CHAOS AND THERMALISATION

DAVID WAKEHAM

Abstract. A quick and not particularly rigorous introduction to chaos and thermalisation in quantum mechanics. The main goals are to motivate (with a minimum of technical fuss) (a) *spectral chaos*, (b) the *Eigenstate Thermalisation Hypothesis*, (c) *out-of-time-order correlators*, and (d) the *chaos bound* on Lyapunov growth.

Contents

1. Introduction

I'm going to be talking about chaos and thermalisation in quantum mechanics, but before I do, I'm going to tell you a classical fable. If we drop a bunch of particles into a box and let it settle down, later measurements will agree with equilibrium statistical mechanics. The system has *thermalised*. But this raises two questions. First, what do we mean by agreement with equilibrium statistical mechanics? Second, what process enables thermalisation, and how quickly does it happen?

Let's start with the first question. By equilibrium statistical mechanics, we usually mean an ensemble, either the *microcanonical ensemble* where we fix energy, or the *canonical ensemble* where we fix temperature. The ensembles are defined as probability distributions over classical phase space:

$$\rho_E(\mathbf{x}, \mathbf{p}) \propto \delta[H(\mathbf{x}, \mathbf{p}) - E], \quad \rho_\beta(\mathbf{x}, \mathbf{p}) \propto e^{-\beta H(\mathbf{x}, \mathbf{p})}.$$

These give the same answers up to corrections suppressed by the system size, which therefore go away in the thermodynamic limit where the systems get large.

Date: December 12, 2018.

The first type of thermalisation we learn about is ergodicity, where the long time average of a measurement equals the ensemble average:

$$\lim_{t \to \infty} \frac{1}{t} \int_0^t dt' \, A(t') = \frac{1}{(2\pi\hbar)^{3N}} \int d\mathbf{x} \, d\mathbf{p} \, \rho A := \langle A \rangle.$$

But it's hard to measure these time averages experimentally, since you would need to continuously track things. Typically, an experimentalist just waits a while, then takes a single measurement. Empirically, we know that most of the time this single measurement agrees with classical statistical mechanics, again up to fluctuations suppressed by the system size (which is the same order as the difference between ensembles):

$$A(t) = \langle A \rangle + \mathcal{O}(N^{-1}) \quad \text{for most } t > t_{\text{relax}}.$$

You can be unlucky and measure a non-thermal result, but these deviations are rare. This concept of individual measurements being thermal is called *strong thermalisation*, because it's stronger than ergodicity, which in contrast we call *weak thermalisation*.

Now, if the system has too many conserved quantities (it's *integrable*) it won't thermalise at all. Instead, it will get restricted to a tiny sliver of phase space consistent with those conserved quantities. Otherwise, the system is *chaotic*. The degrees of freedom interact in an effectively random way, well-described by statistical mechanics, and as a result of these random interactions, the trajectory spreads throughout phase space and can mimic an ensemble.

How quickly does the trajectory spread around? The usual way of capturing this is to look at how quickly nearby trajectories deviate from each other. If the difference between initial conditions at $t = 0$ is $\delta x(0)$, and the difference at later times is $\delta x(t)$, we find that in chaotic systems, the difference grows exponentially quickly:

$$\delta x(t) = e^{\lambda t} \delta x(0).$$

This is called the *butterfly effect* or *sensitivity to initial conditions*, and λ is the *Lyapunov exponent*. The time scale for thermalisation, which requires trajectories to spread through the typical region of phase space, is then $1/\lambda$. Note that this is different from ergodicity! This requires us to explore all of phase even-handedly, including the atypical corners. But to strongly thermalise, we only need to explore the typical region, and we can do that efficiently using chaos.

These are the basic ideas behind classical thermalisation and chaos. The goal of this talk is going to be answer the same two questions for quantum mechanics: what is quantum thermalisation and how does it happen? Since Hilbert space is very different from classical phase space, we can't just port over our classical definitions. We're going to look at two different approaches. The first approach studies pure states, and how they can thermalise under unitary evolution. This culminates in the notions of *spectral chaos* (or random matrix behaviour) and the *eigenstate thermalisation hypothesis*. The second approach considers how perturbations of a thermal state die away with time, which is a distinct notion of thermalisation. This leads to a quantum

version of the butterfly effect, and the *chaos bound* on the corresponding Lyapunov exponents. Throughout, I'll be emphasising ideas over proofs.

