\vec{J} = -\Gamma \vec{\nabla} \mu + \vec{\eta} \quad (6.11)$$

where $\mu(x, t) = \frac{\delta \mathcal{H}}{\delta \phi(x, t)}$ is the local chemical potential. Plugging this into the continuity equation gives

$$\partial_t \phi = \Gamma \vec{\nabla}^2 \frac{\delta \mathcal{H}}{\delta \phi(x, t)} + \zeta \quad (6.12)$$

(which is called the Cahn-Hilliard equation), and $\zeta = \vec{\nabla} \cdot \vec{\eta}$, so assuming the noises η_i were white, we have

$$\langle \zeta(x, t) \zeta(x', t') \rangle = -\Gamma \vec{\nabla}^2 \delta^d(x - x') \delta(t - t'). \quad (6.13)$$

The approach to equilibrium

$$\partial_t \langle \mathcal{H} \rangle = - \int d^d x \left(\vec{\nabla} \frac{\delta \mathcal{H}}{\delta \phi} \right)^2 \leq 0 \quad (6.14)$$

is still monotonic.

An interesting generalization that we will not pursue is if the order parameter is advected by some fluid flow. That means that $\partial_t \phi$ gets replaced by the convective derivative

$$\frac{d\phi}{dt} = \partial_t \phi(x, t) + \vec{v}(x, t) \cdot \vec{\nabla} \phi(x, t) \quad (6.15)$$

where $\vec{v}(x, t)$ is the local fluid velocity.

Let’s think for a moment about the linearized dynamics (set $u = 0$), with a source, $\delta \mathcal{H} = \int h \phi$. Then we can solve by Fourier transform:

$$\partial_t \langle \phi_k \rangle = -\Gamma_k (k^2 + r) \langle \phi_k \rangle + h_k(t) \quad (6.16)$$

and

$$\langle \phi_{\omega,k} \rangle = \tilde{G}(\omega, k) h_{\omega,k}, \quad \tilde{G}(\omega, k) = \frac{1}{-\mathbf{i}\omega + \tau_k^{-1}} \quad (6.17)$$

– each mode has its own relaxation time

$$\tau_k = (\Gamma_k(k^2 + r))^{-1} \sim \begin{cases} (k^2 + \xi^{-2})^{-2}, & \text{model A} \\ (k^2 (k^2 + \xi^{-2}))^{-1}, & \text{model B} \end{cases} \quad (6.18)$$

with $\xi = r^{-1/2}$ is the correlation length. From this we can see that model A has $z = 2$ and model B has $z = 4$. The equilibrium state does not determine even the universal properties of the dynamics.

We can do better than the linearized approximation, and understand when it is correct, by writing an MSR path integral (Cardy calls it ‘response functional formalism’). Any observable made of ϕ at a fixed value of the noise can be written as

$$\int D\phi \delta[\text{EOM} - \zeta] \dots = \int D\phi D\tilde{\phi} e^{-\int d^d x dt \tilde{\phi}(\text{EOM} - \zeta)} \dots \quad (6.19)$$

where

$$\text{EOM} \equiv \partial_t \phi + \Gamma \frac{\delta \mathcal{H}}{\delta \phi} \quad (6.20)$$

is the deterministic part of the equations of motion, and $\tilde{\phi}$ is our friend the response field. Notice that $\tilde{\phi}$ really should run over the imaginary axis, I’m hiding some factors of $\mathbf{i}$. Now we can average this over the noise

$$\int D\zeta P[\zeta] ((6.20)) = \int D\phi D\tilde{\phi} e^{-\int \tilde{\phi}(\text{EOM})} \left\langle e^{-\int \tilde{\phi} \zeta} \right\rangle_{\zeta} \quad (6.21)$$

where the generating functional for the noise distribution with source $\tilde{\phi}$ is

$$\left\langle e^{-\int \tilde{\phi} \zeta} \right\rangle_{\zeta} \equiv \int D\zeta P[\zeta] e^{-\int \tilde{\phi} \zeta} = \int D\zeta e^{-\frac{1}{2} \int \int \zeta \Gamma^{-1} \zeta - \int \tilde{\phi} \zeta} = e^{\frac{1}{2} \int \int \tilde{\phi} \Gamma \tilde{\phi}}. \quad (6.22)$$

Thus, we can compute any noise average of an observable made of ϕ by integrating ϕ and $\tilde{\phi}$ against the measure

$$\langle \dots \rangle_{\zeta} = \int D\phi D\tilde{\phi} e^{-S[\phi, \tilde{\phi}]} \dots, \quad (6.23)$$

$$S[\phi, \tilde{\phi}] = \int d^d x dt \tilde{\phi} \left(\dot{\phi} + \Gamma \frac{\delta \mathcal{H}}{\delta \phi} \right) - \frac{1}{2} \tilde{\phi} \Gamma \tilde{\phi}. \quad (6.24)$$

And recall that the mixed correlator $\left\langle \phi(x, t) \tilde{\phi}(x', t') \right\rangle$ is the response function.

For model A, Γ , the coefficient of the $\tilde{\phi}^2$ term, is well approximated by a number. For model B, $\Gamma = \Gamma_1 \vec{\nabla}^2$, so this looks like an ordinary spatial kinetic term for the field

$\tilde{\phi}$. However you can see that the sign is funny; this is related to the fact that we should really redefine its phase by a factor of $\mathbf{i}$.

One thing we can do right away with this action is understand when mean field theory (the linearized calculation above) is correct. That is, using dimensional analysis we can determine the upper critical dimension, above which MFT is correct. This happens when the nonlinear coupling u becomes dimensionless.

First rescale $\tilde{\phi}$ and the choice of t to remove Γ . The the model-A action is

$$S_A[\tilde{\phi}, \phi] = \int d^d x dt \left(\tilde{\phi} \left(\dot{\phi} - \vec{\nabla}^2 \phi + r\phi + 4u\phi^3 \right) - \frac{1}{2} \tilde{\phi}^2 \right). \quad (6.25)$$

Since the action appears in the exponential each term must be dimensionless. Comparing the first two terms, we see that the units of time and length must be related by $T = L^2$ – this is because $z = 2$. Then we have

$$1 = L^{d+2}[\tilde{\phi}^2] \equiv L^{d+2+2\tilde{\Delta}} \quad (6.26)$$

$$1 = L^{d+2}[\tilde{\phi} \vec{\nabla}^2 \phi] = L^{d+2+\tilde{\Delta}-2+\Delta} \quad (6.27)$$

which says $\Delta = -d - \tilde{\Delta}$ so

$$[\phi] = L^\Delta = L^{1-d/2}, [\tilde{\phi}] = L^{\tilde{\Delta}} = L^{-1-d/2}. \quad (6.28)$$

Therefore we can see that the ‘mass term’ in $r\phi^2$ has dimensions $[r] = L^{-2}$ is a negative power of length; this is a relevant operator in any dimension. The non-linear term has $[u] = L^{d-4}$, which is relevant for $d < 4$ and irrelevant for $d > 4$. $d_{uc}^A = 4$ is the upper critical dimension for model A.

For model B the MSR action is

$$S_B[\tilde{\phi}, \phi] = \int d^d x dt \left(\tilde{\phi} \left(\dot{\phi} - \vec{\nabla}^2 \left(\vec{\nabla}^2 \phi + r\phi + 4u\phi^3 \right) \right) - \frac{1}{2} \tilde{\phi} \vec{\nabla}^2 \tilde{\phi} \right). \quad (6.29)$$

So comparing the first two terms $z = 4$, and

$$1 = L^{d+4}[\tilde{\phi}^2] = L^{d+4+2\tilde{\Delta}} \quad (6.30)$$

$$1 = L^{d+4}[\tilde{\phi} \vec{\nabla}^4 \phi] = L^{d+4+\tilde{\Delta}-4+\Delta} \quad (6.31)$$

which says again $\Delta = -d - \tilde{\Delta}$, so

$$[\phi] \equiv L^\Delta = L^{2-d/2}, [\tilde{\phi}] \equiv L^{\tilde{\Delta}} = L^{-2-d/2}. \quad (6.32)$$

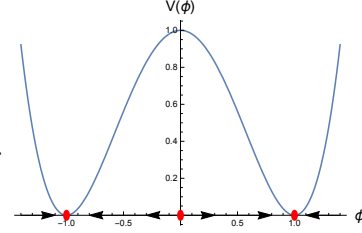
Again we find $[r] = L^{-2}$ in any dimension, and again we find $[u] = L^{d-4}$, the upper critical dimension is still $d_{uc}^B = 4$. In fact the upper critical dimension is a property of the equilibrium state.

6.1 Coarsening

To begin, consider the 0 + 1d differential equation

$$\partial_t \phi(t) = -V'(\phi(t)), \quad V(\phi) = (\phi^2 - 1)^2. \quad (6.33)$$

What is the late-time behavior of its solutions? There are fixed points of the evolution at $\phi = \pm 1$ and $\phi = 0$. The two fixed points at nonzero ϕ are related by the $\phi \rightarrow -\phi$ symmetry of the potential. As you can check by the sign of V' at nearby points, the former are attractive and the latter is repulsive. Thus, at late times, the field goes to $\phi = 1$ (if it started at any $\phi > 0$), or to $\phi = -1$ (if it started at any $\phi < 0$).



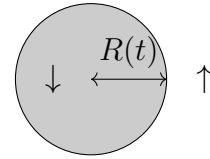
Now we include spatial dependence and noise, and instead consider a field governed by the noise-free model-A dynamics

$$\partial_t \phi(\vec{x}, t) = \vec{\nabla}^2 \phi(\vec{x}, t) - V'(\phi(\vec{x}, t)) + \zeta(\vec{x}, t). \quad (6.34)$$

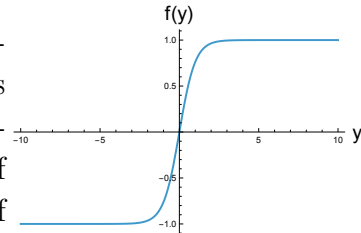
$\phi(\vec{x}, t)$ could be the magnetization of a ferromagnet, where $\phi = \pm 1$ represent all spin up and all spin down respectively. Now, although each individual spin sees the potential above, nearby spins may head in different directions. The result is the formation of a *domain wall* between a region of up spins and a region of down spins. Such a domain wall is a topological object, in the sense that if we know that over here the spins are up, and over there the spins are down, the number of such domain walls in between must be an odd number.

Let's think about a single isolated island of down spins in a sea of up spins. For simplicity we'll think about the case when it is spherical. Suppose the field takes the form

$$\phi(\vec{x}, t) = f(r - R(t)) \quad (6.35)$$



where r is the distance from the origin, $R(t) \geq 0$ is a dynamical variable, and $f(y)$ is a function that approximates $\text{sign}(y)$. This describes a round domain of down spins of radius $R(t)$, surrounded by a sea of up spins. As a function of the initial radius $R(0) = R_0$, how long does the domain of down spins take to shrink?



In d dimensions,

$$\vec{\nabla}^2 = \partial_r^2 + \frac{d-1}{r} \partial_r + \text{angular terms} . \quad (6.36)$$

Plugging the expression for the domain wall configuration into (6.34) gives

$$-\dot{R}f'(r-R) = f''(r-R) + \frac{d-1}{r}f'(r-R) - V'(f). \quad (6.37)$$

But $f'(r-R) \simeq 2\delta(r-R)$, the Dirac delta function. Integrating across $r = R$, the first and last terms on the RHS do not contribute and we get

$$\dot{R} = -\frac{d-1}{R}. \quad (6.38)$$

This equation says

$$R\dot{R} = -(d-1) \implies R(t)^2 - R_0^2 = -(d-1)t \quad (6.39)$$

and hence $t \sim R_0^2$ is the time to shrink.

Notice that the noise was not involved in what I said above. In fact the noise leads the domains to shrink at the same rate. If we keep the noise in (6.34) and average it over the angular directions (so the angular dependence averages to zero and we still get an equation for $f(r-R(t))$, it contributes a term

$$\dot{R} = \dots + \zeta(r, t) \quad (6.40)$$

of the form of 1d Brownian motion in the size of the domain wall, which therefore (even ignoring the deterministic drift) satisfies $R(t) - R(0) \sim t^{1/2}$.

This rule can be shown to be quite general: the size of a domain wall goes like $R(t) \sim t^{1/2}$ where the $1/2$ is called the coarsening exponent. The generalization comes in two steps, which I will just summarize. First, generalizing (6.38), for domain walls of arbitrary shape there is a relation between the local velocity of the domain wall and its curvature:

$$\text{velocity} = -\text{curvature} + \text{noise} \quad (6.41)$$

called the Allen-Cahn equation (see LP Appendix U).

Second, we can define more carefully what we mean by $R(t)$. The better definition is in terms of correlation functions of the order parameter:

$$C(r, t) \equiv \langle \phi(x+r, t)\phi(x, t) \rangle - \langle \phi(x, t) \rangle^2. \quad (6.42)$$

In this problem we are quenching to a state with a nonzero correlation length ξ – it is not scale invariant. This length scale determines the *thickness* of the domain walls. After a little bit of coarsening, as you can see in the images above, the domains themselves become much larger than the thickness of the domain walls $R(t) \gg \xi$, and we can treat them as infinitesimally thin (as we did above) and the problem has only

one length scale $R(t)$. In this regime, because of the absence of other length or time scales in the problem, the correlation function satisfies the scaling relation

$$C(r, t) = b^{2\chi} C(b^n r, bt) \quad (6.43)$$

with $\chi = 0$. $\chi = 0$ because inside the domains, the order parameter is fixed to its equilibrium value $\langle \phi \rangle^2|_{\text{inside}} \approx M_{\text{eq}}^2$ where M_{eq} is determined by equilibrium physics, *i.e.* the temperature, and therefore ϕ cannot scale. (If we were to study the quench to the critical temperature $T = T_c$, then the domains themselves would be fractalicious, and the order parameter would scale under the scale transformation, and we would have $\chi \neq 0$.) Setting $b = 1/t$, we have

$$C(r, t) = f\left(\frac{r}{R(t)}\right) \quad (6.44)$$

with $R(t) = t^n$, n is called the coarsening exponent, which we estimated to be $n = 1/2$ for model A dynamics. $R(t)$ should be regarded as the average size of domain walls. (6.44) predicts that if you plot the correlation function at fixed t for different t versus $r/R(t)$ (with the correct choice of n), they will all collapse onto a single curve. Indeed this happens. Note that in experiments the thing that usually gets measured is the Fourier transform of (6.44).

Two simple generalizations worth mentioning: If we add an external magnetic field, so that we add a linear term, ϕH , to $V(\phi)$, it adds a constant H to the RHS of (6.34). Following this through the analysis, you can see that this adds a constant term to $\dot{R}$ – the domain walls shrink with a constant velocity. This is because such a field is a *pressure* for the domains, which is positive for the favored-sign domains and negative for the disfavored-sign domains.

Finally, the above assumed that $d > 1$. If $d = 1$, the domain walls do not have a curvature and so they don't know which way to go to shrink. The result,

$$\dot{R} = -\frac{1}{I} (U(m)|_0^\infty + \xi) \quad (6.45)$$

is ‘logarithmic coarsening’ (in the absence of noise). A solution with a domain of down spins in a sea of up spins is $m(x) = \tanh \frac{x-R}{\sqrt{2}} - \tanh \frac{x+R}{\sqrt{2}} + 1$, so $m(0) = 1 - 2 \tanh R/\sqrt{2} \stackrel{R \gg 1}{\approx} -1 + 4e^{-\sqrt{2}R}$. This gives

$$\dot{R} = \frac{-16}{I} e^{-2\sqrt{2}R} + \tilde{\xi}(t) . \quad (6.46)$$

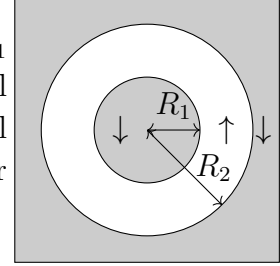
But really this just means that the noise is very important and actually $R(t) \sim \sqrt{t}$.

Ostwald ripening. Now, what if instead we use the model-B dynamics? In this case a big ball of up-spins in a sea of down spins, like we were considering above, is

actually a stationary configuration! However, if there are many domains, a small ball of minority spins will evaporate onto larger balls. The claim is that in $d > 2$, this leads to $R(t) \sim t^{1/3}$.

Let me explain it in steps. In model B dynamics, up spins in a sea of down spins are just like particles that diffuse. The key idea is that the motion of the domain walls is slow compared to the diffusion of these particles. So we let the diffusion relax to the extent possible in a fixed configuration of domain walls, and use the result to determine the force on the domain walls.

A good situation to imagine is a small ball of up spins of radius R_1 surrounded by a shell of down spins out to radius R_2 , and then all up spins from there, for $R > R_2$. The claim is that, under model B dynamics, the small ball of up spins will evaporate onto the inner boundary of the infinite region.



Let me sketch how it goes, we are just trying to get the scaling of R with t . First, the time rate of change of the volume $\dot{V} \sim JA$ is determined by the radial flux integrated over the boundary at R . The LHS is $\dot{V} \sim \dot{R}R^{d-1}$. The RHS is JR^{d-1} . So we learn that

$$\dot{R} \sim J|_R. \quad (6.47)$$

Second, this current is $\vec{J} = -\vec{\nabla}\mu$ where μ is the chemical potential (which is the same as the pressure). What is μ ? μ is determined in the down-spin region by letting the particles diffuse as much as they want, and the final result is

$$0 = \partial_t \mu = \vec{\nabla}^2 \mu. \quad (6.48)$$

So the solution for μ for $r \in (R_1, R_2)$ is $\mu(r) = \frac{A}{r^{d-2}}$ and we must determine A by boundary conditions.

The boundary conditions are interesting and are called the Gibbs-Thomson effect. Think about a soap bubble in air. There is a balance between the pressure difference $\Delta P = P_{in} - P_{out}$ pushing out and the surface tension $\sigma \cdot \text{curvature}$ squeezing the bubble. The result for a spherical bubble is

$$\frac{2\sigma}{R} = \Delta P = \Delta\mu = \mu(R) - \mu_\infty. \quad (6.49)$$

Let's set $\mu_\infty = 0$ and the solution is

$$\mu(r) = \frac{2\sigma R^{d-3}}{r^{d-2}}. \quad (6.50)$$

Putting this together,

$$\dot{R} \sim J|_{r=R} \sim -\vec{\nabla}\mu|_{r=R} = \frac{2\sigma R^{d-3}}{r^{d-1}}|_{r=R} \sim R^{-2}. \quad (6.51)$$

Thus $\dot{R} \sim R^{-2}$ which says $R(t)^3 \sim t$.

[End of Lecture 18]

6.2 Hydrodynamic description of generic non-ergodicity

[He, Jayaprakash, Grinstein] In the preceding subsection, we discussed a continuum description of dynamics of magnets that relax to Gibbs states. But we saw in our discussion of the Toom model above that detailed balance is a real restriction on the possibilities for what can happen. So a natural question is: can we give a continuum description of the Toom model? That is, if we deform model A so that we no longer have $\mathbb{Z}_2$ symmetry but also no longer have detailed balance, can we see that there is an open neighborhood of parameter space where there is still ergodicity breaking, as in the Toom model?

The paper linked above shows that the answer is ‘yes’. The change in the dynamics relative to model A is extremely mild: it is just the addition of a $v (\partial_x^3 \phi + \partial_y^3 \phi)$ term that breaks rotation symmetry and also detailed balance (it is not of the form $\frac{\delta \mathcal{H}}{\delta \phi}$). The authors show that this modification of the evolution rule dramatically changes the coarsening behavior in that domains of minority spins shrink, just as in the Toom rule. The authors estimate that (6.38) is replaced by

$$\frac{dR}{dt} = -\frac{\sigma}{R} - v \quad (6.52)$$

with $v \propto \epsilon$.

