

Chapter 6

Ergodicity and the Approach to Equilibrium

6.1 Equilibrium

Recall that a thermodynamic system is one containing an enormously large number of constituent particles, a typical ‘large number’ being Avogadro’s number, $N_A = 6.02 \times 10^{23}$. 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 path¹. The equilibrium state is time-independent.

6.2 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

$$\frac{dP_i}{dt} = \sum_j (W_{ji} P_j - W_{ij} P_i) . \quad (6.1)$$

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

$$\frac{dP_i}{dt} = - \sum_j \Gamma_{ij} P_j , \quad (6.2)$$

¹Exceptions 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*.

where

$$\Gamma_{ij} = \begin{cases} -W_{ji} & \text{if } i \neq j \\ \sum'_k W_{jk} & \text{if } i = j, \end{cases} \quad (6.3)$$

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

$$W_{ji} = \frac{2\pi}{\hbar} |\langle i | \hat{V} | j \rangle|^2 \rho(E_i), \quad (6.4)$$

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 .

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

$$P_i(t) = (e^{-\Gamma t})_{ij} P_j(0). \quad (6.5)$$

Here we have used the Einstein 'summation convention' in which repeated indices are summed over (in this case, the j index). Note that

$$\sum_i \Gamma_{ij} = 0, \quad (6.6)$$

which says that the total probability $\sum_i P_i$ is conserved:

$$\frac{d}{dt} \sum_i P_i = - \sum_{i,j} \Gamma_{ij} P_j = - \sum_j \left(\sum_i \Gamma_{ij} \right) P_j = 0. \quad (6.7)$$

Suppose we have a time-independent solution to the master equation, P_i^{eq} . Then we must have

$$\Gamma_{ij} P_j^{\text{eq}} = 0 \quad \implies \quad P_j^{\text{eq}} W_{ji} = P_i^{\text{eq}} W_{ij}. \quad (6.8)$$

This is called the condition of *detailed balance*. Assuming $W_{ij} \neq 0$ and $P_j^{\text{eq}} \neq 0$, we can divide to obtain

$$\frac{W_{ji}}{W_{ij}} = \frac{P_i^{\text{eq}}}{P_j^{\text{eq}}}. \quad (6.9)$$

6.2.1 Example: 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

$$W_{nm} = \begin{cases} 0 & \text{if } m \geq n \\ n\gamma & \text{if } m = n - 1 \\ 0 & \text{if } m < n - 1. \end{cases} \quad (6.10)$$

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

$$\frac{dP_n}{dt} = (n+1)\gamma P_{n+1} - n\gamma P_n . \quad (6.11)$$

The interpretation here is as follows: let $|n\rangle$ denote a state in which n atoms are excited. Then $P_n(t) = |\langle \psi(t) | n \rangle|^2$. 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

$$N(t) = \sum_{n=0}^{\infty} n P_n(t) . \quad (6.12)$$

Note that

$$\begin{aligned} \frac{dN}{dt} &= \sum_{n=0}^{\infty} n \left[(n+1)\gamma P_{n+1} - n\gamma P_n \right] \\ &= \gamma \sum_{n=0}^{\infty} \left[n(n-1) P_n - n^2 P_n \right] \\ &= -\gamma \sum_{n=0}^{\infty} n P_n = -\gamma N . \end{aligned} \quad (6.13)$$

Thus,

$$N(t) = N(0) e^{-\gamma t} . \quad (6.14)$$

The relaxation time is $\tau = \gamma^{-1}$, and the equilibrium distribution is

$$P_n^{\text{eq}} = \delta_{n,0} . \quad (6.15)$$

Note that this satisfies detailed balance.

We can go a bit farther here. Let us define

$$P(z, t) \equiv \sum_{n=0}^{\infty} z^n P_n(t) . \quad (6.16)$$

This is sometimes called a *generating function*. Then

$$\begin{aligned} \frac{\partial P}{\partial t} &= \gamma \sum_{n=0}^{\infty} z^n \left[(n+1) P_{n+1} - n P_n \right] \\ &= \gamma \frac{\partial P}{\partial z} - \gamma z \frac{\partial P}{\partial z} . \end{aligned} \quad (6.17)$$

Thus,

$$\frac{1}{\gamma} \frac{\partial P}{\partial t} - (1-z) \frac{\partial P}{\partial z} = 0 . \quad (6.18)$$

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

$$P(z, t) = f(\gamma t - \ln(1 - z)) . \quad (6.19)$$

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

$$P(z, t) = P(1 + (z - 1)e^{-\gamma t}, 0) . \quad (6.20)$$

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

$$N(t) = \sum_{n=0}^{\infty} n P_n(t) = \left. \frac{\partial P}{\partial z} \right|_{z=1} = e^{-\gamma t} P(1, 0) = N(0) e^{-\gamma t} . \quad (6.21)$$

6.2.2 Decomposition of Γ_{ij}

The matrix Γ_{ij} is real but not necessarily symmetric. For such a matrix, the left eigenvectors ϕ_i^α and the right eigenvectors ψ_j^β are not the same: general different:

$$\phi_i^\alpha \Gamma_{ij} = \lambda_\alpha \phi_j^\alpha \quad (6.22)$$

$$\Gamma_{ij} \psi_j^\beta = \lambda_\beta \psi_i^\beta . \quad (6.23)$$

Note that the eigenvalue equation for the right eigenvectors is $\Gamma\psi = \lambda\psi$ while that for the left eigenvectors is $\Gamma^t\phi = \lambda\phi$. The characteristic polynomial is the same in both cases:

$$\det(\lambda - \Gamma) = \det(\lambda - \Gamma^t) , \quad (6.24)$$

which means that the left and right eigenvalues are the same. Multiplying the eigenvector equation for ϕ^α on the right by ψ_j^β and summing over j , and multiplying the eigenvector equation for ψ^β on the left by ϕ_i^α and summing over i , and subtracting the two results yields

$$(\lambda_\alpha - \lambda_\beta) \langle \phi^\alpha | \psi^\beta \rangle = 0 , \quad (6.25)$$

where the inner product is

$$\langle \phi | \psi \rangle = \sum_i \phi_i \psi_i . \quad (6.26)$$

We can now demand

$$\langle \phi^\alpha | \psi^\beta \rangle = \delta_{\alpha\beta} , \quad (6.27)$$

in which case we can write

$$\Gamma = \sum_{\alpha} \lambda_{\alpha} | \psi^{\alpha} \rangle \langle \phi^{\alpha} | \iff \Gamma_{ij} = \sum_{\alpha} \lambda_{\alpha} \psi_i^{\alpha} \phi_j^{\alpha} . \quad (6.28)$$