2. Quantum ergodicity

To begin with, let's just try the simplest approach to thermalisation and see what happens. In classical mechanics, the microcanonical and canonical ensembles are probability distributions over phase space. In quantum mechanics, we don't have a phase space, but we can introduce classical uncertainty using *mixed states* or *density matrices*. You can view these as a probability distribution over pure states:

$$\hat{\rho} = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \quad \sum_i p_i = 1.$$

We can define quantum thermalisation using the mixed state analogues of the microcanonical and canonical ensemble. We say that a pure state thermalises if, after some time has passed, the expectation for macroscopic observables $\hat{O}$ agrees with expectation in the equilibrium mixed state, up to some small corrections. We can generalise all of this very easily to mixed state thermalisation, so we're going to focus on pure states.

Let's explore this idea for the microcanonical ensemble, which corresponds to a uniform probability distribution over all the energy eigenstates states in some energy shell of width 2Δ centred around E, $I = [E - \Delta, E + \Delta]$. The density matrix is

$$\hat{\rho}_{E,\Delta} \propto \sum_{\epsilon_n \in I} |n\rangle\langle n|.$$

This simply projects onto the associated subspace of energy eigenstates $\mathcal{H}_{E,\Delta} = \text{span}_{\epsilon_n \in I}\{|n\rangle\}$, with dimension $D = |\mathcal{H}_{E,\Delta}|$. To get emergent statistical behaviour, we choose E to be large so that levels are closely spaced. We then choose Δ wide enough that I contains many energy levels, but narrow enough that states are hard to distinguish macroscopically.

Let's pick a pure state in this microcanonical shell,

$$|\psi\rangle = \sum_{\epsilon_n \in I} C_n |n\rangle.$$

We can calculate the expectation of an operator $\hat{O}$ with respect to the pure state $|\psi(t)\rangle$:

$$\langle\hat{O}\rangle_t = \langle\psi(t)|\hat{O}|\psi(t)\rangle$$
$$= \sum_{mn} C_m^* C_n e^{i(\epsilon_m - \epsilon_n)t} O_{mn}$$
$$= \sum_m |C_m|^2 O_{mm} + \sum_{m \neq n} C_m^* C_n e^{i(\epsilon_m - \epsilon_n)t} O_{mn}$$

where O_{mn} are the matrix elements

$$O_{mn} = \langle m|\hat{O}|n\rangle.$$

The microcanonical average, in comparison, is

$$\langle \hat{O} \rangle_{E,\Delta} = \text{Tr}[\hat{\rho}_{E,\Delta} \hat{O}] = \frac{1}{D} \sum_m O_{mm}.$$

Of course, we can pick the coefficients C_m arbitrarily, so it seems like there is no way the two can agree.

But there is a trick: if the operators $\hat{O}$ are macroscopic, they cannot distinguish different eigenstates in the shell! So for any eigenstate $|m\rangle$, $O_{mm} \approx \bar{O}_E$ and hence

$$\langle \hat{O} \rangle_t = \bar{O}_E \sum_m |C_m|^2 + \text{off-diagonal} = \langle \hat{O} \rangle_{E,\Delta} + \text{off-diagonal}.$$

Since the off-diagonal terms are oscillating, their long-time average vanishes, and any pure state reproduces the microcanonical average ergodically. It follows that any pure state in the energy shell weakly thermalises. But as von Neumann proved long ago, the off-diagonal matrix elements are typically very small, so every pure state *strongly thermalises* as well. He called this the *quantum ergodic theorem*, though it is really a stronger result than ergodicity.

