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, $N_{\rm A} = 6.02 \times 10^{23}$. 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 $\mathcal{O}\big(10^{23}\big)$ 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

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

Here, W_{ij} is the rate at which j makes a transition to i. We assume the state space is finite and of size N, so that $i, j \in \{1, \ldots, N\}$. 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

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

where $\hat{H}_0\,|\,i\,\rangle = E_i\,|\,i\,\rangle$, $\hat{V}$ is an additional potential which leads to transitions, and $\varrho(E_i)$ 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 $\vec{L} = (1, 1, \ldots, 1)$ is a left eigenvector of Γ with eigenvalue $\lambda = 0$. The corresponding right eigenvector, which we write as $P_i^{\rm eq}$, satisfies $\sum_j \Gamma_{ij} P_j^{\rm eq} = 0$, and is a stationary (*i.e.* time independent) solution to the master equation. In order for probability to be conserved, it must be that $\mathrm{Re}\,\lambda_a \geq 0$ for each eigenvalue λ_a. 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 $P_i^{\rm eq}$ as $t \to \infty$, as shown in the appendix §**??**. 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. In the generic case, we have

$$\Gamma = \sum_{a=1}^{N} \lambda_a \,|\, R^a \,\rangle\langle\, L^a \,| \qquad \Rightarrow \qquad \Gamma_{ij} = \sum_{a=1}^{N} \lambda_a \, R_i^a \, L_j^a \quad , \tag{3.7}$$

with $\langle L^a | R^b \rangle = \delta^{ab}$.

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, $W_{ij}\, P_j^{\rm eq} = W_{ji}\, P_i^{\rm eq}$. Assuming $W_{ij} \neq 0$ and $P_j^{\rm eq} \neq 0$, we can divide to obtain

$$\frac{W_{ji}}{W_{ij}} = \frac{P_j^{\rm eq}}{P_i^{\rm eq}} \quad . \tag{3.8}$$

Note that detailed balance is a stronger condition than that required for a stationary solution to the master equation. When detailed balance holds, then $\Gamma_{ji}\pi_i = \Gamma_{ij}\pi_j \;\forall\; i, j$, where $\pi_i \equiv P_i^{\rm eq}$. But then we have

$$\Delta_{ij} \equiv \pi_i^{-1/2} \Gamma_{ij} \pi_j^{1/2} = \Delta_{ji} \quad . \tag{3.9}$$

The real symmetric matrix Δ has real eigenvalues λ_a and a complete and orthonormal set of left/right eigenvectors ϕ_i^a. We then have $R_j^a = \pi_j^{1/2}\phi_j^a$ and $L_i^a = \pi_i^{-1/2}\phi_i^a$. Since $\sum_i \Gamma_{ij} = 0$, we have that $\pi_i^{1/2}$ is a left eigenvector of Δ with eigenvalue zero. Therefore $R_j^{(1)} = \pi_j$, which is the equilibrium distribution.

If $\Gamma = \Gamma^{\mathsf{T}}$ is symmetric, then, up to normalization, the right eigenvectors and left eigenvectors are transposes of each other, hence $P^{\rm eq} = 1/N$, where N is the matrix dimension of Γ. The system then satisfies the conditions of detailed balance. See §3.6 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

$$\mathcal{H}(t) = \sum_i P_i(t)\ \ln P_i(t) \quad . \tag{3.10}$$

Then

$$\begin{aligned} \frac{d\mathcal{H}}{dt} &= \sum_i \frac{dP_i}{dt}\left(1 + \ln P_i\right) = \sum_i \frac{dP_i}{dt}\ \ln P_i \\ &= -\sum_{i,j} \Gamma_{ij}\, P_j \ln P_i = \sum_{i,j} \Gamma_{ij}\, P_j \big(\ln P_j - \ln P_i\big) \quad , \end{aligned} \tag{3.11}$$

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

$$\frac{d\mathcal{H}}{dt} = \frac{1}{2}\sum_{i,j} \Gamma_{ij}\big(P_i - P_j\big)\,\big(\ln P_i - \ln P_j\big) \quad . \tag{3.12}$$

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

$$(x - y)\,(\ln x - \ln y) \geq 0 \quad , \tag{3.13}$$

we conclude

$$\frac{d\mathcal{H}}{dt} \leq 0 \quad . \tag{3.14}$$

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

$$P_i^{\rm eq} = \frac{1}{\Omega} \quad , \quad \Omega = \sum_i 1 \quad \Longrightarrow \quad \mathcal{H} = -\ln\Omega \quad , \tag{3.15}$$

with $\Omega = N$, the matrix dimension of Γ.

