Contents

List of Tables

List of Figures

Chapter 3

Ergodicity and the Approach to Equilibrium

3.1 References

- R. Balescu, *Equilibrium and Nonequilibrium Statistical Mechanics* (Wiley, 1975)
  An advanced text with an emphasis on fluids and kinetics.

- R. Balian, *From Macrophysics to Microphysics* (2 vols., Springer-Verlag, 2006)
  A very detailed discussion of the fundamental postulates of statistical mechanics and their implications.)

3.2 Modeling the Approach to Equilibrium

3.2.1 Equilibrium

A thermodynamic system typically consists of an enormously large number of constituent particles, a typical 'large number' being Avogadro's number, $N_{\mathrm{A}} = 6.02 \times 10^{23}$. Nevertheless, in *equilibrium*, such a system is characterized by a relatively small number of thermodynamic state variables. Thus, while a complete description of a (classical) system would require us to account for $\mathcal{O}(10^{23})$ evolving degrees of freedom, with respect to the physical quantities in which we are interested, the details of the initial conditions are effectively forgotten over some microscopic time scale τ, called the collision time, and over some microscopic distance scale, ℓ, called the mean free path[1]. The equilibrium state is time-independent.

3.2.2 The Master Equation

Relaxation to equilibrium is often modeled with something called the *master equation*. Let $P_i(t)$ be the probability that the system is in a quantum or classical state i at time t. Then write

$$\frac{dP_i}{dt} = \sum_j \left(W_{ij}\, P_j - W_{ji}\, P_i \right) \quad . \tag{3.1}$$

Here, W_{ij} is the rate at which j makes a transition to i. Note that we can write this equation as

$$\frac{dP_i}{dt} = -\sum_j \Gamma_{ij}\, P_j \quad , \tag{3.2}$$

where

$$\Gamma_{ij} = \begin{cases} -W_{ij} & \text{if } i \neq j \\ \sum_k' W_{kj} & \text{if } i = j \end{cases} \quad , \tag{3.3}$$

where the prime on the sum indicates that $k = j$ is to be excluded. The constraints on the W_{ij} are that $W_{ij} \geq 0$ for all i, j, and we may take $W_{ii} \equiv 0$ (no sum on i). Fermi's Golden Rule of quantum mechanics says that

$$W_{ij} = \frac{2\pi}{\hbar} \left| \langle\, i\, |\, \hat{V}\, |\, j\, \rangle \right|^2 \rho(E_j) \quad , \tag{3.4}$$

where $\hat{H}_0\, |\, i\, \rangle = E_i\, |\, i\, \rangle$, $\hat{V}$ is an additional potential which leads to transitions, and $\rho(E_i)$ is the density of final states at energy E_i. The fact that $W_{ij} \geq 0$ means that if each $P_i(t = 0) \geq 0$, then $P_i(t) \geq 0$ for all $t \geq 0$. To see this, suppose that at some time $t > 0$ one of the probabilities P_i is crossing zero and about to become negative. But then eqn. 3.1 says that $\dot{P}_i(t) = \sum_j W_{ij} P_j(t) \geq 0$. So $P_i(t)$ can never become negative.

---

[1] Exceptions involve quantities which are conserved by collisions, such as overall particle number, momentum, and energy. These quantities relax to equilibrium in a special way called *hydrodynamics*.

3.2.3   Equilibrium distribution and detailed balance

If the transition rates W_{ij} are themselves time-independent, then we may formally write

$$P_i(t) = \left(e^{-\Gamma t}\right)_{ij} P_j(0) \quad . \tag{3.5}$$

Here we have used the Einstein 'summation convention' in which repeated indices are summed over (in this case, the j index). Note that $\sum_i \Gamma_{ij} = 0$, which says that the total probability $\sum_i P_i$ is conserved:

$$\frac{d}{dt} \sum_i P_i = -\sum_{i,j} \Gamma_{ij} P_j = -\sum_j \left(P_j \sum_i \Gamma_{ij} \right) = 0 \quad . \tag{3.6}$$

We conclude that $\vec{\phi} = \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix}$ is a left eigenvector of Γ with eigenvalue $\lambda = 0$. The corresponding right eigenvector, which we write as P_i^{eq}, satisfies $\Gamma_{ij} P_j^{\text{eq}} = 0$, and is a stationary (*i.e.* time independent) solution to the master equation. Generally, there is only one right/left eigenvector pair corresponding to $\lambda = 0$, in which case any initial probability distribution $P_i(0)$ converges to P_i^{eq} as $t \to \infty$, as shown in the appendix §3.7. Note, however, that since the matrix Γ is in general not symmetric, its eigenvectors may not span, which is to say that it may contain nontrivial Jordan blocks when it is brought to canonical form. See the appendix in §3.6 below for a complete discussion.

In equilibrium, the net rate of transitions into a state $|\,i\,\rangle$ is equal to the rate of transitions out of $|\,i\,\rangle$. If, for each state $|\,j\,\rangle$ the transition rate from $|\,i\,\rangle$ to $|\,j\,\rangle$ is equal to the transition rate from $|\,j\,\rangle$ to $|\,i\,\rangle$, we say that the rates satisfy the condition of *detailed balance*. In other words, $W_{ij}\, P_j^{\text{eq}} = W_{ji}\, P_i^{\text{eq}}$. Assuming $W_{ij} \neq 0$ and $P_j^{\text{eq}} \neq 0$, we can divide to obtain

$$\frac{W_{ji}}{W_{ij}} = \frac{P_j^{\text{eq}}}{P_i^{\text{eq}}} \quad . \tag{3.7}$$

Note that detailed balance is a stronger condition than that required for a stationary solution to the master equation.

If $\Gamma = \Gamma^{\text{t}}$ is symmetric, then the right eigenvectors and left eigenvectors are transposes of each other, hence $P^{\text{eq}} = 1/N$, where N is the dimension of Γ. The system then satisfies the conditions of detailed balance. See §3.7.2 below for an example of this formalism applied to a model of radioactive decay.

3.2.4   Boltzmann's $\mathcal{H}$-theorem

Suppose for the moment that Γ is a symmetric matrix, *i.e.* $\Gamma_{ij} = \Gamma_{ji}$. Then construct the function

$$\mathcal{H}(t) = \sum_i P_i(t) \ln P_i(t) \quad . \tag{3.8}$$

Then

$$\frac{d\mathcal{H}}{dt} = \sum_i \frac{dP_i}{dt} \left(1 + \ln P_i\right) = \sum_i \frac{dP_i}{dt} \ln P_i$$

$$= -\sum_{i,j} \Gamma_{ij}\, P_j \ln P_i = \sum_{i,j} \Gamma_{ij}\, P_j \big(\ln P_j - \ln P_i\big) \quad , \tag{3.9}$$

where we have used $\sum_i \Gamma_{ij} = 0$. Now switch $i \leftrightarrow j$ in the above sum and add the terms to get

$$\frac{d\mathcal{H}}{dt} = \frac{1}{2} \sum_{i,j} \Gamma_{ij} \left(P_i - P_j \right) \left(\ln P_i - \ln P_j \right) \quad . \tag{3.10}$$

Note that the $i = j$ term does not contribute to the sum. For $i \neq j$ we have $\Gamma_{ij} = -W_{ij} \leq 0$, and using the result

$$(x - y) \left(\ln x - \ln y \right) \geq 0 \quad , \tag{3.11}$$

we conclude

$$\frac{d\mathcal{H}}{dt} \leq 0 \quad . \tag{3.12}$$

In equilibrium, P_i^{eq} is a constant, independent of i. We write

$$P_i^{\mathrm{eq}} = \frac{1}{\Omega} \quad , \quad \Omega = \sum_i 1 \quad \Longrightarrow \quad \mathcal{H} = -\ln \Omega \quad . \tag{3.13}$$

If $\Gamma_{ij} \neq \Gamma_{ji}$, we can still prove a version of the $\mathcal{H}$-theorem when there is detailed balance. Define a new symmetric matrix

$$\overline{W}_{ij} \equiv W_{ij} \, P_j^{\mathrm{eq}} = W_{ji} \, P_i^{\mathrm{eq}} = \overline{W}_{ji} \quad , \tag{3.14}$$

and the generalized $\mathcal{H}$-function,

$$\mathcal{H}(t) \equiv \sum_i P_i(t) \, \ln \left(\frac{P_i(t)}{P_i^{\mathrm{eq}}} \right) \quad . \tag{3.15}$$

Then

$$\begin{aligned}
\frac{d\mathcal{H}}{dt} &= -\frac{1}{2} \sum_{i,j} \left(W_{ji} P_i - W_{ij} P_j \right) \ln \left(\frac{W_{ji} P_i}{W_{ij} P_j} \right) \\
&= -\frac{1}{2} \sum_{i,j} \overline{W}_{ij} \left(\frac{P_i}{P_i^{\mathrm{eq}}} - \frac{P_j}{P_j^{\mathrm{eq}}} \right) \left[\ln \left(\frac{P_i}{P_i^{\mathrm{eq}}} \right) - \ln \left(\frac{P_j}{P_j^{\mathrm{eq}}} \right) \right] \leq 0 \quad .
\end{aligned} \tag{3.16}$$

3.3 Phase Flows in Classical Mechanics

3.3.1 Hamiltonian evolution

The master equation provides us with a semi-phenomenological description of a dynamical system's relaxation to equilibrium. It explicitly breaks time reversal symmetry. Yet the microscopic laws of Nature are (approximately) time-reversal symmetric. How can a system which obeys Hamilton's equations of motion come to equilibrium?

Let's start our investigation by reviewing the basics of Hamiltonian dynamics. Recall the Lagrangian $L = L(q, \dot{q}, t) = T - V$. The Euler-Lagrange equations of motion for the action $S\big[q(t)\big] = \int dt \, L$ are

$$\dot{p}_\sigma = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_\sigma} \right) = \frac{\partial L}{\partial q_\sigma} \quad , \tag{3.17}$$

where p_σ is the canonical momentum conjugate to the generalized coordinate q_σ, *i.e.* $p_\sigma = \partial L/\partial \dot{q}_\sigma$. Here N is the number of degrees of freedom of the system, which is the total number of generalized coordinates.

The Hamiltonian, $H(q,p)$ is obtained by a Legendre transformation,

$$H(q,p) = \sum_{\sigma=1}^{N} p_\sigma \, \dot{q}_\sigma - L \quad . \tag{3.18}$$

Note that

$$
\begin{aligned}
dH &= \sum_{\sigma=1}^{N} \left(p_\sigma \, d\dot{q}_\sigma + \dot{q}_\sigma \, dp_\sigma - \frac{\partial L}{\partial q_\sigma} \, dq_\sigma - \frac{\partial L}{\partial \dot{q}_\sigma} \, d\dot{q}_\sigma \right) - \frac{\partial L}{\partial t} \, dt \\
&= \sum_{\sigma=1}^{N} \left(\dot{q}_\sigma \, dp_\sigma - \frac{\partial L}{\partial q_\sigma} \, dq_\sigma \right) - \frac{\partial L}{\partial t} \, dt \quad .
\end{aligned}
\tag{3.19}
$$

Thus, we obtain Hamilton's equations of motion,

$$\frac{\partial H}{\partial p_\sigma} = \dot{q}_\sigma \quad , \quad \frac{\partial H}{\partial q_\sigma} = -\frac{\partial L}{\partial q_\sigma} = -\dot{p}_\sigma \tag{3.20}$$

and

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \quad . \tag{3.21}$$

Define the rank $2N$ vector φ by its components,

$$\varphi_i = \begin{cases} q_i & \text{if } 1 \le i \le N \\ p_{i-N} & \text{if } N \le i \le 2N \end{cases} \quad . \tag{3.22}$$

Then we may write Hamilton's equations compactly as

$$\dot{\varphi}_i = J_{ij} \frac{\partial H}{\partial \varphi_j} \quad , \tag{3.23}$$

where

$$J = \begin{pmatrix} 0_{N\times N} & 1_{N\times N} \\ -1_{N\times N} & 0_{N\times N} \end{pmatrix} \tag{3.24}$$

is a rank $2N$ matrix. Note that $J^{\mathrm{t}} = -J$, *i.e.* J is antisymmetric, and that $J^2 = -1_{2N\times 2N}$.

For any function $F(q,p,t)$, the total time derivative is given by

$$
\begin{aligned}
\frac{dF}{dt} &= \frac{\partial F}{\partial t} + \sum_{\sigma=1}^{N} \left(\frac{\partial F}{\partial q_\sigma} \frac{dq_\sigma}{dt} + \frac{\partial F}{\partial p_\sigma} \frac{dp_\sigma}{dt} \right) \\
&= \frac{\partial F}{\partial t} + \{F, H\} \quad ,
\end{aligned}
\tag{3.25}
$$

where $\{\bullet, \bullet\}$ is the *Poisson bracket*,

$$\{A, B\} = \sum_{\sigma=1}^{N} \left(\frac{\partial A}{\partial q_\sigma} \frac{\partial B}{\partial p_\sigma} - \frac{\partial A}{\partial p_\sigma} \frac{\partial B}{\partial q_\sigma} \right) \quad . \tag{3.26}$$

3.3.2   Dynamical systems and the evolution of phase space volumes

Consider a general dynamical system,

$$\frac{d\boldsymbol{\varphi}}{dt} = \boldsymbol{V}(\boldsymbol{\varphi}) \quad , \tag{3.27}$$

where $\boldsymbol{\varphi}(t)$ is a point in an n-dimensional phase space. Consider now a compact[2] region $\mathcal{R}_0$ in phase space, and consider its evolution under the dynamics. That is, $\mathcal{R}_0$ consists of a set of points $\{\boldsymbol{\varphi} \,|\, \boldsymbol{\varphi} \in \mathcal{R}_0\}$, and if we regard each $\boldsymbol{\varphi} \in \mathcal{R}_0$ as an initial condition, we can define the time-dependent set $\mathcal{R}(t)$ as the set of points $\boldsymbol{\varphi}(t)$ that *were* in $\mathcal{R}_0$ at time $t = 0$:

$$\mathcal{R}(t) = \{\boldsymbol{\varphi}(t) \,|\, \boldsymbol{\varphi}(0) \in \mathcal{R}_0\} \quad . \tag{3.28}$$

Now consider the volume $\Omega(t)$ of the set $\mathcal{R}(t)$. We have

$$\Omega(t) = \int_{\mathcal{R}(t)} d\mu \tag{3.29}$$

where $d\mu = d\varphi_1 \, d\varphi_2 \cdots d\varphi_n$ for an n-dimensional phase space. For a Hamiltonian system, $n = 2N$. We then have

$$\Omega(t+dt) = \int_{\mathcal{R}(t+dt)} d\mu' = \int_{\mathcal{R}(t)} d\mu \left| \frac{\partial \varphi_i(t+dt)}{\partial \varphi_j(t)} \right| \quad , \tag{3.30}$$

where

$$\left| \frac{\partial \varphi_i(t+dt)}{\partial \varphi_j(t)} \right| \equiv \frac{\partial(\varphi'_1, \ldots, \varphi'_n)}{\partial(\varphi_1, \ldots, \varphi_n)} \tag{3.31}$$

is a determinant, which is the Jacobean of the transformation from the set of coordinates $\{\varphi_i = \varphi_i(t)\}$ to the coordinates $\{\varphi'_i = \varphi_i(t+dt)\}$. But according to the dynamics, we have

$$\varphi_i(t+dt) = \varphi_i(t) + V_i\big(\boldsymbol{\varphi}(t)\big)\, dt + \mathcal{O}(dt^2) \tag{3.32}$$

and therefore

$$\frac{\partial \varphi_i(t+dt)}{\partial \varphi_j(t)} = \delta_{ij} + \frac{\partial V_i}{\partial \varphi_j}\, dt + \mathcal{O}(dt^2) \quad . \tag{3.33}$$

We now make use of the matrix equality $\ln \det M = \mathsf{Tr}\, \ln M$, which gives us[3], for small ε,

$$\det\big(1 + \varepsilon A\big) = \exp \mathsf{Tr} \ln \big(1 + \varepsilon A\big) = 1 + \varepsilon \ \mathsf{Tr}\, A + \tfrac{1}{2}\, \varepsilon^2\Big(\big(\mathsf{Tr}\, A\big)^2 - \mathsf{Tr}\,\big(A^2\big)\Big) + \ldots \tag{3.34}$$

Thus,

$$\Omega(t+dt) = \Omega(t) + \int_{\mathcal{R}(t)} d\mu \ \boldsymbol{\nabla}\!\cdot\!\boldsymbol{V} \ dt + \mathcal{O}(dt^2) \quad , \tag{3.35}$$

---

[2]'Compact' in the parlance of mathematical analysis means 'closed and bounded'.

[3]The equality $\ln \det M = \mathsf{Tr}\, \ln M$ is most easily proven by bringing the matrix to diagonal form via a similarity transformation, and proving the equality for diagonal matrices.

which says

$$\frac{d\Omega}{dt} = \int\limits_{\mathcal{R}(t)} d\mu \, \boldsymbol{\nabla} \cdot \boldsymbol{V} = \int\limits_{\partial\mathcal{R}(t)} dS \, \hat{\boldsymbol{n}} \cdot \boldsymbol{V} \tag{3.36}$$

Here, the divergence is the *phase space divergence*,

$$\boldsymbol{\nabla} \cdot \boldsymbol{V} = \sum_{i=1}^{n} \frac{\partial V_i}{\partial \varphi_i} \quad , \tag{3.37}$$

and we have used the divergence theorem to convert the volume integral of the divergence to a surface integral of $\hat{\boldsymbol{n}} \cdot \boldsymbol{V}$, where $\hat{\boldsymbol{n}}$ is the surface normal and dS is the differential element of surface area, and $\partial\mathcal{R}$ denotes the boundary of the region $\mathcal{R}$. We see that if $\boldsymbol{\nabla} \cdot \boldsymbol{V} = 0$ everywhere in phase space, then $\Omega(t)$ is a constant, and phase space volumes are *preserved* by the evolution of the system.