We note that $\vec{\phi} = (1, 1, \dots, 1)$ is a left eigenvector 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} . Now let's expand $P_i(t)$ in the right eigenvectors of Γ , writing

$$P_i(t) = \sum_{\alpha} C_{\alpha}(t) \psi_i^{\alpha} . \quad (6.29)$$

Then

$$\begin{aligned}
\frac{dP_i}{dt} &= \sum_{\alpha} \frac{dC_{\alpha}}{dt} \psi_i^{\alpha} \\
&= -\Gamma_{ij} P_j = -\sum_{\alpha} C_{\alpha} \Gamma_{ij} \psi_j^{\alpha} \\
&= -\sum_{\alpha} \lambda_{\alpha} C_{\alpha} \psi_i^{\alpha} .
\end{aligned} \tag{6.30}$$

This allows us to write

$$\frac{dC_{\alpha}}{dt} = -\lambda_{\alpha} C_{\alpha} \quad \Longrightarrow \quad C_{\alpha}(t) = C_{\alpha}(0) e^{-\lambda_{\alpha} t} . \tag{6.31}$$

Hence, we can write

$$P_i(t) = \sum_{\alpha} C_{\alpha}(0) e^{-\lambda_{\alpha} t} \psi_i^{\alpha} . \tag{6.32}$$

It is now easy to see that $\text{Re}(\lambda_{\alpha}) \geq 0$ for all λ , or else the probabilities will become negative. For suppose $\text{Re}(\lambda_{\alpha}) < 0$ for some α . Then as $t \rightarrow \infty$, the sum in eqn. 6.32 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 $\sum_i \psi_i^{\alpha} = 0$ since $|\psi^{\alpha}\rangle$ must be orthogonal to the left eigenvector $\vec{\phi}^{\alpha=0} = (1, 1, \dots, 1)$. The eigenvectors can also be chosen to be real, since Γ_{ij} itself is real. Therefore, at least one component of ψ_i^{α} (*i.e.* for some value of i) must be negative, which means a negative probability!

We conclude that $P_i(t) \rightarrow P_i^{\text{eq}}$ as $t \rightarrow \infty$, relaxing to the $\lambda = 0$ right eigenvector, with $\text{Re}(\lambda_{\alpha}) \geq 0$ for all α .

6.3 Boltzmann's H-theorem

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

$$H(t) = \sum_i P_i(t) \ln P_i(t) . \tag{6.33}$$

Then

$$\begin{aligned}
\frac{dH}{dt} &= \sum_i \frac{dP_i}{dt} (1 + \ln P_i) = \sum_i \frac{dP_i}{dt} \ln P_i \\
&= -\sum_{i,j} \Gamma_{ij} P_j \ln P_i \\
&= \sum_{i,j} \Gamma_{ij} P_j (\ln P_j - \ln P_i) ,
\end{aligned} \tag{6.34}$$

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

$$\frac{dH}{dt} = \frac{1}{2} \sum_{i,j} \Gamma_{ij} (P_i - P_j) (\ln P_i - \ln P_j) . \quad (6.35)$$

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

$$(x - y) (\ln x - \ln y) \geq 0 , \quad (6.36)$$

we conclude

$$\frac{dH}{dt} \leq 0 . \quad (6.37)$$

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

$$P_i^{\text{eq}} = \frac{1}{\Omega} , \quad \Omega = \sum_i 1 \implies H = -\ln \Omega . \quad (6.38)$$

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

$$\bar{W}_{ij} \equiv P_i^{\text{eq}} W_{ij} = P_j^{\text{eq}} W_{ji} = \bar{W}_{ji} , \quad (6.39)$$

and the generalized H -function,

$$H(t) \equiv \sum_i P_i(t) \ln \left(\frac{P_i(t)}{P_i^{\text{eq}}} \right) . \quad (6.40)$$

Then

$$\frac{dH}{dt} = -\frac{1}{2} \sum_{i,j} \bar{W}_{ij} \left(\frac{P_i}{P_i^{\text{eq}}} - \frac{P_j}{P_j^{\text{eq}}} \right) \left[\ln \left(\frac{P_i}{P_i^{\text{eq}}} \right) - \ln \left(\frac{P_j}{P_j^{\text{eq}}} \right) \right] \leq 0 . \quad (6.41)$$

6.4 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 $S[q(t)] = \int dt L$ are

$$\dot{p}_\sigma = \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_\sigma} \right) = \frac{\partial L}{\partial q_\sigma} , \quad (6.42)$$

where p_σ is the canonical momentum conjugate to the generalized coordinate q_σ :

$$p_\sigma = \frac{\partial L}{\partial \dot{q}_\sigma} . \quad (6.43)$$

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

$$H(q, p) = \sum_{\sigma=1}^r p_{\sigma} \dot{q}_{\sigma} - L . \quad (6.44)$$

Note that

$$\begin{aligned} dH &= \sum_{\sigma=1}^r \left(p_{\sigma} dq_{\sigma} + \dot{q}_{\sigma} dp_{\sigma} - \frac{\partial L}{\partial q_{\sigma}} dq_{\sigma} - \frac{\partial L}{\partial \dot{q}_{\sigma}} d\dot{q}_{\sigma} \right) - \frac{\partial L}{\partial t} dt \\ &= \sum_{\sigma=1}^r \left(\dot{q}_{\sigma} dp_{\sigma} - \frac{\partial L}{\partial q_{\sigma}} dq_{\sigma} \right) - \frac{\partial L}{\partial t} dt . \end{aligned} \quad (6.45)$$

Thus, we obtain Hamilton's equations of motion,

$$\frac{\partial H}{\partial p_{\sigma}} = \dot{q}_{\sigma} \quad , \quad \frac{\partial H}{\partial q_{\sigma}} = -\frac{\partial L}{\partial q_{\sigma}} = -\dot{p}_{\sigma} \quad (6.46)$$

and

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} . \quad (6.47)$$

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

$$\varphi_i = \begin{cases} q_i & \text{if } 1 \leq i \leq r \\ p_{i-r} & \text{if } r \leq i \leq 2r . \end{cases} \quad (6.48)$$

Then we may write Hamilton's equations compactly as

$$\dot{\varphi}_i = J_{ij} \frac{\partial H}{\partial \varphi_j} , \quad (6.49)$$

where

$$J = \begin{pmatrix} 0_{r \times r} & 1_{r \times r} \\ -1_{r \times r} & 0_{r \times r} \end{pmatrix} \quad (6.50)$$