6.3 Dynamical HMW theorem and its violation

So far, we’ve been thinking about one-component order parameters where the system has a $\mathbb{Z}_2$ symmetry under $\phi \rightarrow -\phi$. Next let’s consider instead systems with degrees of freedom with more components, $\phi^a, a = 1 \cdots N$, and that can therefore have a continuous symmetry, $\phi^a \rightarrow R^{ab} \phi^b$. The simplest interesting case is $N = 2$, where we can write $\phi^a = M(\cos \theta, \sin \theta)$ at each site. I’ll refer to such a degree of freedom as a ‘pointer’.

A domain wall is a topological object (if we fix $\phi(x > 0) = 1$ and $\phi(x < 0) = -1$, we know that there must be an odd number of domain walls in between) with codimension 1. This fact is partly responsible for the slow dynamics involved in coarsening

described above. Analogously, for $N > 1$, there are interesting topological objects with codimension N . The basic configuration is one where we identify the direction in which the pointers are pointing with the direction in space away from some codimension N locus. For $N = 2$ this is a vortex, for $N = 3$ this is a hedgehog or monopole. These objects similarly get created by a quench from high temperature.

Noisy pointers in a fog. [This nice metaphor is from Toner's book.] Consider a collection of such pointers distributed over space, *i.e.* at each site there's a rotor variable $\theta_i \in [0, 2\pi)$. The rotors are coupled by a local interaction that makes them want to point the same way. And they are subject to a noise force, which we'll assume is Gaussian and white. Let's assume they satisfy

$$\theta_i(t+1) = \langle \theta_i \rangle_{\text{neighbors}} + f_i(t) \quad (6.53)$$

where f_i is some noise and

$$\langle \theta_i \rangle_{\text{neighbors}} \equiv \frac{1}{g} \sum_{\langle i|j \rangle} \theta_j \quad (6.54)$$

is the average of the spins of the neighbors of the site i (g is the number of neighbors). Looked at from far away, the resulting field $\theta(x, t)$ therefore satisfies a Langevin equation

$$\partial_t \theta = \nu \vec{\nabla}^2 \theta + f \quad (6.55)$$

where $\langle f(\mathbf{r}, t) f(\mathbf{r}', t') \rangle = 2D \delta^d(\mathbf{r} - \mathbf{r}') \delta(t - t')$ (and $\nu = a^2/4$ with a the lattice spacing). This satisfies detailed balance with respect to the Hamiltonian

$$H[\theta] = \int d^d x \frac{K}{2} (\vec{\nabla} \theta)^2 = \sum_q \frac{K q^2}{2} |\theta(q)|^2. \quad (6.56)$$

This Hamiltonian is determined by the demands that it is local, depends only on $\vec{\nabla} \theta$ (so that $\theta \rightarrow \theta + \epsilon$ is a symmetry), and is rotation invariant. This last property is an emergent property, either as an enhancement of lattice symmetries, or as an average property of a distribution of random locations for the pointers. Since it's Gaussian, equipartition requires $\langle \frac{K q^2}{2} |\theta(q)|^2 \rangle = T/2$. Comparing to the Langevin equation,

$$\partial_t \theta = -\Gamma \frac{\delta H}{\delta \theta} + f = \Gamma K \vec{\nabla}^2 \theta + f \quad (6.57)$$

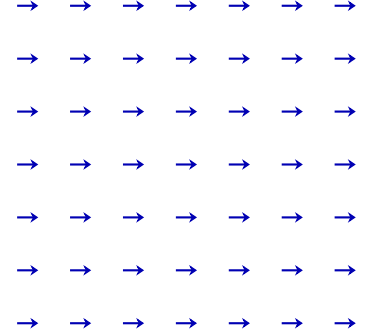
we see the temperature is $T/K = \frac{D}{\nu}$ with $\nu = \Gamma K$. This model is sometimes called time-dependent Landau-Ginzburg XY (TDLGXY) model.

The noise-free dynamics has a conserved current $J_\mu = (\theta, -\vec{\nabla} \theta)_\mu$, in the sense that $\partial_\mu J^\mu = \partial_t \theta + \vec{\nabla} \cdot \vec{J} = f$ (this is the Langevin equation). So, on average, $\partial_\mu \langle J^\mu \rangle =$

$\langle f \rangle = 0$. This means that the value of the ‘magnetization density’ $\int_R \theta(\mathbf{r}, t) d^d r$ can only change by spreading out, *i.e.* by diffusion. For example, if $\theta(\mathbf{r}, 0) = \theta_0 \delta^d(\mathbf{r})$, then

$$\theta(\mathbf{r}, t) = \theta_0 \frac{e^{-\frac{r^2}{4\nu t}}}{(4\pi\nu t)^{d/2}}. \quad (6.58)$$

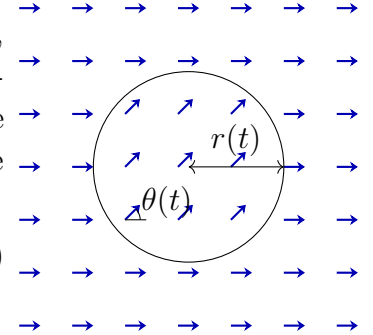
Suppose we start with a configuration where all the pointers point in the same direction, say $\theta = 0$. Without noise, obviously they would just keep pointing that way, with a definite magnetization. What happens with noise? The case of $d = 2$ is special. In this case (or any smaller d), the noise will destroy the long-range order. Here is a heuristic argument for this.



The noise creates errors in the pattern of order, and diffusion slowly dilutes them. Which one wins?

First a sketchy argument. A single isolated error, created by the noise, produces nearby an evolution of the form (6.58). At some small-enough time t , only spins in a neighborhood of radius $r(t) \sim \sqrt{\nu t}$ are affected, call their number $N(r(t))$. We can say that they all have approximately

$$\theta(t) = \frac{\theta_0}{N(r(t))} = \frac{\theta_0}{t^{d/2}}. \quad (6.59)$$



On the other hand, in a time t , the noise creates approximately t errors per pointer. (One meaning we can ascribe to this sentence is that at short times $t < a^2/\nu$ (a is the spacing between pointers), we can ignore the diffusion term and $\partial_t \theta \approx \xi$ says that the variance of θ goes like $2Dt$.) Therefore the total number of errors in our region of radius $r(t) = \sqrt{\nu t}$ is $N(r(t))t = t^{\frac{d+2}{2}}$. So the RMS deviation of the ordering direction is (the sum of K independent random variables goes like $\sqrt{K}$):

$$\sqrt{\langle \theta^2 \rangle - \langle \theta \rangle^2} \sim \sqrt{\frac{\text{total number of errors}}{N(r(t))}} \sim \frac{r^{\frac{d+2}{2}}}{r^d} = r^{1-d/2}, \quad (6.60)$$

which grows with the size of the region if $d < 2$. (If $d = 2$ there is actually a logarithm.)

Here is a more precise argument. Put the system in a box of linear size L (IR cutoff), so the wavenumbers are $\vec{q} = \frac{2\pi}{L} \vec{n}$, and put a UV cutoff on the wavenumber

$|\vec{q}| < \Lambda$. We can solve the Langevin equation in Fourier space, by

$$(-i\omega + \nu q^2)\theta(q, \omega) = f(q, \omega) \quad (6.61)$$

i.e. $\theta(q, \omega) = G(q, \omega)f(q, \omega) = \frac{1}{-i\omega + \nu q^2}f(q, \omega)$. This gives

$$\langle \theta(q, \omega)\theta(q', \omega') \rangle = G(q, \omega)G(q', \omega') \langle f(q, \omega)f(q', \omega') \rangle = C(q, \omega)\delta_{q+q'}\delta_{\omega+\omega'} \quad (6.62)$$

with

$$C(q, \omega) = \langle \theta(q, \omega)\theta(-q, -\omega) \rangle = 2DG(q, \omega)G(-q, -\omega) = \frac{2D}{\omega^2 + \nu^2 q^4}. \quad (6.63)$$

The middle expression here takes advantage of the fact that we are in finite volume so the delta functions are Kronecker deltas. Note that $\theta(-q, -\omega) = \theta(q, \omega)^*$.

This means that the equal-time correlator is

$$\langle \theta(q, t)\theta(q', t) \rangle = \underbrace{\int_{-\infty}^{\infty} d\omega C(q, \omega) \delta_{q+q'}}_{\equiv C_{\text{ET}}(q)} \quad (6.64)$$

where

$$C_{\text{ET}}(q) = \int d\omega \frac{2D}{\omega^2 + \nu^2 q^4} = \frac{2D}{\nu q^2} \frac{\pi}{2\pi} = \frac{D}{\nu q^2}. \quad (6.65)$$

So the position-space correlator is

$$\langle \theta(x, t)\theta(0, t) \rangle = \int d^d q C_{\text{EQ}}(q) e^{iq \cdot x} \propto \int \frac{d^d q}{q^2} e^{iq \cdot x} = \begin{cases} 1/r^{d-2}, & d > 2 \\ \log \frac{r}{\Lambda}, & d = 2 \end{cases} \quad (6.66)$$

or

$$\langle \theta(r, t)^2 \rangle = \frac{D}{\nu} \int_{2\pi/L}^{\Lambda} \frac{d^d q}{q^2} = \frac{D}{\nu} \frac{K_d}{d-2} \left(\Lambda^{d-2} - \left(\frac{L}{2\pi} \right)^{2-d} \right) \stackrel{\Lambda L \gg 1}{\propto} \begin{cases} L^{2-d}, & d < 2 \\ \log \frac{\Lambda L}{2\pi}, & d = 2 \\ L^0, & d > 2 \end{cases}. \quad (6.67)$$

Here $K_d \equiv \frac{\Omega_{d-1}}{(2\pi)^d}$ is the ubiquitous result of the angular integral.

This dynamical argument (which I got from Toner's book) agrees with the usual equilibrium statement of the Hohenberg-Mermin-Wagner-Coleman theorem. The conclusion is that there is no SSB of continuous symmetries in 1+1 dimensions. Notice that if we can somehow change the fact that the Goldstone field θ is a *free* field, we

might evade this conclusion. Indeed this is possible, even in equilibrium with short-ranged interactions. A recent mechanism arises by coupling to some critical degrees of freedom [here, here]. The conclusion of this theorem is also evaded by 2d polymer sheets, whose equilibrium phase diagram includes a well-named “flat phase” where the sheet is flat and therefore picks out an orientation (see §IV in [this review](#) by Leo Radzihovsky or chapter 4 of the book by Nelson and Zhang); again the mechanism is that the would-be Goldstone field is not a free field.

Flocking. [Toner’s book; Nelson-Zhang, *The Renormalization Group and Condensed Matter Physics*, §6.6] The behavior of the order parameter described above, which happens in a model with detailed balance which therefore approaches an equilibrium Gibbs state, is to be contrasted with what can happen in more intrinsically non-equilibrium dynamics, such as the Vicsek model of flocking.

Here are the requirements for a model to be in the universality class of the Vicsek model of flocking. The dofs are a large number (a flock) of “boids”, moving in d dimensions, and attempting at each moment to move in the same direction as their neighbors. That is, each boid is labelled by a pointer θ_i which also has a position variable $\vec{r}_i = (x_i, y_i)$; it moves in the direction that it is pointing. So the key new ingredient is that the pointers are no longer fixed in space, but rather can move around in the fog and find out more about what the other pointers are doing.

The interactions are short-ranged (there is still a fog): each boid follows only those other boids within some fixed distance R_0 . The boids make errors in their headings, modeled as Gaussian white noise. This model has rotation symmetry – there is no special direction in space defined by the model²³. So if the flock of boids heads macroscopically in a particular direction it is spontaneous breaking of rotation symmetry. The very interesting observation is that this can happen even in $d = 2$! The difference from the noisy pointers sitting in a fog that we discussed above is that these pointers can move around and see what is happening elsewhere!

The specific model studied by Vicsek that demonstrated this phenomon has the following update rule for the i th boid:

$$\theta_i(t+1) = \langle \theta_i(t) \rangle_{\text{neighbors}} + f_i(t) \quad (6.68)$$

$$x_i(t+1) = x_i(t) + v_0 \cos \theta(t+1) \quad (6.69)$$

$$y_i(t+1) = y_i(t) + v_0 \sin \theta(t+1) \quad (6.70)$$

²³In the XY model described above, in addition to continuous spatial rotations (which is a symmetry at least in the continuum description we gave; a lattice only preserves a subgroup of this symmetry), there is an independent symmetry acting by $\theta_i \rightarrow \theta_i + \epsilon$. In this model, the rotation symmetry acts simultaneously by $\theta_i \rightarrow \theta_i + \epsilon$ and $r_i \rightarrow R_\epsilon r_i$ where R_ϵ is the spatial rotation by ϵ . Thanks to Julian Dukes for emphasizing this point.

where $f_i(t)$ is white noise, $\langle f_i(t) \rangle = 0$, $\langle f_i(t) f_j(t') \rangle = 2D\delta_{ij}\delta(t-t')$, and $\langle \theta_i(t) \rangle_{\text{neighbors}}$ is the average direction of the boids within a distance R_0 of $(x_i(t), y_i(t))$ ²⁴.

[Here](#) is a javascript simulation of this model with sliders to vary the parameters. ([Here](#) is a simulation of a different model in a similar spirit.) Flocking is defined as “the spontaneous breaking of rotation invariance by collective motion.” You can see quite vividly from even a very simple simulation that this actually happens.

It is possible to understand quite a lot about this phenomenon using the Landau-Ginzburg logic. Toner’s book describes the application of this point of view to this problem in great detail. The key insight is that the slow variables are

1. the density of boids, which is conserved by Vicsek’s dynamics and therefore is slow because it satisfies the continuity equation $\partial_t \rho + \vec{\nabla} \cdot (\rho \vec{v}) = 0$,
2. the velocity field transverse to the flock motion $\delta \vec{v}_\perp$, which is slow because it is the Goldstone mode for SSB of rotation symmetry. The equation for this vector field is a bit complicated, but it is something like

$$\partial_t \vec{v} + \text{convective terms} = -\Gamma \frac{\delta \mathcal{H}[v]}{\delta \vec{v}} + \text{pressure terms} + \vec{f} \quad (6.71)$$

where

$$\mathcal{H} = \int d^d x \left(\frac{D}{2\Gamma} \vec{\nabla} v \cdot \vec{\nabla} v + U(\vec{v}, \rho) \right) \quad (6.72)$$

and $U = a/2|\vec{v}|^2 + b/4|\vec{v}|^4$ so that in the broken phase ($a < 0$), this is a double-well potential for the magnitude of the flock velocity $\vec{v}$. ‘Pressure terms’ means something like $-\nu \vec{\nabla} \rho$. Some effort is required to integrate out the fast mode $\vec{v}_\parallel$. (In the full expression, we have to distinguish parallel and perp directions and there are different couplings for the different directions. I’ve simplified the expression for palatability. Another complication is that the couplings in this equation all depend on the variable ρ .) See Toner chapter 4 for much more detail.

In particular, from this description one can understand why the HMW theorem does not apply. A basic excuse is (again) that the nonlinear terms are not irrelevant so the would-be Goldstone mode is not free. The heuristic argument goes wrong because in addition to diffusing, an “error” (in the language we used above) also moves ballistically, and importantly, not all the pointers move in the same direction. Chapter 7 of Toner’s

²⁴The angles θ_i are only defined mod 2π . If you just average the angles θ , you can sometimes see some weird edge cases where the choice of cut matters. Toner gives a prescription that solves this problem but it’s a little more complicated. In my simulation linked above I didn’t worry about this and it doesn’t seem to matter much.

book attempts a fix of the heuristic argument which predicts the correct upper critical dimension ($d = 4$) and one of the exponents.

This phenomenon of flocking is an example of a more general phenomenon of the formation of consensus without central leadership. Other important examples are the emergence of a common language among early humans, or the emergence of a common belief in the prices in a marketplace.

Mathematicians have proved statements about the formation of flocks in models not too dissimilar to Vicsek's, such as [here](#) and [here](#).

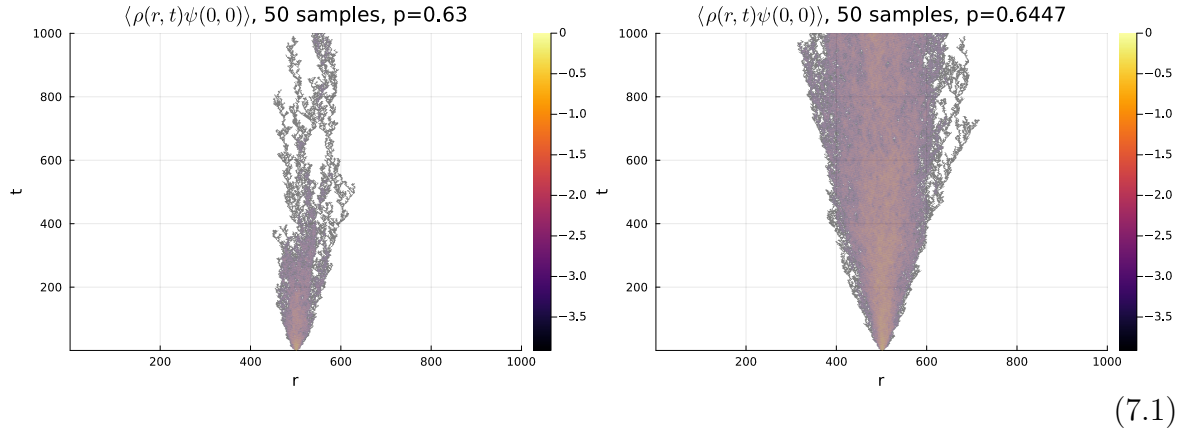
7 Critical phenomena in stochastic dynamics

7.1 Transitions to absorbing states

[This review article is very nice: [Hinrichson, 2000](#), LP §6]

Models of directed percolation. Recall our discussion of population dynamics above: even ignoring space, so that we were studying a $(0 + 1)d$ system, there was a continuous phase transition (as a function of the net birth minus death rate) from an absorbing phase, where everyone goes extinct in a finite time, to an active phase, where there is a finite probability of living descendants in the infinite future. In that case, the critical theory was just Brownian motion of the population.

Directed percolation is in a certain sense the generic non-equilibrium phase transition, at least it is the generic transition to an absorbing state. We'll understand the reason for this below, in terms of the (lack of) symmetries of the associated field theory. A very partial list of realizations includes: forest fires, fluid turbulence, porous media, epidemics. Let's talk about a number of somewhat-microscopic models in this universality class at various levels of description.



These figures are produced by the following procedure. At $t = 0$ a single site is active. At each time step, we loop over active sites; for each active site, we activate each of its two neighbors (independently) with probability p . The figure shows (the log of) the average activation over 50 runs. For p too small (left), everyone always dies out. For p too big, everyone is active. The right figure shows a value near the critical value.

Another model in this universality class is called ‘contact process’. Consider a d -dimensional lattice of sites i . At each site place a two-valued variable, and interpret $s_i(t) = 1$ as ‘the site i is infected’ and $s_i(t) = 0$ as ‘the site i is uninfected’. At each time step, randomly choose a site i , and update by the following reactions: There is a recovery rate, $W(1 \rightarrow 0) = r$, at which an infected site becomes uninfected, and an infection rate

$$W(0 \rightarrow 1, n \equiv \sum_{\langle i|j \rangle} s_j(t)) = \frac{\lambda n}{g}, \quad (7.2)$$

where $\langle i|j \rangle$ means neighbors j of a fixed site i , and g is the coordination number of the site i , the number of neighbors. That is, if the site i is uninfected, it gets infected at a rate proportional to the fraction of its neighbors that are infected.

This model has an absorbing state, $s_i = 0, \forall i$, in which everything ends. Let’s develop a mean field theory of its phase diagram.