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

$$\overline{W}_{ij} \equiv W_{ij}\, P_j^{\rm eq} = W_{ji}\, P_i^{\rm eq} = \overline{W}_{ji} \quad , \tag{3.16}$$

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

$$\mathcal{H}(t) \equiv \sum_i P_i(t)\ \ln\!\left(\frac{P_i(t)}{P_i^{\rm eq}}\right) \quad . \tag{3.17}$$

Then

$$\begin{aligned} \frac{d\mathcal{H}}{dt} &= -\frac{1}{2}\sum_{i,j}\left(W_{ji}P_i - W_{ij}P_j\right)\,\ln\!\left(\frac{W_{ji}P_i}{W_{ij}P_j}\right) \\ &= -\frac{1}{2}\sum_{i,j}\overline{W}_{ij}\left(\frac{P_i}{P_i^{\rm eq}} - \frac{P_j}{P_j^{\rm eq}}\right)\left[\ln\!\left(\frac{P_i}{P_i^{\rm eq}}\right) - \ln\!\left(\frac{P_j}{P_j^{\rm eq}}\right)\right] \leq 0 \quad . \end{aligned} \tag{3.18}$$

3.2.5 Markov chains

A *Markov chain* is a process which describes transitions of a stochastic variable in time. There are four types to consider, depending on whether the state space is discrete (*i.e.* countable) or continuous (and measurable), and whether the process occurs in discrete time or continuous time. Here we will consider a discrete time Markov chain of a discrete state space. Let $P_i(t)$ be the probability that the system is in state i at time t. The evolution equation is

$$P_i(t+1) = \sum_j Q_{ij}\, P_j(t) \quad . \tag{3.19}$$

Note the similarity to the (continuous time) master equation in eqn. 3.2. Indeed, the master equation is an example of a continuous time Markov process. All the elements of the *transition matrix* Q_{ij} are real and nonnegative, and Q_{ij} satisfies $\sum_i Q_{ij} = 1$ so that the total probability $\sum_i P_i(t)$ is conserved. The element Q_{ij} is the *conditional probability* that for the system to evolve to state i at time $t+1$ given that it is in state j at time t, with Q_{ij} independent of t. Interating eqn. 3.19, we have

$$P_i(t+n) = \left[Q^n\right]_{ij} P_j(t) \quad . \tag{3.20}$$

The transition matrix Q is in general non-normal, meaning $Q \neq Q^{\mathsf{T}}$, which means that its left eigenvectors are in general not simply related to the right eigenvectors at the corresponding eigenvalues, and the eigenvalues are not necessarily real. Under generic circumstances, we may write

$$Q = \sum_{a=1}^{n} \lambda_a \,|\, R^a \,\rangle\langle\, L^a \,| \quad , \tag{3.21}$$

with $|\mathrm{Re}\,\lambda_a| \leq 1$ for all a. In nongeneric cases, there may be Jordan blocks in the canonical form of Q, which means the eigenvectors do not span (see appendix 1 of ch. 3 in my PHYS210A notes). From the condition $\sum_i Q_{ij} = 1$, we have that $\vec{L}^{(1)} = (1, 1, \ldots, 1)$ is a left eigenvector corresponding to eigenvalue $\lambda = 1$, and the corresponding right eigenvector $|\, R^{(1)} \,\rangle$ is the asymptotic steady state, with $P_j^{\rm eq} = \langle\, j \,|\, R^1 \,\rangle = R_j^a$. Since $Q \in \mathsf{GL}(N, \mathbb{R})$, where N is the matrix dimension (which we take here to be finite), if $\lambda \in \mathbb{C}$ is an eigenvalue, then λ^* is also an eigenvalue.

Example

Consider a group of Physics graduate students consisting of three theorists and four experimentalists. Within each group, the students are to be regarded as indistinguishable. Together, the students rent two apartments, A and B. Initially the three theorists live in A and the four experimentalists live in B. Each month, a random occupant of A and a random occupant of B exchange domiciles. Compute the transition matrix Q_{ij} for this Markov chain, and compute the average fraction of the time that B contains two theorists and two experimentalists, averaged over the effectively infinite time it takes the students to get their degrees. *Hint:* Q is a 4×4 matrix.

Solution: There are four states available: Now let's compute the transition probabilities. First, we compute the transition probabilities out of state $|\, 1 \,\rangle$, *i.e.* the matrix elements Q_{j1}. Clearly $Q_{21} = 1$

$\vert\, j\,\rangle$	room A	room B	$g_j^{\rm A}$	$g_j^{\rm B}$	$g_j^{\rm TOT}$
$\vert\, 1\,\rangle$	TTT	EEEE	1	1	1
$\vert\, 2\,\rangle$	TTE	EEET	3	4	12
$\vert\, 3\,\rangle$	TEE	EETT	3	6	18
$\vert\, 4\,\rangle$	EEE	ETTT	1	4	4

Table 3.1: States and their degeneracies.

since we must exchange a theorist (T) for an experimentalist (E). All the other probabilities are zero: $Q_{11} = Q_{31} = Q_{41} = 0$. For transitions out of state $\vert\, 2\,\rangle$, the nonzero elements are

$$Q_{12} = \tfrac{1}{4} \times \tfrac{1}{3} = \tfrac{1}{12} \quad , \quad Q_{22} = \tfrac{3}{4} \times \tfrac{1}{3} + \tfrac{1}{4} \times \tfrac{2}{3} = \tfrac{5}{12} \quad , \quad Q_{32} = \tfrac{1}{2} \quad . \tag{3.22}$$