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

6.5 Evolution of Phase Space Volumes

Consider a general dynamical system,

$$\frac{d\varphi}{dt} = V(\varphi) , \quad (6.51)$$

where $\varphi(t)$ is a point in an n -dimensional phase space. Consider now a compact² region \mathcal{R}_0 in phase space, and consider its evolution under the dynamics. That is, \mathcal{R}_0 consists of

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

a set of points $\{\varphi \mid \varphi \in \mathcal{R}_0\}$, and if we regard each $\varphi \in \mathcal{R}_0$ as an initial condition, we can define the time-dependent set $\mathcal{R}(t)$ as the set of points $\varphi(t)$ that *were* in \mathcal{R}_0 at time $t = 0$:

$$\mathcal{R}(t) = \{\varphi(t) \mid \varphi(0) \in \mathcal{R}_0\} . \quad (6.52)$$

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

$$\Omega(t) = \int_{\mathcal{R}(t)} d\mu \quad (6.53)$$

where

$$d\mu = d\varphi_1 d\varphi_2 \cdots d\varphi_n , \quad (6.54)$$

for an n -dimensional phase space. We then have

$$\Omega(t + dt) = \int_{\mathcal{R}(t+dt)} d\mu' = \int_{\mathcal{R}(t)} d\mu \left| \frac{\partial\varphi_i(t + dt)}{\partial\varphi_j(t)} \right| , \quad (6.55)$$

where

$$\left| \frac{\partial\varphi_i(t + dt)}{\partial\varphi_j(t)} \right| \equiv \frac{\partial(\varphi'_1, \dots, \varphi'_n)}{\partial(\varphi_1, \dots, \varphi_n)} \quad (6.56)$$

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

$$\varphi_i(t + dt) = \varphi_i(t) + V_i(\varphi(t)) dt + \mathcal{O}(dt^2) \quad (6.57)$$

and therefore

$$\frac{\partial\varphi_i(t + dt)}{\partial\varphi_j(t)} = \delta_{ij} + \frac{\partial V_i}{\partial\varphi_j} dt + \mathcal{O}(dt^2) . \quad (6.58)$$

We now make use of the equality

$$\ln \det M = \text{Tr} \ln M , \quad (6.59)$$

for any matrix M , which gives us³, for small ε ,

$$\det(1 + \varepsilon A) = \exp \text{Tr} \ln(1 + \varepsilon A) = 1 + \varepsilon \text{Tr} A + \frac{1}{2} \varepsilon^2 \left((\text{Tr} A)^2 - \text{Tr}(A^2) \right) + \dots \quad (6.60)$$

Thus,

$$\Omega(t + dt) = \Omega(t) + \int_{\mathcal{R}(t)} d\mu \nabla \cdot \mathbf{V} dt + \mathcal{O}(dt^2) , \quad (6.61)$$

which says

$$\frac{d\Omega}{dt} = \int_{\mathcal{R}(t)} d\mu \nabla \cdot \mathbf{V} = \int_{\partial\mathcal{R}(t)} dS \hat{\mathbf{n}} \cdot \mathbf{V} \quad (6.62)$$

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

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

$$\nabla \cdot \mathbf{V} = \sum_{i=1}^n \frac{\partial V_i}{\partial \varphi_i}, \quad (6.63)$$

and we have used Stokes' theorem to convert the volume integral of the divergence to a surface integral of $\hat{\mathbf{n}} \cdot \mathbf{V}$, where $\hat{\mathbf{n}}$ 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 $\nabla \cdot \mathbf{V} = 0$ 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,

$$\frac{\partial \varrho}{\partial t} + \nabla \cdot (\varrho \mathbf{V}) = 0. \quad (6.64)$$

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

$$\frac{d}{dt} \int_{\mathcal{R}} d\mu \varrho = - \int_{\mathcal{R}} d\mu \nabla \cdot (\varrho \mathbf{V}) = - \int_{\partial\mathcal{R}} dS \hat{\mathbf{n}} \cdot (\varrho \mathbf{V}). \quad (6.65)$$

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

$$\frac{dQ_{\mathcal{R}}}{dt} = - \int_{\partial\mathcal{R}} dS \hat{\mathbf{n}} \cdot \mathbf{J}, \quad (6.66)$$

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

$$\frac{\partial \varrho}{\partial t} + \mathbf{V} \cdot \nabla \varrho + \varrho \nabla \cdot \mathbf{V} = 0. \quad (6.67)$$

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

$$\frac{D\varrho}{Dt} \equiv \left(\frac{\partial}{\partial t} + \mathbf{V} \cdot \nabla \right) \varrho = 0. \quad (6.68)$$

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

$$\begin{aligned} \frac{d}{dt} \varrho(\boldsymbol{\varphi}(t), t) &= \frac{\partial \varrho}{\partial \varphi_i} \frac{d\varphi_i}{dt} + \frac{\partial \varrho}{\partial t} \\ &= \sum_{i=1}^n V_i \frac{\partial \varrho}{\partial \varphi_i} + \frac{\partial \varrho}{\partial t} = \frac{D\varrho}{Dt}. \end{aligned} \quad (6.69)$$

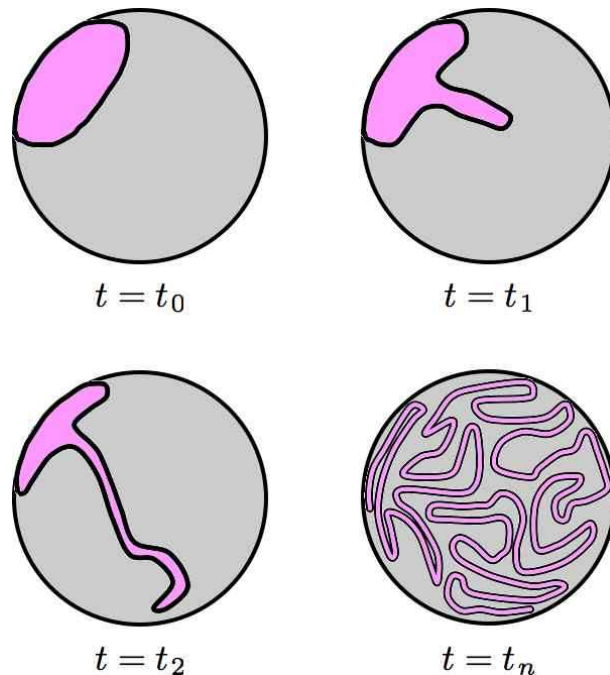


Figure 6.1: Time evolution of two immiscible fluids. The local density remains constant.