Method 1 is to use words. This method is extremely dangerous. Let $\rho(t)$ be a coarse-grained average density (fraction) of active sites, completely ignoring its spatial dependence for now. Then $\partial_t \rho$ gets two contributions. One is from recovery, which is just $-r\rho$. The other is from infection, which has a rate $\rho\lambda$ times $(1 - \rho)$ since the site being infected must be uninfected. So

$$\partial_t \rho = -r\rho + (1 - \rho)\rho\lambda = a\rho - \lambda\rho^2 \quad (7.3)$$

with $a \equiv \lambda - r$. If $\lambda < r$, the absorbing state $\rho^* = 0$ is the only stationary point and it’s stable. If $\lambda > r$, $\rho^* = 0$ is still a stationary state, but it is unstable. There is a new

stationary state $\rho_2^* = a/\lambda = 1 - r/\lambda$ which is stable. So in mean field theory, $(\frac{\lambda}{r})_c = 1$ in any dimension. Simulations give $(\frac{\lambda}{r})_c \approx 3.298$ in $d = 1$.

Let's give a more systematic account in $d = 1$ by writing the equations of motion for the local density of active sites,

$$\rho_i(t) \equiv \langle s_i(t) \rangle. \quad (7.4)$$

Let's ignore the recovery process for now. Then the update rule is: If $s_i(t) = 1$, then $s_{i-1}(t+1) = A_i(t)$, a random variable which is 1 with probability p (related to the rate λ above) and 0 otherwise. And $s_{i+1}(t+1) = B_i(t)$, an independent random variable which is 1 with probability p . A given site s_i at time $t+1$ then fails to be occupied if both of its possible parent sites fail to infect it:

$$1 - s_i(t+1) = \delta(\text{parent at } i-1 \text{ fails}) \cdot \delta(\text{parent at } i+1 \text{ fails}) \quad (7.5)$$

$$= (1 - s_{i-1}(t)A_i(t))(1 - s_{i+1}(t)B_i(t)). \quad (7.6)$$

(Here $\delta(x)$ is the 'indicator function' which is 1 if x is true and 0 otherwise.) So

$$s_i(t+1) = s_{i-1}(t)A_i(t) + s_{i+1}(t)B_i(t) - s_{i-1}(t)s_{i+1}(t)A_i(t)B_i(t). \quad (7.7)$$

Taking expectation values of the BHS gives

$$\rho_i(t+1) = p(\rho_{i-1}(t) + \rho_{i+1}(t)) - p^2 \langle s_{i-1}(t)s_{i+1}(t) \rangle. \quad (7.8)$$

The last term avoids the overcounting that happens if both parents infect the same site – this event is called 'coagulation'. It is not a priori determined by $\rho_i(t)$. This is always what happens if we try to determine the dynamics of n -point functions of an interacting system – they depend on the $n+1$ -point functions. The name for this phenomenon is 'BBGKY hierarchy'. Mean field theory is a way to brutally truncate this infinite chain of dependence by pretending that

$$\langle s_{i-1}(t)s_{i+1}(t) \rangle \stackrel{\text{MFT}}{\approx} \langle s_{i-1}(t) \rangle \langle s_{i+1}(t) \rangle = \rho_{i-1}(t)\rho_{i+1}(t). \quad (7.9)$$

With this approximation, we have

$$\rho_i(t+1) \stackrel{\text{MFT}}{\approx} p(\rho_{i-1}(t) + \rho_{i+1}(t)) - p^2 \rho_{i-1}(t)\rho_{i+1}(t). \quad (7.10)$$

You recognize the first term as describing diffusion. Let's Taylor expand everything in sight:

$$\rho_i(t+1) = \rho(x, t) + \partial_t \rho dt, \quad \rho_{i\pm 1}(t) = \rho(x, t) \pm \partial_x \rho dx + \frac{1}{2} \partial_x^2 \rho dx^2. \quad (7.11)$$

Our mean field equation becomes

$$\partial_t \rho = a\rho - \lambda\rho^2 + D\vec{\nabla}^2 \rho, \quad (7.12)$$

where

$$a = \frac{2p-1}{dt}, \quad \lambda = \frac{p^2}{dt}, \quad D = p \frac{dx^2}{dt}. \quad (7.13)$$

So the mean field critical point is at $p = \frac{1}{2}$. Here I've ignored terms of the form $(\partial_x \rho)^2 dx^2$. Why should I do that? Well, we would like to understand the transition that occurs when $a \rightarrow 0$. Near that transition ρ itself is small. Furthermore, we expect the transition to enjoy scale invariance. Demanding that the three terms $\partial_t \rho, \vec{\nabla}^2 \rho, \rho^2$ are all of the same order, they should all scale the same way. You see that we have to scale time and space differently, in particular the dynamical critical exponent $z = 2$ in mean field theory. So if under a scale transformation $\rho(x, t) \mapsto s^{-\Delta_\rho} \rho(sx, s^2 t)$, we demand that the three leading terms scale the same way, we conclude that $\Delta_\rho = 2$. So the eom scales as

$$s^{-2} a \rho + s^{-4} \left(\partial_t \rho + \lambda \rho^2 - D \vec{\nabla}^2 \rho \right) + s^{-6} \Upsilon (\partial_x \rho)^2. \quad (7.14)$$

That is: the $a\rho$ term is a relevant perturbation of the fixed point, and the term I ignored (with coupling Υ) is an irrelevant perturbation.

Now what are the corrections to this MFT approximation? The thing we are ignoring is fluctuations. One way to fix this is to add them back in and replace our eom with

$$\partial_t \rho = a\rho - \lambda\rho^2 + D\vec{\nabla}^2 \rho + \eta, \quad (7.15)$$

where η is some kind of noise, $\langle \eta \rangle = 0$. What distribution governs the noise η ? It is meant to model the demographic fluctuations of active sites, but such fluctuations are completely absent in the absorbing state where $\rho = 0$. An inspired guess is that we can take it to be Gaussian, but demand that it goes away when $\rho \rightarrow 0$:

$$\langle \eta(t) \eta(t') \rangle = \Gamma \rho(x, t) \delta(x - x') \delta(t - t'). \quad (7.16)$$

(We could have put a more complicated polynomial in ρ , but terms of higher order than linear will be irrelevant.)

To see where this comes from more systematically, let's start again. We coarse-grain the master equation a bit to write it in the form of a reaction-diffusion system. So let $n_i \equiv$ the number of active particles in cell i of our grid, and we will keep track of their probability distribution $P(n_i, t)$. The allowed process are

$$\text{branching: } A \rightarrow A + A, \quad W(n \rightarrow n+1) = an \quad (7.17)$$

$$\text{death: } A \rightarrow \emptyset, \quad W(n \rightarrow n-1) = \mu n \quad (7.18)$$

$$\text{coagulation: } A + A \rightarrow A, \quad W(n \rightarrow n-1) = bn(n-1) \approx bn^2 \quad (7.19)$$

$$\text{diffusion: } A_i \rightarrow A_j, \quad W_{\langle ij \rangle}(n_i, n_j \rightarrow n_i-1, n_j+1) = Dn_i. \quad (7.20)$$

Ignoring space (and therefore the diffusion process) again for a bit, the total rate of increase is $T^+(n) = an$ and the total rate of decrease is $T^-(n) = \mu n + bn(n-1)$. So the master equation is

$$\partial_t P = \sum_{S=\pm 1} \underbrace{(\mathcal{E}^{-S} - 1)}_{=e^{-S\partial_n} - 1} (T^S P) \quad (7.21)$$

Using the Taylor expansion

$$e^{-S\partial_n} - 1 \approx -S\partial_n + \frac{1}{2}S^2\partial_n^2 + \dots, \quad (7.22)$$

this is

$$\partial_t P = -\partial_n (T^+(n) - T^-(n)) P(n) + \frac{1}{2}\partial_n^2 (T^+(n) + T^-(n)) P(n) = -\partial_n (A(n)P(n)) + \partial_n^2 (D(n)P(n)) \quad (7.23)$$

with

$$A(n) = T^+(n) - T^-(n) = (a - \mu + b)n + bn^2 \quad (7.24)$$

$$D(n) = 2(T^+(n) + T^-(n)) = (a + \mu - b)n + bn^2. \quad (7.25)$$

Notice that at $n = 0$, $A = D = 0$, as appropriate for the absorbing state. This FP equation is equivalent to the Langevin equation

$$\partial_t \rho = A(\rho) + \eta, \quad \langle \eta \rangle = 0, \langle \eta(t)\eta(t') \rangle = D(\rho)\delta(t-t'). \quad (7.26)$$

If we put back the spatial dependence, we'll get the diffusion term.

One more model worth mentioning is the Stavskaya automaton. Take a chain of spins, $\eta_i(t) = \pm 1$. The update rule is: with probability $\epsilon \in [0, 1]$, $\eta_i(t+1) = 1$, otherwise $\eta_i(t+1) = \eta_{i-1}(t)\eta_i(t)$. It is interesting because it is simple, it was historically the first model, and it is a sort of edge theory of the Toom rule: it's what you get under the Toom rule evolution for the bottom row if you fix all the spins except for the bottom row to be up.

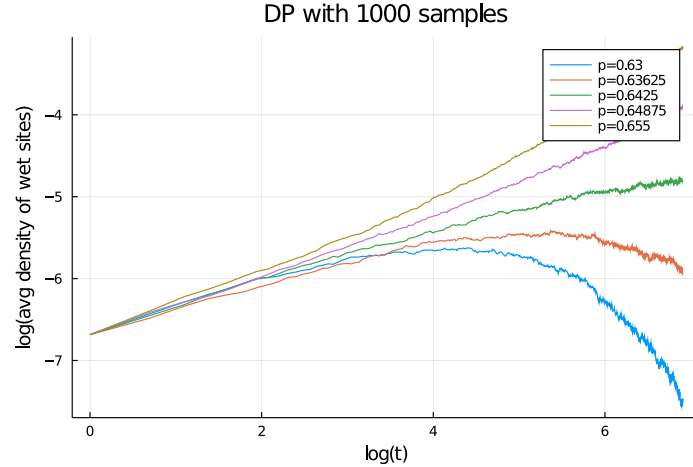
In what sense is DP generic? The field theory with equation of motion

$$\partial_t \rho = D\vec{\nabla}^2 \rho + a\rho - \lambda\rho^2 + \eta \quad (7.27)$$

has no symmetries except translations, and contains all possible relevant and marginal terms involving this single real scalar field. If we impose extra symmetries, there are other universality classes. For example, the class called *directed Ising* has a $\mathbb{Z}_2$ symmetry acting on the field ρ , and hence the ρ^2 term is forbidden, and we must include the ρ^3 term (like in the eom for ϕ^4 theory). Similarly, there are universality classes associated with larger symmetries, analogous to the Potts models.

We'll see below that, surprisingly, the critical theory of DP actually *does* have a $\mathbb{Z}_2$ symmetry.

In the figure at right, you can see the results of a simulation of the model described in (7.1). From this figure we learn at least two things. First, indeed there is a phase transition, where the long-time survival probability goes from zero to nonzero. We should try to understand its universal features. One example of such a universal feature, which we should try to understand, is: At the critical coupling, the average density of surviving sites (starting from a single seed) is a power-law in time.



Field theory of directed percolation. Following the MSR procedure we've already done several times, starting from (7.27), we can write an action for the field theory of DP in terms of the density field ρ and a response field ψ .

$$S[\rho, \psi] = \int dt d^d x \left(\psi \left(\partial_t \rho - D \vec{\nabla}^2 \rho - r \rho - \lambda \rho^2 \right) + \Gamma \rho \psi^2 \right) \quad (7.28)$$

Here Γ comes from the noise kernel, and λ comes from the reaction rate. By rescaling $(\rho, \psi) \rightarrow (K\rho, K^{-1}\psi)$ we can set $\Gamma = \lambda$. Then, you can see that this model has an emergent $\mathbb{Z}_2$ symmetry acting by $(\rho(t), \psi(t)) \mapsto (\psi(-t), \rho(-t))$. I call it emergent because it would be broken if we included irrelevant corrections like $u\rho^3$ in the eom. The symmetry holds not just at the critical point but rather for any r ; however, it is only in the neighborhood of the critical point that we expect to be able to ignore irrelevant terms which will break the symmetry. This symmetry has a bit in common with the time-reversal symmetry (2.161) from which we derived the FDT from the MSR path integral for Brownian motion – indeed it will also relate correlations $C = \langle \rho \rho \rangle$ to response functions $G = \langle \rho \psi \rangle$.

[End of Lecture 19]

What is the meaning of this symmetry? [Hinrichson page 863] $\rho(x, t)$ and $\psi(x, t)$ can be regarded as having dual roles. $\rho(x, t)$ is the density of active sites, now, which must have descended from some lineage of active sites in the past. The

response field $\psi(x, t)$ is instead related to the probability that a particle at (x, t) will have ancestors that survive in the distant future. Let me explain this mysterious statement.

The usual way of thinking about *e.g.* the contact process is forward in time, in terms of descendants: We start with particles and watch them spread, and use $\rho(x, t)$ to keep track of the density. The possible events are branching ($A \rightarrow 2A$) and coagulation ($A + A \rightarrow A$).

Instead, we can think of the history backwards: Start with a particle at time t and trace its lineage backward to time 0. If a particle at t came from a branching event ($A \rightarrow 2A$), its lineage merges with another lineage into one ancestor. If a particle at t came from a coagulation event ($A + A \rightarrow A$), its lineage splits into two ancestors when going backward. The equations of motion for ψ (obtained from $\frac{\delta S}{\delta \rho(x, t)}$) instead

$$-\partial_t \psi = D \vec{\nabla}^2 \psi - r\psi - \lambda \psi^2 + \zeta \quad (7.29)$$

describes evolution back in time. Why should we think about equations of motion going back in time? There are some contexts where this is the natural thing to do: suppose we have some goal for the future and we want to know how to reach it. Suppose further we are in the unlikely circumstance that we know the equations of motion. We should set final-state conditions at our goal and evolve them backwards to find what initial conditions we should choose to make our goal happen.

In the case of DP, a very sensible question that involves backwards evolution is: what is

$$Q(x, t) \equiv \text{Prob} \left(\begin{array}{l} \text{the active particle at } (x, t) \\ \text{is the ancestor of a particle} \\ \text{that survives indefinitely} \\ \text{(or until some fixed goal time)} \end{array} \right) ? \quad (7.30)$$

Just as we derived the eom for the density (*e.g.* for the contact process) by looking at the current state and evolving it forward, we can derive a backwards-in-time eom for $Q(x, t)$ by thinking about the possibilities for who was the ancestor of our particle at (x, t) . This equation, after some coarse-graining, is (7.29).

So this weird emergent symmetry of critical DP is saying that, at the critical point, the statistics of the descendants of a given active particle are the same as those of the statistics of the ancestors.

I can't resist mentioning that one very interesting place where such backwards-in-time evolutions arise is the theory of what are called mean-field games. This is a situation where we have a collection of agents, each trying to optimize their

own outcome (by doing the backward evolution) based on some knowledge of some kind of mean field (which evolves according to some forwards dynamics described by some Fokker-Planck equation for the density of agents). I'm sure you can think of many problems that might usefully be described by such a model. See [this](#) brilliant summary by Terry Tao for a bit more.

The first thing we can do with this action is dimensional analysis to answer the burning question: in which dimensions is the mean field theory we did above going to give the right answer? We will demand that each term of the action is dimensionless (since it gets exponentiated). First let's rescale t to set $D = 1$. Then

$$1 = L^d T[\psi \partial_t \rho] = L^d[\psi \rho], \quad 1 = L^d T[\psi \vec{\nabla}^2 \rho] = L^{d-2} T[\psi \rho] \quad (7.31)$$

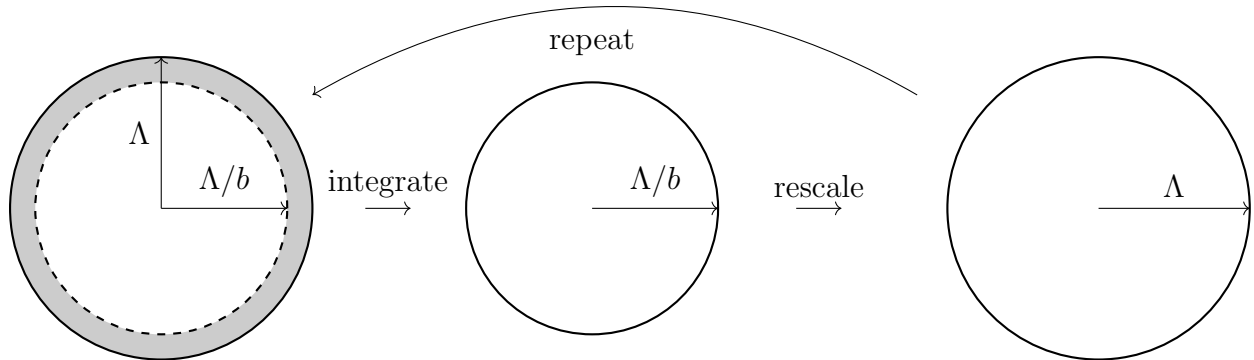
$$1 = L^d T[\psi \rho r], \quad 1 = L^d T[\psi \rho^2 \lambda], \quad 1 = L^d T[\psi^2 \rho \lambda] \quad (7.32)$$

The first two equations tell us that $T = L^2$. Then $[\psi] = L^{-d/2}$, $[\rho] = L^{-d/2}$ and hence $[r] = L^{-2}$ and r is a relevant perturbation in any dimension, and

$$[\lambda] = L^{\frac{d-4}{2}} \quad (7.33)$$

so the upper critical dimension is $d_{uc} = 4$. That is, for $d > 4$, the coupling λ has dimensions of a positive power of length, which means that its effects get smaller and smaller as we ask questions about longer and longer wavelengths – it is irrelevant. This means that for $d > 4$ mean field theory gives the right answer. For $d < 4$, we have to work harder. For d just a little bit less than d_{uc} , $d = 4 - \epsilon$, with ϵ small, the coupling is just a little bit relevant. So it creates a slow, controllable RG flow away from $\lambda = 0$ that we can study using perturbation theory in λ .

Perturbative renormalization group. As in chapter 1, an RG transformation has two parts. First, we integrate out a set of fast modes to get an answer that depends on the slow modes. Then we rescale everything so that we can compare the result to the original description.



We're going to study field theory of $\rho, \psi = \tilde{\rho}$ with a cutoff on the wavenumber: $|\vec{k}| \leq \Lambda$. Think of it as a simplified version of studying the field on a lattice. The fast modes are the ones with wavenumber close to the cutoff. (Let me denote the fields collectively as ϕ .) That is,

$$\phi(x) = \int^{\Lambda} d\vec{k} e^{i\vec{k}x} \phi_k = \phi^< + \phi^> \quad (7.34)$$

where the fast modes are

$$\phi^> \equiv \int_{|\vec{k}| > \Lambda/b}^{|\vec{k}| < \Lambda} d\vec{k} e^{i\vec{k}x} \phi_k \quad (7.35)$$

where $b > 1$ parametrizes how big is our RG step, and the slow modes $\phi^< \equiv \phi - \phi^>$ is everything else.