To compute Q_{12}, we must choose the experimentalist from room A (probability $\frac{1}{3}$) with the theorist from room B (probability $\frac{1}{4}$). For Q_{22}, we can either choose E from A and one of the E's from B, or one of the T's from A and the T from B. This explains the intermediate steps written above. For transitions out of state $\vert\, 3\,\rangle$, the nonzero elements are then $Q_{23} = \frac{1}{3}$, $Q_{33} = \frac{1}{2}$, and $Q_{43} = \frac{1}{6}$. Finally, for transitions out of state $\vert\, 4\,\rangle$, the nonzero elements are $Q_{34} = \frac{3}{4}$ and $Q_{44} = \frac{1}{4}$. The full transition matrix is then

$$Q = \begin{pmatrix} 0 & \frac{1}{12} & 0 & 0 \\ 1 & \frac{5}{12} & \frac{1}{3} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & \frac{3}{4} \\ 0 & 0 & \frac{1}{6} & \frac{1}{4} \end{pmatrix} \quad . \tag{3.23}$$

Note that $\sum_i Q_{ij} = 1$ for all $j = 1, 2, 3, 4$. This guarantees that $L^{(1)} = (1\,,\,1\,,\,1\,,\,1)$ is a left eigenvector of Q with eigenvalue 1. The corresponding right eigenvector is obtained by setting $Q_{ij}\,\psi_j^{(1)} = \psi_i^{(1)}$. Simultaneously solving these four equations and normalizing so that $\sum_j \psi_j^{(1)} = 1$, we easily obtain

$$\psi^{(1)} = \frac{1}{35} \begin{pmatrix} 1 \\ 12 \\ 18 \\ 4 \end{pmatrix} \quad . \tag{3.24}$$

This is the state we converge to after repeated application of the transition matrix Q. If we decompose $Q = \sum_{\alpha=1}^4 \lambda_\alpha\, \vert\, \psi^{(\alpha)}\,\rangle\langle\, L^{(\alpha)}\,\vert$, then in the limit $t \to \infty$ we have $Q^t \approx \vert\, \psi^{(1)}\,\rangle\langle\, L^{(1)}\,\vert$, where $\lambda_1 = 1$, since the remaining eigenvalues are all less than 1 in magnitude[2]. Thus, Q^t acts as a *projector* onto the state $\vert\, \psi^{(1)}\,\rangle$. Whatever the initial set of probabilities $P_j(t=0)$, we must have $\langle\, L^{(1)}\,\vert\, P(0)\,\rangle = \sum_j P_j(0) = 1$. Therefore, $\lim_{t\to\infty} P_j(t) = \psi_j^{(1)}$, and we find $P_3(\infty) = \frac{18}{35}$. Note that the equilibrium distribution satisfies detailed balance:

$$\psi_j^{(1)} = \frac{g_j^{\rm TOT}}{\sum_l g_l^{\rm TOT}} \quad . \tag{3.25}$$

[2]One can check that $\lambda_1 = 1$, $\lambda_2 = \frac{5}{12}$, $\lambda_3 = 0$. and $\lambda_4 = -\frac{1}{4}$.

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

$$\dot{p}_\sigma = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\sigma}\right) = \frac{\partial L}{\partial q_\sigma} \quad , \tag{3.26}$$

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,

$$H(q,p) = \sum_{\sigma=1}^{N} p_\sigma\, \dot{q}_\sigma - L \quad . \tag{3.27}$$

Note that

$$\begin{aligned} dH &= \sum_{\sigma=1}^{N}\left(p_\sigma\, d\dot{q}_\sigma + \dot{q}_\sigma\, dp_\sigma - \frac{\partial L}{\partial q_\sigma}\, dq_\sigma - \frac{\partial L}{\partial \dot{q}_\sigma}\, d\dot{q}_\sigma\right) - \frac{\partial L}{\partial t}\, dt \\ &= \sum_{\sigma=1}^{N}\left(\dot{q}_\sigma\, dp_\sigma - \frac{\partial L}{\partial q_\sigma}\, dq_\sigma\right) - \frac{\partial L}{\partial t}\, dt \quad . \end{aligned} \tag{3.28}$$

Thus, we obtain Hamilton's equations of motion,

$$\frac{\partial H}{\partial p_\sigma} = \dot{q}_\sigma \quad , \quad \frac{\partial H}{\partial q_\sigma} = -\frac{\partial L}{\partial q_\sigma} = -\dot{p}_\sigma \tag{3.29}$$

and

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \quad . \tag{3.30}$$

Define the rank $2N$ vector $\boldsymbol{\varphi}$ by its components,

$$\varphi_i = \begin{cases} q_i & \text{if } 1 \le i \le N \\ p_{i-N} & \text{if } N \le i \le 2N \end{cases} \quad . \tag{3.31}$$

Then we may write Hamilton's equations compactly as

$$\dot{\varphi}_i = J_{ij}\, \frac{\partial H}{\partial \varphi_j} \quad , \tag{3.32}$$

where

$$J = \begin{pmatrix} 0_{N\times N} & 1_{N\times N} \\ -1_{N\times N} & 0_{N\times N} \end{pmatrix} \tag{3.33}$$

is a rank $2N$ matrix. Note that $J^{\mathsf{T}} = -J$, *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

$$\begin{aligned} \frac{dF}{dt} &= \frac{\partial F}{\partial t} + \sum_{\sigma=1}^{N}\left(\frac{\partial F}{\partial q_\sigma}\frac{dq_\sigma}{dt} + \frac{\partial F}{\partial p_\sigma}\frac{dp_\sigma}{dt}\right) \\ &= \frac{\partial F}{\partial t} + \big\{F, H\big\} \quad , \end{aligned} \tag{3.34}$$