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

$$\varrho(\varphi, t = 0) = \begin{cases} 1 & \text{if } \varphi \in \mathcal{R}_0 \\ 0 & \text{otherwise} \end{cases} \quad (6.70)$$

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

$$\dot{q}_i = + \frac{\partial H}{\partial p_i} \quad , \quad \dot{p}_i = - \frac{\partial H}{\partial q_i} \quad (6.71)$$

A point in phase space is specified by r positions q_i and r momenta p_i , hence the dimension of phase space is $n = 2r$:

$$\varphi = \begin{pmatrix} \mathbf{q} \\ \mathbf{p} \end{pmatrix} \quad , \quad \mathbf{V} = \begin{pmatrix} \dot{\mathbf{q}} \\ \dot{\mathbf{p}} \end{pmatrix} = \begin{pmatrix} \partial H / \partial \mathbf{p} \\ -\partial H / \partial \mathbf{q} \end{pmatrix} . \quad (6.72)$$

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

$$\begin{aligned} \nabla \cdot \mathbf{V} &= \sum_{i=1}^r \left\{ \frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} \right\} \\ &= \sum_{i=1}^r \left\{ \frac{\partial}{\partial q_i} \left(\frac{\partial H}{\partial p_i} \right) + \frac{\partial}{\partial p_i} \left(- \frac{\partial H}{\partial q_i} \right) \right\} = 0 . \end{aligned} \quad (6.73)$$

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

$$\frac{D\varrho}{Dt} \equiv \frac{\partial\varrho}{\partial t} + \mathbf{V} \cdot \nabla\varrho = 0, \quad (6.74)$$

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.

6.5.1 Liouville's Equation

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

$$\begin{aligned} \frac{\partial\varrho}{\partial t} &= -\dot{\boldsymbol{\varphi}} \cdot \nabla\varrho = -\sum_{k=1}^r \left(\dot{q}_k \frac{\partial}{\partial q_k} + \dot{p}_k \frac{\partial}{\partial p_k} \right) \varrho \\ &= -i\hat{L}\varrho, \end{aligned} \quad (6.75)$$

where \hat{L} is a differential operator known as the *Liouillian*:

$$\hat{L} = -i \sum_{k=1}^r \left\{ \frac{\partial H}{\partial p_k} \frac{\partial}{\partial q_k} - \frac{\partial H}{\partial q_k} \frac{\partial}{\partial p_k} \right\}. \quad (6.76)$$

Eqn. 6.75, known as *Liouville's equation*, bears an obvious resemblance to the Schrödinger equation from quantum mechanics.

Suppose that $\Lambda_a(\boldsymbol{\varphi})$ is conserved by the dynamics of the system. Typical conserved quantities 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 the Lagrangian is not explicitly time-dependent). Now consider a distribution $\varrho(\boldsymbol{\varphi}, t) = \varrho(\Lambda_1, \Lambda_2, \dots, \Lambda_k)$ which is a function only of these various conserved quantities. Then from the chain rule, we have

$$\dot{\boldsymbol{\varphi}} \cdot \nabla\varrho = \sum_a \frac{\partial\varrho}{\partial\Lambda_a} \dot{\boldsymbol{\varphi}} \cdot \nabla\Lambda_a = 0, \quad (6.77)$$

since for each a we have

$$\frac{d\Lambda_a}{dt} = \sum_{\sigma=1}^r \left(\frac{\partial\Lambda_a}{\partial q_\sigma} \dot{q}_\sigma + \frac{\partial\Lambda_a}{\partial p_\sigma} \dot{p}_\sigma \right) = \dot{\boldsymbol{\varphi}} \cdot \nabla\Lambda_a = 0. \quad (6.78)$$

We conclude that any distribution $\varrho(\boldsymbol{\varphi}, t) = \varrho(\Lambda_1, \Lambda_2, \dots, \Lambda_k)$ which is a function solely of conserved dynamical quantities is a stationary solution to Liouville's equation.

Clearly the microcanonical distribution,

$$\varrho_E(\boldsymbol{\varphi}) = \frac{\delta(E - H(\boldsymbol{\varphi}))}{\Sigma(E)} = \frac{\delta(E - H(\boldsymbol{\varphi}))}{\int d\mu \delta(E - H(\boldsymbol{\varphi}))}, \quad (6.79)$$

is a fixed point solution of Liouville's equation.

6.6 Irreversibility and Poincaré Recurrence

The dynamics of the master equation describe an approach to equilibrium. These dynamics are irreversible: $(dH/dt) \leq 0$. However, the microscopic laws of physics are (almost) time-reversal invariant⁴, 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.

6.6.1 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 the 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.*

$$\int d\mu \delta(E - H(\mathbf{q}, \mathbf{p})) < \infty, \quad (6.80)$$

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

Theorem: In any finite neighborhood \mathcal{R}_0 of phase space there exists a point φ_0 which will return to \mathcal{R}_0 after m applications of g_τ , where m is finite.

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 :

$$\Upsilon = \bigcup_{k=0}^{\infty} g_\tau^k \mathcal{R}_0 \quad (6.81)$$

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

$$\begin{aligned} \text{vol}(\Upsilon) &= \sum_{k=0}^{\infty} \text{vol}(g_\tau^k \mathcal{R}_0) \\ &= \text{vol}(\mathcal{R}_0) \cdot \sum_{k=0}^{\infty} 1 = \infty, \end{aligned} \quad (6.82)$$

since $\text{vol}(g_\tau^k \mathcal{R}_0) = \text{vol}(\mathcal{R}_0)$ 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.

⁴Actually, 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.

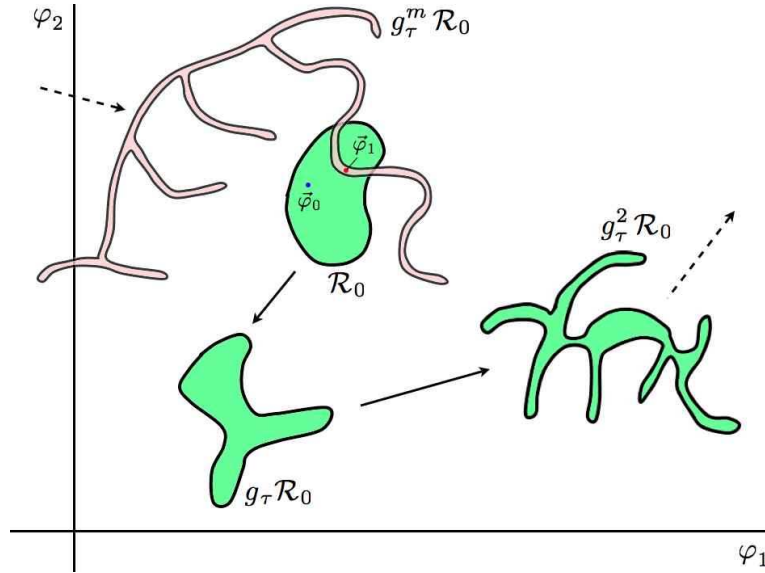


Figure 6.2: Successive images of a set \mathcal{R}_0 under the τ -advance mapping g_τ , projected onto a two-dimensional phase plane. The Poincaré recurrence theorem guarantees that if phase space has finite volume, and g_τ is invertible and volume preserving, then for any set \mathcal{R}_0 there exists an integer m such that $\mathcal{R}_0 \cap g_\tau^m \mathcal{R}_0 \neq \emptyset$.