Let's do the second step first. After integrating out the modes with $|\vec{k}| > \Lambda/b$, we want to rescale wavenumbers so that our cutoff is Λ again. Under this rescaling, the fields will transform:

$$\rho(\vec{x}, t) \mapsto b^{\chi} \rho(b^{-1} \vec{x}, b^{-z} t) \equiv \rho'(\vec{x}, t). \quad (7.36)$$

Here χ and z are exponents to be determined. By the symmetry exchanging ρ and ψ , ψ must transform the same way. The action is

$$S[\rho, \psi] = \int dt d^d x \psi(x) (\tau \partial_t - D \partial_x^2 - r) \rho(x) + \int dt d^d x \lambda \psi(x) \rho(x) (\psi(x) - \rho(x)) \equiv S_0 + S_{\text{int}} \quad (7.37)$$

where I introduced couplings τ and D that can be removed by choosing units for time and space, but will be useful for a while. The free part of the action transforms as (let $\vec{x}' \equiv b^{-1} \vec{x}, t' \equiv b^{-z} t$) – notice that the thing that is changing is the degrees of freedom ψ, ρ , not the coordinates –

$$S_0[\rho, \psi] \mapsto S_0[\rho', \psi'] \quad (7.38)$$

$$= \int dt d^d x b^{2\chi} \psi(x') (\tau \partial_t - D \partial_x^2 - r) \rho(x') \quad (7.39)$$

$$= \int dt' d^d x' b^{d+z+2\chi} \psi(x') (\tau b^{-z} \partial_{t'} - D b^{-2} \partial_{x'}^2 - r) \rho(x') \quad (7.40)$$

$$\stackrel{\text{relabel}}{=} \int dt d^d x \psi(x) (\tau' \partial_t - D' \partial_x^2 - r') \rho(x) \quad (7.41)$$

where

$$\tau' \equiv \tau b^{d+2\chi}, \quad D' \equiv D b^{d+2\chi+z-2}, \quad r' \equiv r b^{d+z+2\chi}. \quad (7.42)$$

The interaction term transforms as

$$S_{\text{int}}[\rho, \psi] \mapsto S_{\text{int}}[\rho', \psi'] = \int dt d^d x \lambda' \psi(x) \rho(x) (\psi(x) - \rho(x)) \quad (7.43)$$

with $\lambda' \equiv b^{3\chi+2+z}\lambda$.

We're going to want to do the RG step just a little at a time, so we let $b \equiv e^\ell \sim 1 + \ell$ be close to 1. Then for $\ell \ll 1$ we have

$$\tau' = \tau(1 + \ell(2\chi + d)) \quad (7.44)$$

$$D' = D(1 + \ell(2\chi + d + z - 2)) \quad (7.45)$$

$$r' = r(1 + \ell(2\chi + d + z)) \quad (7.46)$$

$$\lambda' = \lambda(1 + \ell(3\chi + d + z)) . \quad (7.47)$$

Now we do step 1 in $d = 4 - \epsilon$ dimensions, treating λ as small. Let me denote the fields collectively as ϕ . Then doing the integral over the fast modes gives

$$\int D\phi e^{-S[\phi]} = \int D\phi^< \int D\phi^> e^{-S[\phi^>, \phi^<]} \equiv \int D\phi^< e^{-S_{\text{eff}}[\phi^<]} . \quad (7.48)$$

We are going to find S_{eff} in perturbation theory about the Gaussian theory where $\lambda = 0$. Observe that in the quadratic parts of the action, $\phi^<$ and $\phi^>$ do not mix:

$$S_0[\phi^>, \phi^<] = S_0[\phi^>] + S_0[\phi^<] . \quad (7.49)$$

Let me call $\delta S \equiv S_{\text{int}}[\phi^>, \phi^<] - S_{\text{int}}[\phi^>, 0]$. Then

$$S_{\text{eff}}[\phi^<] = S[\phi^> = 0, \phi^<] + \delta S_{\text{eff}}[\phi^<] \quad (7.50)$$

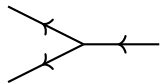
where the first term is just the parts independent of the fast variables and

$$\delta S_{\text{eff}}[\phi^<] = -\log \langle e^{-\delta S} \rangle_> = \langle \delta S \rangle_> - \frac{1}{2} (\langle \delta S \rangle_>^2 - \langle \delta S^2 \rangle_>) + \frac{1}{6} \langle \delta S^3 \rangle + \dots \quad (7.51)$$

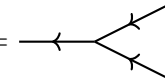
is the bit produced by integrating out the fast modes. This is the cumulant expansion: the terms that appear here are the *connected* correlation functions of δS . They are just Gaussian integrals, which we can compute using the Wick expansion (2.9). The ‘propagator’, *i.e.* the kernel with which we do the Wick contractions, is

$$G_0(\omega, k) \delta(\omega + \omega') \delta^d(k + k') = \langle \rho(\omega, k) \psi(\omega', k') \rangle_0 , \quad G_0(\omega, k) = \frac{1}{i\omega + Dk^2 + r} . \quad (7.52)$$

It is useful to introduce a diagrammatic notation to keep track of the Wick contractions. We denote by an oriented line the propagator $G_0(\omega, k) = \text{---}\leftarrow$. We denote by a 3-point vertex with two arrows out and one arrow in, the interaction $\int \rho \rho \psi$, which produces a factor of $-\lambda =$



two arrows in, associated with $\int \rho \psi \psi$ produces a factor of $+\lambda =$



Incidentally, Heinrichson claims that, in critical DP at least, we can interpret the Feynman diagrams as literally depicting the DP process, specifically along the backbone of the percolation. Do you buy this? (In the case of ordinary (not directed) bond percolation, this analogy is pursued [here](#).)

As long as our initial action had all possible marginal and relevant terms respecting the symmetries, the effective action $S_{\text{eff}}[\phi^<]$ will be of the same form, just with some new values of the couplings. Let me just mention the important contributions; of course irrelevant operators will also be generated but we don't need to worry about them.

The cubic coupling is corrected by

$$\delta S_{\text{eff}}[\phi^<] \ni \int d^d x dt \rho^2(x, t) \psi(x, t) \delta \lambda \quad (7.53)$$

with

$$\delta \lambda = \text{diagram} \quad |_{\text{external wavenumbers}=0} \quad (7.54)$$

$$= -a\lambda^3 \int_{\Lambda/b}^{\Lambda} d^d q \int_{-\infty}^{\infty} d\omega G_0(q, \omega)^2 G_0(-q, -\omega) \quad (7.55)$$

$$= -a\lambda^3 \int d^d q \int d\omega \left(\frac{1}{i\omega\tau + Dq^2 + r} \right)^2 \frac{1}{-i\omega\tau + Dq^2 + r} \quad (7.56)$$

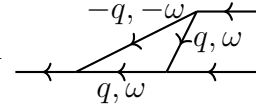
$$= -\frac{a\lambda^3}{\tau} \int d^d q \left(\frac{1}{2(Dq^2 + r)} \right)^2 \quad (7.57)$$

$$\stackrel{\Lambda^2 \gg r}{\equiv} -\frac{a\lambda^3}{4D^2\tau} \underbrace{\int_{\Lambda/b}^{\Lambda} \frac{dq}{q}}_{=\log b} \underbrace{\frac{\Omega_{d-1}}{(2\pi)^d}}_{\equiv K_d} \quad (7.58)$$

$$= -\frac{a\lambda^3}{4D^2\tau} K_d \log b + \mathcal{O}(\Lambda^{-2}) \quad (7.59)$$

We set the external wavenumbers to zero because the terms where they are nonzero just generate terms in S_{eff} involving the same fields but with more derivatives. Here a is a numerical number counting how many contractions lead to this amplitude. (It

is $a = 2$.) We did the frequency integral by residues. The diagram



corrects the $\rho\psi^2$ coupling by the same amount.

So the coupling after steps 1 and 2 is

$$\lambda'' = \lambda' + \delta\lambda + \dots \quad (7.60)$$

Here the leading term in δr is

$$I(0,0) = -\frac{1}{\tau} \int_{\Lambda/b}^{\Lambda} d^d q \frac{1}{2(Dq^2 + r)} = -\frac{K_d}{2\tau D} \int_{\Lambda/b}^{\Lambda} \frac{q^3 dq}{q^2 + r/D} \stackrel{\Lambda^2 \gg r/D}{=} -\frac{K_d}{2\tau D} \Lambda^2 \log b + \dots \quad (7.73)$$

These are the only terms in δS_{eff} that matter to this order.

Now we put together the pieces:

$$\tau'' = \tau' - \frac{\lambda^2}{2} \frac{\tau}{2} \frac{K_d}{2\tau D^2} \log b \quad (7.74)$$

$$D'' = D' + \frac{\lambda^2}{2} \frac{D}{4} \frac{K_d}{2\tau D^2} \log b \quad (7.75)$$

$$r'' = r' + \frac{\lambda^2}{2} \Lambda^2 \frac{K_d}{2\tau D} \log b \quad (7.76)$$

$$\lambda'' = \lambda' - \frac{a\lambda^3}{2D^2} \frac{K_d}{2\tau} \log b \quad (7.77)$$

So using $\partial_\ell = b\partial_b = \partial_{\log b}$,

$$\partial_\ell \tau = \tau \left(2\chi + d - \frac{\lambda^2}{2} \frac{1}{2} \frac{K_d}{2\tau D^2} \right) \quad (7.78)$$

$$\partial_\ell D = D \left(2\chi + d + z - 2 + \frac{\lambda^2}{2} \frac{1}{4} \frac{K_d}{2\tau D^2} \right) \quad (7.79)$$

$$\partial_\ell r = r \left(2\chi + d + z - \frac{\lambda^2}{2} \Lambda^2 \frac{K_d}{2\tau D} \right) \quad (7.80)$$

$$\partial_\ell \lambda = \lambda \left(3\chi + d + z - \frac{a\lambda^2}{2D^2} \frac{K_d}{2\tau} \right). \quad (7.81)$$

In these complicated expressions only two dimensionless combinations appear

$$S_1 \equiv \frac{\lambda^2 K_d}{16\tau D^2}, \quad S_2 \equiv -\frac{\lambda^2 \Lambda^2 K_d}{4\tau D r} \quad (7.82)$$

Now let's fix τ and D to be unity by a choice of time and space units. This implies that $\partial_\ell \tau = 0 = \partial_\ell D$ so

$$2\chi + d - 2S_1 = 0, \quad 2\chi + d + z - 2 - S_1 = 0. \quad (7.83)$$

This says

$$\chi = S_1 - d/2, \quad z = 2 - S_1. \quad (7.84)$$

The remaining equations say

$$\partial_\ell r = r(2\chi + d + z - S_2) = r(S_1 - S_2 + 2) \quad (7.85)$$

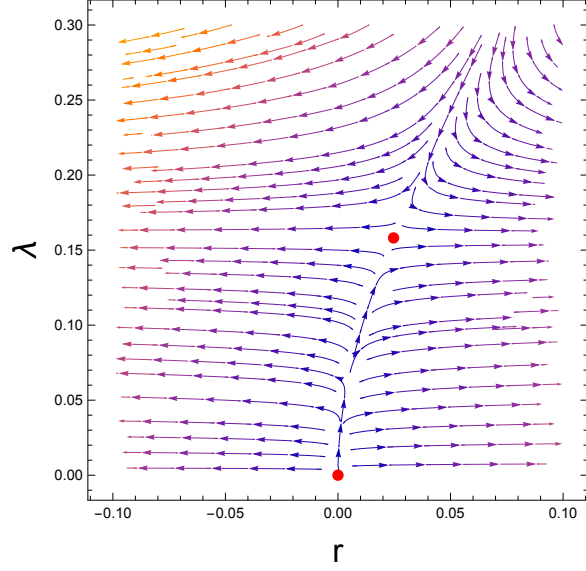
$$\partial_\ell \lambda = \lambda(3\chi + d + z - 4a S_1) = \lambda(2 - d/2 - (4a - 2)S_1) = \lambda(\epsilon/2 - 6S_1). \quad (7.86)$$

These equations have fixed points at $\lambda = r = 0$ (the absorbtive fixed point) and

$$S_1^* = \frac{\epsilon}{12}, \quad S_2^* = 2 + S_1^* \quad (7.87)$$

Expanding around this point, you can see (see the figure below) that there is one unstable direction (mostly r), so this is the critical point. Plugging this into (7.84) gives

$$\chi = -2 - \epsilon \frac{7}{12}, \quad z = 2 - \frac{\epsilon}{12}. \quad (7.88)$$



The stability of the fixed points is determined by the eigenvalues $\lambda_{1,2}$ of the Jacobian matrix

$$M_{IJ}(r^*, \lambda^*) \equiv \frac{\partial \beta^I}{\partial g^J} \big|_{g=g^*}. \quad (7.89)$$

Because this matrix is upper triangular, its eigenvalues are the diagonal entries:

$$\partial_r r (2 + S_1 - S_2) \big|_* = 2 - \epsilon/4 + \mathcal{O}(\epsilon^2), \quad \partial_\lambda \lambda (\epsilon/2 - 6S_1) \big|_* = -\epsilon + \mathcal{O}(\epsilon^2). \quad (7.90)$$

The positive eigenvalue y_r corresponds to the direction we have to tune to reach the critical point and therefore determines the correlation-length critical exponent of the DP critical point

$$\nu = 1/y_r = 2 + \epsilon/4 + \mathcal{O}(\epsilon^2). \quad (7.91)$$

Another question we can answer after all this heavy lifting is: what is the critical exponent for the fraction of active sites at time t , starting from a single seed at time $t = 0$? The probability that the site x is active at time t , given that we started from a single seed at $x = 0, t = 0$ is the response function $G(x, t) = \langle \rho(x, t) \psi(0, 0) \rangle$. So the number of active sites at time t , starting from a single seed is the integral of the response function:

$$N(t) = \int d^d x \langle \rho(x, t) \psi(0, 0) \rangle. \quad (7.92)$$

At the critical point, this is scale invariant, *i.e.* invariant under the transformation (7.36):

$$N(t) = \int d^d x \langle b^x \rho(b^{-1}x, b^{-z}t) b^x \psi(0, 0) \rangle = b^{d+2\chi} \int d^d x' \langle \rho(x', b^{-z}t) \psi(0, 0) \rangle = b^{d+2\chi} N(b^{-z}t) \quad (7.93)$$

so the power law must satisfy²⁵

$$N(t) \sim t^\eta = b^{d+2\chi} (b^{-z}t)^\eta = b^{d+2\chi-\eta z} t^\eta \quad (7.95)$$

which requires the exponent to be²⁶

$$\eta = \frac{d+2\chi}{z}. \quad (7.96)$$

Excitable media. To help justify how much time we've spent on directed percolation, let's talk about another microscopic model that exhibits a continuous transition in this class, called Greenberg-Hastings automaton or SIR model. It is a model of *excitable media*. Each site of some graph has $n+1$ possible states, $s_i \in \{0, 1, \dots, n\}$ to which we can assign the following interpretations:

²⁵Actually this relation $N(b^z t') = b^{d-2\chi} N(t')$ implies that $N(t)$ is a power law: Set $t' = 1$ and $t = b^z$ (so $b = t^{1/z}$):

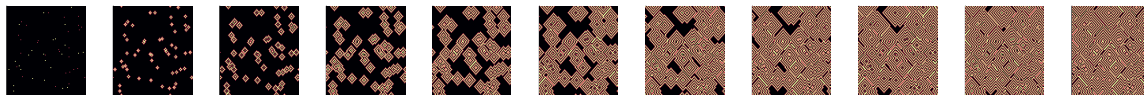
$$N(t) \sim (t^{1/z})^{d-2\chi} = t^{\frac{d-2\chi}{z}}. \quad (7.94)$$

²⁶Beware that you may find someone else call $-\chi$ what I've called χ here.

System/ Medium	State 0: Resting (Susceptible)	State 1: Excited (Active)	State 2...n: Refrac- tory (Recovering)
Forest Fires	Unburnt Tree (Fuel) The tree is grown, dry, and ready to ignite if a neighbor burns.	Burning (Fire) The tree is combust- ing and actively re- leasing heat to ignite neighbors.	Ash / Char Burnt remains. No fuel is left. The site must regrow before it can burn again.
Neurons & Cardiac Tis- sue	Resting Potential Cell is polarized ($\approx -70\text{mV}$). Sodium channels are ready to open.	Action Potential Voltage spikes. The cell fires, transmit- ting electrical signals to neighbors.	Hyperpolarized Ion channels are in- activated. The cell cannot physically fire again until reset.
Epidemics (SIR Model)	Susceptible Healthy individual with no immunity to the pathogen.	Infected The individual is sick and actively spreading the virus to contacts.	Recovered / Re- moved Immune (antibodies) or deceased. Cannot catch or spread the disease.
Chemical Waves (BZ Reaction)	Reagents The mix is in a re- duced state, calm and ready for oxidation.	Reacting Front Rapid autocatalytic oxidation occurs; the color shifts abruptly.	Oxidized Tail Reagents are ex- hausted or inhibitor concentration is high. Must chemically reset.

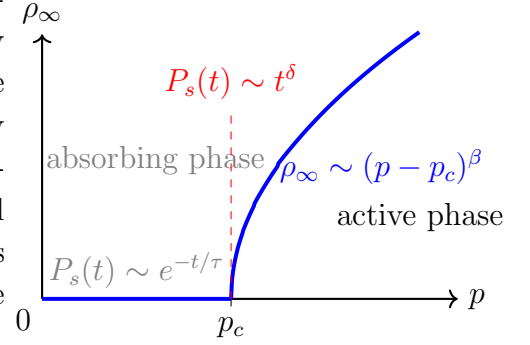
The update rule is very simple. If a site is in state $k > 1$ it goes (deterministically) to state $k + 1 \bmod n$. Think of ($k = 1 \rightarrow k = 2$) as the fire burns out or trees grow back. If a site i is in state 0, $s_i(t) = 0$, we look at the neighboring sites (there is some choice to make about who counts as a neighbor); if at least one neighbor is active (in state 1), $s_i(t + 1)$ is set to 1 with probability p . The transmission probability p is a parameter of the model.

Here's what happens. For p large, any initial activity produces dramatic spreading waves of activity which annihilate when they collide. Here is $p = 1$:



A defect of some kind (like a site with a different value of p) produces spiraling waves. These patterns, at least superficially, look a lot like things that happen in the heart and in the brain. For smaller p , the randomness messes up the nice waves and produces a big mess of activity. Below some critical p all the activity dies. Notice that $s_i = 0$, $\forall i$ is an absorbing state.

Here is the phase diagram as a function of p . An order parameter, as above, is just the density of active sites (say sites in state 1). In the NESS, the density of each state is equal. In the subcritical phase the survival probability is $P(t) \sim e^{-t/\tau}$, in the supercritical phase, it (also exponentially) approaches a constant nonzero value. Near the critical point (from $p > p_c$) the steady-state density ρ_∞ behaves as $(p_c - p)^\beta$ with $\beta \simeq 0.58$ in $d = 2$. At the critical point, the survival probability is a power law in time $P(t) \sim t^\delta$.



The RG calculation of β follows from what we've already done. Let $\Delta = p - p_c$. Then

$$\rho_\infty(\Delta) = b^{-\chi} \rho(b^{y_r} \Delta) \quad (7.97)$$

where $y_r = 1/\nu$ is the eigenvalue of M in the relevant direction and ν is the correlation length exponent. It is $y_r = 2 - \epsilon/4 + \mathcal{O}(\epsilon^2)$ through one loop. Now choose the scaling parameter b so that $b^{y_r} \Delta = 1$, which gives

$$\rho_\infty(\Delta) = (\Delta^{-1/y_r})^{-\chi} \sim \Delta^{-\chi/y_r}. \quad (7.98)$$

So $\beta = -\chi/y_r$. Using our one-loop results $\chi = -2 + 7\epsilon/12$ and $y_r = 2 - \epsilon/4$ we estimate $\beta = 1 + \frac{17}{48}\epsilon + \mathcal{O}(\epsilon^2)$.

[This](#) paper has some nice simulations and shows the robustness of the DP criticality.

Now here comes a puzzle: I listed above a number of naturally-occurring systems that show features related to this model. And indeed many of them (forest fires, brains) do exhibit power laws in time. Does this mean that they are somehow automatically at the critical point? Let's look at the examples I mentioned in turn. The heart we definitely want to be supercritical $p > p_c$ and not in the absorbing phase of no activity. Under various circumstances, the parameter p drops close to p_c or below, there are spiral waves, or the waves simply die out and it's no good. For epidemics, the parameter p is determined by what people called R_0 , and public health officials work hard to keep the system in the subcritical phase. In forest fires, power laws are indeed observed in the number of fires of a given size. There is a story one can tell about self-tuning to criticality – in a version of the model where p is itself dynamical²⁷.