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

$$\big\{A, B\big\} = \sum_{\sigma=1}^{N}\left(\frac{\partial A}{\partial q_\sigma}\frac{\partial B}{\partial p_\sigma} - \frac{\partial A}{\partial p_\sigma}\frac{\partial B}{\partial q_\sigma}\right) \quad . \tag{3.35}$$

3.3.2 Dynamical systems and the evolution of phase space volumes

Consider a general dynamical system,

$$\frac{d\boldsymbol{\varphi}}{dt} = \boldsymbol{V}(\boldsymbol{\varphi}) \quad , \tag{3.36}$$

where $\boldsymbol{\varphi}(t)$ is a point in an n-dimensional phase space. Consider now a compact[3] 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 $\big\{\boldsymbol{\varphi}\,\big|\,\boldsymbol{\varphi}\in\mathcal{R}_0\big\}$, 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 $t=0$:

$$\mathcal{R}(t) = \big\{\boldsymbol{\varphi}(t)\,\big|\,\boldsymbol{\varphi}(0)\in\mathcal{R}_0\big\} \quad . \tag{3.37}$$

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

$$\Omega(t) = \int\limits_{\mathcal{R}(t)} d\mu \tag{3.38}$$

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

$$\Omega(t+dt) = \int\limits_{\mathcal{R}(t+dt)} d\mu' = \int\limits_{\mathcal{R}(t)} d\mu \left|\frac{\partial\varphi_i(t+dt)}{\partial\varphi_j(t)}\right| \quad , \tag{3.39}$$

where

$$\left|\frac{\partial\varphi_i(t+dt)}{\partial\varphi_j(t)}\right| \equiv \frac{\partial(\varphi_1',\ldots,\varphi_n')}{\partial(\varphi_1,\ldots,\varphi_n)} \tag{3.40}$$

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

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

$$\varphi_i(t+dt) = \varphi_i(t) + V_i\big(\boldsymbol{\varphi}(t)\big)\,dt + \mathcal{O}(dt^2) \tag{3.41}$$

and therefore

$$\frac{\partial\varphi_i(t+dt)}{\partial\varphi_j(t)} = \delta_{ij} + \frac{\partial V_i}{\partial\varphi_j}\,dt + \mathcal{O}(dt^2) \quad . \tag{3.42}$$

We now make use of the matrix equality $\ln\det M = \mathsf{Tr}\,\ln M$, which gives us[4], for small ε,

$$\det\big(1+\varepsilon A\big) = \exp\mathsf{Tr}\,\ln\big(1+\varepsilon A\big) = 1 + \varepsilon\,\mathsf{Tr}\,A + \tfrac{1}{2}\,\varepsilon^2\Big(\big(\mathsf{Tr}\,A\big)^2 - \mathsf{Tr}\,(A^2)\Big) + \ldots \tag{3.43}$$

Thus,

$$\Omega(t+dt) = \Omega(t) + \int\limits_{\mathcal{R}(t)} d\mu\;\boldsymbol{\nabla}\cdot\boldsymbol{V}\;dt + \mathcal{O}(dt^2) \quad , \tag{3.44}$$

which says

$$\frac{d\Omega}{dt} = \int\limits_{\mathcal{R}(t)} d\mu\;\boldsymbol{\nabla}\cdot\boldsymbol{V} = \int\limits_{\partial\mathcal{R}(t)} dS\;\hat{\boldsymbol{n}}\cdot\boldsymbol{V} \tag{3.45}$$

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

$$\boldsymbol{\nabla}\cdot\boldsymbol{V} = \sum_{i=1}^{n}\frac{\partial V_i}{\partial\varphi_i} \quad , \tag{3.46}$$

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,

$$\frac{\partial\varrho}{\partial t} + \boldsymbol{\nabla}\cdot(\varrho\boldsymbol{V}) = 0 \quad . \tag{3.47}$$

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

$$\frac{d}{dt}\int\limits_{\mathcal{R}} d\mu\;\varrho = -\int\limits_{\mathcal{R}} d\mu\;\boldsymbol{\nabla}\cdot(\varrho\boldsymbol{V}) = -\int\limits_{\partial\mathcal{R}} dS\;\hat{\boldsymbol{n}}\cdot(\varrho\boldsymbol{V}) \quad . \tag{3.48}$$

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

$$\frac{dQ_{\mathcal{R}}}{dt} = -\int\limits_{\partial\mathcal{R}} dS\;\hat{\boldsymbol{n}}\cdot\boldsymbol{J} \quad , \tag{3.49}$$

[4]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.