For an alternative derivation, consider a function $\varrho(\boldsymbol{\varphi}, t)$ which is defined to be the *density* of some collection of points in phase space at phase space position $\boldsymbol{\varphi}$ and time t. This must satisfy the continuity equation,

$$\frac{\partial \varrho}{\partial t} + \boldsymbol{\nabla} \cdot (\varrho \boldsymbol{V}) = 0 \quad . \tag{3.38}$$

This is called the *continuity equation*. It says that 'nobody gets lost'. If we integrate it over a region of phase space $\mathcal{R}$, we have

$$\frac{d}{dt} \int\limits_{\mathcal{R}} d\mu \, \varrho = - \int\limits_{\mathcal{R}} d\mu \, \boldsymbol{\nabla} \cdot (\varrho \boldsymbol{V}) = - \int\limits_{\partial\mathcal{R}} dS \, \hat{\boldsymbol{n}} \cdot (\varrho \boldsymbol{V}) \quad . \tag{3.39}$$

It is perhaps helpful to think of ϱ as a charge density, in which case $\boldsymbol{J} = \varrho \boldsymbol{V}$ is the current density. The above equation then says

$$\frac{dQ_{\mathcal{R}}}{dt} = - \int\limits_{\partial\mathcal{R}} dS \, \hat{\boldsymbol{n}} \cdot \boldsymbol{J} \quad , \tag{3.40}$$

where $Q_{\mathcal{R}}$ is the total charge contained inside the region $\mathcal{R}$. In other words, the rate of increase or decrease of the charge within the region $\mathcal{R}$ is equal to the total integrated current flowing in or out of $\mathcal{R}$ at its boundary.

The Leibniz rule lets us write the continuity equation as

$$\frac{\partial \varrho}{\partial t} + \boldsymbol{V} \cdot \boldsymbol{\nabla} \varrho + \varrho \, \boldsymbol{\nabla} \cdot \boldsymbol{V} = 0 \quad . \tag{3.41}$$

But now suppose that the phase flow is divergenceless, *i.e.* $\boldsymbol{\nabla} \cdot \boldsymbol{V} = 0$. Then we have

$$\frac{D\varrho}{Dt} \equiv \left(\frac{\partial}{\partial t} + \boldsymbol{V} \cdot \boldsymbol{\nabla} \right) \varrho = 0 \quad . \tag{3.42}$$

The combination inside the brackets above is known as the *convective derivative*. It tells us the total rate of change of ϱ for an observer *moving with the phase flow*. That is

$$\frac{d}{dt} \varrho(\boldsymbol{\varphi}(t), t) = \frac{\partial \varrho}{\partial \varphi_i} \frac{d\varphi_i}{dt} + \frac{\partial \varrho}{\partial t}$$

$$= \sum_{i=1}^{n} V_i \frac{\partial \rho}{\partial \varphi_i} + \frac{\partial \varrho}{\partial t} = \frac{D\varrho}{Dt} \quad . \tag{3.43}$$
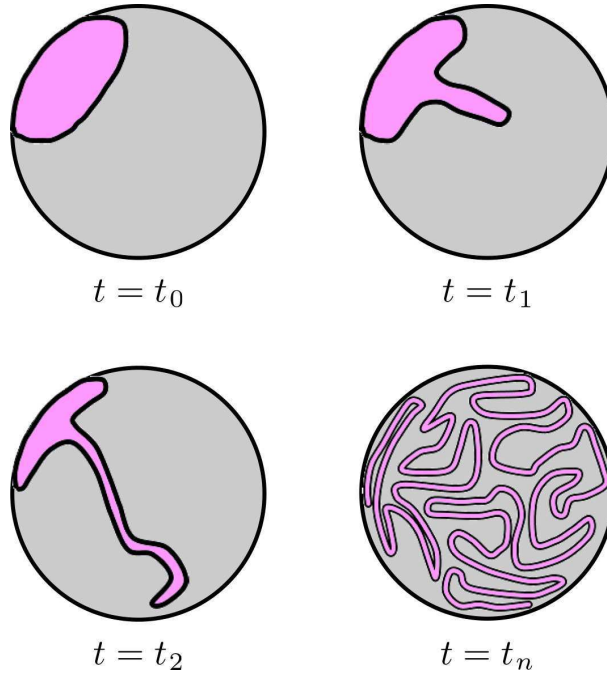
Figure 3.1: Time evolution of two immiscible fluids. The local density remains constant.

If $D\varrho/Dt = 0$, the local density remains the same during the evolution of the system. If we consider the 'characteristic function'

$$\varrho(\boldsymbol{\varphi}, t = 0) = \begin{cases} 1 & \text{if } \boldsymbol{\varphi} \in \mathcal{R}_0 \\ 0 & \text{otherwise} \end{cases} \tag{3.44}$$

then the vanishing of the convective derivative means that the image of the set $\mathcal{R}_0$ under time evolution will always have the same volume.

Hamiltonian evolution in classical mechanics is volume preserving. The equations of motion are

$$\dot{q}_\sigma = + \frac{\partial H}{\partial p_\sigma} \qquad , \qquad \dot{p}_\sigma = - \frac{\partial H}{\partial q_\sigma} \tag{3.45}$$

A point in phase space is specified by N positions q_σ and N momenta p_σ, hence the dimension of phase space is $n = 2N$:

$$\boldsymbol{\varphi} = \begin{pmatrix} \boldsymbol{q} \\ \boldsymbol{p} \end{pmatrix} \qquad , \qquad \boldsymbol{V} = \begin{pmatrix} \dot{\boldsymbol{q}} \\ \dot{\boldsymbol{p}} \end{pmatrix} = \begin{pmatrix} +\partial H/\partial \boldsymbol{p} \\ -\partial H/\partial \boldsymbol{q} \end{pmatrix} \qquad . \tag{3.46}$$

Hamilton's equations of motion guarantee that the phase space flow is divergenceless:

$$\begin{aligned} \nabla \cdot \boldsymbol{V} &= \sum_{\sigma=1}^N \left\{ \frac{\partial \dot{q}_\sigma}{\partial q_\sigma} + \frac{\partial \dot{p}\sigma}{\partial p_\sigma} \right\} \\ &= \sum_{\sigma=1}^N \left\{ \frac{\partial}{\partial q_\sigma} \left(\frac{\partial H}{\partial p_\sigma} \right) + \frac{\partial}{\partial p_\sigma} \left(-\frac{\partial H}{\partial q_\sigma} \right) \right\} = 0 \qquad . \end{aligned} \tag{3.47}$$

Thus, we have that the convective derivative vanishes, *viz.*

$$\frac{D\varrho}{Dt} \equiv \frac{\partial\varrho}{\partial t} + \boldsymbol{V}\cdot\boldsymbol{\nabla}\varrho = 0 \quad, \tag{3.48}$$

for any distribution $\varrho(\boldsymbol{\varphi}, t)$ on phase space. Thus, the value of the density $\varrho(\boldsymbol{\varphi}(t), t)$ is constant, which tells us that the phase flow is *incompressible*. In particular, phase space volumes are preserved.

3.3.3 Liouville's equation and the microcanonical distribution

Let $\varrho(\boldsymbol{\varphi}, t) = \varrho(\boldsymbol{q}, \boldsymbol{p}, t)$ be a distribution on phase space. Assuming the evolution is Hamiltonian, we can write

$$\frac{\partial\varrho}{\partial t} = -\dot{\boldsymbol{\varphi}}\cdot\boldsymbol{\nabla}\varrho = -\{\varrho, H\} \quad. \tag{3.49}$$

We may also write this as $\partial_t\varrho + i\hat{L}\varrho = 0$, where $\hat{L}$ is a differential operator known as the *Liouvillian*:

$$i\hat{L} = \sum_{\sigma=1}^{N}\left\{\frac{\partial H}{\partial p_\sigma}\frac{\partial}{\partial q_\sigma} - \frac{\partial H}{\partial q_\sigma}\frac{\partial}{\partial p_\sigma}\right\} \quad. \tag{3.50}$$

Eqn. 3.49, known as *Liouville's equation*. Note $i\hat{L}F = \{F, H\}$ for any function $F(\boldsymbol{\varphi}, t)$.

Recall that the evolution of quantum mechanical density matrices satisfies

$$\frac{\partial\hat{\varrho}}{\partial t} = \frac{i}{\hbar}\left[\hat{\varrho}, \hat{H}\right] \quad, \tag{3.51}$$

whence we infer the correspondence

$$\{\varrho, H\} \rightarrow \frac{1}{i\hbar}\left[\hat{\varrho}, \hat{H}\right] \quad. \tag{3.52}$$

Suppose that there is a family of conserved quantities $\Lambda_a(\boldsymbol{\varphi})$, with $a \in \{1, \ldots, k\}$, each of which is conserved by the dynamics of the system. Such conserved quantities might include the components of the total linear momentum (if there is translational invariance), the components of the total angular momentum (if there is rotational invariance), and the Hamiltonian itself (if it is not explicitly time-dependent). Now consider a distribution $\varrho(\boldsymbol{\varphi}) = \varrho(\Lambda_1, \Lambda_2, \ldots, \Lambda_k)$ which is a function only of these various conserved quantities. That Λ_a is conserved entails $\dot{\boldsymbol{\varphi}}\cdot\boldsymbol{\nabla}\Lambda_a = \{\Lambda_a, H\} = 0$. Then from the chain rule, we have

$$\dot{\boldsymbol{\varphi}}\cdot\boldsymbol{\nabla}\varrho = \sum_{a=1}^{k}\frac{\partial\varrho}{\partial\Lambda_a}\dot{\boldsymbol{\varphi}}\cdot\boldsymbol{\nabla}\Lambda_a = 0 \quad. \tag{3.53}$$

We conclude that any distribution $\varrho(\boldsymbol{\varphi}) = \varrho(\Lambda_1, \Lambda_2, \ldots, \Lambda_k)$ which is a function solely of conserved dynamical quantities is a stationary solution to Liouville's equation.

Clearly the microcanonical distribution,

$$\varrho_E(\boldsymbol{\varphi}) = \frac{\delta(E - H(\boldsymbol{\varphi}))}{D(E)} = \frac{\delta(E - H(\boldsymbol{\varphi}))}{\int d\mu\,\delta(E - H(\boldsymbol{\varphi}))} \quad, \tag{3.54}$$

is a fixed point solution of Liouville's equation. If there were a second conserved quantity, $\Lambda(\boldsymbol{\varphi})$, the generalized microcanonical distribution,

$$\varrho_{E,Q}(\boldsymbol{\varphi}) = \frac{\delta\big(E - H(\boldsymbol{\varphi})\big)\,\delta\big(Q - \Lambda(\boldsymbol{\varphi})\big)}{D(E,Q)} = \frac{\delta\big(E - H(\boldsymbol{\varphi})\big)\,\delta\big(Q - \Lambda(\boldsymbol{\varphi})\big)}{\int d\mu\,\delta\big(E - H(\boldsymbol{\varphi})\,\delta\big(Q - \Lambda(\boldsymbol{\varphi})\big)} \quad , \tag{3.55}$$

would be a solution to Liouville's equation for arbitrary E and Q.

Similarly, the Gibbs distribution,

$$\varrho_\beta(\boldsymbol{\varphi}) = \frac{1}{Z(\beta)}\,e^{-\beta H(\boldsymbol{\varphi})} \quad , \tag{3.56}$$

where $Z(\beta) = \mathsf{Tr}\,e^{-\beta H(\boldsymbol{\varphi})}$ is the partition function, satisfies $\{\varrho_\beta, H\} = 0$. In the presence of multiple conserved quantities, one defines the *generalized Gibbs distribution*,

$$\varrho_{\boldsymbol{\beta}}(\boldsymbol{\varphi}) = \frac{1}{Z(\boldsymbol{\beta})}\,\exp\Big(-\sum_{a=1}^{k}\beta_a\,\Lambda_a(\boldsymbol{\varphi})\Big) \quad , \tag{3.57}$$

where the $\{\Lambda_a(\boldsymbol{\varphi})\}$, with $a \in \{1,\dots,k\}$, are the conserved quantities, including among them H itself. The coefficients $\{\beta_a\}$ are k Lagrange multipliers enforcing the k conservation constraints $\Lambda_a(\boldsymbol{\varphi}) = Q_a$.

3.4 Irreversibility and Poincaré Recurrence

The dynamics of the master equation describe an approach to equilibrium. These dynamics are irreversible: $dH/dt \le 0$, where H is Boltzmann's H-function. However, the microscopic laws of physics are (almost) time-reversal invariant[4], so how can we understand the emergence of irreversibility? Furthermore, any dynamics which are deterministic and volume-preserving in a finite phase space exhibits the phenomenon of *Poincaré recurrence*, which guarantees that phase space trajectories are arbitrarily close to periodic if one waits long enough.

3.4.1 Poincaré recurrence theorem

The proof of the recurrence theorem is simple. Let g_τ be the 'τ-advance mapping' which evolves points in phase space according to Hamilton's equations. Assume that g_τ is invertible and volume-preserving, as is the case for Hamiltonian flow. Further assume that phase space volume is finite. Since energy is preserved in the case of time-independent Hamiltonians, we simply ask that the volume of phase space *at fixed total energy E* be finite, *i.e.*

$$\int d\mu\,\delta\big(E - H(\boldsymbol{q},\boldsymbol{p})\big) < \infty \quad , \tag{3.58}$$

where $d\mu = d\boldsymbol{q}\,d\boldsymbol{p}$ is the phase space uniform integration measure.

Theorem: In any finite neighborhood $\mathcal{R}_0$ of phase space there exists a point $\boldsymbol{\varphi}_0$ which will return to $\mathcal{R}_0$ after m applications of g_τ, where m is finite.

---

[4]Actually, the microscopic laws of physics are *not* time-reversal invariant, but rather are invariant under the product PCT, where P is parity, C is charge conjugation, and T is time reversal.
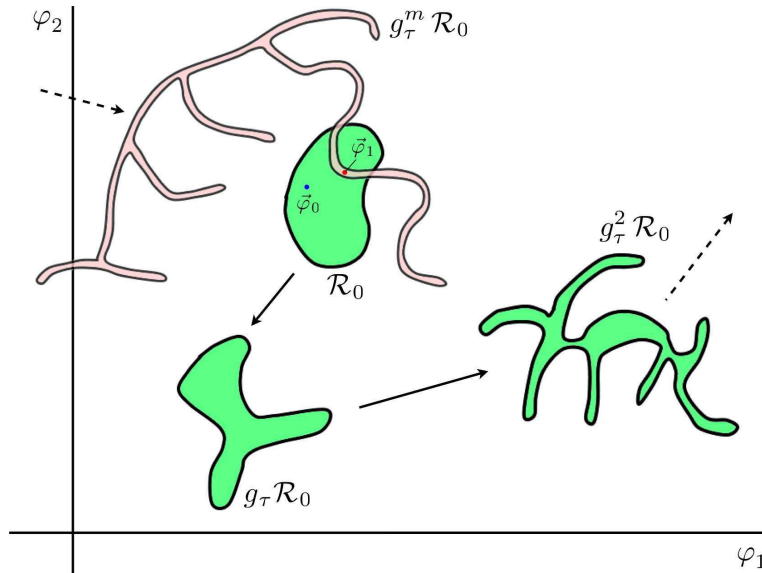
Figure 3.2: Successive images of a set $\mathcal{R}_0$ under the τ-advance mapping g_τ, projected onto a two-dimensional phase plane. The Poincaré recurrence theorem guarantees that if phase space has finite volume, and g_τ is invertible and volume preserving, then for any set $\mathcal{R}_0$ there exists an integer m such that $\mathcal{R}_0 \cap g_\tau^m \mathcal{R}_0 \neq \emptyset$.

Proof: Assume the theorem fails; we will show this assumption results in a contradiction. Consider the set Υ formed from the union of all sets $g_\tau^k \mathcal{R}$ for all m:

$$\Upsilon = \bigcup_{k=0}^{\infty} g_\tau^k \mathcal{R}_0 \tag{3.59}$$

We assume that the set $\{g_\tau^k \mathcal{R}_0 \mid k \in \mathbb{N}_0\}$ is disjoint[5]. The volume of a union of disjoint sets is the sum of the individual volumes. Thus,

$$\mathrm{vol}(\Upsilon) = \sum_{k=0}^{\infty} \mathrm{vol}\big(g_\tau^k \mathcal{R}_0\big) = \mathrm{vol}(\mathcal{R}_0) \cdot \sum_{k=0}^{\infty} 1 = \infty \quad , \tag{3.60}$$

since $\mathrm{vol}\big(g_\tau^k \mathcal{R}_0\big) = \mathrm{vol}\big(\mathcal{R}_0\big)$ from volume preservation. But clearly Υ is a subset of the entire phase space, hence we have a contradiction, because by assumption phase space is of finite volume.

Thus, the assumption that the set $\{g_\tau^k \mathcal{R}_0 \mid k \in \mathbb{N}_0\}$ is disjoint fails. This means that there exists some pair of integers k and l, with $k \neq l$, such that $g_\tau^k \mathcal{R}_0 \cap g_\tau^l \mathcal{R}_0 \neq \emptyset$. Without loss of generality we may assume $k < l$. Apply the inverse g_τ^{-1} to this relation k times to get $g_\tau^{l-k} \mathcal{R}_0 \cap \mathcal{R}_0 \neq \emptyset$. Now choose any point $\boldsymbol{\varphi}_1 \in g_\tau^m \mathcal{R}_0 \cap \mathcal{R}_0$, where $m = l - k$, and define $\boldsymbol{\varphi}_0 = g_\tau^{-m} \boldsymbol{\varphi}_1$. Then by construction both $\boldsymbol{\varphi}_0$ and $g_\tau^m \boldsymbol{\varphi}_0$ lie within $\mathcal{R}_0$ and the theorem is proven.

Poincaré recurrence has remarkable implications. Consider a bottle of perfume which is opened in an otherwise evacuated room, as depicted in fig. 3.3. The perfume molecules evolve according to

---

[5]The natural numbers $\mathbb{N}_0$ is the set of non-negative integers $\{0, 1, 2, \ldots\}$.