3. Random matrix elements

One way to see that off-diagonal elements are typically small is to choose the operator $\hat{O}$ randomly. (This is different from the method that von Neumann used, but it will be thematically apposite.) The operator has a spectral decomposition

$$\hat{O} = \sum_\lambda O_\lambda |\lambda\rangle\langle\lambda|,$$

with random eigenvectors $|\lambda\rangle$ in the energy basis. This means that, on average, each $|\lambda\rangle$ is equally distributed between energy eigenstates, with

$$|\langle n|\lambda\rangle| \sim \frac{1}{\sqrt{D}},$$

We can treat the overlaps $\langle n|\lambda\rangle =: \psi_{\lambda n}$ as independent random variables for different λ and n. Hence,

$$\overline{\psi_{\lambda m}\psi_{\lambda n}^*} = \frac{1}{D}\delta_{mn}.$$

We can then compute the average matrix element:

$$\overline{O_{mn}} = \overline{\langle m|\hat{O}|n\rangle} = \sum_\lambda O_\lambda \overline{\langle m|\lambda\rangle\langle\lambda|n\rangle} = \frac{1}{D}\delta_{mn} \sum_\lambda O_\lambda = \bar{O}\delta_{mn}.$$

This confirms our earlier claim that the diagonal elements typically give the same answer $\bar{O}$, while off-diagonal elements vanish.

We can take this calculation further, and determine the *fluctuations* in these matrix elements. I won't go through the details here, but it's straightforward to show that both diagonal and off-diagonal elements have small fluctuations, suppressed by the size of the system:

$$\delta O_{mn}^2 \sim \frac{1}{D}\overline{O^2}.$$

Combining these observations, the matrix elements of a random operator should take the form:

$$O_{mn} = \bar{O}\delta_{mn} + \sqrt{\frac{\overline{O^2}}{D}}R_{mn},$$

where R_{mn} is a random matrix with zero mean and unit variance, built out of the overlap variables $\psi_{\lambda n}$.

If we return to the statement of quantum ergodicity, we note that the off-diagonal term has the form

$$\sum_{m\neq n} C_m^* C_n e^{i(\epsilon_m - \epsilon_n)t} O_{mn} \sim \sqrt{\frac{\overline{O^2}}{D}}\sum_{m\neq n} C_m^* C_n e^{i(\epsilon_m - \epsilon_n)t}.$$

At $t = 0$, the coefficients C_m can be chosen so that the sum is actually order D, and the off-diagonal elements are *not* small! However, after some time evolution, if the energies are somewhat random the sum will partially dephase and it will go from an $\mathcal{O}(D)$ number to $\mathcal{O}(1)$ number. This means that the state has thermalised. People often say that the role of time evolution is not to *create* thermalisation, but to *reveal*, since it is already present and appears once enough dephasing has occurred.

4. Spectral chaos and ETH

But this raises the question of how eigenenergies are distributed. This is presumably going to determine how long it takes a state to thermalise, or even if it thermalises at all, since you can imagine picking specific eigenenergies to ensure the fluctuations remain large. But the distribution of eigenenergies is a property of the system, not the state and not the operator.

To get random energies, let's pick a *random Hamiltonian* $\hat{H}$, for instance, as a Hermitian matrix with Gaussian entries. The eigenvectors of an operator $\hat{O}$ will still look random in the energy basis, so our matrix element calculation goes through. But now we'll have a probability distribution for the energy eigenvalues ϵ_m themselves, called the *level statistics* for the system. I should emphasise that our earlier argument applied to a randomly chosen macroscopic observable in any system (with a few nice properties). We're now considering a very special sort of system, motivated by the queston of when energies look random.

Let's do a quick example. Consider a 2×2 Hermitian matrix, where we can easily calculate the eigenvalues:

$$\hat{H} = \begin{bmatrix} \alpha & V^*/\sqrt{2} \\ V/\sqrt{2} & \beta \end{bmatrix}, \quad \epsilon_{1,2} = \frac{\alpha + \beta}{2} \pm \frac{1}{2}\sqrt{(\alpha - \beta)^2 + 2|V|^2}.$$

Suppose we draw $\alpha, \beta, \Im(V), \Re(V)$ from a Gaussian distribution with zero mean and variance σ^2. The average energy is Gaussian, which is a bit boring, but after a little work, you find that the gap between energy levels $\omega := \epsilon_1 - \epsilon_2$ has a probability distribution

$$P(\omega) \propto \omega^2 \exp\left[-\frac{\omega^2}{4\sigma^2}\right].$$

The spacing ω exhibits *level repulsion*, since the ω^2 term forbids energy levels to be close. Eigenvalues are effectively repelled! So in the case of a random Hamiltonian, you can usually assume there is no degeneracy.