²⁷This version of the model not totally unrelated to the Drossel-Schwabl Forest Fire

If the forest is too sparse ($p < p_c$), fires don't spread and so trees keep growing and increasing the density (which increases p). Eventually it reaches above p_c and fires burn everything and we start over. This leads to oscillations around p_c . The brain is famously observed to sit at the “edge of chaos”. This means that brain activity does not happen by a constantly-active state like the supercritical behavior nor by total silence like in the absorbing state. Instead, there are bursts or avalanches of activity. Apparently, as a function of the total number S of neurons involved and the duration T , these do satisfy power law distributions $P(S) \sim S^{-1.5}$, $P(T) \sim T^{-2.0}$. And again there is a story we can tell about the tuning to criticality. Here there is definitely tuning. The story is something like: if the brain is too active (like seizures), synapses weaken to lower the activity, decreasing p ; if it is too quiet, synapses get stronger. And there are some information-theoretic statements that the critical distribution of firings somehow optimizes information processing.

Let me elaborate a bit on these statements about the optimality of criticality for information processing.

One concrete statement is that the critical point maximizes the dynamic range [Kinouchi-Copelli (and for an advertisement for this paper see [here](#)), Larremore et al]. Think of the excitable medium we are discussing as a *sensor*, like a bunch of neurons in your eye. Write a Langevin description of the system in terms of the activity density $\rho(x, t)$:

$$\partial_t \rho = -\vec{\nabla}^2 \rho + r\rho - \lambda\rho^2 + \xi + S. \quad (7.99)$$

Here I've added to the noise an external *stimulus* term S (which maybe doesn't average to zero). Think of it like the lightning starting forest fires.

To get started, let's regard $R \equiv \rho_{ss}$, the steady-state activity density, as the response of the sensor. And further, let's ignore space, so we throw away the $\vec{\nabla}^2 \rho$ term, and we also ignore the noise. Then $S = -r\rho_{ss} + \lambda\rho_{ss}^2$. In the subcritical regime, $r < 0$, $S = |r|\rho_{ss} + \lambda\rho_{ss}^2 \simeq |r|\rho_{ss}$, and we have $R = S/|r|$, a linear response law. At the critical point $r = 0$, in this mean-field description, $S \sim \lambda\rho_{ss}^2$ and

$$R \sim \sqrt{S/\lambda} \sim S^{1/2}, \quad \delta = \delta_{MF} = 2. \quad (7.100)$$

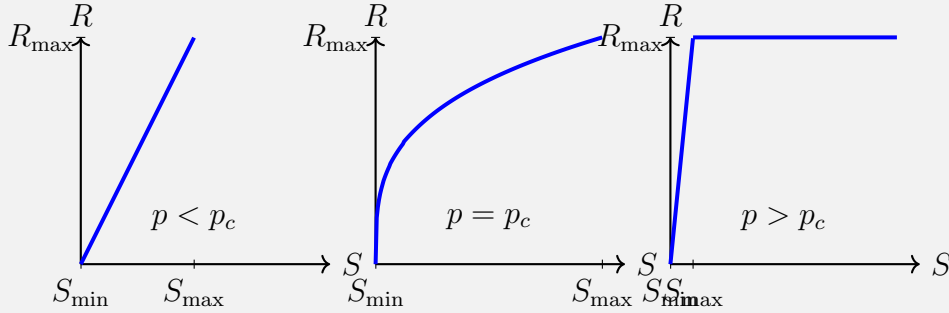
The actual behavior at the critical point is $R \sim S^{1/\delta}$ where δ is another critical exponent that we can compute in the ϵ expansion^a. δ is always bigger than one. Finally, in the supercritical regime $r > 0$, even without the stimulus, when $S = 0$, there is a nonzero activity $\rho_{ss} = r/\lambda > 0$, and the effect of S is a

Model or these papers: <https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.102.118110>
<https://arxiv.org/abs/0712.1003>.

small perturbation of this. The dynamic range is defined as the range of (log of) stimulus levels where dR/dS is above some threshold:

$$\Delta = \log S_{\max}/S_{\min} \quad (7.102)$$

where $S_{\max}$ is the value of the output that saturates the response, and $S_{\min}$ is the smallest value of the output that is detectable above the noise. In the supercritical range, any stimulus already maxes out the response. In the subcritical range, the response is linear, so $S_{\max}/S_{\min} = R_{\max}/R_{\min}$. But because $\delta > 1$, the response curve at the critical point bends down and we fit more values of S before we reach $S_{\max}$.



You might complain that in the subcritical regime R is only linear in S if we leave the stimulus on the whole time. What if, more realistically for the example I mentioned, the stimulus is transient? To do better, we can regard the total activity $R \equiv \int_0^\infty dt \rho(t)$ as the response of the sensor. Then in the subcritical regime, the activity in response to a perturbation at $t = 0$ is $\rho(t) \sim S e^{-t/\tau}$, and each seed creates a small isolated blob of activity, independent from all the others. So $R \propto S \times (\text{blob size})$. In contrast, at the critical point, an isolated input produces a fractal avalanche which spreads all over. This means that multiple inputs will overlap and try to redundantly activate the same sites (recall that this is the role of the nonlinear $-\lambda \rho^2$ term in the DP equation). The result is that R is sublinear in S . At $r > r_c$, $R = \int dt \rho(t)$ is infinite for arbitrarily small input.

You can see that the above desirable property of the critical point is related to the divergence of the susceptibility $\chi = \frac{\partial R}{\partial S}|_{S=0}$.

There are some other putative information-theoretic virtues of the critical point. The fact that it maximizes mutual information between different parts of the system can be useful. It maximizes the “statistical complexity” which is some measure of ability to do computation, which requires both stable storage of

information (optimized by the dead phase where nothing happens, but not by the active phase where everything is constantly being overwritten) and convenient transmission of information (which doesn't happen at all in the dead phase but happens well in the active phase).

^aIn fact, the RG calculation of δ is quite simple. The perturbation by the stimulus is a new operator in the MSR action:

$$S[\rho, \psi] \ni \int d^d x dt S\psi(x, t). \quad (7.101)$$

You can check that the new coupling S is not corrected by any diagrams, and only changes under step 2 of the RG by $S'' = Sb^{x-d-z}$. Using this, we find $\delta = 2 + \epsilon/6$.

Another reason for us to be interested in this possibility of SOC or *generic scale invariance* is that a generically-large correlation length would justify our use of a field theory description even when no one is doing fine tuning to stay near a critical point.

Beware Griffiths effects. However, there is a very important caveat about these observed power laws and possible implications about criticality. That is the phenomenon of *Griffiths* or *Rare-region* effects. The claim is that this can lead to power laws that are not because of criticality. (Warning: this is an area of physics with a lot of “~”s and not a lot of “=”s.)

Suppose that instead of having a uniform value of p for each cell in our simulation, we choose p for each cell from a broad distribution (could be iid) with some average value p_0 . This is called quenched disorder, quenched just means that the randomly-chosen ps are fixed while the dynamics happens. It could be patches of wet trees, or scarred regions of heart tissue, or the variation of synapse strengths over the brain.

Suppose that $p_0 < p_c$, so we might expect to be in the cold, dead phase where nothing happens and everyone dies out. Indeed most of the system will have subcritical couplings for which the activity dies out quickly. But there will be rare regions of supercritical couplings. Activity in such a region can last for a long time – let's estimate how long.

The idea is that we can get a power law from a sum of exponentials²⁸. For many choices of distribution of the ps (including iid), the probability of a rare region of linear size L is

$$P(L) \sim e^{-c_1 L^d}. \quad (7.103)$$

(For the iid case, this is because $P(L) = w^{\text{number of sites}}$ where $w = \int_{p>p_c} P(p)dp$ is the

²⁸Another context where the same thing happens is the mechanism of so-called ‘two-level systems’ used by Anderson, Halperin and Varma to explain the thermal properties of glasses; see *e.g.* [this paper](#) for a summary.

probability that one site is supercritical.) The time it takes the activity in the rare region to die out is

$$\tau(L) \sim e^{+c_2 L^d}; \quad (7.104)$$

if it reached the thermodynamic limit, it would be in the supercritical phase and last forever. (The fact that it is an exponential in the volume of the region follows because the death of the region requires a tunneling event where the sites in the rare region collectively fluctuate to zero, so the rate is determined by a saddle point of the action (as in Kramers' barrier escape problem) which is extensive.) So the survival probability of activity in a given region goes like $e^{-t/\tau(L)}$. The density of active sites can then be estimated as

$$\rho(t) \sim \int dL P(L) e^{-t/\tau(L)} \quad (7.105)$$

To see how the power law comes about from (7.105), notice that this integral is a huge struggle between the rapidly decaying $P(L)$ and the rapidly rising $e^{-t/\tau(L)}$, and so there is an optimal region size that dominates the integral, satisfying ($x \equiv L^d$)

$$0 = \partial_x (-c_1 x - t e^{-c_2 x}) \implies e^{-c_2 x_*} = \frac{c_1}{c_2 t}. \quad (7.106)$$

Notice that this saddle point condition says that the second term $t e^{-c_2 x_*} = \frac{c_1}{c_2}$ is independent of t . So the result is just

$$\rho(t) \sim t^{-\delta}, \quad \delta = \frac{c_1}{c_2}. \quad (7.107)$$

Notice that the exponent is determined by garbage non-universal quantities.

Given a critical point arising in a system without such quenched disorder, there is a criterion to decide whether the IR behavior is stable to making the couplings random, or whether instead it will be governed by rare-region effects. This *Harris criterion* for the clean critical point to remain stable is

$$d\nu > 2 \quad (7.108)$$

where $\nu = y_r^{-1}$ is the correlation length critical exponent of the clean critical point²⁹ For DP, $\nu \approx 1.097, 0.734, 0.581$ in $d = 1, 2, 3$. None of these passes this criterion. This

²⁹Here is a sketch of an argument. The idea is to think about a region whose linear size is the correlation length ξ , inside this region the spins basically agree, and compare the average distance to the critical point with the size of the fluctuations in the tuning parameter. If the latter is bigger, different parts of the system are in different phases and the randomness wins. The definition of the correlation length critical exponent ν is the power by which the correlation length diverges at the critical point $\xi \sim t^{-\nu}$ where $t = \frac{r-r_c}{r_c}$. So on the one hand, the distance to the critical point is $r - r_c \sim \xi^{-\nu}$. On the other hand, in a block of volume ξ^d , there are roughly $N \sim \xi^d$ independent samples of the disorder. The CLT says that the variance in the block couplings is then $\Delta r = \sigma_r / \sqrt{N}$,

means that in the presence of disorder, the critical point smears out into a Griffiths phase. In this phase, the activity decays like a power law $\rho(t) \sim t^{-\delta}$, but this exponent δ is not a universal constant of the critical point, but is determined by combinations of horrible details of the disorder distribution (like c_1, c_2 above).

The Griffiths phase still has power laws, just with non-universal exponents. Is the Griffiths phase optimal in the information theory senses we mentioned above? At least the statement about dynamic range is not true: it depends on divergence of the susceptibility at the critical point. But in the Griffiths phase the average susceptibility is finite, and the rare regions of activity are isolated. This means $R \sim S$. I'm not sure about other senses of the claims about the "edge of chaos".

7.2 "Self-organized criticality"

Recall that in the first chapter we talked about the two kinds of fixed points of the RG, ones with $\xi = 0$ and ones with $\xi = \infty$. Usually, a *stable* fixed point, one with no relevant operators has $\xi = 0$, no local degrees of freedom left. A *good* meaning of "self-organized criticality" would be an infinite-correlation-length fixed-point with no relevant parameters. Such things could exist! (The monster CFT is an example in $d = 1 + 1$. Maybe you could say a Fermi liquid is such a thing, if you ignore the BCS instability to superconductivity.) Certainly if we impose a certain symmetry, we know many examples of CFTs with no *symmetric* relevant operators, which are therefore stable as long as the symmetry is present (an example is an algebraic spin liquid phase in 2+1d, described by some number of Dirac fermions coupled to a U(1) gauge field).

In practice, this is not what it means. It is instead a brand name for a certain set of, nevertheless interesting, phenomena. A good definition, I think, is: a system where the parameter you are tuning to make it critical is not obvious. In ordinary critical points, the value of the critical coupling can be hard to find (Kramers and Wannier famously predict T_c for the square lattice Ising model using the duality transformation – that's a nontrivial thing). SOC is a name for a system where the critical point occurs where some parameter (one that you might not even talk about in the description of the model) is zero. So this makes it seem like no tuning is required.

Let me give two examples. The first is sandpiles. The degrees of freedom are an

where N is the number of sites in the block and σ_r is the variance of the disorder distribution. So the criterion for disorder to lose the competition is $\Delta r < \bar{r} - r_c$ or

$$\frac{\sigma_r}{\xi^{d/2}} < \xi^{-1/\nu} \quad (7.109)$$

which for large ξ gives (7.108).

integer height z_i of a pile of sand grains above each point in a graph; the update rule is: if any pile exceeds its neighbors by some threshold z_c , then it *topples*, meaning that the excess above its neighbors is redistributed to neighboring piles³⁰. This process then repeats until no more piles need to be toppled. At that point another sand grain is dropped on a random pile, which may or may not begin a new avalanche.

Sandpile models are the poster children for SOC. The distribution of avalanche sizes is a power law. In this case, the parameter that is secretly being tuned to a critical value is the *rate* at which the sand is deposited. If we don't wait for all avalanche activity to finish before depositing the next grain of sand, we get a supercritical mess of avalanches which do not exhibit nice power laws.

The second example is *invasion percolation*. The degree of freedom is some binary variable on a graph indicating whether a site is occupied by fluid or not. Each edge of the graph is assigned a 'reluctance' indicating how difficult it is for the fluid to move along that edge. The model proceeds by the following non-local update rule: at each step, look along the boundary of the set of occupied sites for the edge with the smallest reluctance, and occupy the site at the other end of it. The claim is that the resulting region occupied by the fluid is fractalicious.

In invasion percolation, the parameter that is being tuned is the equilibration rate of the fluid pressure. Only if the fluid is at uniform pressure at each time step will it manage to find the absolute weakest link along the invasion front. If the equilibration time is finite, the fluid will instead follow the local pressure gradients, which lead to a different phenomenon called viscous fingering.

Many examples with generic scale invariance share this property of involving an infinite separation of time scales. It seems like there is no tuning because all the experimenter has to do to achieve the tuning is *wait* long enough between time steps. Having said all this I should emphasize that earthquakes and other related natural phenomena actually do exhibit power laws over a long range of timescales! So assuming this separation of timescales is sometimes not such a bad approximation.

SOC is sometimes described as an absorbing-state transition with infinitely-many

³⁰There are various important specifications to be made of the model. The Bak-Tang-Wiesenfeld model I believe is: set z_c equal to the coordination number and topple exactly z_c grains, one to each neighbor. This model has the property that it does not matter the order in which one topples sites above the threshold; such a model is called an *abelian sandpile*. A different choice is z_c less than the coordination number and we introduce another stochastic element (in addition to the deposition location) to decide how to distribute the excess amongst the neighbors. Finally, on a finite domain, we also need to choose boundary conditions. One interesting choice in 1d is to never topple grains off the left edge, and let grains fall off the right end of the world. This will lead to a net flow of sand from left to right.

absorbing states. Any configuration of sand that is stable (i.e., every grain is below the toppling threshold) is an absorbing state for the avalanche dynamics. But (in the thermodynamic limit) there is a continuum of such stable states. A pile with a total slope of 15° is absorbing. So is a pile with 15.1° , 15.2° , and so on, all the way up to the critical slope. SOC is a system that, in a certain sense, because of this assumed separation of time scales, tunes itself to the edge of this infinite set of absorbing states.

However, an infinite set of absorbing states is not a sufficient condition for a new and interesting universality class beyond DP. Consider the Contact Process without diffusion or *pair contact process* (PCP). Sites on a grid can be active or inactive. An active particle can be annihilated, or it can interact with a neighbor to create a new particle. The state with no particles is an absorbing state. But so is any configuration of isolated particles. Since they are isolated, they can't interact to create new particles, and they can't be annihilated (in this model's rules). So the PCP has infinitely many absorbing states, not related to each other by any symmetry. Despite having $N = \infty$ asymmetric absorbing states, the Pair Contact Process transition is in the DP universality class.

You might reasonably ask: I said something about a universality class with $N = 1$ absorbing state and about one with $N = \infty$ absorbing states, what about other values of N ? If we assume a symmetry amongst the N absorbing states, there are distinct classes, called Directed Ising ($N = 2$) and Directed Potts ($N > 2$) classes.

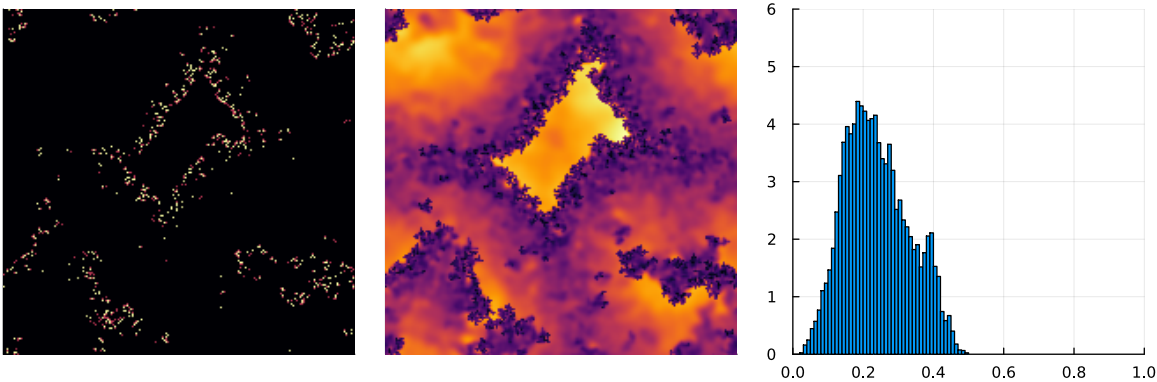
There is more to say about sandpile models, and more generally about driven noisy systems like them. I will say something about it in the section on surface growth.

SOC by homeostatic plasticity? There is a different sense in which criticality can be self-organized that is unique to dynamical systems – it is a mechanism that is not available at all to equilibrium systems. One could imagine that the dynamical rules have the property that the coupling that needs to be tuned is itself dynamical, and that the update rules are such that deviations from the critical point are corrected away by the dynamics. This is in line with the words I said when we first discussed the generic scale invariance of forest fires and brain waves and other things. So we could try to make these ideas more concrete by studying an understandable model that implements them. For example, we can think about the Greenberg-Hastings automaton with its parameter p that decides how excitable a site is. Let's introduce some 'plasticity' to this parameter. That is, we let each site have its own excitability p , and we update it at each time step according to the current situation: when the site i is active ($s_i > 0$), its p_i goes down, think of it as a depletion of resources; when the site is inactive $s_i = 0$,

its p_i grows, like some recovery of resources:

$$p_i(t+1) = p_i(t) + \underbrace{\frac{1}{\tau}(1 - p_i(t))}_{\text{Slow Recovery}} - \underbrace{u \cdot \delta_{s_i(t), \text{Active}} \cdot p_i(t)}_{\text{Fast Depletion}}. \quad (7.110)$$

If we do this, we indeed find that the distribution of the p_i approaches some steady-state distribution not too far from the critical value of p (or rather, its time-average is near the critical value – the distribution of p_i actually oscillates quite a lot). I am not sure whether the resulting non-equilibrium steady state is actually critical. (From my simulations, it looks like the distribution of the p_i s is broad, so even if we see power laws, it is hard to rule out Griffiths physics.