Thus, the assumption that the set $\{g_\tau^k \mathcal{R}_0 \mid k \in \mathbb{Z}_+\}$ 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 $\varphi_1 \in g_\tau^m \mathcal{R}_0 \cap \mathcal{R}_0$, where $m = l - k$, and define $\varphi_0 = g_\tau^{-m} \varphi_1$. Then by construction both φ_0 and $g_\tau^m \varphi_0$ 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. 6.3. The perfume molecules evolve according to 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 $T(\mathbf{p}) > E_{\text{TOT}}$, where $T(\mathbf{p})$ is the single particle kinetic energy function⁵. 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! In this case, we could define the region \mathcal{R}_0 as

$$\mathcal{R}_0 = \{(q_1, \dots, q_r, p_1, \dots, p_r) \mid |q_i - q_i^0| \leq \Delta q \text{ and } |p_j - p_j^0| \leq \Delta p \forall i, j\}, \quad (6.83)$$

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

Each of the three central assumptions – finite phase space, invertibility, and volume preser-

⁵In the nonrelativistic limit, $T = p^2/2m$. For relativistic particles, we have $T = (p^2 c^2 + m^2 c^4)^{1/2} - mc^2$.

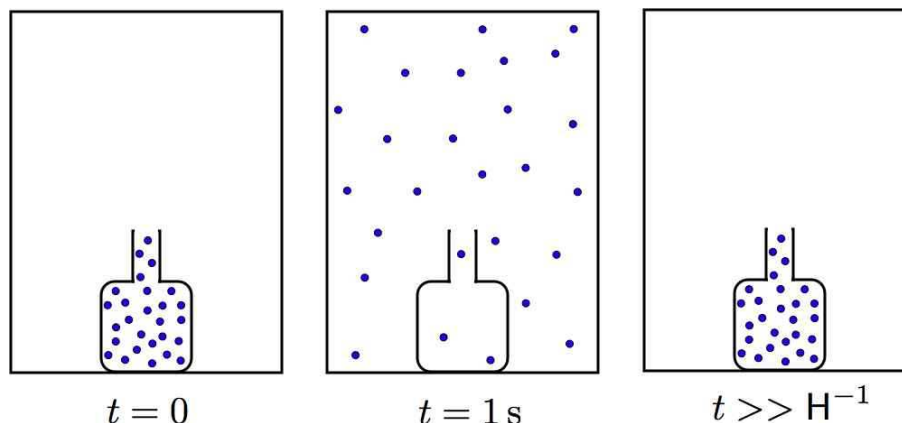


Figure 6.3: Poincaré recurrence guarantees that if we remove the cap from a bottle of perfume in an otherwise evacuated room, all the perfume molecules will eventually return to the bottle!

vation – 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 $f: \mathbb{R} \rightarrow \mathbb{R}$ which takes any real number to its fractional part. Thus, $f(\pi) = 0.14159265\dots$. 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 $\nabla \cdot \mathbf{V} = -2\beta < 0$, since $\beta > 0$ for physical damping. Thus the convective derivative is $D_t \varrho = -(\nabla \cdot \mathbf{V})\varrho = 2\beta\varrho$ 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: $\Omega(t) = e^{-2\beta t} \Omega(0)$, and are not preserved by the dynamics. The proof of recurrence therefore 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 $\text{vol}(g_\tau^n \mathcal{R}_0) = e^{-2n\beta\tau} \text{vol}(\mathcal{R}_0)$. A damped pendulum, released from rest at some small angle θ_0 , will not return arbitrarily close to these initial conditions.

6.7 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

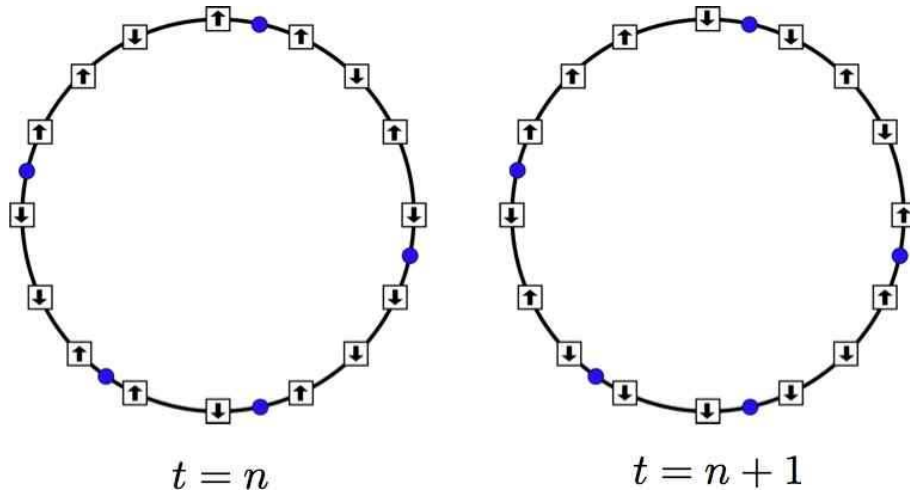


Figure 6.4: Left: A configuration of the Kac ring with $N = 16$ sites and $F = 4$ flippers. The flippers, which live on the links, are represented by blue dots. Right: The ring system after one time step. Evolution proceeds by clockwise rotation. Spins passing through flippers are flipped.