This random Hamiltonian model turns out to beautifully match the level statistics of a messy system. In fact, Wigner invented random matrix theory to explain the messy spectra of atomic nuclei! Random Hamiltonians couple degrees of freedom in a random way; this should remind us of the random classical interactions we used to characterise chaos earlier. Pursuing this analogy, we *define* a quantum system as chaotic if it exhibits random level statistics. This type of chaos is called *spectral chaos* to distinguish it from *dynamical chaos* or the butterfly effect. As a simple check, if you take a classically chaotic system and quantise it, then (apart from a small set of pathological counterexamples) you always get random level statistics. Conversely, if you take a non-chaotic classical system and quantise it, you don't get random level statistics. So spectral chaos seems to be tracking something real!

Random matrix theory gives us an ansatz for matrix elements in the microcanonical energy shell. But in real life, finite systems are usually at fixed temperature rather than fixed energy. This is easier to set up experimentally. Fluctuations will allow us to leave the energy shell, and to have matrix elements between eigenstates with macroscopically different energies. We need to go beyond random matrix theory to figure out what these fluctuations look like. For two eigenstates $|m\rangle, |n\rangle$ with energies ϵ_m, ϵ_n, finite temperature fluctuations should depend on the average energy $\bar{E}$ and the energy difference ω,

$$\bar{E} := \frac{\epsilon_m + \epsilon_n}{2}, \quad \omega := |\epsilon_m - \epsilon_n|.$$

Of course, fluctuations also depend on the observable $\hat{O}$. Earlier, we had the xpectation $\overline{O^2}$, but more complicated sorts of dependence may be possible as well.

We can throw all this dependence into an *envelope function* $f_O(\bar{E}, \omega)$ which modifies the fluctuations. This leads to a new ansatz for matrix elements at finite temperature called the *eigenstate thermalisation hypothesis*:

$$O_{mn} = O(\bar{E})\delta_{mn} + e^{-S(\bar{E})/2} f_O(\bar{E}, \omega) R_{mn},$$

since e^S is density of states, and $e^{-S/2}$ replaces $1/\sqrt{D}$. This is expected to hold for systems with a classically chaotic counterpart, for high energy eigenstates and for observables which are sufficiently local. Part of the hypothesis is that the functions $f_O(\bar{E}, \omega)$ and $O(\bar{E})$ are *smooth*, with chaotic interactions among many levels ironing out any kinks.

5. Fluctuation-dissipation

We're going to look at one application of the ETH, which is slightly technical but has a nice payoff. It turns out that chaotic eigenstates not only reproduce the equilibrium behaviour of observables, but also the *fluctuation-dissipation theorem*. Loosely speaking, thermodynamic systems like to oscillate at certain frequencies both in equilibrium and out of equilibrium, with the spectrum of thermal fluctuations determining how the system dissipates energy in the mildly non-equilibrium regime where it is perturbed.

As a warm-up, let's consider fluctuations of an observable $\hat{O}$ in the energy eigenstate $|n\rangle$, in other words, the variance. Assuming the state and observable satisfy the ETH, we get the sum of off-diagonal terms

$$\delta O_n^2 := \langle \hat{O}^2 \rangle_n - \langle \hat{O} \rangle_n^2 = \sum_{m \neq n} e^{-S(\bar{E})} |f_O(\bar{E}, \omega)|^2 |R_{mn}|^2.$$

Since R_{mn} is unit variance, we are going to make the replacement $|R_{mn}|^2 \to 1$. Using the smoothness of the envelope function f_O and assuming high level density, we can rewrite the sum as an integral over $\omega = \epsilon_m - \epsilon_n$ (the signed difference between energies):

$$\delta O_n^2 \approx \int d\omega \, e^{S(E+\omega) - S(E+\omega/2)} |f_O(E + \omega/2, \omega)|^2,$$

inserting the density of states $e^{S(E+\omega)}$ and rewriting $\bar{E} = E + \omega/2$. For few-body observables, f_O should drop off rapidly as a function of ω, so we can treat ω as small and Taylor expand the entropy S and the envelope function:

$$S(E + \omega) - S(E + \omega/2) \approx \frac{\omega}{2} \frac{\partial S}{\partial E} \approx \frac{\beta \omega}{2}, \quad f_O(E + \omega/2, \omega) \approx f_O(E, \omega) + \frac{\partial f_O}{\partial E} \frac{\omega}{2}.$$