At left is an instantaneous configuration of s with $N = 3$ in what seems to be the dynamical steady state; in the middle is the corresponding instantaneous configuration of p_i ; at right is a normalized histogram of the p_i . The critical value of p for $N = 3$ in $d = 2$ is a little less than 0.2, which is indeed where the peak of the distribution of p_i is.) Some papers about this kind of thing, but with a lot more words about biology and neuroscience, are the ones by Levina Herrmann Geisel I cited above in footnote 27, which claim that this mechanism really works, but they are talking about models of neurons that I haven't understood yet. The Drossel-Schwabl Forest Fire Model does not quite achieve this goal; it introduces some mechanisms that tune the activity but the resulting criticality is ordinary (undirected) site percolation (if it is [indeed critical](#)).

So let me conclude this discussion with a list of possible answers to the question: why do we see generic scale invariance in some non-equilibrium systems?

1. The system flows to an infinite-correlation-length fixed point with no relevant operators. With Lorentz symmetry, such a thing is called a ‘[dead-end CFT](#)’.
2. Someone is tuning a parameter to a critical value. Perhaps the critical value is at some special place like $g = 0$ so that the existence of the parameter is not obvious, as in the examples above where it is a ratio of time scales.

3. SSB of a continuous symmetry. This is definitely what happens in surface growth as described by the Edwards-Wilkinson model or KPZ as described in §7.5 below. The idea is that the dynamical variable is the height $h(x, t)$ of some interface, and in the absence of *e.g.* Earth gravity, the dynamics is invariant under a uniform shift $h(x, t) \rightarrow h(x, t) + \epsilon$. This has the consequence that $h(x, t)$ only appears in the equations of motion via its derivatives.
4. Griffiths physics (*e.g.* (7.105)). This can definitely produce power laws with non-universal exponents. It seems to be what happens if we perturb the DP critical point by quenched disorder.
5. Homeostatic plasticity, as we just discussed. This is argued to happen in models of the brain; it would be nice to see a simpler model that realizes this mechanism.

It seems to me that different explanations are appropriate to different systems. I would love to have some more unified understanding.

[\[End of Lecture 20\]](#)

Let me describe the model studied by Levina, Herrmann and Geisel. It is very much like a sandpile model, but with continuous heights h_i at each site of an N -site graph. We take the graph to be fully connected, and associated a matrix of synapse strengths J_{ij} to each link. Note that J_{ij} and J_{ji} are independent. h_i and J_{ij} are the dofs. There are several parts of the dynamics:

- Drive: There is a slow drive by which the h_i increase. Let's think about the case where we randomly pick an i and increment $h_i \rightarrow h_i + \delta$.
- Toppling: If any $h_i > \theta$, we set $h_i = 0$ and increase its neighbors:

$$h_j \rightarrow h_j + J_{ji}. \quad (7.111)$$

This may push the neighbors over the threshold, and lead to an avalanche of activity.

- Depletion: Whenever a link J_{ij} is used to increase a neighbor, it gets depleted:

$$J_{ji} \rightarrow J_{ji} - u J_{ji}. \quad (7.112)$$

- Recovery: There is also a slow regeneration of resources:

$$\dot{J}_{ij} = (J_{\max} - J_{ij})/\tau. \quad (7.113)$$

So the model has parameters $N, u, J_{\max}, \tau, \delta, \theta$. In my simulations, I set

Parameter	Description
$N = 500$	Number of sites
$u = 0.2$	Depletion parameter (fraction lost per firing event)
$J_{\max} = 1.0$	Max synaptic strength
$\tau = 5000.0$	Recovery timescale (Needs to be slow relative to drive)
$\theta = 1.0$	Firing threshold
$\delta = 0.2$	Amount of “sand” added per deposition event

but the claim is that the critical behavior happens as long as $u\tau$ is large and we wait until the avalanche is over to do the next deposition event. Here is a mean field theory, using the method of words, that illustrates the mechanism: The synaptic dynamics reaches a steady state if

$$0 = \dot{J} = \text{recovery} - \text{depletion} \quad (7.114)$$

$$= (J_{\max} - J)/\tau - (uJR) \quad (7.115)$$

where R is the firing rate. From this we learn

$$R(J) = \frac{J_{\max} - J}{u\tau J}. \quad (7.116)$$

What is the firing rate? It satisfies

$$R = \text{events directly from drive} + \text{induced events} \quad (7.117)$$

$$= h_{\text{ext}} + \sigma(J)R \quad (7.118)$$

where h_{ext} is the driving rate, and $\sigma(J)$ is the *branching parameter*, defined to be the expected number of descendants of a firing event. We can estimate $\sigma(J) \sim \langle J \rangle K/\theta$ where K is the number of neighbors, but better is just to solve the self-consistency equation (7.118) for R :

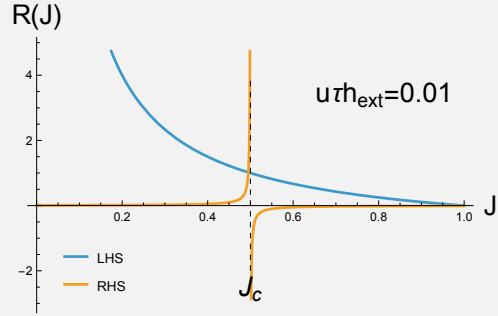
$$R(J) = \frac{h_{\text{ext}}}{1 - \sigma(J)} \quad (7.119)$$

from which we see that the critical value of the branching parameter is $\sigma = 1$ – at this value, an arbitrarily small drive produces an arbitrarily large response. Near the critical value, we have $\sigma(J) \sim J/J_c$.

Equating these two expressions for R , (7.116) and (7.119),

$$\frac{J_{\max} - J}{u\tau J} = \frac{h_{\text{ext}}}{1 - J/J_c} \quad (7.120)$$

which looks like the figure at right.



You see that when the external drive h_{ext} is small, the only way to equate the finite LHS with the RHS is to sit near the pole, which happens precisely at the critical point for J .

So this mechanism is actually some combination of items 1 and 5: there is definitely some mechanism of plasticity, but in order for it to produce criticality it requires tuning a ratio of timescales.

Now, you could very reasonably ask: can we apply the same mean field strategy to the “plastic Greenberg-Hastings model”, the model I discussed above that is obtained by making dynamical the infection rate p of the Greenberg-Hastings automaton? This model differs mainly in that there is a notion of locality (less importantly, the site variables are discrete). So let R be the activity, meaning the density of sites with $s > 0$; in that case the equation from demanding a steady state of the dynamics of p is identical to (7.116). But the equation for the activity is a bit different, because I didn’t include any driving. The density of sites with value 1 satisfies

$$\dot{\rho}_1 = \text{inflow from 0} - \text{outflow to 2} \quad (7.121)$$

$$= \rho_0 P(\text{infection}) + h_{\text{ext}} - \rho_1. \quad (7.122)$$

Here $\rho_0 = 1 - R$, $P(\text{infection}) = Kp\rho_1$ where K is the number of neighbors, and $\rho_1 = R/n$. Demanding a steady state gives

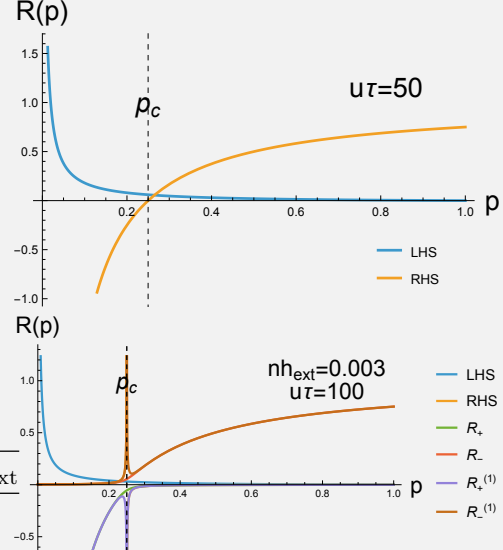
$$0 = (1 - R)RKp/n - Rn + h_{\text{ext}}. \quad (7.123)$$

The external drive h_{ext} is a very singular perturbation of the equation in its absence. If we set $h_{\text{ext}} = 0$, we have two solutions, $R = 0$ (the dead phase) and (for $p > p_c$),

$$R = 1 - \frac{1}{Kp} = 1 - \frac{p_c}{p}. \quad (7.124)$$

Equating our two expression for R , we find that again when $u\tau \gg 1$, they intersect at p_c . What happens when $h_{\text{ext}} \neq 0$? In that case, there is no dead phase, and the two solutions are

$$\begin{aligned} R_{\pm} &= \frac{-1 + Kp \pm \sqrt{(1 - Kp)^2 + 4Kph_{\text{ext}}}}{2Kp} \\ &\simeq -\frac{1 - Kp + |Kp - 1|}{2Kp} \pm \frac{h_{\text{ext}}n}{|Kp - 1|} + \mathcal{O}(h_{\text{ext}})^2 \\ &\equiv R_{\pm}^{(1)} + \mathcal{O}(h_{\text{ext}})^2. \end{aligned} \quad (7.125)$$



7.3 Reaction-diffusion problems and spin chains

The dynamics of extensive systems with a finite number of states at each site can be mapped to a (generally non-Hermitian) spin chain Hamiltonian problem, a context very familiar to quantum condensed matter physicists. Some such systems can be solved by the Jordan-Wigner map to fermions. Let me describe these steps in the context of a nice model.

The SSEP (symmetric simple exclusion process) has as its dofs spins on a chain $s_i = \pm 1$. We'll interpret $s_i = -1 = \downarrow$ as a particle and $s_i = +1 = \uparrow$ as a hole or absence of a particle. The reactions are

- Hopping $\uparrow_i \downarrow_{i+1} \leftrightarrow \downarrow_i \uparrow_{i+1}$ at rate r
- Pair annihilation $\downarrow_i \downarrow_{i+1} \leftrightarrow \uparrow_i \uparrow_{i+1}$ at rate γ .

It's called an 'exclusion process' because at most one particle can sit at each site. The update rule violates DB since the time-reverse of pair annihilation is pair creation which happens with rate 0. And the vacuum $\uparrow\uparrow\uparrow \dots$ is an absorbing state, since none of the reactions leave it.

The master equation is

$$\partial_t P(s, t) = -W P(s, t) \quad (7.126)$$

where the Liouvillian $W = \sum_I W_I$ is a sum over the possible reactions of a term that implements the reactions, $s = s_1 s_2 \dots s_L$ labels the spin configuration, and $\sum_s P(s, t) = 1, \forall t$. We can rewrite this as

$$\partial_t |P(t)\rangle = -H |P(t)\rangle \quad (7.127)$$

where

$$|P(t)\rangle \equiv \sum_s P(s, t) |s\rangle \quad (7.128)$$

is a state in the Hilbert space of a spin chain. Observables are

$$\overline{O} \equiv \sum_s O(s) P(s, t) = \langle 1 | O | P(t) \rangle \quad (7.129)$$

where the all-ones vector is

$$\langle 1 | = \otimes_i (\langle \uparrow | + \langle \downarrow |) \quad (7.130)$$

the product of $X \equiv \sigma^x \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ eigenstates. In particular, the normalization condition is

$$1 = \langle 1 | P(t) \rangle = \sum_s P(s, t). \quad (7.131)$$

Let's figure out what are the terms in the Hamiltonian. Notice that the particle number at site j is measured by the operator $\frac{1-Z_j}{2}$ where $Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. The operator that takes the state $|\downarrow_j \uparrow_{j+1}\rangle$ to $|\uparrow_j \downarrow_{j+1}\rangle$, where a particle hops one site to the right is

$$\sigma_j^+ \sigma_{j+1}^- \quad (7.132)$$

where $\sigma^\pm = \frac{1}{2}(X \pm iY)$ are the spin raising and lowering operators

$$\sigma^+ |\downarrow\rangle = |\uparrow\rangle, \sigma^+ |\uparrow\rangle = 0, \sigma^- |\downarrow\rangle = 0, \sigma^- |\uparrow\rangle = |\downarrow\rangle. \quad (7.133)$$

Similarly, to hop a particle to the left,

$$\sigma_j^- \sigma_{j+1}^+ |\uparrow_j \downarrow_{j+1}\rangle = |\downarrow_j \uparrow_{j+1}\rangle. \quad (7.134)$$

Notice that if the state is either $\uparrow\uparrow$ or $\downarrow\downarrow$ these operators will just give zero. So we've learned that the off-diagonal terms in H_{hop} are

$$-\tau \sum_j (\sigma_j^+ \sigma_{j+1}^- + \sigma_j^- \sigma_{j+1}^+) = \frac{\tau}{2} \sum_j (X_j X_{j+1} + Y_j Y_{j+1}). \quad (7.135)$$

This is called the XX Hamiltonian. The diagonal term in H_{hop} describes the rate at which probability leaves the state. Probability can only leave a state by a particle hopping from j to $j+1$ if j is occupied and $j+1$ is empty, so the diagonal term is ($-r$ times) the projector

$$n_j(1 - n_{j+1}) + (1 - n_j)n_{j+1} = \frac{1}{2}(1 - Z_j Z_{j+1}). \quad (7.136)$$

Altogether,

$$H_{\text{hop}} = \frac{r}{2} \sum_j (X_j X_{j+1} + Y_j Y_{j+1} + Z_j Z_{j+1} - 1), \quad (7.137)$$

this is the Heisenberg Hamiltonian, which actually has $\text{SU}(2)$ symmetry. The $\text{U}(1)$ subgroup generated by $\prod_j Z_j$ is associated with the fact that the hopping process conserves particle number. The rest of the symmetry is mysterious to me. Under the Jordan-Wigner transformation, this model maps to an interacting fermion system. But it is solvable by Bethe ansatz.

The analogous manipulation for the Glauber dynamics is maybe a little easier (see the homework), where it leads to a Liouvillian similar to the transverse-field Ising model (but non-Hermitian). In the 1d case, this method immediately leads to a solution of the problem because the resulting spin chain is solvable by a Jordan-Wigner transformation.

7.4 Second quantization for reaction-diffusion problems

[Cardy, cond-mat/9607163] Suppose we have a species of particles that can annihilate each other at some rate λ if they are on top of each other: $A + A \xrightarrow{\lambda} \emptyset$. If we consider a single site, so they are always on top of each other, the master equation is

$$\partial_t P(n, t) = (\mathcal{E}^2 - 1) (W_a P) = \lambda ((n+2)(n+1)P(n+2, t) - n(n-1)P(n, t)), \quad (7.138)$$

where $W_a(n) = \lambda n(n-1)$ is the rate. We've seen a number of parallels between the master equation and the Schrödinger equation, where the probability maps to the quantum amplitude – in particular, they are both linear and they are both first-order in time derivatives. For quantum mechanical problems involving processes where the number of particles can change, it is useful to introduce creation and annihilation operators that act between sectors with different number of particles. So it's not surprising that such a formalism would be useful here too. (Note that here the name 'second quantization' is even stupider than usual, since there was never even any first quantization.) It is sometimes named after Doi and Peliti.

So, for each species of particle, introduce an oscillator mode a_j , with $[a_j^\dagger, a_k] = \delta_{jk}$, and write a state as

$$|\Psi(t)\rangle = \sum_{\{n_1, n_2, \dots\}} p(n_1, n_2, \dots; t) \left(a_1^\dagger\right)^{n_1} \left(a_2^\dagger\right)^{n_2} \cdots |0\rangle \quad (7.139)$$

where $a_j |0\rangle = 0, \forall$ species j . (Note that the j label could be a site index or some other label.) Time evolution, *i.e.* the master equation, is then

$$\partial_t |\Psi(t)\rangle = -H |\Psi(t)\rangle, \quad (7.140)$$

and a given set of stoichiometry vectors S_I can be translated into a term in the Liou-villian operator H made of particular combination of a s and $a^\dagger$ s.

Before we do that, we must confront some weirdnesses. Since $|\Psi\rangle$ is linear in the probability, an expectation value is not of the usual QM form $\langle\Psi|O|\Psi\rangle$. Rather, I claim that

$$\overline{O} = \sum_n p(n) O(n) = \langle 1 | \hat{O} | \Psi \rangle \quad (7.141)$$

where

$$\langle 1 | \equiv \langle 0 | e^{\sum_j a_j} \quad (7.142)$$

is the all-ones vector. In particular, the normalization condition is

$$1 = \overline{1} = \sum_n p(n) = \langle 1 | \Psi \rangle = \langle 0 | e^{\sum_j a_j} | \Psi \rangle. \quad (7.143)$$

Another important example is the operator a itself. Let's think about the case with just one species:

$$\bar{a} = \langle 0 | e^a a \sum_n p(n) (a^\dagger)^n | 0 \rangle \quad (7.144)$$

$$= \sum_n \sum_m \frac{1}{m!} \underbrace{\langle 0 | a^m (a^\dagger)^n | 0 \rangle}_{=\delta_{nm}n} = \sum_n p(n)n = \langle n \rangle, \quad (7.145)$$

– its expectation value is the density. The condition that a given H conserves probability is (not that it is Hermitian but instead) the strange-looking condition

$$0 = \langle 1 | H = \langle 0 | e^{\sum_j a_j} H. \quad (7.146)$$

For the example of the annihilation reaction with a single site, I claim that

$$H = H_a = -\lambda \left(a^2 - (a^\dagger)^2 a^2 \right). \quad (7.147)$$

Let's check that we get the master equation right:

$$\partial_t |\Psi\rangle = \lambda \left(a^2 - (a^\dagger)^2 a^2 \right) \sum_n p(n) (a^\dagger)^n |0\rangle \quad (7.148)$$

$$= \lambda \left(\sum_n p(n) a^2 (a^\dagger)^2 (a^\dagger)^{n-2} + \sum_n p(n) (a^\dagger)^2 a^2 (a^\dagger)^n \right) |0\rangle \quad (7.149)$$

$$\stackrel{*}{=} \lambda \left(\sum_n p(n) (a^\dagger a + 2) (a^\dagger a + 1) (a^\dagger)^{n-2} |0\rangle + \sum_n p(n) (a^\dagger a - 1) a^\dagger a (a^\dagger)^n |0\rangle \right) \\ = \lambda \left(\sum_n p(n) (n)(n-1) (a^\dagger)^{n-2} |0\rangle - \sum_n p(n) (n-1)n (a^\dagger)^n |0\rangle \right) \quad (7.150)$$

$$= \lambda \left(\sum_{m \equiv n-2} (m+2)(m+1)p(m+2) (a^\dagger)^m |0\rangle - \sum_n n(n-1)p(n) (a^\dagger)^n |0\rangle \right) \\ = \lambda \sum_n ((n+2)(n+1)p(n+2) - n(n-1)p(n)) (a^\dagger)^n |0\rangle \quad (7.151)$$

where at the step labelled $*$ I used the identities $a^2 (a^\dagger)^2 = (a^\dagger a + 2)(a^\dagger a + 1)$ and $(a^\dagger)^2 a^2 = (a^\dagger a - 1)a^\dagger a$ which follow from $[a, a^\dagger] = 1$. Comparing the coefficient of $(a^\dagger)^n |0\rangle$ on both sides gives the master equation (7.138). Note that we don't need to bother with normalizing $(a^\dagger)^n |0\rangle$ with the usual $n!$; nevertheless I've resisted the temptation to call this state $|n\rangle$ to avoid clashing with the usual normalization.