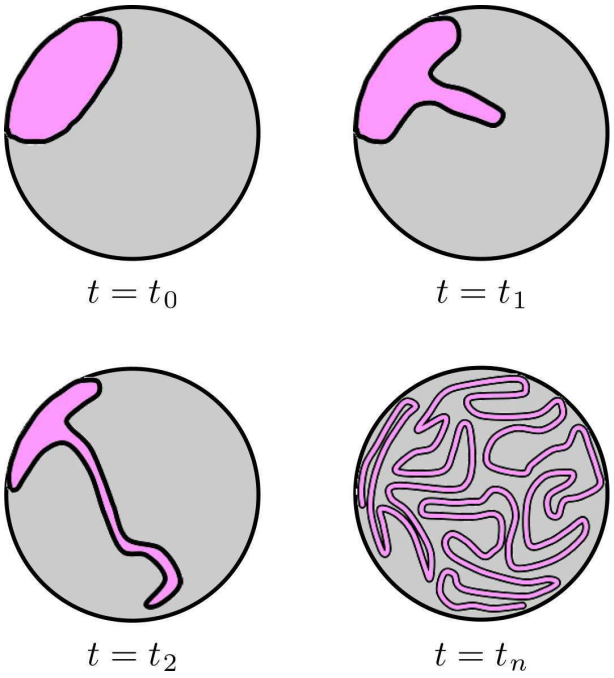

Figure 3.1: Time evolution of two immiscible fluids. The local density remains constant.

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

$$\frac{\partial \varrho}{\partial t} + \boldsymbol{V}\cdot\boldsymbol{\nabla}\varrho + \varrho\,\boldsymbol{\nabla}\cdot\boldsymbol{V} = 0 \quad . \tag{3.50}$$

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

$$\frac{D\varrho}{Dt} \equiv \left(\frac{\partial}{\partial t} + \boldsymbol{V}\cdot\boldsymbol{\nabla}\right)\varrho = 0 \quad . \tag{3.51}$$

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

$$\begin{aligned}\frac{d}{dt}\,\varrho\big(\boldsymbol{\varphi}(t),t\big) &= \frac{\partial\varrho}{\partial\varphi_i}\,\frac{d\varphi_i}{dt} + \frac{\partial\varrho}{\partial t} \\ &= \sum_{i=1}^{n} V_i\,\frac{\partial\varrho}{\partial\varphi_i} + \frac{\partial\varrho}{\partial t} = \frac{D\varrho}{Dt} \quad .\end{aligned} \tag{3.52}$$

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

$$\varrho(\boldsymbol{\varphi}, t=0) = \begin{cases} 1 & \text{if } \boldsymbol{\varphi}\in\mathcal{R}_0 \\ 0 & \text{otherwise} \end{cases} \tag{3.53}$$

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

$$\dot{q}_\sigma = +\frac{\partial H}{\partial p_\sigma} \quad , \quad \dot{p}_\sigma = -\frac{\partial H}{\partial q_\sigma} \tag{3.54}$$

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

$$\boldsymbol{\varphi} = \begin{pmatrix} \boldsymbol{q} \\ \boldsymbol{p} \end{pmatrix} \quad , \quad \boldsymbol{V} = \begin{pmatrix} \dot{\boldsymbol{q}} \\ \dot{\boldsymbol{p}} \end{pmatrix} = \begin{pmatrix} +\partial H/\partial \boldsymbol{p} \\ -\partial H/\partial \boldsymbol{q} \end{pmatrix} \quad . \tag{3.55}$$

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

$$\begin{aligned} \nabla \cdot \boldsymbol{V} &= \sum_{\sigma=1}^{N} \left\{ \frac{\partial \dot{q}_\sigma}{\partial q_\sigma} + \frac{\partial \dot{p}\sigma}{\partial p_\sigma} \right\} \\ &= \sum_{\sigma=1}^{N} \left\{ \frac{\partial}{\partial q_\sigma}\left(\frac{\partial H}{\partial p_\sigma}\right) + \frac{\partial}{\partial p_\sigma}\left(-\frac{\partial H}{\partial q_\sigma}\right) \right\} = 0 \quad . \end{aligned} \tag{3.56}$$

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

$$\frac{D\varrho}{Dt} \equiv \frac{\partial \varrho}{\partial t} + \boldsymbol{V} \cdot \nabla \varrho = 0 \quad , \tag{3.57}$$

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

$$\frac{\partial \varrho}{\partial t} = -\dot{\boldsymbol{\varphi}} \cdot \boldsymbol{\nabla} \varrho = -\big\{ \varrho, H \big\} \quad . \tag{3.58}$$

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*:

$$i\hat{L} = \sum_{\sigma=1}^{N} \left\{ \frac{\partial H}{\partial p_\sigma}\frac{\partial}{\partial q_\sigma} - \frac{\partial H}{\partial q_\sigma}\frac{\partial}{\partial p_\sigma} \right\} \quad . \tag{3.59}$$

Eqn. 3.58, known as *Liouville's equation*. Note $i\hat{L}F = \big\{F, H\big\}$ for any function $F(\boldsymbol{\varphi}, t)$.