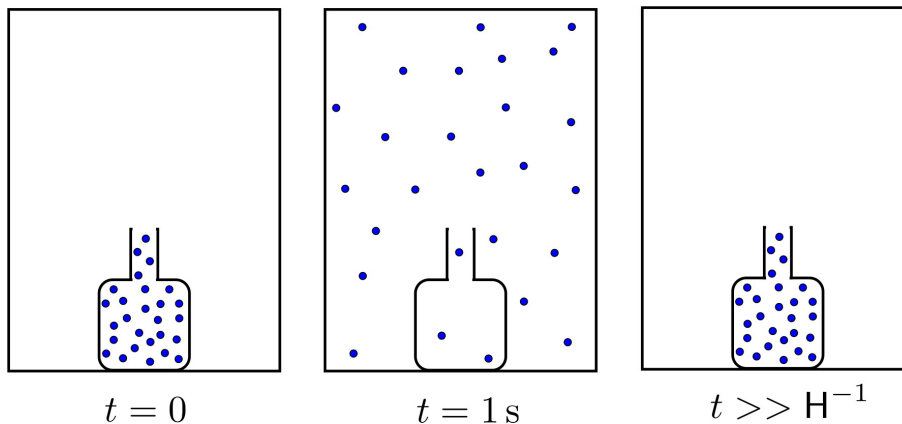$$t = 0 \qquad\qquad t = 1\,\mathrm{s} \qquad\qquad t >> \mathsf{H}^{-1}$$

Figure 3.3: Poincaré recurrence guarantees that if we remove the cap from a bottle of perfume in an otherwise evacuated room, all the perfume molecules will eventually return to the bottle! (Here H is the Hubble constant.)

Hamiltonian evolution. The positions are bounded because physical space is finite. The momenta are bounded because the total energy is conserved, hence no single particle can have a momentum such that $T(\boldsymbol{p}) > E_{\mathrm{TOT}}$, where $T(\boldsymbol{p})$ is the single particle kinetic energy function[6]. Thus, phase space, however large, is still bounded. Hamiltonian evolution, as we have seen, is invertible and volume preserving, therefore the system is recurrent. All the molecules must eventually return to the bottle. What's more, they all must return with momenta arbitrarily close to their initial momenta![7] In this case, we could define the region $\mathcal{R}_0$ as

$$\mathcal{R}_0 = \big\{ (q_1, \ldots, q_r, p_1, \ldots, p_r) \,\big|\, |q_i - q_i^0| \le \Delta q \text{ and } |p_j - p_j^0| \le \Delta p \,\forall\, i, j \big\} \quad , \qquad (3.61)$$

which specifies a hypercube in phase space centered about the point $(\boldsymbol{q}^0, \boldsymbol{p}^0)$.

Each of the three central assumptions – finite phase space, invertibility, and volume preservation – is crucial. If any one of these assumptions does not hold, the proof fails. Obviously if phase space is infinite the flow needn't be recurrent since it can keep moving off in a particular direction. Consider next a volume-preserving map which is not invertible. An example might be a mapping $f \colon \mathbb{R} \to \mathbb{R}$ which takes any real number to its fractional part. Thus, $f(\pi) = 0.14159265 \ldots$. Let us restrict our attention to intervals of width less than unity. Clearly f is then volume preserving. The action of f on the interval $[2, 3)$ is to map it to the interval $[0, 1)$. But $[0, 1)$ remains fixed under the action of f, so no point within the interval $[2, 3)$ will ever return under repeated iterations of f. Thus, f does not exhibit Poincaré recurrence.

Consider next the case of the damped harmonic oscillator. In this case, phase space volumes contract. For a one-dimensional oscillator obeying $\ddot{x} + 2\beta\dot{x} + \Omega_0^2\, x = 0$ one has $\nabla \cdot \boldsymbol{V} = -2\beta < 0$, since $\beta > 0$ for physical damping. Thus the convective derivative is $D_t\varrho = -(\nabla \cdot \boldsymbol{V})\varrho = 2\beta\varrho$ which says that the density increases exponentially in the comoving frame, as $\varrho(t) = e^{2\beta t}\,\varrho(0)$. Thus, phase space volumes collapse: $\Omega(t) = e^{-2\beta 2}\,\Omega(0)$, and are not preserved by the dynamics. The proof of recurrence therefore

---

[6]In the nonrelativistic limit, $T = p^2/2m$. For relativistic particles, we have $T = (p^2c^2 + m^2c^4)^{1/2} - mc^2$.

[7]Actually, what the recurrence theorem guarantees is that there is a configuration arbitrarily close to the initial one which recurs, to within the same degree of closeness.
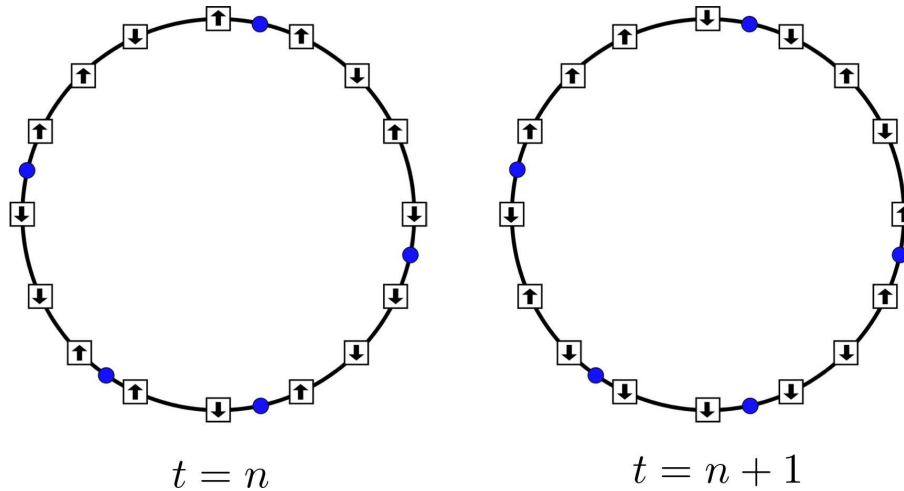
$t = n$ $t = n + 1$

Figure 3.4: Left: A configuration of the Kac ring with $N = 16$ sites and $F = 4$ flippers. The flippers, which live on the links, are represented by blue dots. Right: The ring system after one time step. Evolution proceeds by clockwise rotation. Spins passing through flippers are flipped.

fails. In this case, it is possible for the set Υ to be of finite volume, even if it is the union of an infinite number of sets $g_\tau^k \mathcal{R}_0$, because the volumes of these component sets themselves decrease exponentially, as $\mathrm{vol}(g_\tau^n \mathcal{R}_0) = e^{-2n\beta\tau} \mathrm{vol}(\mathcal{R}_0)$. A damped pendulum, released from rest at some small angle θ_0, will not return arbitrarily close to these initial conditions.

3.4.2 Kac ring model

The implications of the Poincaré recurrence theorem are surprising – even shocking. If one takes a bottle of perfume in a sealed, evacuated room and opens it, the perfume molecules will diffuse throughout the room. The recurrence theorem guarantees that after some finite time T all the molecules will go back inside the bottle (and arbitrarily close to their initial velocities as well). The hitch is that this could take a very long time, *e.g.* much much longer than the age of the Universe.

On less absurd time scales, we know that most systems come to thermodynamic equilibrium. But how can a system both exhibit equilibration *and* Poincaré recurrence? The two concepts seem utterly incompatible!

A beautifully simple model due to Kac shows how a recurrent system can exhibit the phenomenon of equilibration. Consider a ring with N sites. On each site, place a 'spin' which can be in one of two states: up or down. Along the N links of the system, F of them contain 'flippers'. The configuration of the flippers is set at the outset and never changes. The dynamics of the system are as follows: during each time step, every spin moves clockwise a distance of one lattice spacing. Spins which pass through flippers reverse their orientation: up becomes down, and down becomes up.

The 'phase space' for this system consists of 2^N discrete configurations. Since each configuration maps onto a unique image under the evolution of the system, phase space 'volume' is preserved. The evolution is invertible; the inverse is obtained simply by rotating the spins counterclockwise. Figure 3.4
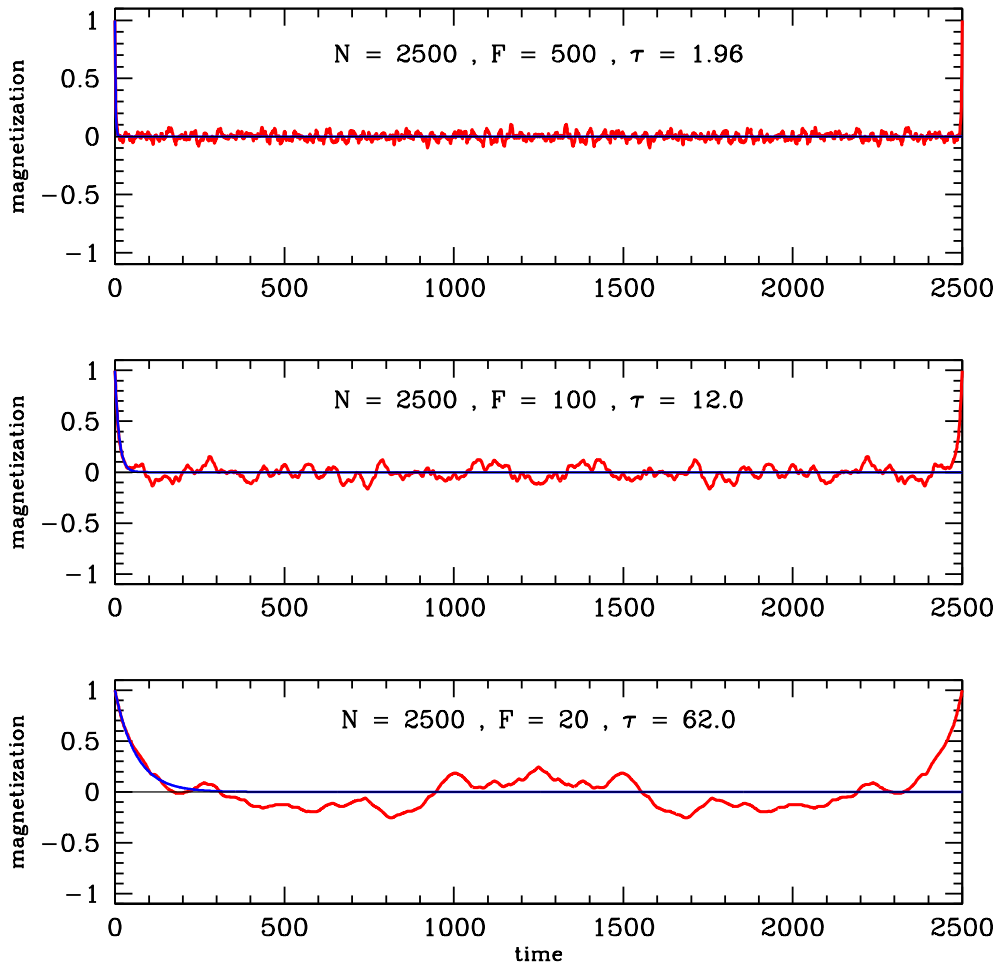
Figure 3.5: Three simulations of the Kac ring model with $N = 2500$ sites and three different concentrations of flippers. The red line shows the magnetization as a function of time, starting from an initial configuration in which 100% of the spins are up. The blue line shows the prediction of the *Stosszahlansatz*, which yields an exponentially decaying magnetization with time constant τ.

depicts an example configuration for the system, and its first iteration under the dynamics.

Suppose the flippers were not fixed, but moved about randomly. In this case, we could focus on a single spin and determine its configuration probabilistically. Let p_n be the probability that a given spin is in the up configuration at time n. The probability that it is up at time $(n + 1)$ is then

$$p_{n+1} = (1 - x)\, p_n + x\, (1 - p_n) \quad , \tag{3.62}$$

where $x = F/N$ is the fraction of flippers in the system. In words: a spin will be up at time $(n + 1)$ if it was up at time n and did not pass through a flipper, or if it was down at time n and did pass through a flipper. If the flipper locations are randomized at each time step, then the probability of flipping is simply $x = F/N$. Equation 3.62 can be solved immediately:

$$p_n = \tfrac{1}{2} + (1 - 2x)^n\, (p_0 - \tfrac{1}{2}) \quad , \tag{3.63}$$
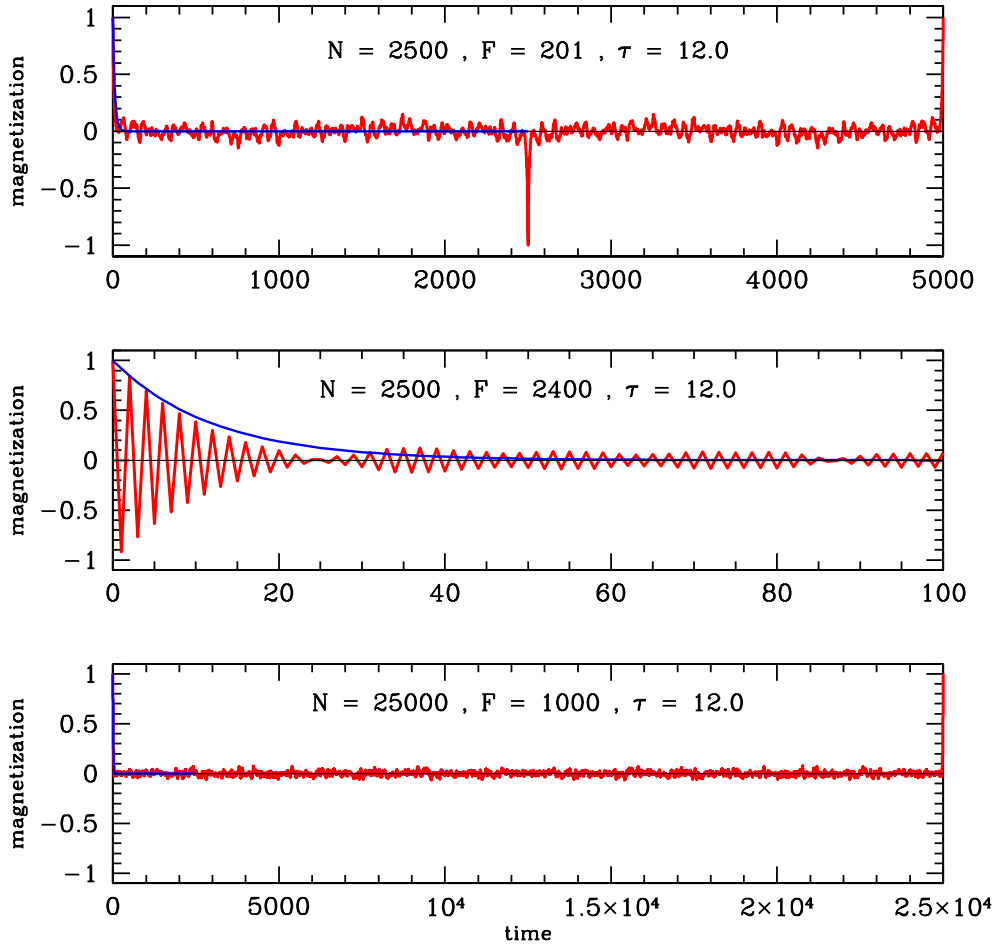
Figure 3.6: Simulations of the Kac ring model. Top: $N = 2500$ sites with $F = 201$ flippers. After 2500 iterations, each spin has flipped an odd number of times, so the recurrence time is $2N$. Middle: $N = 2500$ with $F = 2400$, resulting in a near-complete reversal of the population with every iteration. Bottom: $N = 25000$ with $N = 1000$, showing long time equilibration and dramatic resurgence of the spin population.

which decays exponentially to the equilibrium value of $p_{eq} = \frac{1}{2}$ with time scale

$$\tau(x) = -\frac{1}{\ln|1 - 2x|} \quad . \tag{3.64}$$

We identify $\tau(x)$ as the microscopic relaxation time over which local equilibrium is established. If we define the magnetization $m \equiv (N_\uparrow - N_\downarrow)/N$, then $m = 2p - 1$, so $m_n = (1 - 2x)^n m_0$. The equilibrium magnetization is $m_{eq} = 0$. Note that for $\frac{1}{2} < x < 1$ that the magnetization reverses sign each time step, as well as decreasing exponentially in magnitude.

The assumption that leads to equation 3.62 is called the *Stosszahlansatz*[8], a long German word mean-

---

[8]Unfortunately, many important physicists were German and we have to put up with a legacy of long German words like *Gedankenexperiment*, *Zitterbewegung*, *Brehmsstrahlung*, *Stosszahlansatz*, *Kartoffelsalat*, *etc.*

ing, approximately, 'assumption on the counting of hits'. The resulting dynamics are irreversible: the magnetization inexorably decays to zero. However, the Kac ring model is purely deterministic, and the *Stosszahlansatz* can at best be an approximation to the true dynamics. Clearly the *Stosszahlansatz* fails to account for correlations such as the following: if spin i is flipped at time n, then spin $i+1$ will have been flipped at time $n-1$. Also if spin i is flipped at time n, then it also will be flipped at time $n+N$. Indeed, since the dynamics of the Kac ring model are invertible and volume preserving, it must exhibit Poincaré recurrence. We see this most vividly in figs. 3.5 and 3.6.

The model is trivial to simulate. The results of such a simulation are shown in figure 3.5 for a ring of $N = 1000$ sites, with $F = 100$ and $F = 24$ flippers. Note how the magnetization decays and fluctuates about the equilibrium value $m_{\text{eq}} = 0$, but that after N iterations m recovers its initial value: $m_N = m_0$. The recurrence time for this system is simply N if F is even, and $2N$ if F is odd, since every spin will then have flipped an even number of times.

In figure 3.6 we plot two other simulations. The top panel shows what happens when $x > \frac{1}{2}$, so that the magnetization wants to reverse its sign with every iteration. The bottom panel shows a simulation for a larger ring, with $N = 25000$ sites. Note that the fluctuations in m about equilibrium are smaller than in the cases with $N = 1000$ sites. Why?

3.5   Remarks on Ergodic Theory

3.5.1   Definition of ergodicity

A mechanical system evolves according to Hamilton's equations of motion. We have seen how such a system is *recurrent* in the sense of Poincaré.