We can study a full reaction-diffusion problem in this language by including site labels on the a s. Then if we also include the hopping reaction: $A_i \xrightarrow{D} A_j$ with rate D

between neighbors $\langle ij \rangle$, the Hamiltonian is

$$H = D \sum_{\langle ij \rangle} \left(a_i^\dagger - a_j^\dagger \right) \left(a_i - a_j \right) - \lambda \sum_i \left(a_i^2 - \left(a_i^\dagger \right)^2 a_i^2 \right). \quad (7.152)$$

Now you could reasonably ask what we gain from doing this. With a single site, it's not worth it. But for problems where the particles can diffuse as well as interact, it is quite nice. One virtue is that it makes it very easy to write a continuum limit:

$$H = \int d^d x \left(D \vec{\nabla} a^\dagger \cdot \vec{\nabla} a - \lambda (a^2 - (a^\dagger)^2 a^2) \right). \quad (7.153)$$

We can develop a path integral formulation using coherent states, and from there develop both mean field descriptions and a full renormalization group account. So this provides an alternative approach to MSR. Notice by the way that in the case of an exclusion process, where $n_j = 0, 1$, we might have also considered treating the particles as fermions and using anti-commuting creation and annihilation operators.

We can construct the path integral for an observable O in the usual way, using the basis of coherent states, which satisfy

$$a |z\rangle = z |z\rangle, \langle z| a^\dagger = \langle z| \bar{z}, \langle z_1| z_2\rangle = e^{\bar{z}_1 z_2}, \mathbb{1} = \int \frac{dz d\bar{z}}{\pi} e^{-\bar{z}z} |z\rangle\langle z|. \quad (7.154)$$

Start from

$$\overline{O} = \langle 1| O e^{-t \sum_j h_j(a, a^\dagger)} |\psi(0)\rangle, \quad (7.155)$$

divide up $e^{-tH} = (e^{-dtH})^{N=t/dt}$, and in between each insert the resolution of the identity above, at each point in space (which I suppress in the following). The result is

$$\overline{O} = \int \prod_{i=1}^N dz_i d\bar{z}_i e^{z_N} O e^{-dt \sum_i h(z_i, \bar{z}_i)} e^{\sum_j \bar{z}_{i+1} z_i} \langle z_1 | \psi(0) \rangle \quad (7.156)$$

$$= \int D z D \bar{z} e^{-S[z, \bar{z}] + z(t_f) + \log \langle z(t_0) | \psi(0) \rangle} O \quad (7.157)$$

with

$$S[z, \bar{z}] = \int dt d^d x \left(\bar{z} \partial_t z + H(z, \bar{z}) \right). \quad (7.158)$$

The boundary terms in (7.157) specify the initial conditions on the path integral. Having derived this path integral, we are confronted with the fact that it is not gaussian if $H(z, \bar{z})$ is not quadratic. So there are two questions I'd like to answer. One is to make contact with other ways of constructing such a path integral, like

the MSR formalism we talked about earlier in the single-particle case. The other questions is how to treat the interactions. The answer to the latter is to use the whatever methods of quantum field theory we can, the most immediate of which is to do perturbation theory in λ . This is the right thing to do in an expansion in $\epsilon = d_c - d$ where d_c is the upper critical dimension where the coupling λ becomes marginal.

Regarding the former, you can't stop me from defining new variables

$$\bar{z} \equiv 1 + \psi, z \equiv \rho. \quad (7.159)$$

ψ will play the role of the response field, and since $\rho = \bar{n} = z$, ρ is the density. Let's think about the diffusion-limited annihilation case, where (after an IBP),

$$\begin{aligned} S[z = \rho, \bar{z} = 1 + \psi] &= \int dt d^d x \left(\bar{z} \partial_t z + D \vec{\nabla} \bar{z} \cdot \vec{\nabla} z - \lambda (z^2 - \bar{z}^2 z^2) \right) \quad (7.160) \\ &= \int dt d^d x \left(\partial_t \rho + \psi \partial_t \rho - D \psi \vec{\nabla}^2 \rho - 2\lambda \psi \rho^2 - \rho^2 \psi^2 \right). \end{aligned}$$

The first term is a total derivative and I will put it in my closet with the other boundary terms.

We can use the Hubbard-Stratonovich trick to *decouple* the quartic term, site by site:

$$e^{-\lambda \rho^2 \psi^2} = \int d\eta P(\eta) e^{-\eta \psi} \quad (7.161)$$

where $P(\eta) \propto e^{-\frac{\eta^2}{4\lambda}}$ (at each point in space and time). This gives

$$\int D\rho D\psi e^{-S[\rho, \psi]} \dots = \int D\eta P[\eta] \int D\rho \int D\psi e^{\int d^d x dt \psi (\partial_t \rho - D \vec{\nabla}^2 \rho + 2\lambda \rho^2 - \eta)} \dots \quad (7.162)$$

and you see that the integral over $\bar{z}$ is a delta function that imposes the familiar equation

$$0 = -\partial_t \rho + D \vec{\nabla}^2 \rho - 2\lambda \rho^2 + \eta. \quad (7.163)$$

Thus the equation for DP, right *at* the critical point, describes the relaxation of this diffusion-limited annihilation process, at least until we think about the noise. Why didn't we have to do any tuning to get there? I think it's essentially an accident of studying such a simple model, with out any birth or death processes. See the next subsection for more on such questions.

So we have recovered the MSR path integral that we would have written starting with this DP equation. The only funny thing is I didn't say what was the

distribution for the noise η . In order to get the right answer in (7.161), we need $\Gamma = -4\lambda\rho^2$, *i.e.*

$$\langle \eta(x, t) \eta(x', t') \rangle = -4\lambda\rho^2 \delta^d(x - x') \delta(t - t') \quad (7.164)$$

and we recognize the familiar absorptive nature of the state with $\rho = 0$; the fact that ρ^2 appears and not ρ is also because this model is fine tuned. The sign, however, is pretty funny – it means that η appearing in (7.163) is imaginary. This sign is a special property of this weird model of diffusion-limited annihilation. To see this, let's redo the analysis for the contact process (describing generic points in the DP class).

$$\text{branching: } A \rightarrow A + A, \quad H_{\text{branch}} = -\sigma \int d^d x (a^{\dagger 2} - a^{\dagger}) a \quad (7.165)$$

$$\text{death: } A \rightarrow \emptyset, \quad H_d = \int d^d x D \nabla a^{\dagger} \cdot \nabla a \quad (7.166)$$

$$\text{coagulation: } A + A \rightarrow A, \quad H_c = \lambda \int d^d x (a^{\dagger 2} a^2 - a^{\dagger} a^2) \quad (7.167)$$

$$\text{diffusion: } A_i \rightarrow A_j, \quad H_D = \int d^d x D \nabla a^{\dagger} \cdot \nabla a \quad (7.168)$$

By the same logic as above, the action is

$$\begin{aligned} S &= \int dt \int d^d x \bar{z} \partial_t z + \int dt \sum_I H_I(\bar{z}, z) \\ &= \int dt d^d x [\bar{z} \partial_t z + D \nabla \bar{z} \cdot \nabla z + \mu(\bar{z} - 1)z - \sigma(\bar{z}^2 - \bar{z})z + \lambda(\bar{z}^2 z^2 - \bar{z} z^2)] \end{aligned} \quad (7.169)$$

As before, define $\bar{z} = 1 + \psi, z = \rho$. The resulting MSR action is

$$S = \int [\partial_t \rho + \psi \partial_t \rho - D \psi \nabla^2 \rho + \mu \psi \rho - \sigma(\psi + \psi^2) \rho + \lambda(\psi + \psi^2) \rho^2] \quad (7.170)$$

$$= \int [\partial_t \rho + \psi (\partial_t \rho - D \nabla^2 \rho + (\mu - \sigma) \rho + \lambda \rho^2) + \psi^2 (-\sigma \rho + \lambda \rho^2)] \quad (7.171)$$

$$= \int [\psi(\text{EOM}) + \psi^2(\text{Noise Kernel})]. \quad (7.172)$$

Here EOM is exactly the Langevin equation for DP. And the noise kernel (which determines the correlation function of the response field with itself, which you'll recall was a delta function in the case of white noise) is

$$B = -\sigma \rho + \lambda \rho^2. \quad (7.173)$$

To understand what this means, observe that the probability distribution for $\rho(x)$ is

$$P[\rho] \propto \int D\psi e^{-S[\rho, \psi]} \quad (7.174)$$

$$= \int D\psi \exp \left(- \int [\psi \cdot (\text{EOM}) + \psi^2 \cdot (B)] \right) \quad (7.175)$$

$$\propto \exp \left(\int \frac{(\text{EOM})^2}{4B} \right) \equiv e^{-S_{\text{eff}}[\rho]} \quad (7.176)$$

with

$$S_{\text{eff}} = + \int \frac{(\text{EOM})^2}{4(\sigma\rho - \lambda\rho^2)} . \quad (7.177)$$

Comparing to the way we got this distribution from ρ by plugging in the solution of the Langevin equation for the noise distribution, the noise variance is $2(2\sigma\rho - \lambda\rho^2) \approx 2\sigma\rho > 0$:

$$\langle \eta(x, t) \eta(x', t') \rangle = 2\sigma\rho \delta(x - x') \delta(t - t'). \quad (7.178)$$

Notice that if the action is real, either the noise distribution is weird or the response field distribution is weird. Only if we use a factor of $\mathbf{i}$, the same one that we should put in $\int d\psi e^{\mathbf{i}\psi x} = \delta(x)$, can they both have $-\psi^2$ and $-\eta^2$ terms in their respective distributions. I am saying: it is best to think of ψ as imaginary.

Returning to diffusion-limited annihilation, the late-time behavior of this model is no mystery: all the particles eventually annihilate each other (assuming you start with an even number, otherwise one lonely particle will be left to diffuse forever on its own). But the rate of approach to that eventual absorbing state of extinction is quite interesting: the upper critical dimension is $d = 2$, meaning that for $d > 2$, the mean field answer is correct. (The mean field answer is $\partial_t \rho \sim \rho^2 \implies \rho \sim 1/t$.) But for $d < 2$, the noise is a relevant perturbation. This we can see from the action by the same scaling analysis we did above: we found that $\rho(x, t) \rightarrow \lambda^{-2} \rho(\lambda x, \lambda^2 t)$. Under this transformation, we see from (7.16) that the noise scales as $\eta \sim \sqrt{\rho \delta^d(x) \delta(t)} \sim \sqrt{\lambda^{-2-d-2}} = \lambda^{-d/2-2}$. Comparing this with the marginal terms, which scale like λ^{-4} , we see that for $d > 2$ we can ignore the noise term at long wavelengths.

One conclusion that one can derive using this machinery (bringing to bear the usual apparatus of the RG, which I exemplified above with directed percolation) is that in $d < 2$ space dimensions,

$$\langle n(t) \rangle = \frac{A}{t^{d/2}} \quad (7.179)$$

where the constant A is independent of the initial density, and can be computed analytically in an expansion in $\epsilon = 2 - d$ about the upper critical dimension.

Suppose instead we have two species of particles, and they can only co-annihilate, *i.e.* the reaction is $A + B \rightarrow \emptyset$. It's still true that the final state, when starting with equal numbers of A s and B s is boring emptiness, but the approach is interesting and different from the above. Initially all the A s next to B s eat each other and we are left with domains of A s and domains of B s. After this the only action is at the boundary. The upper critical dimension for this case is $d = 4$!

To see this, observe that this system involves two fields, $\rho_S = \rho_A + \rho_B$, the total density and $\psi \equiv \rho_A - \rho_B$. The latter is conserved by the reaction, and so only evolves by diffusion, $\partial_t \psi = D \nabla^2 \psi$. Once we get to a state with domains of A s ($\psi > 0$) and of B s ($\psi < 0$), the reaction only proceeds at the domain walls where $\psi = 0$, and the rate of decrease of ρ_S is limited by particles *diffusing* to this boundary to replace the ones that have been annihilated.

How long does this take? I claim that in d spatial dimensions $\psi_{\text{rms}}(t) \sim t^{-d/4}$. Starting from a given initial configuration $\psi_0(x)$, the later configuration of ψ is

$$\psi(x, t) = \int d^d y G(x - y) \psi_0(y) \quad (7.180)$$

where $G(x) = \frac{e^{-\frac{|x|^2}{4Dt}}}{(4\pi Dt)^{d/2}}$ is the Green's function for the diffusion equation. If we assume that the initial conditions are white noise: $\langle \psi_0 \rangle = 0$ and

$$\langle \psi_0(y) \psi_0(y') \rangle = c \delta^d(y - y') \quad (7.181)$$

then

$$\langle \psi(x, t)^2 \rangle = c \int d^d y G(x - y)^2 = c \int d^d x \frac{e^{-\frac{|x|^2}{2Dt}}}{(4\pi Dt)^d} \propto t^{-d/2}. \quad (7.182)$$

Therefore $\rho_S \sim \psi_{\text{rms}} \sim t^{-d/4}$, according to this mechanism. But mean field theory predicts $\rho_S \sim t^{-1}$, and these cross when $d = 4$.

I recommend the book by Krapivsky et al for more on these questions.

7.5 Surface growth and roughening

[[The KPZ paper](#), Toner's book, this beautiful review by [Takeuchi](#) (an experimentalist!), [Hwa-Kardar](#)]

There are many physical phenomena where the relevant degrees of freedom specify a mobile interface between two media. This is a class of problems with “generic scale

invariance”. Shouldn’t we call them critical phenomena? It doesn’t seem like we are tuning any parameters. This paper [Hwa Kardar PRL 1989] claims that the reason is simply SSB of continuous symmetry. That is, the dofs enjoy a shift symmetry that forces their evolution to depend only on their derivatives, which therefore makes it slow in the long-wavelength limit.

Edwards-Wilkinson model of surface growth. Consider the following model of an interface $h(\vec{x}, t)$. We will ignore the possibility of overhangs, so h is a function of x . Demand time-translation symmetry, spatial translation symmetry in $\vec{x}$, spatial rotation symmetry, and an invariance under $h \rightarrow h + h_0$, this is spatial translation invariance in h . The resulting Langevin equation seems to be

$$\partial_t h(\vec{x}, t) = v_0 + \nu \vec{\nabla}^2 h + \eta(\vec{x}, t) \quad (7.183)$$

where v_0 is a uniform velocity that we can forthwith remove by a Galilean boost $h \rightarrow h + v_0 t$. When we set $v_0 = 0$ we also have $h \rightarrow -h$ symmetry. The surface tension term $\nu \vec{\nabla}^2 h$ will serve to smooth out irregularities of the surface. The last term is some noise, which we will take to be Gaussian and white.

This equation is linear! We can solve it using Fourier transforms. What might we want to compute? We would like to understand the statistics of $\delta h(x, t) \equiv h(x, t) - \langle h(x, t) \rangle$, an interesting one of which is

$$C_h(\ell, t) \equiv \langle (h(x + \ell, t + \tau) - h(x, \tau))^2 \rangle_\eta. \quad (7.184)$$

Because of the scale invariance under

$$\delta h(x, t) \mapsto b^\chi \delta h(bx, b^z t), \quad (7.185)$$

this observable behaves as

$$C_h(\ell, t) = t^{2\chi/z} F_h(\ell t^{-1/z}) \sim \begin{cases} \ell^{2\chi}, & \ell \ll \xi(t) \\ t^{2\chi/z}, & \ell \gg \xi(t) \end{cases}, \quad (7.186)$$

with $\xi(t) \sim t^{1/z}$. Here z is the dynamical exponent which we’ve met before, and χ is called the *roughening exponent*. Demanding that the EW equation (7.183) is invariant under the scale transformation (7.185), we have $z = 2$, $\chi = \frac{2-d}{2}$. So this is another example of a system with ‘generic scale invariance’. I’m not sure where the secret tuning parameter is in this one.

Since the EW model is Gaussian, we have no excuse for not being explicit. The

result is

$$C_h(r, t) = 2 \int \mathrm{d}\omega \int \mathrm{d}^d q C(q, \omega) (1 - \cos(\vec{q} \cdot \vec{r} - \omega t)) \quad (7.187)$$

$$= 4D \int \mathrm{d}\omega \int \mathrm{d}^d q \frac{1 - \cos(\vec{q} \cdot \vec{r} - \omega t)}{\omega^2 + \nu^2 q^4} \quad (7.188)$$

where I used the fact that the EW equation has the same form as the TDLGXY equation, for which we already computed the correlations in (6.63). By scaling $q \equiv Q/r, \omega \equiv \nu\Omega/r^2$, this is

$$C_r(r, t) = r^{2-d} F_h(t/r^2) = \begin{cases} r^{2\chi}, & r^z \gg t \\ |t|^{2\chi/z}, & r^z \ll t \end{cases} \quad (7.189)$$

which comparing the first expression to (7.186), gives $\chi = \chi_{\text{EW}} = \frac{2-d}{2}, z = z_{\text{EW}} = 2$. The last expression for $r^z \gg t$ follows from analyzing the explicit integral above.

The FP equation corresponding to the Langevin equation (7.183) is

$$\partial_t P(h(x), t) = - \int d^d x \frac{\delta}{\delta h(x, t)} \left(\nu \vec{\nabla}^2 h P(h(x), t) \right) + \frac{D}{2} \int d^d x \frac{\delta^2}{\delta h(x)^2} P(h(x), t). \quad (7.190)$$

The functional derivative can be written as $\frac{\delta}{\delta h} \equiv \partial_h - \vec{\nabla} \cdot \left(\frac{\partial}{\partial \vec{\nabla} h} \right)$. (Just think of $h(x_i)$ as a collection of variables and apply the relations we found above.) This has a stationary solution

$$P_{\text{eq}}[h(x)] \propto e^{-\int d^d x \frac{\nu}{D} (\vec{\nabla} h)^2}. \quad (7.191)$$

For $d = 1$, this is Brownian motion with x playing the role of time and h playing the role of position, with diffusion constant $\frac{D}{4\nu}$, so

$$\langle (h(x + \ell) - h(\ell))^2 \rangle_{\text{eq}} \propto \ell \quad (7.192)$$

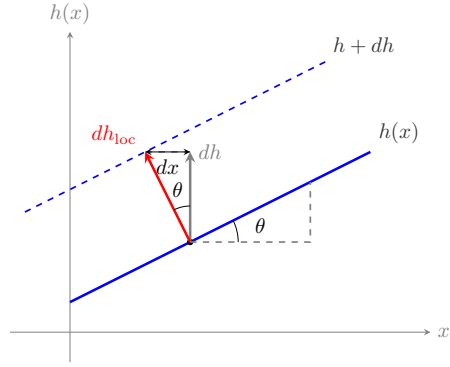
which is consistent with $\chi = \frac{1}{2}$.

So why aren't we done with this problem? Well, the EW equation describes growth only by *vertical* deposition, where the h direction is special.

The more naturally-occurring process is where the surface grows along the *local normal*. Using $\tan \theta = |\vec{\nabla} h|$,

$$dh = \frac{dh_{\text{loc}}}{\cos \theta} = dh_{\text{loc}} \sqrt{1 + (\vec{\nabla} h)^2}, \quad (7.193)$$

$$dx = dx_{\text{loc}} \cos \theta = \frac{dx_{\text{loc}}}{\sqrt{1 + (\vec{\nabla} h)^2}}. \quad (7.194)$$



Inserting this into the EW evolution of $h_{\text{loc}}, x_{\text{loc}}$ gives the KPZ equation

$$\partial_t h(\vec{x}, t) = \nu \vec{\nabla}^2 h + \frac{\lambda}{2} (\vec{\nabla} h)^2 + \eta(x, t). \quad (7.195)$$

The nonlinear term is not of the form $\frac{\delta}{\delta h(x, t)} \mathcal{F}[h]$ for any local functional $\mathcal{F}$, so there is no obvious notion of equilibrium. Also because of this, the translation symmetry $h \rightarrow h + \epsilon$ does not imply a conservation law for some integral involving h , unlike in equilibrium where Noether's theorem would imply such a statement³¹.