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 6.4 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

$$p_{n+1} = (1 - x)p_n + x(1 - p_n), \quad (6.84)$$

where $x = F/N$ is the fraction of flippers in the system. In words: a spin will be up at

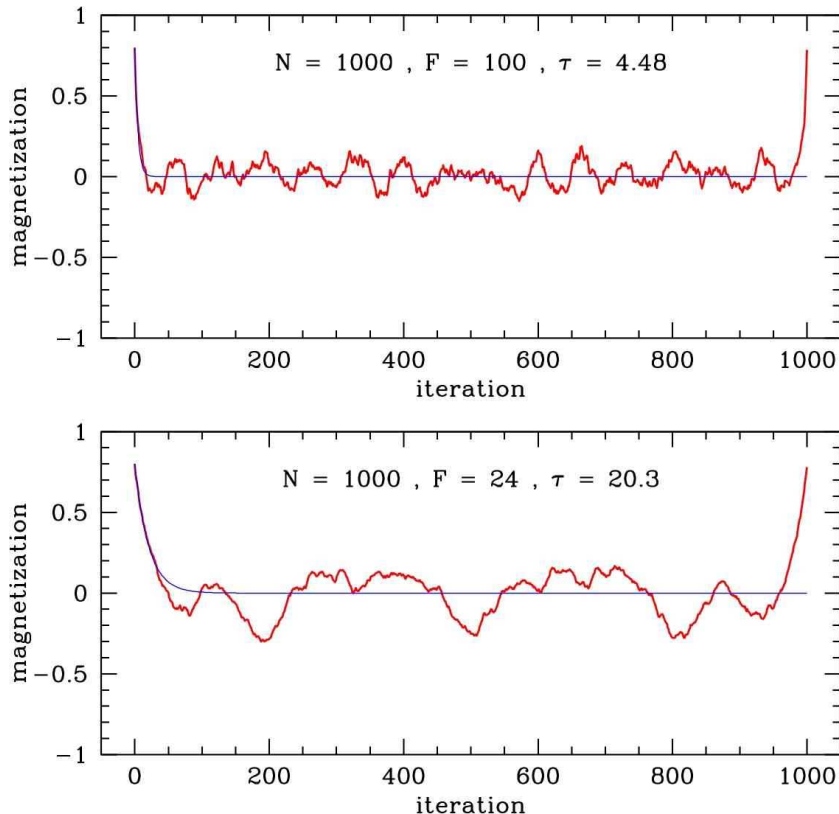


Figure 6.5: Two simulations of the Kac ring model, each with $N = 1000$ sites and with $F = 100$ flippers (top panel) and $F = 24$ flippers (bottom panel). The red line shows the magnetization as a function of time, starting from an initial configuration in which 90% of the spins are up. The blue line shows the prediction of the *Stosszahlansatz*, which yields an exponentially decaying magnetization with time constant τ .

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 6.84 can be solved immediately:

$$p_n = \frac{1}{2} + (1 - 2x)^n (p_0 - \frac{1}{2}) , \quad (6.85)$$

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

$$\tau(x) = -\frac{1}{\ln|1 - 2x|} . \quad (6.86)$$

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 $m_{\text{eq}} = 0$. Note that for $\frac{1}{2} < x < 1$ that the magnetization reverses sign each time step, as well as decreasing exponentially in magnitude.

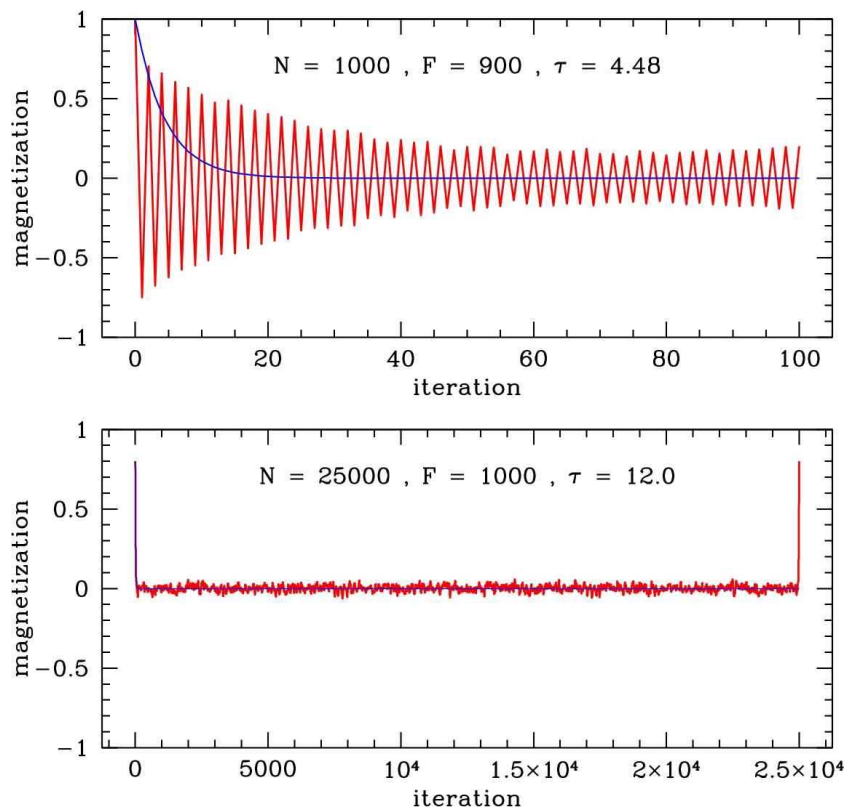


Figure 6.6: Simulations of the Kac ring model. Top: $N = 1000$ sites with $F = 900$ flippers. The flipper density $x = F/N$ is greater than $\frac{1}{2}$, so the magnetization reverses sign every time step. Only 100 iterations are shown, and the blue curve depicts the absolute value of the magnetization within the *Stosszahlansatz*. Bottom: $N = 25,000$ sites with $F = 1000$ flippers. Note that the fluctuations about the ‘equilibrium’ magnetization $m = 0$ are much smaller than in the $N = 1000$ site simulations.

The assumption that leads to equation 6.84 is called the *Stosszahlansatz*⁶, a long German word meaning, 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. 6.5 and 6.6.

The model is trivial to simulate. The results of such a simulation are shown in figure 6.5 for a ring of $N = 1000$ sites, with $F = 100$ and $F = 24$ flippers. Note how the magnetization

⁶Unfortunately, many important physicists were German and we have to put up with a legacy of long German words like *Gedankenexperiment*, *Zitterbewegung*, *Bremsstrahlung*, *Stosszahlansatz*, *Kartoffelsalat*, etc.

decays and fluctuates about the equilibrium value $e_{\text{q}} = 0$, 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 6.6 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?