Recall that the evolution of quantum mechanical density matrices satisfies

$$\frac{\partial \hat{\varrho}}{\partial t} = \frac{i}{\hbar}\big[\, \hat{\varrho}, \hat{H} \,\big] \quad , \tag{3.60}$$

whence we infer the correspondence

$$\big\{\varrho, H\big\} \to \frac{1}{i\hbar}\,\big[\,\hat{\varrho}, \hat{H}\,\big] \quad . \tag{3.61}$$

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 $\dot{\boldsymbol{\varphi}} \cdot \boldsymbol{\nabla}\Lambda_a = \big\{\Lambda_a, H\big\} = 0$. Then from the chain rule, we have

$$\dot{\boldsymbol{\varphi}} \cdot \boldsymbol{\nabla}\varrho = \sum_{a=1}^{k} \frac{\partial \varrho}{\partial \Lambda_a}\, \dot{\boldsymbol{\varphi}} \cdot \boldsymbol{\nabla}\Lambda_a = 0 \quad . \tag{3.62}$$

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,

$$\varrho_E(\boldsymbol{\varphi}) = \frac{\delta\big(E - H(\boldsymbol{\varphi})\big)}{D(E)} = \frac{\delta\big(E - H(\boldsymbol{\varphi})\big)}{\int d\mu\, \delta\big(E - H(\boldsymbol{\varphi})\big)} \quad , \tag{3.63}$$

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

$$\varrho_{E,Q}(\boldsymbol{\varphi}) = \frac{\delta\big(E - H(\boldsymbol{\varphi})\big)\,\delta\big(Q - \Lambda(\boldsymbol{\varphi})\big)}{D(E,Q)} = \frac{\delta\big(E - H(\boldsymbol{\varphi})\big)\,\delta\big(Q - \Lambda(\boldsymbol{\varphi})\big)}{\int d\mu\, \delta\big(E - H(\boldsymbol{\varphi})\,\delta\big(Q - \Lambda(\boldsymbol{\varphi})\big)} \quad , \tag{3.64}$$

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

Similarly, the Gibbs distribution,

$$\varrho_\beta(\boldsymbol{\varphi}) = \frac{1}{Z(\beta)}\, e^{-\beta H(\boldsymbol{\varphi})} \quad , \tag{3.65}$$

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

$$\varrho_{\boldsymbol{\beta}}(\boldsymbol{\varphi}) = \frac{1}{Z(\boldsymbol{\beta})}\, \exp\Big(-\sum_{a=1}^{k} \beta_a\, \Lambda_a(\boldsymbol{\varphi})\Big) \quad , \tag{3.66}$$

where the $\{\Lambda_a(\boldsymbol{\varphi})\}$, with $a \in \{1, \ldots, k\}$, 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: $dH/dt \leq 0$, where H is Boltzmann's H-function. However, the microscopic laws of physics are

(almost) time-reversal invariant[5], 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.*

$$\int d\mu\, \delta\big(E - H(\boldsymbol{q},\boldsymbol{p})\big) < \infty \quad , \tag{3.67}$$

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.

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:

$$\Upsilon = \bigcup_{k=0}^{\infty} g_\tau^k\,\mathcal{R}_0 \tag{3.68}$$

We assume that the set $\{g_\tau^k\,\mathcal{R}_0 \,|\, k\in\mathbb{N}_0\}$ is disjoint[6]. The volume of a union of disjoint sets is the sum of the individual volumes. Thus,

$$\mathrm{vol}(\Upsilon) = \sum_{k=0}^{\infty} \mathrm{vol}\big(g_\tau^k\,\mathcal{R}_0\big) = \mathrm{vol}(\mathcal{R}_0)\cdot\sum_{k=0}^{\infty} 1 = \infty \quad , \tag{3.69}$$

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 $\{g_\tau^k\,\mathcal{R}_0 \,|\, k\in\mathbb{N}_0\}$ 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]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.

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

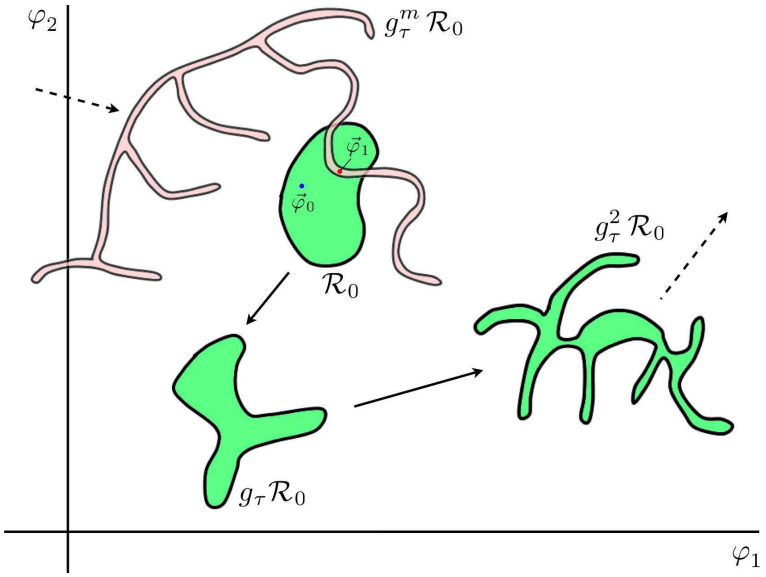

Figure 3.2: Successive images of a set $\mathcal{R}_0$ under the τ-advance mapping g_τ, projected onto a two-dimensional phase plane. The Poincaré recurrence theorem guarantees that if phase space has finite volume, and g_τ is invertible and volume preserving, then for any set $\mathcal{R}_0$ there exists an integer m such that $\mathcal{R}_0 \cap g_\tau^m\,\mathcal{R}_0 \neq \emptyset$.