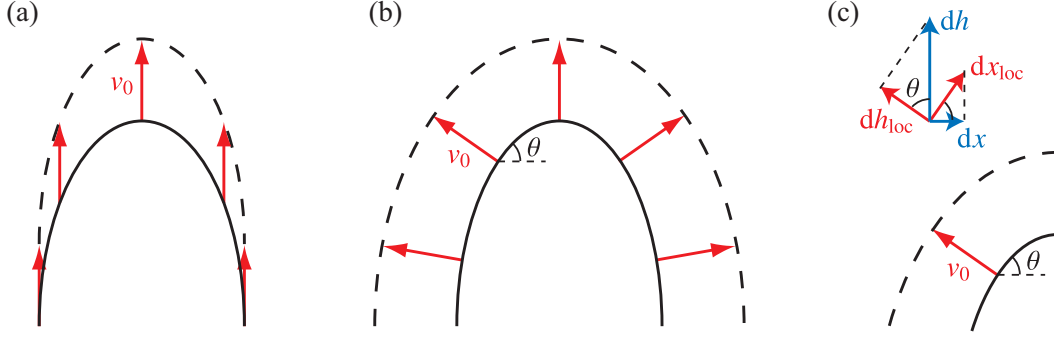
The distinction between EW growth (a) and KPZ growth (b) is made very clear in

³¹Using the MSR description, Noether's theorem *does* apply. That is, because $S[h(x, t) + \epsilon] = S[h(x, t)]$ is invariant under the transformation with constant ϵ , we can write the variation of the MSR action under the 'local' transformation $h(x) \rightarrow h(x) + \epsilon(x, t)$ in the form

$$\delta S = \int \epsilon \text{EOM} + \int \partial_\mu \epsilon j^\mu. \quad (7.196)$$

(I'm ignoring boundary terms.) Since, for a configuration satisfying the EOM, $\delta S = 0$ for *any* variation, we must have $\partial_\mu j^\mu = 0$. If you apply this procedure (named after Emmy Noether) to the above shift symmetry of KPZ, you'll find that the conserved current is $j^\mu = (\tilde{h}, \nu \vec{\nabla} \tilde{h} + \lambda \tilde{h} \vec{\nabla} h)^\mu$. (Exercise: check that using the backwards EOM ($0 = \frac{\delta S}{\delta h}$) this satisfies $\partial_\mu j^\mu = 0$.) So the quantity $\int d^d x \tilde{h}$ integrated over all space is time-independent. But the time evolution of h itself is not constrained by this! What is the physical meaning of this conservation law? Here's my best attempt: Recall that the role of $\tilde{h}$ is to be the *source* for h , so $\langle h \tilde{h} \rangle$ is the response function. The fact that $\int \tilde{h}$ is time-independent reflects the fact that if we apply a *uniform* impulse to the system, the resulting change in the height is remembered forever.

this figure from Takeuchi's review:



Naively the KPZ equation has four parameters: ν, λ, D and the UV cutoff on wavenumbers, Λ . In fact by scaling, the physics only depends on the one combination

$$g \equiv K_d \frac{\lambda^2 D \Lambda^{d-2}}{\nu^3} \quad (7.197)$$

where $K_d \equiv \frac{\Omega_d}{(2\pi)^d}$ is there for convenience. Any other value of the parameters is related by the rescalings $r \rightarrow r/\Lambda$ (to set $\Lambda = 1$) $t \rightarrow \alpha t, h \rightarrow \beta h$ to set $\nu = 1, D = 1$ with $\alpha = (\nu \Lambda^2)^{-1}, \beta = \sqrt{\frac{D \Lambda^{d-2}}{\nu}}$, and finally $\nu \rightarrow \frac{\alpha}{\beta} \nu$.

So the EW theory is to KPZ as the Gaussian fixed point is to the Wilson-Fisher Ising fixed point. As you can see from (7.197), the nonlinear coupling λ has dimension $[\lambda] = L^{1-d/2}$. The nonlinear term is a relevant perturbation of the EW theory for $d < 2$, *i.e.* $d = 2$ is the upper critical dimension. To see this more explicitly, consider the scale transformation (7.185), and note that we must scale $\eta \sim b^{-\frac{d+z}{2}}$. Thus $\partial_t h \sim b^{\chi-z} \stackrel{\text{EW}}{=} b^{-1-d/2}, \vec{\nabla}^2 h \sim b^{\chi-2} \stackrel{\text{EW}}{=} b^{-1-d/2}, \eta \sim b^{-(d+z/2)}$ and $(\vec{\nabla} h)^2 \sim b^{-2+2\chi} = b^{-d}$, so it is relevant when

$$-d \geq -1 - d/2 \quad \Leftrightarrow \quad d \leq 2, \quad (7.198)$$

that is, for $d < 2$, this term is more important than the linear terms at large length scales b .

For $d = 1$ it is nevertheless solvable and the exponents are $z = 3/2, \chi = 1/2$. Some properties of KPZ that lead to this conclusion.

1. It is related to the noisy Burgers' equation describing a randomly-stirred viscous fluid. Take $\vec{\nabla}(\text{KPZ})$ and let $\vec{v}(x, t) \equiv -\lambda \vec{\nabla} h(x, t)$. The result is

$$\partial_t \vec{v} + (\vec{v} \cdot \vec{\nabla}) \vec{v} = \nu \vec{\nabla}^2 \vec{v} - \lambda \vec{\nabla} \eta(x, t). \quad (7.199)$$

Like any good hydrodynamic equation, this has Galilean symmetry, $\vec{v}(x, t) \rightarrow \vec{v}_0 + \vec{v}(x - v_0 t, t)$ for constant $\vec{v}_0$. (Note that the noise distribution is also preserved.)

In terms of the KPZ equation for h , this symmetry acts by

$$h(x, t) \mapsto h(x, t) + \vec{s} \cdot x - \frac{\lambda}{2} s^2 t = h_{\text{new}}(x_{\text{new}}, t), x_{\text{new}} \equiv x - \lambda s t \quad (7.200)$$

with the same noise. Notice that this boost transformation (7.200), which is an exact symmetry of the KPZ equation (and its MSR action), depends explicitly on λ . This has a very strong consequence: it means that λ cannot run!

This in turn implies (just plug in the scale transformation into the $\partial_t h$ and non-linear terms with fixed λ) the following relation between exponents:

$$\chi + z = 2 \quad (7.201)$$

for any d .

2. The FP equation is

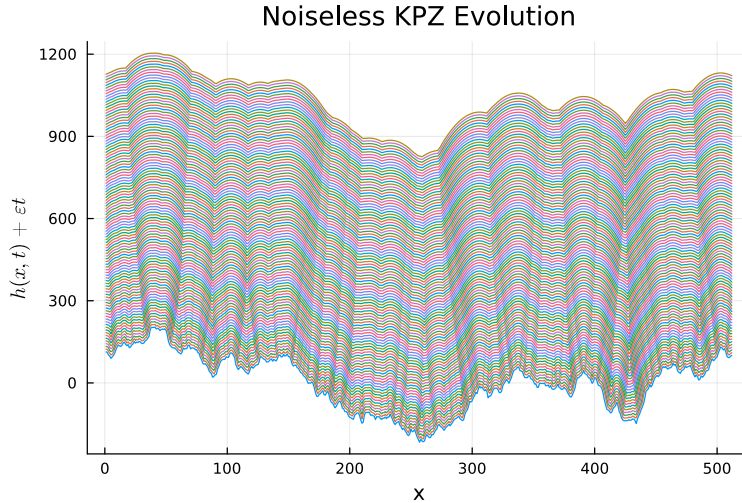
$$\partial_t P(h(x), t) = - \int d^d x \frac{\delta}{\delta h(x, t)} \left((\nu \vec{\nabla}^2 h + \frac{\lambda}{2} (\vec{\nabla} h)^2) P(h(x), t) \right) + \frac{D}{2} \int d^d x \frac{\delta^2}{\delta h(x)^2} P(h(x), t). \quad (7.202)$$

For $d = 1$, this has the *same* stationary solution as for the EW equation. This implies that $\int dx (\partial_x h)^2$ is scale invariant, which says $\chi = \frac{1}{2}$ and using (7.201) $z = 3/2$ (for $d = 1$).

3. Let $z(\vec{x}, t) \equiv e^{\frac{\lambda}{2\nu} h(\vec{x}, t)}$. This is called a Cole-Hopf transformation. The KPZ equation for h says that the variable z satisfies

$$\partial_t z(\vec{x}, t) = \nu \vec{\nabla}^2 z(\vec{x}, t) + \frac{\lambda}{2\nu} z(\vec{x}, t) \circ \eta(x, t) \quad (7.203)$$

a *linear* equation in z (!) but with multiplicative noise ($\odot$).

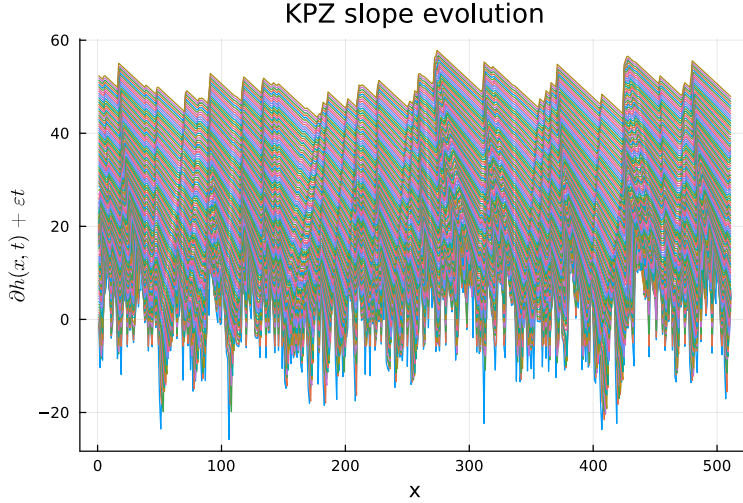


Using this result, we can find the exact evolution for the noise-free case at small ν . It is piecewise quadratic. The idea is just that in the noise-free case $z(x, t)$ solves the diffusion equation, so

$$h(\vec{x}, t) = \frac{2\nu}{\lambda} \log \int d^d x' e^{-\frac{|\vec{x} - \vec{x}'|^2}{2\nu t} + \frac{\lambda}{2\nu} h_0(\vec{x}')} . \quad (7.204)$$

In the small- ν limit, we can do the integral by saddle point and the result is

$$h(x, t) = \max_{\vec{x}'} \left(h_0(\vec{x}') - \frac{|\vec{x} - \vec{x}'|^2}{2\lambda t} \right) . \quad (7.205)$$



It is just like geometric optics. The slope (the thing related to the fluid velocity) rapidly becomes piecewise linear.

To do better, we can again expand around the upper critical dimension. In the literature, you will find that in addition to the method using the MSR path integral that we used above for KPZ, another way in which to frame the calculation is just directly using the Langevin equation. All the answers should be the same. Let me give a brief summary of this other perspective for KPZ.

Dynamical RG. Start with a stochastic equation with all terms consistent with the symmetries, with a set of couplings g_A . Again there are two steps.

1. Integrate out the short-wavelength modes, $h^>$. Here short-wavelength means the wavenumber is in the momentum shell near the cutoff, $|q| \in [b^{-1}\Lambda, \Lambda]$. Here integrate out just means solve their equations of motion
 - (a) as a functional of the long-wavelength modes $h^<$, which are held fixed here
 - (b) perturbatively in λ

and plug back in to the equations for the long-wavelength modes $h^<$. Claim: Because we included all possible terms consistent with the symmetries, the result has the same form as in the initial equations, with some couplings $g_A^I(\{g_A\}) = g_A + \delta g_A$ ('I' is for 'intermediate'), and with Λ replaced by Λ/b . Here δg_A is proportional to positive powers of the nonlinearity λ , without which modes of different wavelength do not couple.

2. Rescale $\vec{x}, t$ and the field(s) to restore the original cutoff Λ :

$$\vec{q} = b^{-1}\vec{q}', \vec{x} = b\vec{x}', t = b^z t', h(x, t) = b^x h'(x', t') \quad (7.206)$$

for some z, χ to be determined. Now the couplings have exactly the same form but with

$$g_A^R(\{g_A\}) = b^{\Delta_A}(g_A + \delta g_A), \quad (7.207)$$

where ‘ R ’ is for ‘renormalized’.

This procedure defined by steps 1 and 2 can be iterated. It produces a flow on the space of couplings, an auxiliary dynamical system. As always the first question we should ask is about fixed points: a choice of g_A, z, χ such that $g^R(g^*) = g^*$. The step of integrating out the long-wavelength modes sounds scary, but the idea is just to solve the equation iteratively: The equation is

$$(-i\omega + Dk^2) h_k = \zeta_k - \int d^q \vec{q} \cdot (\vec{q} + \vec{k}) h_{-q} h_{q+k} \quad (7.208)$$

so the solution is

$$h_k = G(\omega, k) \left(\zeta_k - \int d^q \vec{q} \cdot (\vec{q} + \vec{k}) h_{-q} h_{q+k} \right) \quad (7.209)$$

with $G(\omega, k) = (-i\omega + Dk^2)^{-1}$. Of course this isn’t really a solution because h also appears on the RHS. But we can plug this solution back in on the RHS and postpone the problem a bit more. Then we can keep doing this, and develop a solution as an infinite series of tree-like terms. As the preceding words suggest, we can introduce a diagrammatic notation for this which is isomorphic to the Feynman diagrams we developed using the other method.

But there is an important extra wrinkle: we must average over the noise. The effect of the averaging, because the distribution is Gaussian, is to join together the trees.

For KPZ, however you do it, the resulting RG equations are

$$\frac{d\lambda}{d\ell} = \lambda(\ell) (z + \chi - 2) \quad (7.210)$$

$$\frac{dD}{d\ell} = D(\ell) \left(z - 2\chi - d + \frac{g(\ell)}{4} + \mathcal{O}(g^2) \right) \quad (7.211)$$

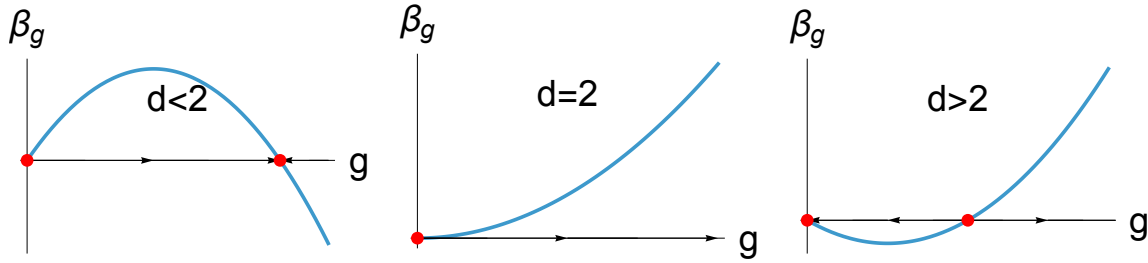
$$\frac{d\nu}{d\ell} = \nu(\ell) \left(z - 2 + \left(\frac{2-d}{4d} \right) g(\ell) + \mathcal{O}(g^2) \right) \quad (7.212)$$

The first equation is exact by the boost symmetry, the other two are one-loop results. These can be combined into an RG equation for the coupling $g \propto \frac{\lambda^2 D}{\nu^3}$:

$$\frac{dg}{d\ell} = g(\ell) \left(2 - d + \left(1 - \frac{3}{2d} \right) g(\ell) + \mathcal{O}(g^2) \right). \quad (7.213)$$

You can see that the upper critical dimension is $d = 2$ – when $d = 2$ the g^0 term, which comes from dimensional analysis, is absent. For $d > 2$, for sufficiently small bare coupling, the coupling flows to zero, which means that the linearized EW description is correct at long wavelength. If the bare coupling is large enough there is believed to be another (“rough”) strongly-coupled phase not directly accessible to this analysis. If we take seriously the one loop term, you see that for $d > 2$ there is another fixed point (the KPZ fixed point at $g_* = \frac{2d(d-2)}{2d-3}$) which is unstable – it is a critical point separating the EW phase of smooth surfaces from the rough phase. For $d = 2 + \epsilon$ with small ϵ , this calculation is reliable.

For $d < 2$ there are also two fixed points, but now it is the nontrivial KPZ fixed point that is stable.



In the right plot, the stable fixed point corresponding to the rough phase is off to the right of the plot and not visible.

So why are these examples critical? One possible answer is that the deterministic (noise-free) equations have a shift symmetry under $h(x, t) \rightarrow h(x, t) + \epsilon$. This forbids a term proportional to h without any derivatives in the Langevin equation. Such a term (a mass term in the QFT language) is the simplest way to produce a nonzero correlation length and correlation time.

[Hwa-Kardar; Kardar, *Fields* §9.5] Let's briefly consider the following seemingly-minor modification of the KPZ equation, obtained simply by picking out a special (downhill) direction and thereby breaking the full rotation symmetry. Think of it as a model of the flow of water along a river or traffic along a highway.

$$\partial_t h(x, t) = \nu_{\parallel} \partial_{\parallel}^2 h + \nu_{\perp} \partial_{\perp}^2 h - \frac{\lambda}{2} \partial_{\parallel} (h^2) + \eta(x, t) \quad (7.214)$$

The leading nonlinearity has one fewer derivative than that of the KPZ equation. In the absence of noise the evolution equation is of the form of a continuity equation $\partial_t h + \vec{\nabla} \cdot \vec{j} = 0$. The noise (drainage/rain or exits/entrances on the freeway) need not be conservative (i.e. involve extra derivatives which would be required to include it in the current) and we'll just take it to be white and Gaussian.

This is a hydrodynamic description of the following kind of sandpile model: the dofs are an integer height variable $z(x)$ at each site x . If for a pair of neighbors $z(x) - z(x+a) > z_c$, we do $z(x) \rightarrow z(x) - 1, z(x+a) \rightarrow z(x+a) + 1$. Sand is randomly added and allowed to escape from one of the boundaries, which establishes a transport direction $\hat{T}$. $\vec{x}_{\parallel} = (\hat{T} \cdot \hat{x}) \hat{T}$ and $\vec{x}_{\perp} = \vec{x} - \vec{x}_{\parallel}$.

In this example something very interesting happens in the RG equations. As for KPZ, the model has a Galilean symmetry

$$h(\vec{x}, t) \mapsto h(x_{\parallel} - \epsilon \lambda t, t) + \epsilon \quad (7.215)$$

that depends on λ which again means that λ is not renormalized. Secondly, because of the $\parallel$ derivative in the nonlinearity, neither $\nu_{\perp}$ nor D are renormalized by loops; dimensional analysis gives the right answer for their scaling. So the only coupling that runs nontrivially is $\nu_{\parallel}$, and the three exponents z, χ, ζ appearing in the scale transformation

$$h(\vec{x}_{\parallel}, \vec{x}_{\perp}, t) \mapsto b^{\chi} h(b^{-1} \vec{x}_{\parallel}, b^{-\zeta} \vec{x}_{\perp}, b^{-z} t) \quad (7.216)$$

are known exactly below the upper critical dimension $d \leq 4$: $\chi = \frac{1-d}{7-d}, z = \frac{6}{7-d}, \zeta = \frac{3}{7-d}$.

This is the end of the lecture notes. Here are a few loose threads I'd like to record:

- I'm frustrated that I didn't find a clear unified understanding of generic scale invariance. I'd like to understand better the work of Levina-Herrmann-Geisel on homeostatic plasticity.

- I wanted to say something about predator-prey systems including spatial information. Simulations of these exhibit some amazing patterns and there is some quantitative discussion in the book by Tauber and *e.g.* [here](#).
- I wanted to make contact between the KPZ discussion and Tracy-Widom statistics.
- It would have been nice to talk about fluid turbulence. There is some nice discussion in the new book by Nelson and Zhang.

Some other topics for the future:

- Mean-field games
- Spontaneous stochasticity
- Statistical physics of machine learning