We will keep only leading order terms for simplicity. This gives us an approximate expression for fluctuations in terms of f_O:

$$\delta O_n^2 \approx \int d\omega \, e^{\beta \omega/2} |f_O(E, \omega)|^2.$$

To find the power spectrum of fluctuations in $\hat{O}$, we need to do something a bit fancier, and look at the overlap between $\hat{O}$ at $t = 0$ and $\hat{O}$ at some later time t. Fourier-transforming that will give us the spectrum of fluctuations. This overlap, more precisely, is a connected 2-point function, or covariance:

$$C_O(t) := \langle \hat{O}(t) \hat{O}(0) \rangle_n - \langle \hat{O}(t) \rangle_n \langle \hat{O}(0) \rangle_n,$$

where $\hat{O}(t) = e^{i\hat{H}t} \hat{O} e^{-i\hat{H}t}$ is the time-evolved operator in the Heisenberg picture. This covariance is calculated in a similar way to the variance, except that the evolution operators yield a factor of $e^{i\omega t}$. After some cancellations, we end up with the expression

$$C_O(t) = \int d\omega \, e^{-i\omega t + \beta \omega/2} |f_O(E, \omega)|^2.$$

But this is manifestly a Fourier transform! The energy gap ω is playing the role of a frequency, so we can directly read off the spectrum:

$$\tilde{C}_O(\omega) = 2\pi e^{\beta \omega/2} |f_O(E, \omega)|^2.$$

To get a real-valued quantity, it's conventional to look instead at the *symmetrised* correlation functions:

$$\tilde{C}_O^+(\omega) := \tilde{C}_O(\omega) + \tilde{C}_O(-\omega) = 4\pi \cosh(\beta \omega/2) |f_O(E, \omega)|^2.$$

This is the fluctuation part of the theorem.

The second part of the theorem involves the *linear response* χ_O of the system. Specifically, we will perturb the Hamiltonian with a current $\phi\hat{O}$, and see how measurements of $\hat{O}$ change later in time. The *susceptibility* can be defined as a functional derivative, or equivalently, as the kernal appearing in a convolutional integral for the system's response to the perturbation:

$$\delta\langle\hat{O}(t)\rangle = \langle\hat{O}(t)\rangle_{\hat{H}'} - \langle\hat{O}(t)\rangle_{\hat{H}} = \int d\tau\,\phi(\tau)\chi_O(t-\tau)$$

$$\iff \quad \frac{\delta\langle\hat{O}(t)\rangle}{\delta\phi(\tau)} = \chi_O(t)\delta(t-\tau).$$

Kubo's formula tells us that we can compute the susceptibility by looking at the commmutator of the measurement and the perturbing operator, so in our case, we can directly write down the Fourier transformed susceptibility:

$$\tilde{\chi}_O(\omega) = i\int_0^\infty dt\,e^{i\omega t}\langle[\hat{O}(t),\hat{O}(0)]\rangle_n = i\int_0^\infty dt\,e^{i\omega t}\left[C_O(t) - C_O(-t)\right],$$

since $\langle\hat{O}(0)\hat{O}(t)\rangle_c = \langle\hat{O}(-t)\hat{O}(0)\rangle_c = C_O(-t)$. Inverting the Fourier transform and doing a little algebra, we find that

$$\tilde{\chi}_O(\omega) = 2\pi i\sinh(\beta\omega/2)|f_O|^2 + \text{real part}.$$

The imaginary part of the susceptibility is called the *dissipation*, since it's that bit that leads to energy losses when we perturb. Comparing the power spectrum and the dissipation, we reproduce the famous quantum fluctuation-dissipation relation, but now for individual eigenstates:

$$\tilde{C}_O^+(\omega) = 2\coth(\beta\omega/2)\Im[\tilde{\chi}_O(\omega)].$$

6. Scrambling and butterflies

The fluctuation-dissipation theorem tells us about the system's response when we perturb and then measure the same operator. But what if we want to perturb with one operator, and measure with another? It turns out that this requires a nontrivial modification of the ETH. Even then, the response is tied up in the envelope functions f_O, which we don't know explicitly. The ETH is great for understanding in fine detail how pure states thermalise, but it's too delicate to give robust or generic answers about perturbing thermal states. This would be like trying to do classical statistical mechanics using a single chaotic trajectory! The point of thermalisation is to understand when we can ignore these fine details. To start exploring generic perturbations, let's abandon the ETH and move to the canonical ensemble. We can treat this as a mixed state with density

$$\hat{\rho}_\beta = \frac{1}{Z[\beta]}e^{-\beta\hat{H}}.$$

Alternatively, this arises as a pure state $|\beta\rangle$ on the system plus a reservoir; we recover the density when we trace out the reservoir degrees of freedom.