There is a level beyond recurrence called *ergodicity*. In an ergodic system, time averages over intervals $[0, T]$ with $T \to \infty$ may be replaced by phase space averages. The time average of a function $f(\varphi)$ is defined as

$$\langle f(\varphi) \rangle_t = \lim_{T \to \infty} \frac{1}{T} \int_0^T dt \, f\big(\varphi(t)\big) \quad . \tag{3.65}$$

For a Hamiltonian system, the *phase space average* of the same function is defined by

$$\langle f(\varphi) \rangle_{\mu\text{CE}} = \int d\mu \, f(\varphi) \, \delta\big(E - H(\varphi)\big) \bigg/ \int d\mu \, \delta\big(E - H(\varphi)\big) \quad , \tag{3.66}$$

where $H(\varphi) = H(\boldsymbol{q}, \boldsymbol{p})$ is the Hamiltonian, and where $\delta(x)$ is the Dirac δ-function. The energy is fixed to be $E = H\big(\varphi(t=0)\big)$. Thus,

$$\text{ergodicity} \quad \Longleftrightarrow \quad \langle f(\varphi) \rangle_t = \langle f(\varphi) \rangle_{\mu\text{CE}} \quad , \tag{3.67}$$

for all smooth functions $f(\varphi)$ for which $\langle f(\varphi) \rangle_{\mu\text{CE}}$ exists and is finite. Note that we do not average over *all* of phase space. Rather, we average only over a hypersurface along which $H(\varphi) = E$ is fixed, *i.e.* over one of the *level sets* of the Hamiltonian function. This is because the dynamics *preserves the energy*.

Ergodicity means that almost all points φ will, upon Hamiltonian evolution, move in such a way as to eventually pass through every finite neighborhood on the energy surface, and will spend equal time in equal regions of phase space.

Let $\chi_{\mathcal{R}}(\varphi)$ be the characteristic function of a region $\mathcal{R}$:

$$\chi_{\mathcal{R}}(\varphi) = \begin{cases} 1 & \text{if } \varphi \in \mathcal{R} \\ 0 & \text{otherwise,} \end{cases} \tag{3.68}$$

where $H(\varphi) = E$ for all $\varphi \in \mathcal{R}$, so $\dim \mathcal{R} = 2N - 1$. Then

$$\langle \chi_{\mathcal{R}}(\varphi) \rangle_t = \lim_{T \to \infty} \left(\frac{\text{time spent in } \mathcal{R}}{T} \right) \ . \tag{3.69}$$

If the system is ergodic, then

$$\langle \chi_{\mathcal{R}}(\varphi) \rangle_t = P(\mathcal{R}) = \frac{D_{\mathcal{R}}(E)}{D(E)} \ , \tag{3.70}$$

where $P(\mathcal{R})$ is the *a priori* probability to find $\varphi \in \mathcal{R}$, based solely on the relative volumes of $\mathcal{R}$ and of the entire energy-restricted phase space. Here,

$$D(E) = \int d\mu \, \delta\big(E - H(\varphi)\big) = \int_{\mathcal{S}_E} d\Sigma_E \ , \tag{3.71}$$

called the *density of states*, is the surface area of phase space at energy E. The hypersurface $\mathcal{S}_E$ is the set of points φ satisfying $H(\varphi) = E$, and the *invariant differential surface element* $d\Sigma_E$ is defined as follows. We can write the differential phase space volume $d\mu$ as the product $d\mu = dS_E \, d\zeta_E$, where dS_E is the differential surface element for the level set $\mathcal{S}_E$ and ζ_E is a phase space coordinate locally perpendicular to $\mathcal{S}_E$. We then define[9]

$$d\Sigma_E \equiv \frac{dS_E}{|\boldsymbol{\nabla} H|}\bigg|_{H(\varphi)=E} \ , \tag{3.72}$$

and we may now write $d\mu = dE \, d\Sigma_E$. Note that we may also express $D(E)$ as

$$D(E) = \frac{d}{dE} \int d\mu \, \Theta\big(E - H(\varphi)\big) \equiv \frac{d\Omega(E)}{dE} \ , \tag{3.73}$$

where $\Omega(E) = \int d\mu \, \Theta\big(E - H(\varphi)\big)$ is the volume of phase space over which $H(\varphi) < E$. The density of states for the subset $\mathcal{R}$ is defined as

$$D_{\mathcal{R}}(E) = \int_{\mathcal{R}} d\Sigma_E \ . \tag{3.74}$$

Note that $\mathcal{R} \subset \mathcal{S}_E$.

---

[9]Recall that the phase space coordinates don't all have the same units! N of the coordinates have units of position and N have units of momentum. Furthermore, some may be angles and some angular momenta. However in any case $d\mu$ has units A^N, where A stands for action, *i.e.* $[d\mu] = \mathrm{ML}^2/\mathrm{T}$. Thus while the product $d\mu = dS_E \, d\zeta_E$ has units of A^N, individually the units of dS_E and $d\zeta_E$ vary along the hypersurface $\mathcal{S}_E$! However, the invariant differential surface element $d\Sigma_E$ always has units of A^N/E. To resolve any confusion, one may choose to rescale so that all phase space coordinates are dimensionless.
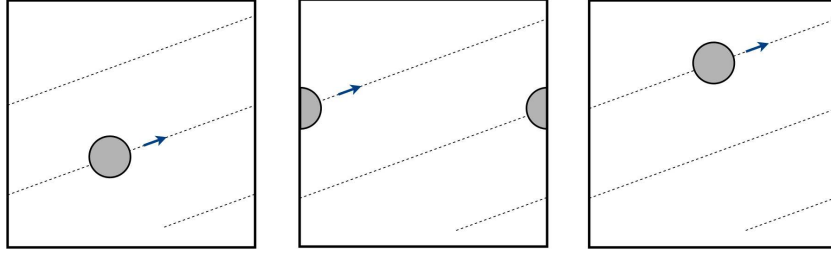
Figure 3.7: Constant phase space velocity at an irrational angle over a toroidal phase space is ergodic, but not mixing. A circle remains a circle, and a blob remains a blob.

3.5.2   The microcanonical ensemble

The distribution,

$$\varrho_E(\boldsymbol{\varphi}) = \frac{\delta\big(E - H(\boldsymbol{\varphi})\big)}{D(E)} = \frac{\delta\big(E - H(\boldsymbol{\varphi})\big)}{\int d\mu \,\delta\big(E - H(\boldsymbol{\varphi})\big)} \quad , \tag{3.75}$$

defines the *microcanonical ensemble* (μCE) of Gibbs. We could also write

$$\big\langle f(\boldsymbol{\varphi})\big\rangle_{\mu\text{CE}} = \frac{1}{D(E)} \int\limits_{\mathcal{S}_E} d\Sigma_E \, f(\boldsymbol{\varphi}) \quad , \tag{3.76}$$

integrating over the hypersurface $\mathcal{S}_E$ rather than the entire phase space.

3.5.3   Ergodicity and mixing

Just because a system is ergodic, it doesn't necessarily mean that $\varrho(\varphi, t) \to \varrho^{\text{eq}}(\boldsymbol{\varphi})$, for consider the following motion on the toroidal space $\big(\boldsymbol{\varphi} = (q,p)\,\big|\,0 \leq q < 1 \;\;,\; 0 \leq p < 1\big\}$, where we identify opposite edges, *i.e.* we impose periodic boundary conditions. We also take q and p to be dimensionless, for simplicity of notation. Let the dynamics be given by $\dot{q} = 1$ and $\dot{p} = \alpha$. The motion is then $q(t) = q_0 + t$ and $p(t) = p_0 + \alpha t$. Thus the phase curves are given by $p = p_0 + \alpha(q - q_0)$.

Now consider the average of some function $f(q, p)$. We can write $f(q, p)$ in terms of its Fourier transform,

$$f(q, p) = \sum_{m,n} \hat{f}_{m,n} \, e^{2\pi i(mq + np)} \quad . \tag{3.77}$$

We have, then,

$$f\big(q(t), p(t)\big) = \sum_{m,n} \hat{f}_{m,n} \, e^{2\pi i(mq_0 + np_0)} \, e^{2\pi i(m + \alpha n)t} \quad . \tag{3.78}$$

We can now perform the time average of f:

$$\big\langle f(q,p) \big\rangle_t = \hat{f}_{0,0} + \lim_{T \to \infty} \frac{1}{T} \sideset{}{'}\sum_{m,n} \hat{f}_{m,n} \, e^{2\pi i(mq_0 + np_0)} \, \frac{e^{2\pi i(m + \alpha n)T} - 1}{2\pi i(m + \alpha n)}$$

$$= \hat{f}_{0,0} \quad \text{if } \alpha \notin \mathbb{Q} \quad . \tag{3.79}$$

Clearly,

$$\left\langle f(q,p)\right\rangle_{\mu\text{CE}} = \int\limits_0^1 dq \int\limits_0^1 dp\, f(q,p) = \hat{f}_{0,0} = \left\langle f(q,p)\right\rangle_t \quad , \tag{3.80}$$

so the system is ergodic. However, if $\alpha = r/s$ with $r, s \in \mathbb{Z}$ and $\gcd(r,s) = 1$ (*i.e.* r and s are relatively prime), then $\exp\{2\pi i(m + \alpha n)t\} = 1$ whenever $(m,n) = (kr, -ks)$ for any $k \in \mathbb{Z}$. Thus,

$$\left\langle f(q,p)\right\rangle_t = \sum_{k=-\infty}^{\infty} \hat{f}_{kr,-ks}\, e^{2\pi i k(rq_0 - sp_0)} \quad , \tag{3.81}$$

which is not the same as $\left\langle f(q,p)\right\rangle_{\mu\text{CE}} = \hat{f}_{0,0}$.

The situation is depicted in fig. 3.7. If we start with the characteristic function of a disc,

$$\varrho(q,p,t=0) = \Theta\big(a^2 - (q-q_0)^2 - (p-p_0)^2\big) \quad , \tag{3.82}$$

then it remains the characteristic function of a disc:

$$\varrho(q,p,t) = \Theta\big(a^2 - (q-q_0-t)^2 - (p-p_0-\alpha t)^2\big) \quad , \tag{3.83}$$

For an example of a transition to ergodicity in a simple dynamical Hamiltonian model, see §3.8.

A stronger condition one could impose is the following. Let A and B be subsets of $\mathcal{S}_E$. Define the *measure*

$$\nu(A) = \int d\Sigma_E\, \chi_A(\boldsymbol{\varphi}) \Big/ \int d\Sigma_E = \frac{D_A(E)}{D(E)} \quad , \tag{3.84}$$

where $\chi_A(\boldsymbol{\varphi})$ is the characteristic function of A. The measure of a set A is the fraction of the energy surface $\mathcal{S}_E$ covered by A. This means $\nu(\mathcal{S}_E) = 1$, since $\mathcal{S}_E$ is the entire phase space at energy E. Now let g be a volume-preserving map on phase space. Given two measurable sets A and B, we say that a system is *mixing* if

$$\text{mixing} \quad \Longleftrightarrow \quad \lim_{n\to\infty} \nu\big(g^n A \cap B\big) = \nu(A)\,\nu(B) \quad . \tag{3.85}$$

In other words, the fraction of B covered by the nth iterate of A, *i.e.* $g^n A$, is, as $n \to \infty$, simply the fraction of $\mathcal{S}_E$ covered by A. The iterated map g^n distorts the region A so severely that it eventually spreads out 'evenly' over the entire energy hypersurface. Of course by 'evenly' we mean 'with respect to any finite length scale', because at the very smallest scales, the phase space density is still locally constant as one evolves with the dynamics.

Mixing means that

$$\left\langle f(\boldsymbol{\varphi})\right\rangle = \int d\mu\, \varrho(\boldsymbol{\varphi},t)\, f(\boldsymbol{\varphi}) \xrightarrow[t\to\infty]{} \int d\mu\, f(\boldsymbol{\varphi})\, \delta\big(E - H(\boldsymbol{\varphi})\big) \Big/ \int d\mu\, \delta\big(E - H(\boldsymbol{\varphi})\big)$$
$$\equiv \mathsf{Tr}\Big[f(\boldsymbol{\varphi})\, \delta\big(E - H(\boldsymbol{\varphi})\big)\Big] \Big/ \mathsf{Tr}\Big[\delta\big(E - H(\boldsymbol{\varphi})\big)\Big] \quad . \tag{3.86}$$

Physically, we can imagine regions of phase space being successively stretched and folded. During the stretching process, the volume is preserved, so the successive stretch and fold operations map phase space back onto itself.
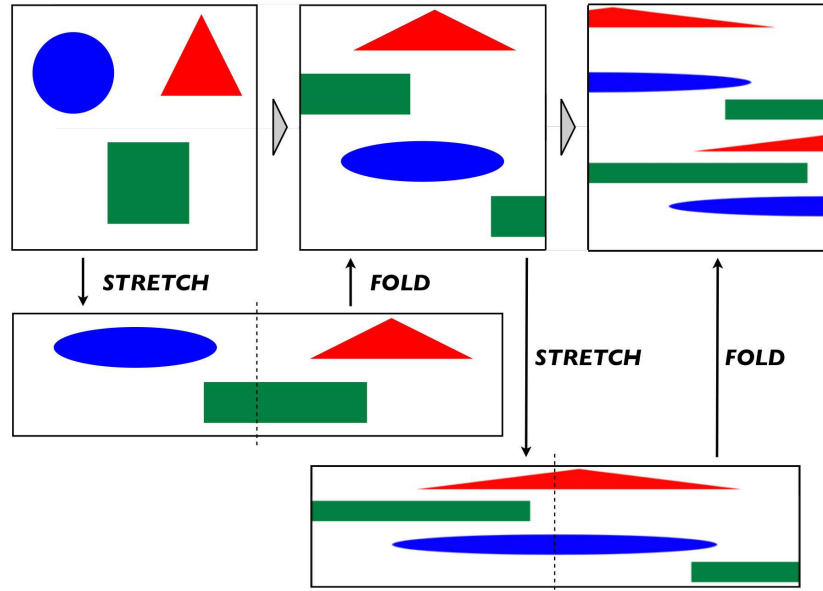
Figure 3.8: The baker's transformation is a successive stretching, cutting, and restacking.

An example of a mixing system is the *baker's transformation*, depicted in fig. 3.8, and defined by

$$g(q,p) = \begin{cases} \left(2q\,,\,\tfrac{1}{2}p\right) & \text{if} \quad 0 \le q < \tfrac{1}{2} \\ \left(2q-1\,,\,\tfrac{1}{2}p+\tfrac{1}{2}\right) & \text{if} \quad \tfrac{1}{2} \le q < 1 \quad. \end{cases} \tag{3.87}$$

Note that g is invertible and volume-preserving. The baker's transformation consists of an initial stretch in which q is expanded by a factor of two and p is contracted by a factor of two, which preserves the total volume. The system is then mapped back onto the original area by cutting and restacking, which we can call a 'fold'. The inverse transformation is accomplished by stretching first in the vertical (p) direction and squashing in the horizontal (q) direction, followed by a slicing and restacking. Explicitly,

$$g^{-1}(q,p) = \begin{cases} \left(\tfrac{1}{2}q\,,\,2p\right) & \text{if} \quad 0 \le p < \tfrac{1}{2} \\ \left(\tfrac{1}{2}q+\tfrac{1}{2}\,,\,2p-1\right) & \text{if} \quad \tfrac{1}{2} \le p < 1 \quad. \end{cases} \tag{3.88}$$

Another example of a mixing system is Arnold's 'cat map'[10]

$$g(q,p) = \left(\,[q+p]\,,\,[q+2p]\,\right) \quad, \tag{3.89}$$

where $[x]$ denotes the fractional part of x. One can write this in matrix form as

$$\begin{pmatrix} q' \\ p' \end{pmatrix} = \overbrace{\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}}^{M} \begin{pmatrix} q \\ p \end{pmatrix} \bmod \mathbb{Z}^2 \quad. \tag{3.90}$$

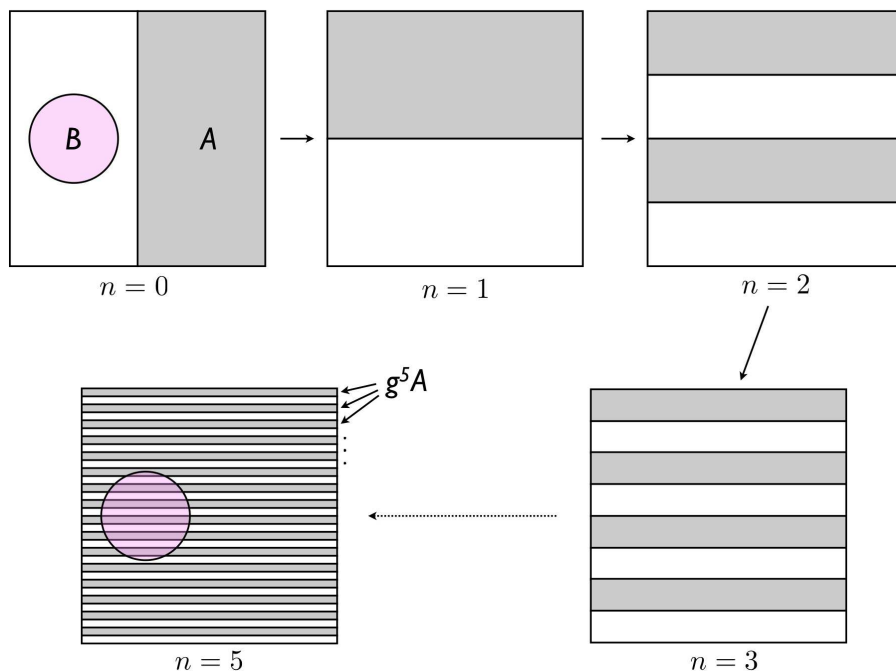[10]The cat map gets its name from its initial application, by Arnold, to the image of a cat's face.

Figure 3.9: The multiply iterated baker's transformation. The set A covers half the phase space and its area is preserved under the map. Initially, the fraction of B covered by A is zero. After many iterations, the fraction of B covered by g^nA approaches $\frac{1}{2}$.