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 $T(\boldsymbol{p}) > E_{\rm TOT}$, where $T(\boldsymbol{p})$ is the single particle kinetic energy function[7]. 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![8] In this case, we could define the region $\mathcal{R}_0$ as

$$\mathcal{R}_0 = \left\{(q_1,\ldots,q_r,p_1,\ldots,p_r)\;\middle|\;|q_i-q_i^0|\le\Delta q \text{ and } |p_j-p_j^0|\le\Delta p\;\forall\,i,j\right\}\quad , \tag{3.70}$$

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.

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

[8] 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.

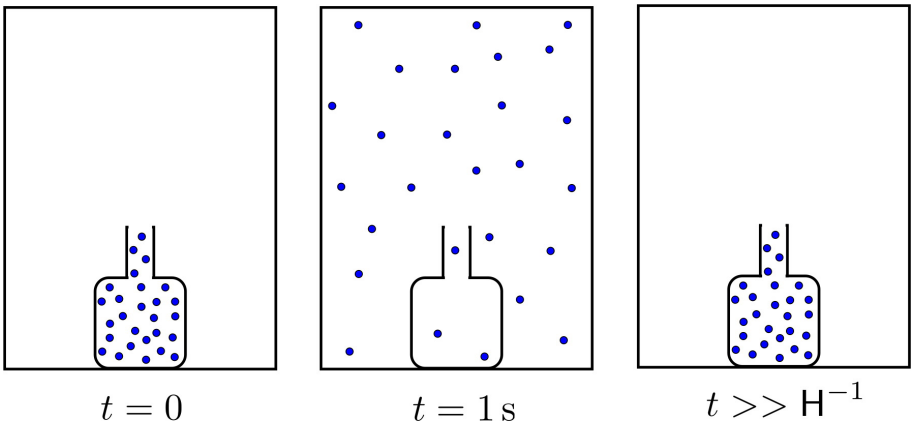

Figure 3.3: Poincaré recurrence guarantees that if we remove the cap from a bottle of perfume in an otherwise evacuated room, all the perfume molecules will eventually return to the bottle! (Here H is the Hubble constant.)

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 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.

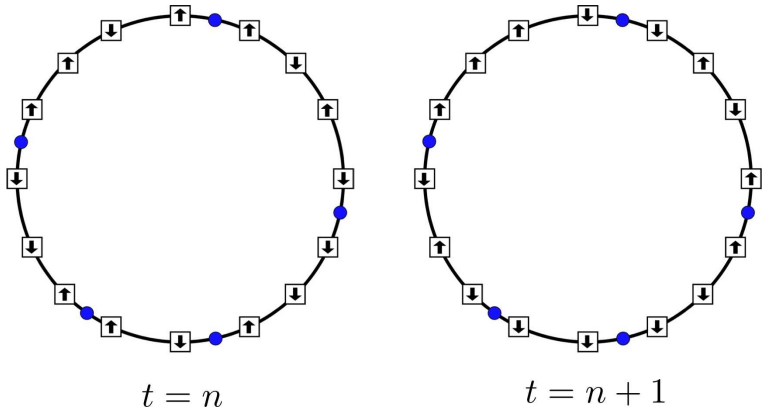

Figure 3.4: Left: A configuration of the Kac ring with $N = 16$ sites and $F = 4$ flippers. The flippers, which live on the links, are represented by blue dots. Right: The ring system after one time step. Evolution proceeds by clockwise rotation. Spins passing through flippers are flipped.

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 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

$$p_{n+1} = (1-x)\,p_n + x\,(1-p_n) \quad , \tag{3.71}$$

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 3.71 can be solved immediately:

$$p_n = \tfrac{1}{2} + (1-2x)^n\,(p_0 - \tfrac{1}{2}) \quad , \tag{3.72}$$

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

$$\tau(x) = -\frac{1}{\ln|1-2x|} \quad . \tag{3.73}$$

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 $m_{\rm eq} = 0$. 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 3.71 is called the *Stosszahlansatz*[9], a long German word meaning, approximately, 'assumption on the counting of hits'. The resulting dynamics are irreversible: the