If we poke the thermal state, it can "forget" the perturbation in stronger or weaker ways. A weak type of forgetting is *Ruelle behaviour*, where the perturbation leaks into

harmonics of the system, a bit like hitting a gong. This is reflected in exponentially decaying two-point functions. A second, much stronger notion is *scrambling*, where a local perturbation spreads around and can only be revealed by access to an extensive chunk of the system's degrees of freedom. It will be invisible to local measurements. For instance, if you take a thesis and put it in a document shredder, that information is scrambled and hard to recover using a local operation like reading. You need to assemble some extensive fraction of shreds in order to figure out what the thesis is about.

Although both Ruelle behaviour and scrambling are types of thermalisation, I'm going to focus instead on the related notion of *scrambling as dynamical chaos*. Let's first go back to the classical butterfly effect, where the distance between nearby classical trajectories grows exponentially:

$$\frac{\partial q(t)}{\partial q(0)} \approx \frac{\delta q(t)}{\delta q(0)} = e^{\lambda t},$$

where λ is our classical Lyapunov exponent. We can rewrite this derivative using the classical Poisson bracket with respect to canonical variables at $t = 0$:

$$\frac{\partial q(t)}{\partial q(0)} = \{q(t), p(0)\}.$$

If we start the system in a thermal state, and perturb it in slightly different ways, the resulting probability distributions move apart chaotically as well. We can measure this spread by squaring the Poisson bracket (to make it positive definite) and taking the thermal average:

$$\left\langle \left(\frac{\partial q(t)}{\partial q(0)} \right)^2 \right\rangle_\beta = \left\langle \{q(t), p(0)\}^2 \right\rangle_\beta \sim e^{2\lambda t}.$$

We now just blindly quantise this by changing the bracket to a commutator, and promoting canonical variables to Heisenberg-picture operators:

$$\left\langle \{q(t), p(0)\}^2 \right\rangle_\beta \quad \Longrightarrow \quad -\langle [\hat{q}(t), \hat{p}(0)]^2 \rangle_\beta.$$

We're going to cross our fingers and hope this means something interesting quantum-mechanically.

But first of all, we need to modify this a bit to make it more appropriate to a quantum many-body system. In something like a spin system, the simple observables are not positions and momenta, but operators involving a few spins. More generally, we would like to replace $\hat{p}$ and $\hat{q}$ with *approximately local observables* V and W involving only a few degrees of freedom. If we plug V and W into our earlier expression, we get what's called the *double commutator*:

$$D(t) := -\langle [W(t), V(0)]^2 \rangle_\beta.$$

If V and W act on different spins, for instance, they will initially commute, so $D(t)$ should be small near $t = 0$. But as t increases, the perturbation V will spread, provided the system scrambles effectively, and eventually the commutator with $W(t)$ will increase.

We can make this story much more precise. First, expand $D(t)$ and tidy up the terms a little:

$$D(t) = \langle W(t)VVW(t)\rangle_\beta + \langle VW(t)W(t)V\rangle_\beta - \langle W(t)VW(t)V\rangle_\beta - \langle VW(t)VW(t)\rangle_\beta$$
$$= 2 - 2\Re\langle W(t)V(0)W(t)V(0)\rangle_\beta.$$

Here we used the fact that the first two terms are just norms of states (assuming V and W are Hermitian), while the last two are conjugate. The term whose real part we are taking is called the *out-of-time-ordered correlator* (OTOC):

$$\mathrm{OTOC}(t) = \langle W(t)V(0)W(t)V(0)\rangle_\beta = \langle\beta|W(t)V(0)W(t)V(0)|\beta\rangle.$$

It is "out-of-time order" since the operators zig-zag backwards and forwards in time.