The matrix M is very special because it has integer entries and its determinant is $\det M = 1$. This means that the inverse also has integer entries. The inverse transformation is then

$$\begin{pmatrix} q \\ p \end{pmatrix} = \overbrace{\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}}^{M^{-1}} \begin{pmatrix} q' \\ p' \end{pmatrix} \mod \mathbb{Z}^2 \quad . \tag{3.91}$$

Now for something cool. Suppose that our image consists of a set of discrete points located at $(n_1/k\,,\,n_2/k)$, where the denominator $k \in \mathbb{Z}$ is fixed, and where n_1 and n_2 range over the set $\{1,\ldots,k\}$. Clearly g and its inverse preserve this set, since the entries of M and M^{-1} are integers. If there are two possibilities for each pixel (say off and on, or black and white), then there are $2^{(k^2)}$ possible images, and the cat map will map us invertibly from one image to another. Therefore it must exhibit Poincaré recurrence! This phenomenon is demonstrated vividly in fig. 3.10, which shows a $k = 150$ pixel (square) image of a cat subjected to the iterated cat map. The image is stretched and folded with each successive application of the cat map, but after 300 iterations the image is restored! How can this be if the cat map is mixing? The point is that only the discrete set of points $(n_1/k\,,\,n_2/k)$ is periodic. Points with different denominators will exhibit a different periodicity, and points with irrational coordinates will in general never return to their exact initial conditions, although recurrence says they will come arbitrarily close, given enough iterations. The baker's transformation is also different in this respect, since the denominator of the p coordinate is doubled upon each successive iteration.

The student should now contemplate the hierarchy of dynamical systems depicted in fig. 3.13, under-
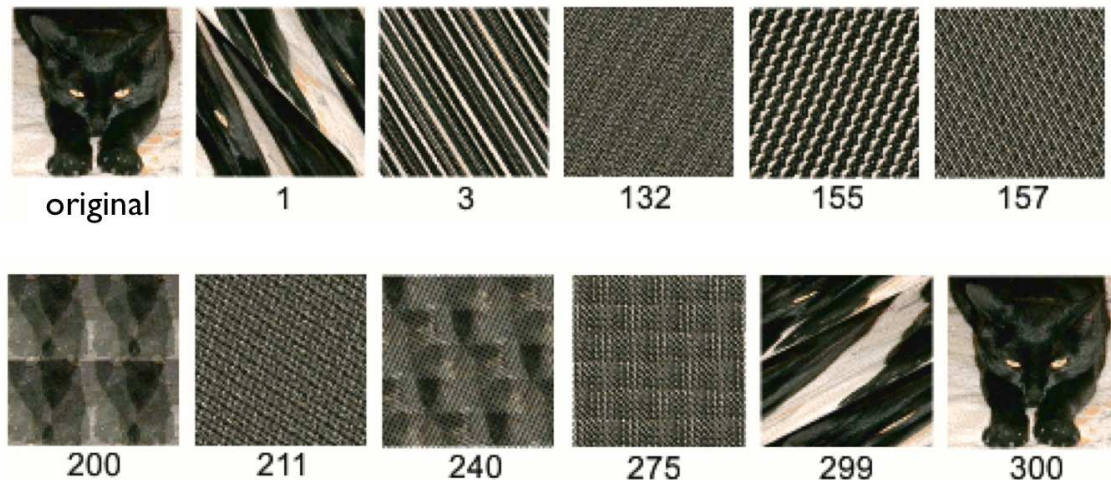
Figure 3.10: The Arnold cat map applied to an image of 150×150 pixels. After 300 iterations, the image repeats itself. (*Source: Wikipedia*)

standing the characteristic features of each successive refinement[11].

3.6   Appendix I : Normal matrices and Jordan canonical form

If a matrix A is normal, meaning $\left[A, A^\dagger\right] = 0$, then it may be diagonalized by a unitary transformation. Indeed, any $n \times n$ matrix A is diagonalizable by a unitary transformation if and only if A is normal. When A is normal, the eigenspaces span, and we may choose $\langle\!\langle\, L^\alpha \,\|\, R^\beta \,\rangle\!\rangle = \delta^{\alpha\beta}$, using Gram-Schmidt in the case of degeneracies.

When A is not normal, while the sum of the dimensions of its eigenspaces generically is equal to its dimension $\dim(A) = n$, this is not guaranteed, and it may be less than n. What is true is that any non-normal complex matrix A can be brought to Jordan canonical form by a similarity transformation $\tilde{A} = Q^{-1} A\, Q$, where Q is invertible and

$$\tilde{A} = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_b \end{pmatrix} \quad . \tag{3.92}$$

---

[11]There is something beyond mixing, called a *K-system*. A *K*-system has positive Kolmogorov-Sinai entropy. For such a system, closed orbits separate exponentially in time, and consequently the Liouvillian L has a Lebesgue spectrum with denumerably infinite multiplicity.

Here b is the number of Jordan blocks, where each block J_α is of the form

$$
J_\alpha = \begin{pmatrix} \lambda_\alpha & 1 & & \\ & \lambda_\alpha & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_\alpha \end{pmatrix} \quad . \tag{3.93}
$$

Thus each J_α is tridiagonal, with diagonal elements all given by λ_α and each element directly above the diagonal equal to one. We denote the right and left eigenvectors of $\tilde{A}$ as $\| \widetilde{R^\alpha} \rangle\!\rangle = Q^{-1} \| R^\alpha \rangle\!\rangle$ and $\langle\!\langle \widetilde{L^\alpha} \| = \langle\!\langle L^\alpha \| Q$, respectively. Each J_α has only one right eigenvector, $\psi_j^\alpha = \delta_{j,1}$, whose corresponding left eigenvector is $\chi_j^\alpha = \delta_{j,n_\alpha}$, where $n_\alpha = \mathrm{rank}(J_\alpha)$. Note $n = \mathrm{rank}(A)$ is the sum of the dimensions of the Jordan blocks, *i.e.* $n = \sum_{\alpha=1}^b n_\alpha$. When $n_\alpha = 1$, the Jordan block is the 1×1 matrix λ_α. For a non-normal matrix A, its eigenvalues λ_α may be complex. However, if all the elements of A are real, then any complex eigenvalues must occur in complex conjugate pairs, because the characteristic polynomial $P(\lambda) = \det(\lambda - A)$ satisfies $\big[P(\lambda) \big]^* = P(\lambda^*)$.

When $\lambda_\alpha \neq \lambda_\beta$, we have $\langle\!\langle L^\alpha \| R^\beta \rangle\!\rangle = \langle\!\langle \widetilde{L^\alpha} \| \widetilde{R^\beta} \rangle\!\rangle = 0$. For eigenspaces with $n_\alpha = 1$, we may choose $\langle\!\langle L^\alpha \| R^\beta \rangle\!\rangle = \langle\!\langle \widetilde{L^\alpha} \| \widetilde{R^\beta} \rangle\!\rangle = \delta^{\alpha\beta}$, but for the nontrivial Jordan blocks with $n_\alpha > 1$ we have $\langle\!\langle \widetilde{L^\alpha} \| \widetilde{R^\alpha} \rangle\!\rangle = 0$, as we have seen in the previous paragraph, and therefore $\langle\!\langle L^\alpha \| R^\alpha \rangle\!\rangle = 0$. Real symmetric matrices are all normal, with no Jordan blocks. For complex symmetric matrices, we may have nontrivial Jordan blocks.

Since $\langle\!\langle L \| R \rangle\!\rangle = \langle\!\langle \widetilde{L} \| \widetilde{R} \rangle\!\rangle = 0$, we may write

$$
A = \bigoplus_{n_\alpha = 1} \lambda_\alpha \| R^\alpha \rangle\!\rangle \langle\!\langle L^\alpha \| \bigoplus_{n_\beta > 1} Q J_\beta Q^{-1} \tag{3.94}
$$

and raising A to the k power yields

$$
A^k = \bigoplus_{n_\alpha = 1} \lambda_\alpha^k \| R^\alpha \rangle\!\rangle \langle\!\langle L^\alpha \| \bigoplus_{n_\beta > 1} Q J_\beta^k Q^{-1} \quad . \tag{3.95}
$$

Note that J_β^k is upper triangular with all diagonal elements given by λ_k^β. Note that for complex symmetric matrices, the left and right eigenvectors are identical and we may write

$$
L_j^\alpha = R_j^\alpha \equiv \Psi_j^\alpha \qquad \Rightarrow \qquad \langle\!\langle L^\alpha \| = \langle\!\langle \Psi^\alpha \| \quad , \quad \| R^\alpha \rangle\!\rangle = \| \Psi^\alpha \rangle\!\rangle \quad , \tag{3.96}
$$

with no complex conjugation, *i.e.* $\langle\!\langle L^\alpha \| j \rangle\!\rangle = \langle\!\langle j \| R^\alpha \rangle\!\rangle \equiv \Psi_j^\alpha$.

3.6.1 Contrast with singular value decomposition

We now remark upon the difference between the decomposition into Jordan canonical form and the singular value decomposition (SVD), in which we write an $m \times n$ matrix A as $A = U D V^\dagger$, where U is $m \times k$, V is $n \times k$ (hence $V^\dagger$ is $k \times n$), $U^\dagger U = V^\dagger V = \mathbb{I}_{k \times k}$, and $D = \mathrm{diag}(d_1, \ldots, d_k)$ is $k \times k$ with

$k \leq \min(m, n)$ and each $d_j > 0$. The elements d_j are the singular values and the rows of U and V are the singular vectors. Note that $A^\dagger A = V D^2 V^\dagger$ is $n \times n$ and $AA^\dagger = UD^2 U^\dagger$ is $m \times m$. If we define

$$R(\lambda) = \prod_{j=1}^{k} \left(\lambda - d_j^2 \right) \quad , \tag{3.97}$$

Then

$$P(\lambda) \equiv \det(\lambda - A^\dagger A) = \lambda^{n-k} R(\lambda) \quad , \qquad Q(\lambda) \equiv \det(\lambda - AA^\dagger) = \lambda^{m-k} R(\lambda) \quad . \tag{3.98}$$

For any square $n \times n$ complex matrix A we therefore have two decompositions, via JCF and SVD, *viz.*

$$A = QJ Q^{-1} = UD V^\dagger \quad , \tag{3.99}$$

where J is the Jordan canonical form of A. When A is normal, $k = n$ and $U = V = Q$, *i.e.* the two decompositions are equivalent.

3.6.2   Example

Consider the real asymmetric matrix

$$A = \begin{pmatrix} 2a & -4a^2 \\ \frac{1}{4} & 0 \end{pmatrix} \quad , \tag{3.100}$$

where $a \in \mathbb{Z}$ is any real number. The characteristic polynomial is $F(\lambda) = \det(\lambda - A) = (\lambda - a)^2$ and there is a single eigenvalue, $\lambda = a$. The right and left eigenvectors are found to be

$$\| R \rangle\!\rangle = \begin{pmatrix} 4a \\ 1 \end{pmatrix} \quad , \qquad \langle\!\langle L \| = \begin{pmatrix} \frac{1}{4} & -a \end{pmatrix} \quad , \tag{3.101}$$

where the normalization is arbitrary. Note $\langle\!\langle L \| R \rangle\!\rangle = 0$. The matrix A is brought to JCF by the similarity transformation $\tilde{A} = Q^{-1} A Q$ with

$$Q = \begin{pmatrix} 4a & 4 \\ 1 & 0 \end{pmatrix} \quad , \qquad Q^{-1} = \begin{pmatrix} 0 & 1 \\ \frac{1}{4} & -a \end{pmatrix} \quad , \qquad \tilde{A} = Q^{-1} A Q = \begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix} \quad . \tag{3.102}$$

Note that

$$\| \tilde{R} \rangle\!\rangle = Q^{-1} \| R \rangle\!\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad , \qquad \langle\!\langle \tilde{L} \| = \langle\!\langle L \| Q = \begin{pmatrix} 0 & 1 \end{pmatrix} \quad , \tag{3.103}$$

and that $\langle\!\langle \tilde{L} \| \tilde{R} \rangle\!\rangle = 0$.

Adding another row and column to our matrix A, consider the matrix

$$B = \begin{pmatrix} 2a & -4a^2 & 0 \\ \frac{1}{4} & 0 & 0 \\ 0 & 0 & b \end{pmatrix} \quad , \tag{3.104}$$

where both a and b are arbitrary real numbers. Since $\det(\lambda - B) = (\lambda - a)^2 (\lambda - b)$, B thus has two eigenvalues: $\lambda_1 = a$ and $\lambda_2 = b$. The decomposition of B is then

$$B = Q J_1 Q^{-1} \bigoplus \lambda_2 \, \| \, R_2 \, \rangle\!\rangle \langle\!\langle \, L_2 \, \| \quad , \tag{3.105}$$

where $\langle\!\langle \, L_2 \, \| = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$ and $\| \, R_2 \, \rangle\!\rangle = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}^{\mathrm{t}}$. Of course, we could mix up the various elements of B by applying a general similarity transformation $B \to B' \equiv SBS^{-1}$, but the JCF of B' would be the same.

3.7 Appendix II : Formal Solution of the Master Equation

Recall the master equation $\dot{P}_i = -\Gamma_{ij} P_j$. The matrix Γ_{ij} is real but not necessarily symmetric. For such a matrix, the left eigenvectors L_i^α and the right eigenvectors R_j^β are in general not related by a simple transpose operation:

$$\begin{aligned} L_i^\alpha \, \Gamma_{ij} &= \lambda_\alpha \, L_j^\alpha \\ \Gamma_{ij} \, R_j^\beta &= \lambda_\beta \, R_i^\beta \quad . \end{aligned} \tag{3.106}$$

We denote the right and eigenvectors by $\| \, R^\alpha \, \rangle\!\rangle$ and $\langle\!\langle \, L^\alpha \, \|$, which are column vectors and row vectors, respectively. Thus $\Gamma \, \| \, R^\alpha \, \rangle\!\rangle = \lambda_\alpha \| \, R^\alpha \, \rangle\!\rangle$ and $\langle\!\langle \, L^\alpha \, \| \, \Gamma = \langle\!\langle \, L^\alpha \, \| \, \lambda_\alpha$, the second of which may be written as $\Gamma^{\mathrm{t}} \, \| \, L^\alpha \, \rangle\!\rangle = \lambda_\alpha \, \| \, L^\alpha \, \rangle\!\rangle$, where the column vector $\| \, L^\alpha \, \rangle\!\rangle$ is the transpose of the row vector $\langle\!\langle \, L^\alpha \, \|$ and Γ^{t} is the matrix transpose of Γ. The characteristic polynomial is the same in both cases:

$$F(\lambda) \equiv \det (\lambda - \Gamma) = \det (\lambda - \Gamma^{\mathrm{t}}) \quad , \tag{3.107}$$

which means that the left and right eigenvalues are the same. Note also that $\big[F(\lambda)\big]^* = F(\lambda^*)$, hence the eigenvalues are either real or appear in complex conjugate pairs. Multiplying the eigenvector equation for L^α on the right by R_j^β and summing over j, and multiplying the eigenvector equation for R^β on the left by L_i^α and summing over i, and subtracting the two results yields $(\lambda_\alpha - \lambda_\beta) \, \langle\!\langle \, L^\alpha \, \| \, R^\beta \, \rangle\!\rangle = 0$, where the inner product is now $\langle\!\langle \, L^\alpha \, \| \, R^\beta \, \rangle\!\rangle = \sum_i L_i^\alpha \, R_i^\beta$ with no complex conjugation on the bra vector. We may now demand $\langle\!\langle \, L^\alpha \, \| \, R^\beta \, \rangle\!\rangle = \delta_{\alpha\beta}$, which is our eigenvector normalization condition. As discussed above in §3.6, in the event that Γ contains nontrivial Jordan blocks, its eigenvectors do not span. However, this is a nongeneric state of affairs, and here we assume that Γ contains no nontrivial Jordan blocks.

We have seen that $\langle\!\langle \, L \, \| = \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix}$ is a left eigenvector of the matrix Γ with eigenvalue $\lambda = 0$, since $\sum_i \Gamma_{ij} = 0$. We do not know *a priori* the corresponding right eigenvector, which depends on other details of Γ_{ij}. Generically, a matrix which is not normal has spanning eigenvectors, *i.e.* the existence of nontrivial Jordan blocks is nongeneric. Assuming that the eigenvectors of Γ span, then, let's expand the probability distribution $P_i(t)$ in the right eigenvectors of Γ, writing

$$P_i(t) = \sum_\alpha C_\alpha(t) \, R_i^\alpha \quad , \tag{3.108}$$

where $\alpha \in \{0, 1, \dots, n-1\}$, where n is the rank of Γ. Then

$$\frac{dP_i}{dt} = \sum_\alpha \frac{dC_\alpha}{dt}\, R_i^\alpha = -\Gamma_{ij}\, P_j = -\sum_\alpha C_\alpha\, \Gamma_{ij}\, R_j^\alpha = -\sum_\alpha \lambda_\alpha\, C_\alpha\, R_i^\alpha \quad . \tag{3.109}$$

This allows us to write

$$\frac{dC_\alpha}{dt} = -\lambda_\alpha\, C_\alpha \qquad \Longrightarrow \qquad C_\alpha(t) = C_\alpha(0)\, e^{-\lambda_\alpha t} \quad . \tag{3.110}$$

Hence, we can write

$$P_i(t) = \sum_\alpha C_\alpha(0)\, e^{-\lambda_\alpha t}\, R_i^\alpha \quad . \tag{3.111}$$

Let $\alpha = 1$ correspond to the left eigenvector $\langle\!\langle\, L^1\, \| = \begin{pmatrix} 1 & 1 & \cdots & 1 \end{pmatrix}$. The corresponding eigenvalue is $\lambda_1 = 1$. It is now easy to see that $\mathrm{Re}\,(\lambda_\alpha) > 0$ for all $\alpha > 1$, or else the probabilities will become negative[12]. For suppose $\mathrm{Re}\,(\lambda_\alpha) < 0$ for some α. Then as $t \to \infty$, the sum in eqn. 3.111 will be dominated by the term for which λ_α has the largest negative real part; all other contributions will be subleading. But we must have $\sum_i R_i^\alpha = 0$ since $\| R^\alpha \rangle\!\rangle$ must be orthogonal to the left eigenvector $\langle\!\langle\, L^0\, \|$. Therefore, at least one component of R_i^α (*i.e.* for some value of i) must have a negative real part, which means a negative probability![13] As we have already proven that an initial nonnegative distribution $\{P_i(t = 0)\}$ will remain nonnegative under the evolution of the master equation, we conclude that $P_i(t) \to P_i^{\mathrm{eq}}$ as $t \to \infty$, relaxing to the $\lambda = 0$ right eigenvector, with $\mathrm{Re}\,(\lambda_\alpha) \geq 0$ for all α.