6.8 Remarks on Ergodic Theory

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 $T \rightarrow \infty$ may be replaced by phase space averages. The time average of a function $f(\varphi)$ is defined as

$$\langle f(\varphi) \rangle_T = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T dt f(\varphi(t)) . \quad (6.87)$$

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

$$\langle f(\varphi) \rangle_S = \int d\mu f(\varphi) \delta(E - H(\varphi)) \Big/ \int d\mu \delta(E - H(\varphi)) , \quad (6.88)$$

where $H(\varphi) = H(\mathbf{q}, \mathbf{p})$ is the Hamiltonian, and where $\delta(x)$ is the Dirac δ -function. Thus,

$$\text{ergodicity} \iff \langle f(\varphi) \rangle_T = \langle f(\varphi) \rangle_S , \quad (6.89)$$

for all smooth functions $f(\varphi)$ for which $\langle f(\varphi) \rangle_S$ 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} :

$$\chi_{\mathcal{R}}(\varphi) = \begin{cases} 1 & \text{if } \varphi \in \mathcal{R} \\ 0 & \text{otherwise,} \end{cases} \quad (6.90)$$

where $H(\varphi) = E$ for all $\varphi \in \mathcal{R}$. Then

$$\langle \chi_{\mathcal{R}}(\varphi) \rangle_T = \lim_{T \rightarrow \infty} \left(\frac{\text{time spent in } \mathcal{R}}{T} \right) . \quad (6.91)$$

If the system is ergodic, then

$$\langle \chi_{\mathcal{R}}(\varphi) \rangle_T = P(\mathcal{R}) = \frac{\Sigma_{\mathcal{R}}(E)}{\Sigma(E)}, \quad (6.92)$$

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 phase space. The latter is given by

$$\Sigma(E) = \int d\mu \delta(E - H(\varphi)) \quad (6.93)$$

is the surface area of phase space at energy E , and

$$\Sigma_{\mathcal{R}}(E) = \int_{\mathcal{R}} d\mu \delta(E - H(\varphi)). \quad (6.94)$$

is the surface area of phase space at energy E contained in \mathcal{R} .

Note that

$$\Sigma(E) \equiv \int d\mu \delta(E - H(\varphi)) = \int_{\mathcal{S}_E} \frac{dS}{|\nabla H|} \quad (6.95)$$

$$= \frac{d}{dE} \int d\mu \Theta(E - \mathcal{H}(\varphi)) = \frac{d\Omega(E)}{dE}. \quad (6.96)$$

Here, dS is the differential surface element, \mathcal{S}_E is the constant H hypersurface $H(\varphi) = E$, and $\Omega(E)$ is the volume of phase space over which $H(\varphi) < E$. Note also that we may write

$$d\mu = dE d\Sigma_E, \quad (6.97)$$

where

$$d\Sigma_E = \left. \frac{dS}{|\nabla H|} \right|_{H(\varphi)=E} \quad (6.98)$$

is the the *invariant surface element*.

6.8.1 The microcanonical ensemble

The distribution,

$$\varrho_E(\varphi) = \frac{\delta(E - H(\varphi))}{\Sigma(E)} = \frac{\delta(E - H(\varphi))}{\int d\mu \delta(E - H(\varphi))}, \quad (6.99)$$

defines the *microcanonical ensemble* (μ CE) of Gibbs.

We could also write

$$\langle f(\varphi) \rangle_S = \frac{1}{\Sigma(E)} \int_{\mathcal{S}_E} d\Sigma_E f(\varphi), \quad (6.100)$$

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

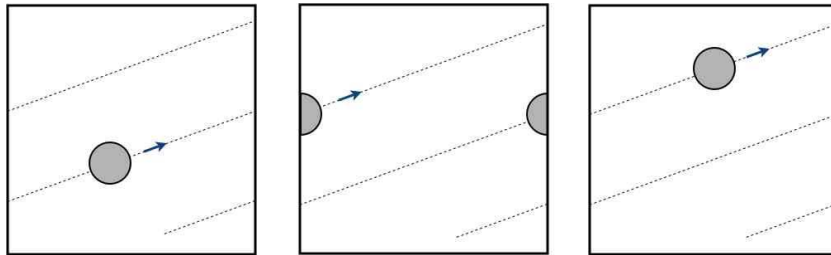


Figure 6.7: Constant phase space velocity at an irrational angle over a toroidal phase space is ergodic, but not mixing. A circle remains a circle, and a blob remains a blob.

6.8.2 Ergodicity and mixing

Just because a system is ergodic, it doesn't necessarily mean that $\varrho(\varphi, t) \rightarrow \varrho^{\text{eq}}(\varphi)$, for consider the following motion on the toroidal space $(\varphi = (q, p) \mid 0 \leq q < 1, 0 \leq p < 1)$, 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

$$\dot{q} = 1 \quad , \quad \dot{p} = \alpha . \quad (6.101)$$

The solution is

$$q(t) = q_0 + t \quad , \quad p(t) = p_0 + \alpha t , \quad (6.102)$$

hence the phase curves are given by

$$p = p_0 + \alpha(q - q_0) . \quad (6.103)$$

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

$$f(q, p) = \sum_{m, n} \hat{f}_{mn} e^{2\pi i(mq + np)} . \quad (6.104)$$

We have, then,

$$f(q(t), p(t)) = \sum_{m, n} \hat{f}_{mn} e^{2\pi i(mq_0 + np_0)} e^{2\pi i(m + \alpha n)t} . \quad (6.105)$$

We can now perform the time average of f :

$$\begin{aligned} \langle f(q, p) \rangle_T &= \hat{f}_{00} + \lim_{T \rightarrow \infty} \frac{1}{T} \sum'_{m, n} e^{2\pi i(mq_0 + np_0)} \frac{e^{2\pi i(m + \alpha n)T} - 1}{2\pi i(m + \alpha n)} \\ &= \hat{f}_{00} \quad \text{if } \alpha \text{ irrational.} \end{aligned} \quad (6.106)$$

Clearly,

$$\langle f(q, p) \rangle_S = \int_0^1 dq \int_0^1 dp f(q, p) = \hat{f}_{00} = \langle f(q, p) \rangle_T , \quad (6.107)$$

so the system is ergodic.

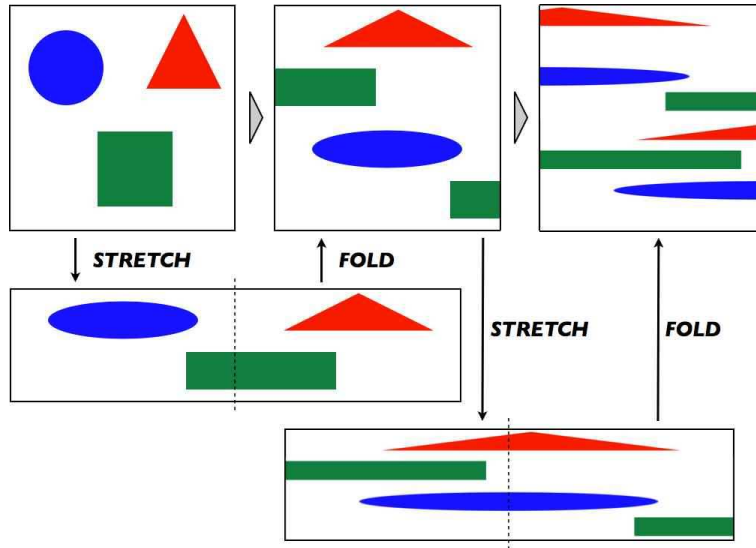


Figure 6.8: The baker's transformation is a successive stretching, cutting, and restacking.