The OTOC computes the quantum equivalent of the butterfly effect in a rather beautiful way. Remember that time evolution is unitary, so nearby states remain nearby. We can't get sensitive dependence by comparing two trajectories in Hilbert space, since they just stick together forever. However, we can get sensitive dependence by asking about how measurements of V and W affect each other. Imagine we evolve the state forward in time, do a small perturbation W, and evolve backwards in time to where we started. This is the same as applying $W(t) = e^{i\hat{H}t}We^{-i\hat{H}t}$. The backwards evolution can take us to a state which, although close to the thermal state in Hilbert space (by unitarity) has a very different operator structure, with W spread throughout the system. This new state will respond very differently to a second perturbation V.

We can either do V before the W perturbation, in which case it goes at the top of the diagram, or after the W perturbation, in which case it goes at the bottom. These two states can be completely different! This failure of local measurements to commute when one of them is scrambled is the quantum version of dynamical chaos or the butterfly effect. We can easily measure how different these two states are by taking the overlap. This is precisely what the OTOC computes, with

$$\mathrm{OTOC}(t) = \langle\psi_2|\psi_1\rangle, \quad |\psi_1\rangle := W(t)V(0)|\beta\rangle, \quad |\psi_2\rangle := V(0)W(t)|\beta\rangle.$$

7. Bound on chaos

In the last section, we're going to think a bit more about the mechanism for operator spreading. This will give us some insights into the time-dependence of dynamical chaos. Let's consider how a local perturbation grows with time in the Heisenberg picture. Time evolution can be written as a power series of nested commutators, using either the Baker-Campbell-Hausdorff formula, or exponentiating the Liouvillian:

$$W(t) = e^{i\hat{H}t}We^{-i\hat{H}t} = e^{i\mathcal{L}}W = \sum_{n\geq 0}\frac{(-it)^n}{n!}\overbrace{[\hat{H},\cdots[\hat{H},W]\cdots]}^{n\text{ times}}.$$

The point is that each time we commute, we add more and more local operators to $W(t)$ until it saturates the entire system, at which point we say that $W(t)$ has been *scrambled*. The time it takes for this to happen is called the *scrambling time t_**.

We can actually estimate the scrambling time for a simple system using dimensional analysis. Imagine that each site in our lattice has k neighbours in the graph of interactions. At each effective time step, the operator spreads to the neighbours of sites already involved, so the number of sites (or local operators) after n steps is roughly k^n. If the total system has N degrees of freedom, it then takes $\sim \log_k N$ time steps to completely scramble a local operator. It remains to determine the size of an effective time step. In a thermal state of energy $1/\beta$ and no other energy scales, the answer has to be $\sim \hbar\beta$ by dimensional analysis. Putting it all together, we find a scrambling time

$$t_* \sim \hbar\beta \log N.$$

This simple model of operator growth also tells us how the OTOC should grow with time. Early on, the overlap between the states $|\psi_1\rangle$ and $|\psi_2\rangle$ is close to unity, but as the perturbation spreads exponentially through the system, the overlaps drops exponentially quickly:

$$\text{OTOC}(t) \approx 1 - e^{\lambda t},$$

for some *quantum Lyapunov exponent* λ. It follows that the double commutator obeys Lyapunov-like exponential growth until the scrambling time:

$$D(t) \approx 2e^{\lambda t}, \quad t < t_*.$$

This strengthens the resemblance to the butterfly effect further.

The dimensional argument suggests that when temperature is the only energy scale, the Lyapunov exponent is

$$\lambda \sim \frac{1}{\hbar\beta} = \frac{k_B T}{\hbar}.$$

But most systems have scales! These can come from a lattice spacing, interaction strengths, and so on. Once you have structure, you can get *structural bottlenecks* like separation of scales, and these restrict the flow of information in the system. It seems reasonable to expect that extra structure slows scrambling down, and the answer from dimensional analysis is a maximum.

In fact, it's possible to *prove* that this is true using results from complex analysis and some clever manipulation of the correlation functions. The precise form of the bound on the Lyapunov exponent is as follows:

$$\lambda \leq \frac{2\pi k_B T}{\hbar}.$$

This is just our dimensional guess with an extra factor of 2π. Since λ controls the scrambling, this is called the *chaos bound*. The chaos bound has fascinating connections to black holes and string theory, but that is a story for another time.

8. References

(1) To do!