3.7.1 Detailed balance

Consider an arbitrary nonnegative real upper triangular matrix T with $T_{ij} \geq 0$ for all $1 \leq i < j \leq n$. Let π_i be a normalized distribution, *i.e.* $\pi_i \geq 0$ for all $i \in \{1, \dots, n\}$ with $\sum_i \pi_i = 1$. Now define the nonnegative matrix

$$W_{ij} = \begin{cases} \pi_i\, T_{ij} & \text{if } i < j \\ \pi_j\, T_{ji} & \text{if } i > j \end{cases} \quad . \tag{3.112}$$

and take this to be the matrix of transition rates so that the master equation is as in Eqn. 3.1

$$\frac{dP_i}{dt} = \sum_j \left(W_{ij}\, P_j - W_{ji}\, P_i \right) \quad . \tag{3.113}$$

Since $W_{ij}/W_{ji} = \pi_i/\pi_j$, the matrix W satisfies detailed balance relative to the distribution π. With $\Gamma_{ij} \equiv W_{ij}$ for $i \neq j$ and $\Gamma_{ii} \equiv \sum_k' W_{ki}$ (with $k = i$ excluded from the sum) as before, we recover the form of the master equation $\dot{P}_i = -\sum_j \Gamma_{ij}\, P_j$.

How many parameters does it take to describe a general $n \times n$ transition matrix W_{ij} satisfying detailed balance? Since there are $\frac{1}{2}n(n-1)$ freedoms in T and $n-1$ freedoms in π, we conclude that $\frac{1}{2}(n-1)(n+2)$

---

[12]We presume that the eigenvalue $\lambda = 0$ is nondegenerate.

[13]Since the probability $P_i(t)$ is real, if the eigenvalue with the smallest (*i.e.* largest negative) real part is complex, there will be a corresponding complex conjugate eigenvalue, and summing over all eigenvectors will result in a real value for $P_i(t)$.

parameters are required to specify W_{ij}. But if we drop the constraint of detailed balance, then all the elements of W_{ij} not lying on the diagonal are independent, corresponding to $n(n-1)$ parameters. Note that we may set $W_{ii} = 0$ for all i.

3.7.2 Example : Poisson processes and radioactive decay

Here we consider two examples where the state labels of the master equation denote a number, corresponding to the discrete population of some group. The master equation is

$$\frac{dP_n}{dt} = \sum_m \left(W_{nm} P_m - W_{mn} P_n \right) \quad . \tag{3.114}$$

Thus W_{nm} is the transition rate for the process $|m\rangle \to |n\rangle$.

We first consider the *Poisson process*, for which

$$W_{nm} = \begin{cases} \lambda & \text{if } n = m+1 \\ 0 & \text{if } n \neq m+1 \, . \end{cases} \tag{3.115}$$

We then have

$$\frac{dP_n}{dt} = \lambda \left(P_{n-1} - P_n \right) . \tag{3.116}$$

The *generating function* $P(z,t) = \sum_{n=0}^{\infty} z^n P_n(t)$ then satisfies

$$\frac{\partial P}{\partial t} = \lambda (z-1) P \quad \Rightarrow \quad P(z,t) = e^{(z-1)\lambda t} P(z,0) . \tag{3.117}$$

If the initial distribution is $P_n(0) = \delta_{n,0}$, then

$$P_n(t) = \frac{(\lambda t)^n}{n!} e^{-\lambda t} , \tag{3.118}$$

which is known as the *Poisson distribution*. If we define $\alpha \equiv \lambda t$, then from $P_n = \alpha^n e^{-\alpha}/n!$ we have

$$\langle n^k \rangle = e^{-\alpha} \left(\alpha \frac{\partial}{\partial \alpha} \right)^k e^{\alpha} . \tag{3.119}$$

Thus, $\langle n \rangle = \alpha$, $\langle n^2 \rangle = \alpha^2 + \alpha$, *etc.*

3.7.3 Radioactive decay

Consider a group of atoms, some of which are in an excited state which can undergo nuclear decay. Let $P_n(t)$ be the probability that n atoms are excited at some time t. We then model the decay dynamics by

$$W_{nm} = \begin{cases} 0 & \text{if } n \geq m \\ m\gamma & \text{if } n = m-1 \\ 0 & \text{if } n < m-1 \quad . \end{cases} \tag{3.120}$$

Here, γ is the decay rate of an individual atom, which can be determined from quantum mechanics. The master equation then tells us

$$\frac{dP_n}{dt} = (n+1)\,\gamma\,P_{n+1} - n\,\gamma\,P_n \quad . \tag{3.121}$$

The interpretation here is as follows: let $|\,n\,\rangle$ denote a state in which n atoms are excited. Then $P_n(t) = \left|\langle\,n\,|\,P(t)\,\rangle\right|^2$. Then $P_n(t)$ will increase due to spontaneous transitions from $|\,n{+}1\,\rangle$ to $|\,n\,\rangle$, and will decrease due to spontaneous transitions from $|\,n\,\rangle$ to $|\,n{-}1\,\rangle$.

The average number of particles in the system is $N(t) = \sum_{n=0}^{\infty} n\,P_n(t)$. Note that

$$\begin{aligned}
\frac{dN}{dt} &= \sum_{n=0}^{\infty} n\Big[(n+1)\,\gamma\,P_{n+1} - n\,\gamma\,P_n\Big] \\
&= \gamma \sum_{n=0}^{\infty}\Big[n(n-1)\,P_n - n^2 P_n\Big] = -\gamma \sum_{n=0}^{\infty} n\,P_n = -\gamma\,N \quad .
\end{aligned} \tag{3.122}$$

Thus, $N(t) = N(0)\,e^{-\gamma t}$. The relaxation time is $\tau = \gamma^{-1}$, and the equilibrium distribution is $P_n^{\mathrm{eq}} = \delta_{n,0}$. Note that this satisfies detailed balance.

We can go a bit farther here. Let us define

$$P(z,t) \equiv \sum_{n=0}^{\infty} z^n\,P_n(t) \quad . \tag{3.123}$$

This is sometimes called a *generating function*. Then

$$\begin{aligned}
\frac{\partial P}{\partial t} &= \gamma \sum_{n=0}^{\infty} z^n\Big[(n+1)\,P_{n+1} - n\,P_n\Big] \\
&= \gamma\,\frac{\partial P}{\partial z} - \gamma z\,\frac{\partial P}{\partial z} \quad .
\end{aligned} \tag{3.124}$$

Thus,

$$\frac{1}{\gamma}\frac{\partial P}{\partial t} - (1-z)\frac{\partial P}{\partial z} = 0 \quad . \tag{3.125}$$

We now see that any function $f(\xi)$ satisfies the above equation, where $\xi = \gamma t - \ln(1-z)$. Thus, we can write

$$P(z,t) = f\big(\gamma t - \ln(1-z)\big) \quad . \tag{3.126}$$

Setting $t=0$ we have $P(z,0) = f\big(-\ln(1-z)\big)$, and inverting this result we obtain $f(u) = P(1 - e^{-u}, 0)$, which entails

$$P(z,t) = P\big(1 + (z-1)\,e^{-\gamma t},\,0\big) \quad . \tag{3.127}$$

The total probability is $P(z{=}1,t) = \sum_{n=0}^{\infty} P_n$, which clearly is conserved: $P(1,t) = P(1,0)$. The average particle number is

$$N(t) = \sum_{n=0}^{\infty} n\,P_n(t) = \frac{\partial P}{\partial z}\bigg|_{z=1} = e^{-\gamma t}\,P(1,0) = N(0)\,e^{-\gamma t} \quad . \tag{3.128}$$

3.8   Appendix III: Transition to Ergodicity in a Simple Model

A ball of mass m executes perfect one-dimensional motion along the symmetry axis of a piston. Above the ball lies a mobile piston head of mass M which slides frictionlessly inside the piston. Both the ball and piston head execute ballistic motion, with two types of collision possible: (i) the ball may bounce off the floor, which is assumed to be infinitely massive and fixed in space, and (ii) the ball and piston head may engage in a one-dimensional elastic collision. The Hamiltonian is

$$H = \frac{P^2}{2M} + \frac{p^2}{2m} + MgX + mgx \quad ,$$

where X is the height of the piston head and x the height of the ball. Another quantity is conserved by the dynamics: $\Theta(X - x)$. I.e., the ball always is below the piston head.

(a) Choose an arbitrary length scale L, and then energy scale $E_0 = MgL$, momentum scale $P_0 = M\sqrt{gL}$, and time scale $\tau_0 = \sqrt{L/g}$. Show that the dimensionless Hamiltonian becomes

$$\bar{H} = \tfrac{1}{2}\bar{P}^2 + \bar{X} + \frac{\bar{p}^2}{2r} + r\bar{x} \quad ,$$

with $r = m/M$, and with equations of motion $dX/dt = \partial\bar{H}/\partial\bar{P}$, etc. (Here the bar indicates dimensionless variables: $\bar{P} = P/P_0$, $\bar{t} = t/\tau_0$, etc.) What special dynamical consequences hold for $r = 1$?

(b) Compute the microcanonical average piston height $\langle X \rangle$. The analogous dynamical average is

$$\langle X \rangle_t = \lim_{T \to \infty} \frac{1}{T} \int_0^T dt\, X(t) \quad .$$

When computing microcanonical averages, it is helpful to use the Laplace transform, discussed in §4.2.2 of the notes. (It is possible to compute the microcanonical average by more brute force methods as well.)

(c) Compute the microcanonical average of the rate of collisions between the ball and the floor. Show that this is given by

$$\Big\langle \sum_i \delta(t - t_i) \Big\rangle = \big\langle \Theta(v)\, v\, \delta(x - 0^+) \big\rangle \quad .$$

The analogous dynamical average is

$$\langle \gamma \rangle_t = \lim_{T \to \infty} \frac{1}{T} \int_0^T dt \sum_i \delta(t - t_i) \quad ,$$

where $\{t_i\}$ is the set of times at which the ball hits the floor.

(d) How do your results change if you do not enforce the dynamical constraint $X \geq x$?

(e) Write a computer program to simulate this system. The only input should be the mass ratio r (set $\bar{E} = 10$ to fix the energy). You also may wish to input the initial conditions, or perhaps to choose the initial conditions randomly (all satisfying energy conservation, of course!). Have your program compute

the microcanonical as well as dynamical averages in parts (b) and (c). Plot out the Poincaré section of P vs. X for those times when the ball hits the floor. Investigate this for several values of r. Just to show you that this is interesting, I've plotted some of my own numerical results in fig. 3.11.

Solution:

(a) Once we choose a length scale L (arbitrary), we may define $E_0 = M\,gL$, $P_0 = M\sqrt{gL}$, $V_0 = \sqrt{gL}$, and $\tau_0 = \sqrt{L/g}$ as energy, momentum, velocity, and time scales, respectively, the result follows directly. Rather than write $\bar{P} = P/P_0$ etc., we will drop the bar notation and write

$$H = \tfrac{1}{2}P^2 + X + \frac{p^2}{2r} + rx \quad .$$

(b) What is missing from the Hamiltonian of course is the interaction potential between the ball and the piston head. We assume that both objects are impenetrable, so the potential energy is infinite when the two overlap. We further assume that the ball is a point particle (otherwise reset ground level to minus the diameter of the ball). We can eliminate the interaction potential from H if we enforce that each time $X = x$ the ball and the piston head undergo an elastic collision. From energy and momentum conservation, it is easy to derive the elastic collision formulae

$$P' = \frac{1-r}{1+r}\,P + \frac{2}{1+r}\,p$$

$$p' = \frac{2r}{1+r}\,P - \frac{1-r}{1+r}\,p \quad .$$

We can now answer the last question from part (a). When $r = 1$, we have that $P' = p$ and $p' = P$, i.e. the ball and piston simply exchange momenta. The problem is then equivalent to two identical particles elastically bouncing off the bottom of the piston, and moving through each other as if they were completely transparent. When the trajectories cross, however, the particles exchange identities.

Averages within the microcanonical ensemble are normally performed with respect to the phase space distribution

$$\varrho(\varphi) = \frac{\delta\big(E - H(\varphi)\big)}{\mathsf{Tr}\,\delta\big(E - H(\varphi)\big)} \quad ,$$

where $\varphi = (P, X, p, x)$, and

$$\mathsf{Tr}\,F(\varphi) = \int\limits_{-\infty}^{\infty} dP \int\limits_{0}^{\infty} dX \int\limits_{-\infty}^{\infty} dp \int\limits_{0}^{\infty} dx\; F(P, X, p, x) \quad .$$

Since $X \geq x$ is a dynamical constraint, we should define an appropriately restricted microcanonical average:

$$\big\langle F(\varphi)\big\rangle_{\mu ce} \equiv \widetilde{\mathsf{Tr}}\left[F(\varphi)\,\delta\big(E - H(\varphi)\big)\right]\bigg/\widetilde{\mathsf{Tr}}\,\delta\big(E - H(\varphi)\big)$$

where

$$\widetilde{\mathsf{Tr}}\,F(\varphi) \equiv \int\limits_{-\infty}^{\infty} dP \int\limits_{0}^{\infty} dX \int\limits_{-\infty}^{\infty} dp \int\limits_{0}^{X} dx\; F(P, X, p, x)$$

is the modified trace. Note that the integral over x has an upper limit of X rather than ∞, since the region of phase space with $x > X$ is *dynamically inaccessible*.

When computing the traces, we shall make use of the following result from the theory of Laplace transforms. The Laplace transform of a function $K(E)$ is

$$\widehat{K}(\beta) = \int\limits_0^\infty dE\, K(E)\, e^{-\beta E} \quad .$$

The inverse Laplace transform is given by

$$K(E) = \int\limits_{c-i\infty}^{c+i\infty} \frac{d\beta}{2\pi i}\, \widehat{K}(\beta)\, e^{\beta E} \quad ,$$

where the integration contour, which is a line extending from $\beta = c - i\infty$ to $\beta = c + i\infty$, lies to the right of any singularities of $\widehat{K}(\beta)$ in the complex β-plane. For this problem, all we shall need is the following:

$$K(E) = \frac{E^{t-1}}{\Gamma(t)} \quad \Longleftrightarrow \quad \widehat{K}(\beta) = \beta^{-t} \quad .$$

For a proof, see §4.2.2 of the lecture notes.

We're now ready to compute the microcanonical average of X. We have

$$\langle X \rangle = \frac{N(E)}{D(E)} \quad ,$$

where $N(E) = \widetilde{\text{Tr}}\left[X\,\delta(E - H)\right]$ and $D(E) = \widetilde{\text{Tr}}\,\delta(E - H)$. Let's first compute $D(E)$. To do this, we compute the Laplace transform $\widehat{D}(\beta)$:

$$\widehat{D}(\beta) = \widetilde{\text{Tr}}\, e^{-\beta H}$$

$$= \int\limits_{-\infty}^\infty dP\, e^{-\beta P^2/2} \int\limits_{-\infty}^\infty dp\, e^{-\beta p^2/2r} \int\limits_0^\infty dX\, e^{-\beta X} \int\limits_0^X dx\, e^{-\beta r x}$$

$$= \frac{2\pi\sqrt{r}}{\beta} \int\limits_0^\infty dX\, e^{-\beta X}\left(\frac{1 - e^{-\beta r X}}{\beta r}\right) = \frac{\sqrt{r}}{1+r} \cdot \frac{2\pi}{\beta^3} \quad .$$

Similarly for $\widehat{N}(\beta)$ we have

$$\widehat{N}(\beta) = \widetilde{\text{Tr}}\, X\, e^{-\beta H}$$

$$= \int\limits_{-\infty}^\infty dP\, e^{-\beta P^2/2} \int\limits_{-\infty}^\infty dp\, e^{-\beta p^2/2r} \int\limits_0^\infty dX\, X\, e^{-\beta X} \int\limits_0^X dx\, e^{-\beta r x}$$

$$= \frac{2\pi\sqrt{r}}{\beta} \int\limits_0^\infty dX\, X\, e^{-\beta X}\left(\frac{1 - e^{-\beta r X}}{\beta r}\right) = \frac{(2+r)\,r^{3/2}}{(1+r)^2} \cdot \frac{2\pi}{\beta^4} \quad .$$

Taking the inverse Laplace transform, we then have

$$D(E) = \frac{\sqrt{r}}{1+r} \cdot \pi E^2 \qquad , \qquad N(E) = \frac{(2+r)\sqrt{r}}{(1+r)^2} \cdot \tfrac{1}{3}\pi E^3 \quad .$$

We then have

$$\langle X \rangle = \frac{N(E)}{D(E)} = \left(\frac{2+r}{1+r}\right) \cdot \tfrac{1}{3}E \quad .$$

The 'brute force' evaluation of the integrals isn't so bad either. We have

$$D(E) = \int\limits_{-\infty}^{\infty} dP \int\limits_{0}^{\infty} dX \int\limits_{-\infty}^{\infty} dp \int\limits_{0}^{X} dx \; \delta\big(\tfrac{1}{2}P^2 + \tfrac{1}{2r}p^2 + X + rx - E\big) \quad .$$