The situation is depicted in fig. 6.7. If we start with the characteristic function of a disc,

$$\varrho(q, p, t = 0) = \Theta(a^2 - (q - q_0)^2 - (p - p_0)^2) , \quad (6.108)$$

then it remains the characteristic function of a disc:

$$\varrho(q, p, t) = \Theta(a^2 - (q - q_0 - t)^2 - (p - p_0 - \alpha t)^2) , \quad (6.109)$$

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

$$\nu(A) = \int d\Sigma_E \chi_A(\varphi) / \int d\Sigma_E = \frac{\Sigma_A(E)}{\Sigma(E)} , \quad (6.110)$$

where $\chi_A(\varphi)$ 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

$$\text{mixing} \iff \lim_{n \rightarrow \infty} \nu(g^n A \cap B) = \nu(A) \nu(B) . \quad (6.111)$$

In other words, the fraction of B covered by the n^{th} iterate of A , *i.e.* $g^n A$, is, as $n \rightarrow \infty$, 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.

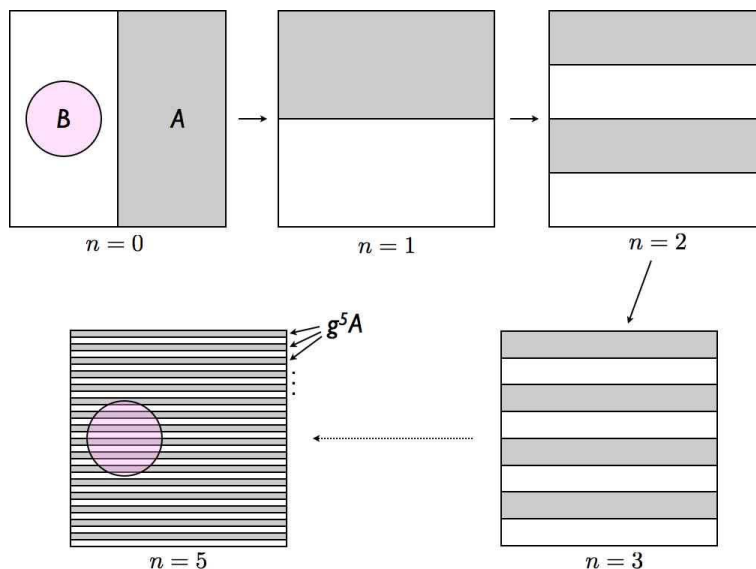


Figure 6.9: The multiply iterated baker's transformation. The set A covers half the phase space and its area is preserved under the map. Initially, the fraction of B covered by A is zero. After many iterations, the fraction of B covered by $g^n A$ approaches $\frac{1}{2}$.

Mixing means that

$$\begin{aligned}
 \langle f(\varphi) \rangle &= \int d\mu \varrho(\varphi, t) f(\varphi) \\
 &\xrightarrow{t \rightarrow \infty} \int d\mu f(\varphi) \delta(E - H(\varphi)) / \int d\mu \delta(E - H(\varphi)) \\
 &\equiv \text{Tr} \left[f(\varphi) \delta(E - H(\varphi)) \right] / \text{Tr} \left[\delta(E - H(\varphi)) \right].
 \end{aligned} \tag{6.112}$$

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.

An example of a mixing system is the *baker's transformation*, depicted in fig. 6.8. The baker map is defined by

$$g(q, p) = \begin{cases} (2q, \frac{1}{2}p) & \text{if } 0 \leq q < \frac{1}{2} \\ (2q - 1, \frac{1}{2}p + \frac{1}{2}) & \text{if } \frac{1}{2} \leq q < 1. \end{cases} \tag{6.113}$$

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



Figure 6.10: The Arnold cat map applied to an image of 150×150 pixels. After 300 iterations, the image repeats itself. (Source: Wikipedia)

(q) direction, followed by a slicing and restacking. Explicitly,

$$g^{-1}(q, p) = \begin{cases} (\frac{1}{2}q, 2p) & \text{if } 0 \leq p < \frac{1}{2} \\ (\frac{1}{2}q + \frac{1}{2}, 2p - 1) & \text{if } \frac{1}{2} \leq p < 1. \end{cases} \quad (6.114)$$

Another example of a mixing system is Arnold's 'cat map',⁷

$$g(q, p) = ([q + p], [q + 2p]), \quad (6.115)$$

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

$$\begin{pmatrix} q' \\ p' \end{pmatrix} = \overbrace{\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}}^M \begin{pmatrix} q \\ p \end{pmatrix} \pmod{\mathbb{Z}^2}. \quad (6.116)$$

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

$$\begin{pmatrix} q \\ p \end{pmatrix} = \overbrace{\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}}^{M^{-1}} \begin{pmatrix} q' \\ p' \end{pmatrix} \pmod{\mathbb{Z}^2}. \quad (6.117)$$

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 $\{1, \dots, k\}$. Clearly g and its inverse preserve this set, since the entries of M and M^{-1} are

⁷The cat map gets its name from its initial application, by Arnold, to the image of a cat's face.

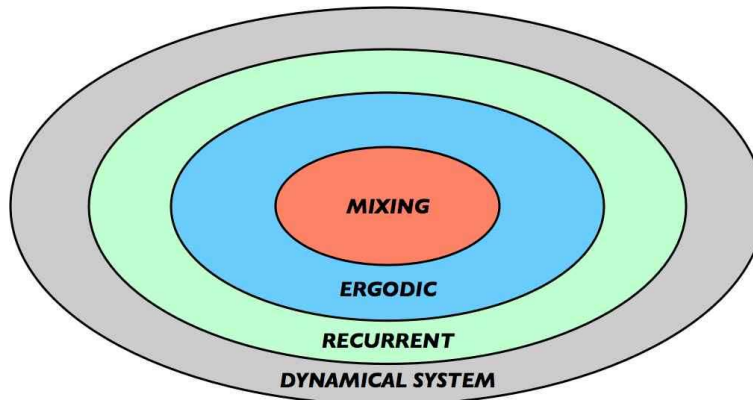


Figure 6.11: The hierarchy of dynamical systems.

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. 6.10, 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. 6.11, understanding the characteristic features of each successive refinement⁸.

⁸There 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 Liouvillean L has a Lebesgue spectrum with denumerably infinite multiplicity.