To evaluate, define $P = \sqrt{2}\,u_x$ and $p = \sqrt{2r}\,u_y$. Then we have $dP\,dp = 2\sqrt{r}\,du_x\,du_y$ and $\tfrac{1}{2}P^2 + \tfrac{1}{2r}p^2 = u_x^2 + u_y^2$. Now convert to 2D polar coordinates with $w \equiv u_x^2 + u_y^2$. Thus,

$$D(E) = 2\pi\sqrt{r} \int\limits_{0}^{\infty} dw \int\limits_{0}^{\infty} dX \int\limits_{0}^{X} dx \; \delta\big(w + X + rx - E\big)$$

$$= \frac{2\pi}{\sqrt{r}} \int\limits_{0}^{\infty} dw \int\limits_{0}^{\infty} dX \int\limits_{0}^{X} dx \; \Theta(E - w - X)\,\Theta(X + rX - E + w)$$

$$= \frac{2\pi}{\sqrt{r}} \int\limits_{0}^{E} dw \int\limits_{\frac{E-w}{1+r}}^{E-w} dX = \frac{2\pi\sqrt{r}}{1+r} \int\limits_{0}^{E} dq\,q = \frac{\sqrt{r}}{1+r} \cdot \pi E^2 \quad ,$$

with $q = E - w$. Similarly,

$$N(E) = 2\pi\sqrt{r} \int\limits_{0}^{\infty} dw \int\limits_{0}^{\infty} dX\,X \int\limits_{0}^{X} dx \; \delta\big(w + X + rx - E\big)$$

$$= \frac{2\pi}{\sqrt{r}} \int\limits_{0}^{\infty} dw \int\limits_{0}^{\infty} dX\,X \int\limits_{0}^{X} dx \; \Theta(E - w - X)\,\Theta(X + rX - E + w)$$

$$= \frac{2\pi}{\sqrt{r}} \int\limits_{0}^{E} dw \int\limits_{\frac{E-w}{1+r}}^{E-w} dX\,X = \frac{2\pi}{\sqrt{r}} \int\limits_{0}^{E} dq \left(1 - \frac{1}{(1+r)^2}\right) \cdot \tfrac{1}{2}q^2 = \left(\frac{2+r}{1+r}\right) \cdot \frac{\sqrt{r}}{1+r} \cdot \tfrac{1}{3}\pi E^3 \quad .$$

(c) Using the general result

$$\delta\big(F(x) - A\big) = \sum_i \frac{\delta(x - x_i)}{|F'(x_i)|} \quad ,$$

where $F(x_i) = A$, we recover the desired expression. We should be careful not to double count, so to avoid this difficulty we can evaluate $\delta(t - t_i^+)$, where $t_i^+ = t_i + 0^+$ is infinitesimally later than t_i. The point
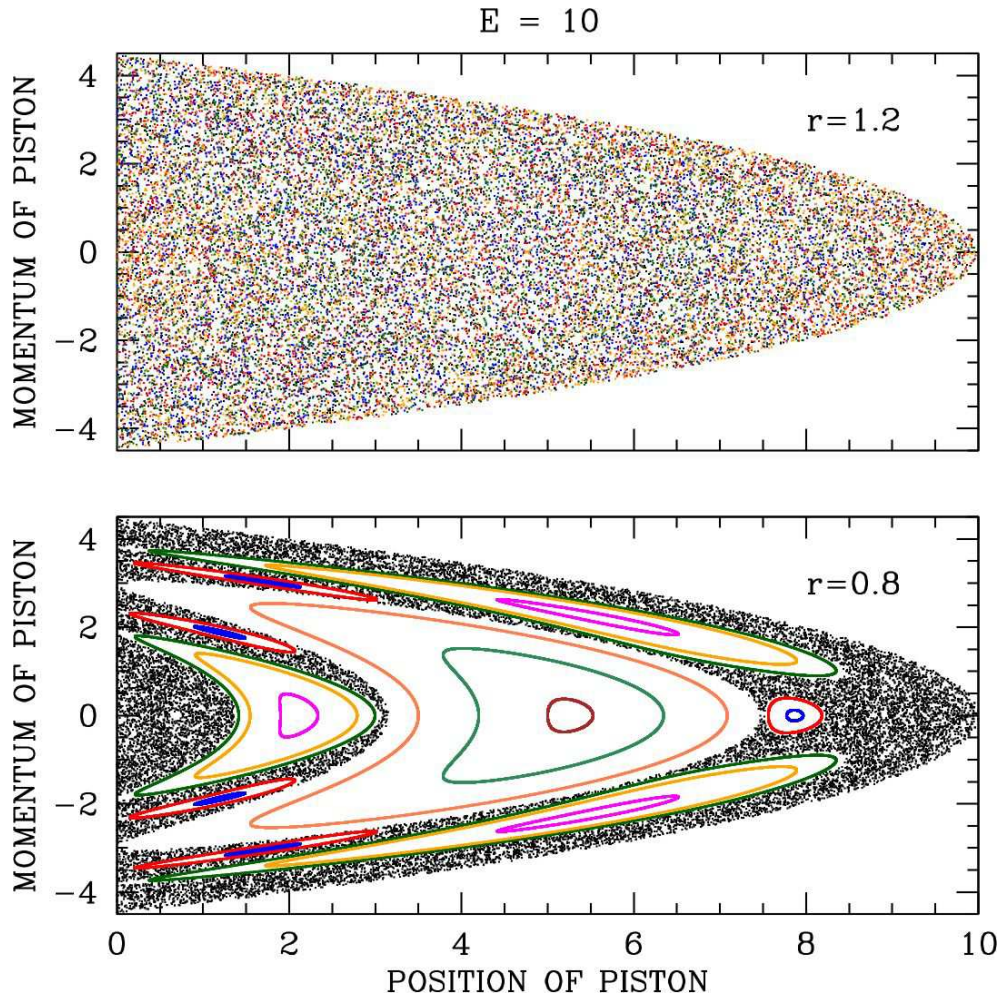
Figure 3.11: Poincaré sections for the ball and piston head problem. Each color corresponds to a different initial condition. When the mass ratio $r = m/M$ exceeds unity, the system apparently becomes ergodic.

here is that when $t = t_i^+$ we have $p = r\,v > 0$ (*i.e.* just after hitting the bottom). Similarly, at times $t = t_i^-$ we have $p < 0$ (*i.e.* just prior to hitting the bottom). Note $v = p/r$. Again we write $\gamma(E) = N(E)/D(E)$, this time with

$$N(E) = \widetilde{\mathsf{Tr}}\left[\Theta(p)\,r^{-1}p\,\delta(x - 0^+)\,\delta(E - H)\right] \quad .$$

The Laplace transform is

$$\widehat{N}(\beta) = \int\limits_{-\infty}^{\infty} dP\,e^{-\beta P^2/2}\int\limits_{0}^{\infty} dp\,r^{-1}\,p\,e^{-\beta p^2/2r}\int\limits_{0}^{\infty} dX\,e^{-\beta X}$$

$$= \sqrt{\frac{2\pi}{\beta}}\cdot\frac{1}{\beta}\cdot\frac{1}{\beta} = \sqrt{2\pi}\,\beta^{-5/2} \quad .$$
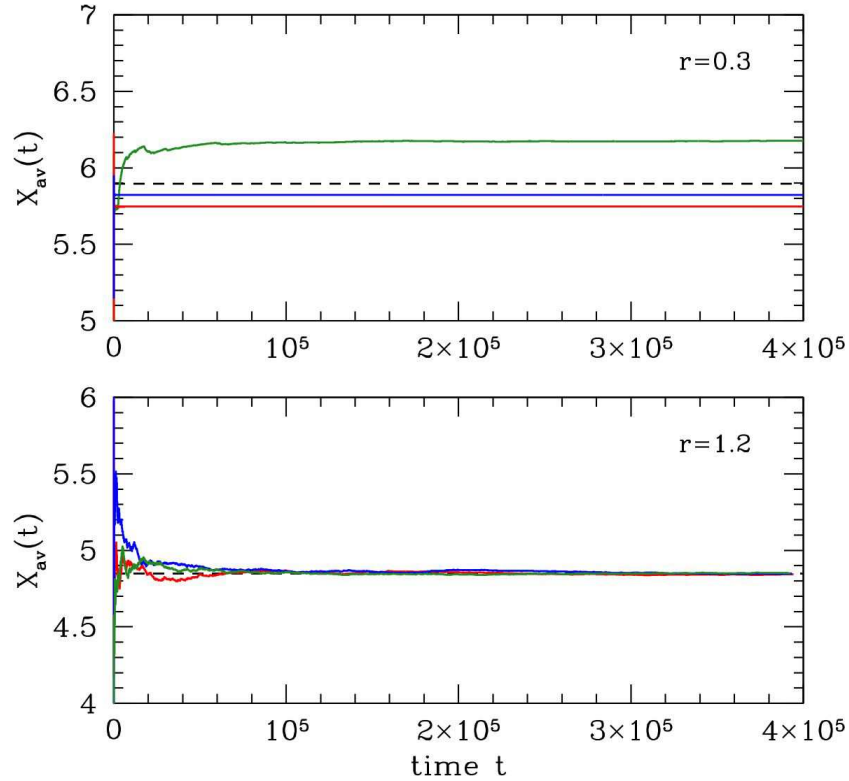
Figure 3.12: Long time running numerical averages $X_{\mathrm{av}}(t) \equiv t^{-1} \int_0^t dt'\, X(t')$ for $r = 0.3$ (top) and $r = 1.2$ (bottom), each for three different initial conditions, with $E = 10$ in all cases. Note how in the $r = 0.3$ case the long time average is dependent on the initial condition, while the $r = 1.2$ case is ergodic and hence independent of initial conditions. The dashed black line shows the restricted microcanonical average, $\langle X \rangle_{\mu\mathrm{ce}} = \frac{(2+r)}{(1+r)} \cdot \frac{1}{3}E$.

Thus, $N(E) = \frac{4\sqrt{2}}{3}\, E^{3/2}$ and

$$\langle \gamma \rangle = \frac{N(E)}{D(E)} = \tfrac{4\sqrt{2}}{3\pi}\left(\frac{1+r}{\sqrt{r}}\right) E^{-1/2} \quad .$$

(d) When the constraint $X \geq x$ is removed, we integrate over all phase space. We then have

$$\widehat{D}(\beta) = \mathsf{Tr}\; e^{-\beta H}$$

$$= \int\limits_{-\infty}^{\infty} dP\, e^{-\beta P^2/2} \int\limits_{-\infty}^{\infty} dp\, e^{-\beta p^2/2r} \int\limits_{0}^{\infty} dX\, e^{-\beta X} \int\limits_{0}^{\infty} dx\, e^{-\beta r x} = \frac{2\pi\sqrt{r}}{\beta^3} \quad .$$

For part (b) we would then have

$$\widehat{N}(\beta) = \mathsf{Tr}\; X\, e^{-\beta H}$$

$$= \int\limits_{-\infty}^{\infty} dP\, e^{-\beta P^2/2} \int\limits_{-\infty}^{\infty} dp\, e^{-\beta p^2/2r} \int\limits_{0}^{\infty} dX\, X\, e^{-\beta X} \int\limits_{0}^{\infty} dx\, e^{-\beta r x} = \frac{2\pi\sqrt{r}}{\beta^4} \quad .$$

r	$X(0)$	$\langle X(t)\rangle$	$\langle X\rangle_{\mu ce}$	$\langle\gamma(t)\rangle$	$\langle\gamma\rangle_{\mu ce}$	r	$X(0)$	$\langle X(t)\rangle$	$\langle X\rangle_{\mu ce}$	$\langle\gamma(t)\rangle$	$\langle\gamma\rangle_{\mu ce}$
0.3	0.1	6.1743	5.8974	0.5283	0.4505	1.2	0.1	4.8509	4.8545	0.3816	0.3812
0.3	1.0	5.7303	5.8974	0.4170	0.4505	1.2	1.0	4.8479	4.8545	0.3811	0.3812
0.3	3.0	5.7876	5.8974	0.4217	0.4505	1.2	3.0	4.8493	4.8545	0.3813	0.3812
0.3	5.0	5.8231	5.8974	0.4228	0.4505	1.2	5.0	4.8482	4.8545	0.3813	0.3812
0.3	7.0	5.8227	5.8974	0.4228	0.4505	1.2	7.0	4.8472	4.8545	0.3808	0.3812
0.3	9.0	5.8016	5.8974	0.4234	0.4505	1.2	9.0	4.8466	4.8545	0.3808	0.3812
0.3	9.9	6.1539	5.8974	0.5249	0.4505	1.2	9.9	4.8444	4.8545	0.3807	0.3812

Table 3.1: Comparison of time averages and microcanonical ensemble averages for $r = 0.3$ and $r = 0.9$. Initial conditions are $P(0) = x(0) = 0$, with $X(0)$ given in the table and $E = 10$. Averages were performed over a period extending for $N_b = 10^7$ bounces.

The respective inverse Laplace transforms are $D(E) = \pi\sqrt{r}\,E^2$ and $N(E) = \frac{1}{3}\pi\sqrt{r}\,E^3$. The microcanonical average of X would then be $\langle X\rangle = \frac{1}{3}E$. Using the restricted phase space, we obtained a value which is greater than this by a factor of $(2+r)/(1+r)$. That the restricted average gives a larger value makes good sense, since X is not allowed to descend below x in that case. For part (c), we would obtain the same result for $N(E)$ since $x = 0$ in the average. We would then obtain

$$\langle\gamma\rangle = \tfrac{4\sqrt{2}}{3\pi}\,r^{-1/2}\,E^{-1/2} \quad .$$

The restricted microcanonical average yields a rate which is larger by a factor $1 + r$. Again, it makes good sense that the restricted average should yield a higher rate, since the ball is not allowed to attain a height greater than the instantaneous value of X.

(e) It is straightforward to simulate the dynamics. So long as $0 < x(t) < X(t)$, we have

$$\dot{X} = P \quad , \quad \dot{P} = -1 \quad , \quad \dot{x} = \frac{p}{r} \quad , \quad \dot{p} = -r \quad .$$

Starting at an arbitrary time t_0, these equations are integrated to yield

$$X(t) = X(t_0) + P(t_0)\,(t - t_0) - \tfrac{1}{2}(t - t_0)^2$$
$$P(t) = P(t_0) - (t - t_0)$$
$$x(t) = x(t_0) + \frac{p(t_0)}{r}\,(t - t_0) - \tfrac{1}{2}(t - t_0)^2$$
$$p(t) = p(t_0) - r(t - t_0) \quad .$$

We must stop the evolution when one of two things happens. The first possibility is a bounce at $t = t_b$, meaning $x(t_b) = 0$. The momentum $p(t)$ changes discontinuously at the bounce, with $p(t_b^+) = -p(t_b^-)$, and where $p(t_b^-) < 0$ necessarily. The second possibility is a collision at $t = t_c$, meaning $X(t_c) = x(t_c)$.

r	$X(0)$	N_{b}	$\langle X(t)\rangle$	$\langle X\rangle_{\mu\text{ce}}$	$\langle\gamma(t)\rangle$	$\langle\gamma\rangle_{\mu\text{ce}}$
1.2	7.0	10^4	4.8054892	4.8484848	0.37560388	0.38118510
1.2	7.0	10^5	4.8436969	4.8484848	0.38120356	0.38118510
1.2	7.0	10^6	4.8479414	4.8484848	0.38122778	0.38118510
1.2	7.0	10^7	4.8471686	4.8484848	0.38083749	0.38118510
1.2	7.0	10^8	4.8485825	4.8484848	0.38116282	0.38118510
1.2	7.0	10^9	4.8486682	4.8484848	0.38120259	0.38118510
1.2	1.0	10^9	4.8485381	4.8484848	0.38118069	0.38118510
1.2	9.9	10^9	4.8484886	4.8484848	0.38116295	0.38118510

Table 3.2: Comparison of time averages and microcanonical ensemble averages for $r = 1.2$, with N_{b} ranging from 10^4 to 10^9.

Integrating across the collision, we must conserve both energy and momentum. This means

$$P(t_{\text{c}}^+) = \frac{1-r}{1+r}\,P(t_{\text{c}}^-) + \frac{2}{1+r}\,p(t_{\text{c}}^-)$$

$$p(t_{\text{c}}^+) = \frac{2r}{1+r}\,P(t_{\text{c}}^-) - \frac{1-r}{1+r}\,p(t_{\text{c}}^-)\quad.$$

In the following tables I report on the results of numerical simulations, comparing dynamical averages with (restricted) phase space averages within the microcanonical ensemble. For $r = 0.3$ the microcanonical averages poorly approximate the dynamical averages, and the dynamical averages are dependent on the initial conditions, indicating that the system is not ergodic. For $r = 1.2$, the agreement between dynamical and microcanonical averages generally improves with averaging time. Indeed, it has been shown by N. I. Chernov, *Physica D* **53**, 233 (1991), building on the work of M. P. Wojtkowski, *Comm. Math. Phys.* **126**, 507 (1990) that this system is ergodic for $r > 1$. Wojtkowski also showed that this system is equivalent to the *wedge billiard*, in which a single point particle of mass m bounces inside a two-dimensional wedge-shaped region $\{(\mathsf{x},\mathsf{y})\,|\,\mathsf{x} \geq 0\,,\,\mathsf{y} \geq \mathsf{x}\,\text{ctn}\,\phi\}$ for some fixed angle $\phi = \tan^{-1}\sqrt{\frac{m}{M}}$. To see this, pass to relative ($\mathcal{X}$) and center-of-mass ($\mathcal{Y}$) coordinates,

$$\mathcal{X} = X - x \qquad\qquad \mathcal{P}_x = \frac{mP - Mp}{M + m}$$

$$\mathcal{Y} = \frac{MX + mx}{M + m} \qquad\qquad \mathcal{P}_y = P + p\quad.$$

Then

$$H = \frac{(M+m)\,\mathcal{P}_x^2}{2Mm} + \frac{\mathcal{P}_y^2}{2(M+m)} + (M+m)\,g\mathcal{Y}\quad.$$

There are two constraints. One requires $X \geq x$, *i.e.* $\mathcal{X} \geq 0$. The second requires $x > 0$, *i.e.*

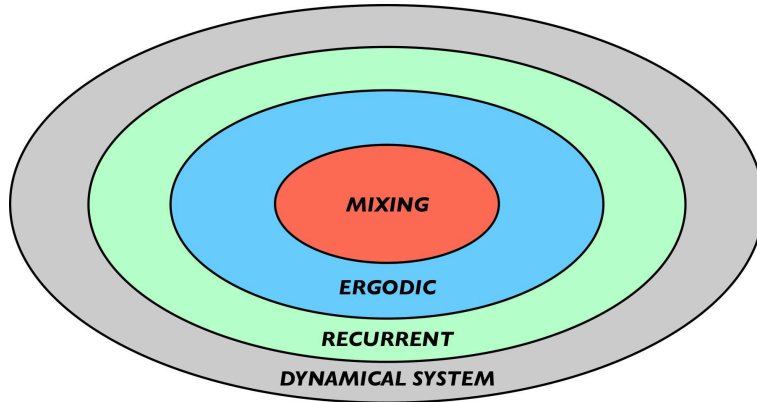$$x = \mathcal{Y} - \frac{M}{M+m}\,\mathcal{X} \geq 0\quad.$$

Figure 3.13: The hierarchy of dynamical systems.

Now define $x \equiv \mathcal{X}$, $p_x \equiv \mathcal{P}_x$, and rescale $y \equiv \frac{M+m}{\sqrt{Mm}} \mathcal{Y}$ and $p_y \equiv \frac{\sqrt{Mm}}{M+m} \mathcal{P}_y$ to obtain

$$H = \frac{1}{2\mu} \left(p_x^2 + p_y^2 \right) + M\, g\, y$$

with $\mu = \frac{Mm}{M+m}$ the familiar reduced mass and $M = \sqrt{Mm}$. The constraints are then $x \geq 0$ and $y \geq \sqrt{\frac{M}{m}}\, x$

3.9 Appendix IV : Thermalization of Quantum Systems

3.9.1 Quantum dephasing

Thermalization of quantum systems is fundamentally different from that of classical systems. Whereas time evolution in classical mechanics is in general a nonlinear dynamical system, the Schrödinger equation for time evolution in quantum mechanics is linear: $i\hbar\partial\Psi/\partial t = \hat{H}\Psi$, where $\hat{H}$ is a many-body Hamiltonian. In classical mechanics, the thermal state is constructed by time evolution – this is the content of the ergodic theorem. In quantum mechanics, as we shall see, the thermal distribution must be encoded in the eigenstates themselves.

Let us assume an initial condition at $t = 0$ with $|\Psi(0)\rangle = \sum_\alpha C_\alpha |\Psi_\alpha\rangle$, where $\{|\Psi_\alpha\rangle\}$ is an orthonormal eigenbasis for $\hat{H}$ satisfying $\hat{H}|\Psi_\alpha\rangle = E_\alpha |\Psi_\alpha\rangle$. The expansion coefficients satisfy $C_\alpha = \langle\Psi_\alpha|\Psi(0)\rangle$ and $\sum_\alpha |C_\alpha|^2 = 1$. Normalization requires $\langle\Psi(0)|\Psi(0)\rangle = \sum_\alpha |C_\alpha|^2 = 1$.

The time evolution of $|\Psi\rangle$ is then given by

$$|\Psi(t)\rangle = \sum_\alpha C_\alpha\, e^{-iE_\alpha t/\hbar} |\Psi_\alpha\rangle \quad . \tag{3.129}$$

The energy is distributed according to the time-independent function

$$P(E) = \langle\Psi(t)|\,\delta(E - \hat{H})\,|\Psi(t)\rangle = \sum_\alpha |C_\alpha|^2\, \delta(E - E_\alpha) \quad . \tag{3.130}$$

Thus, the average energy is time-independent and is given by

$$\langle E \rangle = \langle\, \Psi(t)\, |\, \hat{H}\, |\, \Psi(t)\, \rangle = \int\limits_{-\infty}^{\infty} dE\, P(E)\, E = \sum_{\alpha} |C_{\alpha}|^2\, E_{\alpha} \quad . \tag{3.131}$$

The root mean square fluctuations of the energy are given by

$$(\Delta E)_{\mathrm{rms}} = \left\langle \left(E - \langle E \rangle \right)^2 \right\rangle^{1/2} = \sqrt{ \sum_{\alpha} |C_{\alpha}|^2\, E_{\alpha}^2 - \left(\sum_{\alpha} |C_{\alpha}|^2\, E_{\alpha} \right)^2 } \quad . \tag{3.132}$$

Typically we assume that the distribution $P(E)$ is narrowly peaked about $\langle E \rangle$, such that $(\Delta E)_{\mathrm{rms}} \ll \langle E \rangle - E_0$, where E_0 is the ground state energy. Note that $P(E) = 0$ for $E < E_0$, *i.e.* the eigenspectrum of $\hat{H}$ is bounded from below.

Now consider a general quantum observable described by an operator $\mathcal{A}$. We have

$$\langle \mathcal{A}(t) \rangle = \langle\, \Psi(t)\, |\, \mathcal{A}\, |\, \Psi(t)\, \rangle = \sum_{\alpha,\beta} C_{\alpha}^* \, C_{\beta}\, e^{i(E_{\alpha} - E_{\beta})t/\hbar}\, \mathcal{A}_{\alpha\beta} \quad , \tag{3.133}$$

where $\mathcal{A}_{\alpha\beta} = \langle\, \Psi_{\alpha}\, |\, \mathcal{A}\, |\, \Psi_{\beta}\, \rangle$. In the limit of large times, we have

$$\langle \mathcal{A} \rangle_t \equiv \lim_{T \to \infty} \frac{1}{T} \int\limits_{0}^{T} dt\, \langle \mathcal{A}(t) \rangle = \sum_{\alpha} |C_{\alpha}|^2\, \mathcal{A}_{\alpha\alpha} \quad . \tag{3.134}$$

Note that this implies that all coherence between different eigenstates is lost in the long time limit, due to *dephasing*.

3.9.2   Eigenstate thermalization hypothesis

The essential ideas behind the *eigenstate thermalization hypothesis* (ETH) were described independently by J. Deutsch (1991) and by M. Srednicki (1994). The argument goes as follows. If the total energy is the only conserved quantity, and if $\mathcal{A}$ is a local, translationally-invariant, few-body operator, then the time average $\langle \mathcal{A} \rangle$ is given by its microcanonical value,

$$\langle \mathcal{A} \rangle_t = \sum_{\alpha} |C_{\alpha}|^2\, \mathcal{A}_{\alpha\alpha} = \frac{\sum_{\alpha} \mathcal{A}_{\alpha\alpha}\, \Theta(E_{\alpha} \in I)}{\sum_{\alpha} \Theta(E_{\alpha} \in I)} \equiv \langle \mathcal{A} \rangle_E, \tag{3.135}$$

where $I = \left[E, E + \Delta E \right]$ is an energy interval of width ΔE. So once again, time averages are micro canonical averages.

But how is it that this is the case? The hypothesis of Deutsch and of Srednicki is that thermalization in isolated and bounded quantum systems occurs *at the level of individual eigenstates*. That is, for all eigenstates $|\Psi_{\alpha}\rangle$ with $E_{\alpha} \in I$, one has $\mathcal{A}_{\alpha\alpha} = \langle \mathcal{A} \rangle_{E_{\alpha}}$. This means that *thermal information is encoded in each eigenstate*. This is called the *eigenstate thermalization hypothesis* (ETH).

An equivalent version of the ETH is the following scenario. Suppose we have an infinite or extremely large quantum system U (the 'universe') fixed in an eigenstate $|\Psi_\alpha\rangle$. Then form the projection operator $P_\alpha = |\Psi_\alpha\rangle\langle\Psi_\alpha|$. Projection operators satisfy $P^2 = P$ and their eigenspectrum consists of one eigenvalue 1 and the rest of the eigenvalues are zero[14]. Now consider a partition of $U = W \cup S$, where $W \gg S$. We imagine S to be the 'system' and W the 'world'. We can always decompose the state $|\Psi_\alpha\rangle$ in a complete product basis for W and S, *viz.*

$$|\Psi_\alpha\rangle = \sum_{p=1}^{N_W} \sum_{j=1}^{N_S} \mathcal{Q}_{pj}^\alpha \, |\psi_p^W\rangle \otimes |\psi_j^S\rangle \quad . \tag{3.136}$$

Here $N_{W/S}$ is the size of the basis for W/S. The *reduced density matrix* for S is defined as

$$\rho_S = \underset{W}{\mathrm{Tr}} \, P_\alpha = \sum_{j,j'=1}^{N_S} \left(\sum_{p=1}^{N_W} \mathcal{Q}_{pj}^\alpha \, \mathcal{Q}_{pj'}^{\alpha*} \right) |\psi_j^S\rangle\langle\psi_{j'}^S| \quad . \tag{3.137}$$

The claim is that ρ_S approximates a *thermal density matrix* on S, *i.e.*

$$\rho_S \approx \frac{1}{Z_S} \, e^{-\beta \hat{H}_S} \quad , \tag{3.138}$$

where $\hat{H}_S$ is some Hamiltonian on S, and $Z_S = \mathrm{Tr} \, e^{-\beta\hat{H}_S}$, so that $\mathrm{Tr} \, \rho_S = 1$ and ρ_S is properly normalized. A number of issues remain to be clarified:

(i) What do we mean by "approximates"?

(ii) What do we mean by $\hat{H}_S$?

(iii) What do we mean by the temperature T?

We address these in reverse order. The temperature T of an eigenstate $|\Psi_\alpha\rangle$ of a Hamiltonian $\hat{H}$ is defined by setting its energy density E_α/V_U to the thermal energy density, *i.e.*

$$\frac{E_\alpha}{V} = \frac{1}{V} \frac{\mathrm{Tr} \, \hat{H} \, e^{-\beta\hat{H}}}{\mathrm{Tr} \, e^{-\beta\hat{H}}} \quad . \tag{3.139}$$

Here, $\hat{H} = \hat{H}_U$ is the full Hamiltonian of the universe $U = W \cup S$, and $V = V_U$. Our intuition is that $\hat{H}_S$ should reflect a restriction of the original Hamiltonian $\hat{H}_U$ to the system S. What should be done, though, about the interface parts of $\hat{H}_U$ which link S and W? For lattice Hamiltonians, we can simply but somewhat arbitrarily cut all the bonds coupling S and W. But we could easily imagine some other prescription, such as halving the coupling strength along all such interface bonds. Indeed, the definition of H_S is somewhat arbitrary. However, so long as we use ρ_S to compute averages of *local operators* which lie sufficiently far from the boundary of S, the precise details of how we truncate $\hat{H}_U$ to $\hat{H}_S$ are

---

[14]More generally, we could project onto a K-dimensional subspace, in which case there would be K eigenvalues of $+1$ and $N - K$ eigenvalues of 0, where N is the dimension of the entire vector space.

unimportant. This brings us to the first issue: the approximation of ρ_S by its Gibbs form in eqn. 3.138 is only valid when we consider averages of local operators lying within the bulk of S. This means that we must only examine operators whose support is confined to regions greater than some distance ξ_T from ∂S, where ξ_T is a *thermal correlation length*. This, in turn, requires that $L_S \gg \xi_T$, *i.e.* the region S is very large on the scale of ξ_T. How do we define ξ_T? For a model such as the Ising model, it can be taken to be the usual correlation length obtained from the spin-spin correlation function $\langle \sigma_{\boldsymbol{r}} \sigma_{\boldsymbol{r}'} \rangle_T$. More generally, we may choose the largest correlation length from among the correlators of all the independent local operators in our system. Again, the requirement is that $\exp(-d_{\partial}(\boldsymbol{r})/\xi_T) \ll 1$, where $d_{\partial}(\boldsymbol{r})$ is the shortest distance from the location of our local operator $\mathcal{O}_{\boldsymbol{r}}$ to the boundary of S. At criticality, the exponential is replaced by a power law $(d_{\partial}(\boldsymbol{r})/\xi_T)^{-p}$, where p is a critical exponent. Another implicit assumption here is that $V_S \ll V_W$.

3.9.3 More precise formulation

More precisely (Srednicki, 1999), ETH is formulated in terms of general matrix elements of local observables in the energy eigenbasis, *viz.*[15]

$$A_{mn} = \langle m \,|\, \hat{A} \,|\, n \rangle = A(E)\,\delta_{mn} + e^{-S(E)/2}\,f_A(E,\omega)\,R_{mn} \tag{3.140}$$

where $E = \frac{1}{2}(E_m + E_n)$ is an average of the energy eigenvalues, $\omega = E_m - E_n$ is their difference, R_{mn} is a random matrix with $\langle R_{mn} \rangle = 0$ and $\mathsf{var}(R_{mn}) = 1$, $S(E)$ is the thermodynamic entropy, with $S(E) \sim \log D_{\mathcal{H}}$ when E lies in the middle of the spectrum ($D_{\mathcal{H}}$ is the Hilbert space dimension), and $A(E)$ and $f(E,\omega)$ are smooth functions of their arguments. Additionally, one has

$$\begin{aligned} A_{mn} \in \mathbb{R} &\quad : \quad R_{nm} = R_{mn} \quad, \quad f_A(E,-\omega) = f_A(E,\omega) \\ A_{mn} \in \mathbb{C} &\quad : \quad R_{nm} = R^*_{mn} \quad, \quad f_A(E,-\omega) = f^*_A(E,\omega) \quad, \end{aligned} \tag{3.141}$$

for systems with and without time-reversal symmetry, respectively.

ETH reduces to random matrix theory within a given small energy window. Eqn. 3.140 is to be contrasted with the corresponding result for random matrix theory,

$$A_{mn} = \langle \hat{A} \rangle \, \delta_{mn} + \sqrt{\frac{\langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2}{D_{\mathcal{H}}}}\, R_{mn} \quad, \tag{3.142}$$

where $\langle \hat{A} \rangle \equiv \mathsf{Tr}\,\hat{A}$ and $\langle \hat{A}^2 \rangle \equiv \mathsf{Tr}\,(\hat{A}^2)$ so that the averages are over the entire spectrum. ETH thus reduces to RMT within any small window which contains $\mathcal{O}(D_{\mathcal{H}})$ states.

A consequence of ETH is that the expectation value of an operator in a Gibbs state $\rho = Z^{-1}\exp(-\beta H)$ can be replaced by its expectation value in any eigenstate whose energy eigenvalue corresponds to the peak in $g(E)\exp(-\beta E)$, where $g(E)$ is the density of states. Such an eigenstate is then called a *representative pure state*[16]. To see this, note

$$\mathsf{Tr}\,(\rho \hat{A}) \approx \frac{1}{Z}\int dE\, g(E)\, e^{-\beta E}\, A(E) \approx A(E^*) \quad . \tag{3.143}$$

---

[15]See the review article by L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, *Adv. Phys.*, **65**, 239 (2016) and references therein.

[16]See V. Khemani, A. Chandran, H. Kim, and S. L. Sondhi, *Phys. Rev. E* **90**, 052133 (2014).

Here it is assumed that the function $A(E)$ in Eqn. 3.140 is a smooth function of E, and that the distribution $g(E)\exp(-\beta E)$ has a narrow peak centered at $E = E^*$. For example, if

$$g(E) \sim \exp\left(-\frac{(E - E_{\mathrm{c}})^2}{2\sigma^2}\right) \quad , \tag{3.144}$$

where the energy variance is extensive, *i.e.* $\sigma^2 \approx wN$ with w a constant and N the total number of particles or the system volume in microscopic units, then $E^*(\beta) \approx E_{\mathrm{c}} - \beta\sigma^2$ and the energy density is

$$\varepsilon^*(\beta) \equiv \frac{E^*}{N} \approx \varepsilon_{\mathrm{c}} - \beta w \quad , \tag{3.145}$$

where ε_{c} is the energy density at the center of the spectrum. Eigenstates of H in the vicinity of energy density $\varepsilon^*(\beta)$ are thus representative pure states of the Gibbs density matrix which reproduce expectation values of few body operators. Note that we may obtain $\varepsilon^*(\beta)$ from the expression

$$\varepsilon^*(\beta) = \frac{1}{NZ}\,\mathsf{Tr}\,(\hat{H}e^{-\beta\hat{H}}) \quad . \tag{3.146}$$

3.9.4 When is the ETH true?

There is no rigorous proof of the ETH. Deutsch showed that the ETH holds for the case of an integrable Hamiltonian weakly perturbed by a single Gaussian random matrix. Horoi *et al.* (1995) showed that nuclear shell model wavefunctions reproduce thermodynamic predictions. Recent numerical work by M. Rigol and collaborators has verified the applicability of the ETH in small interacting boson systems. ETH fails for so-called integrable models, where there are a large number of conserved quantities, which commute with the Hamiltonian. Integrable models are, however, quite special, and as Deutsch showed, integrability is spoiled by weak perturbations, in which case ETH then applies.

ETH also fails in the case of noninteracting disordered systems which exhibit *Anderson localization*. Single particle energy eigenstates ψ_j whose energies ε_j the localized portion of the eigenspectrum decay exponentially, as $|\psi_j(\boldsymbol{r})|^2 \sim \exp(-|\boldsymbol{r} - \boldsymbol{r}_j|/\xi(\varepsilon_j))$, where $\boldsymbol{r}_j$ is some position in space associated with ψ_j and $\xi(\varepsilon_j)$ is the *localization length*. Within the localized portion of the spectrum, $\xi(\varepsilon)$ is finite. As ε approaches a *mobility edge*, $\xi(\varepsilon)$ diverges as a power law. In the *delocalized regime*, eigenstates are spatially *extended* and typically decay at worst as a power law[17]. Exponentially localized states are unable to thermalize with other distantly removed localized states. Of course, all noninteracting systems will violate ETH, because they are integrable. The interacting version of this phenomenon, *many-body localization* (MBL), is a topic of intense current interest in condensed matter and statistical physics. MBL systems also exhibit a large number of conserved quantities, but in contrast to the case of integrable systems, where each conserved quantity is in general expressed in terms of an integral of a local density, in MBL systems the conserved quantities are themselves local, although *emergent*. The emergent nature of locally conserved quantities in MBL systems means that they are not simply expressed in terms of the original local operators of the system, but rather are arrived at via a sequence of local unitary transformations.

Note again that in contrast to the classical case, time evolution of a quantum state does not create the thermal state. Rather, it *reveals* the thermal distribution which is encoded in all eigenstates after sufficient

---

[17]Recall that in systems with no disorder, eigenstates exhibit Bloch periodicity in space.

time for dephasing to occur, so that correlations between all the wavefunction expansion coefficients $\{C_\alpha\}$ for $\alpha \neq \alpha'$ are all lost.