[9]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.*

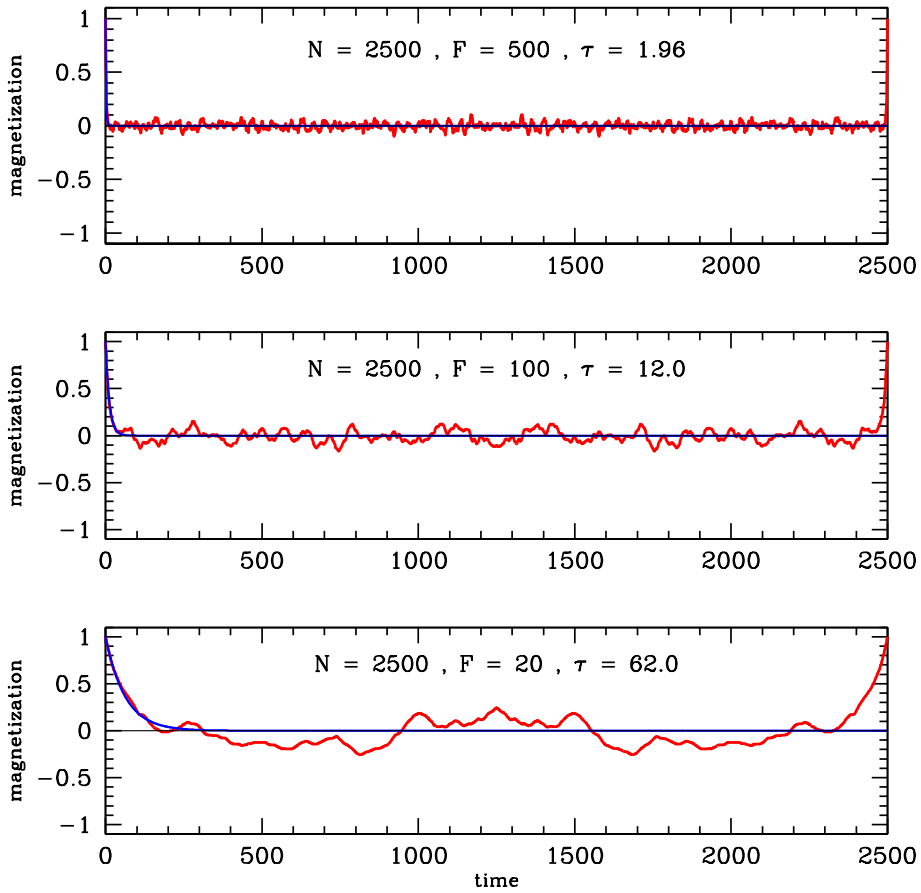

Figure 3.5: Three simulations of the Kac ring model with $N = 2500$ sites and three different concentrations of flippers. The red line shows the magnetization as a function of time, starting from an initial configuration in which 100% of the spins are up. The blue line shows the prediction of the *Stosszahlansatz*, which yields an exponentially decaying magnetization with time constant τ.

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 $m_{\rm eq} = 0$, 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

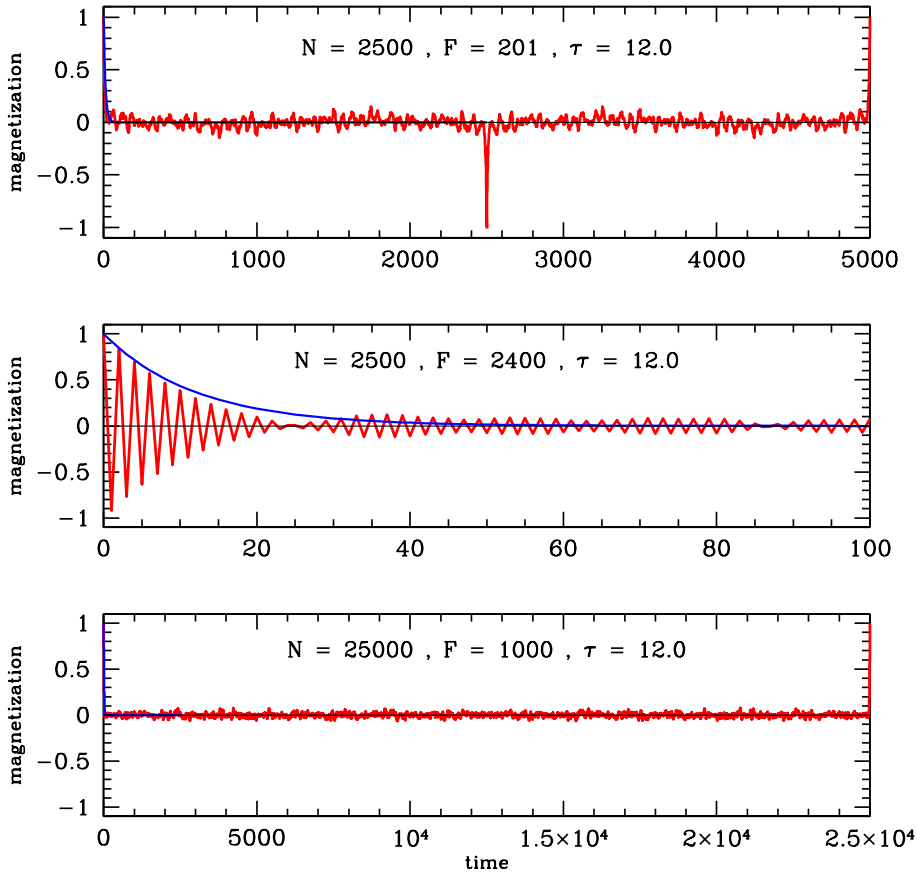

Figure 3.6: Simulations of the Kac ring model. Top: $N = 2500$ sites with $F = 201$ flippers. After 2500 iterations, each spin has flipped an odd number of times, so the recurrence time is $2N$. Middle: $N = 2500$ with $F = 2400$, resulting in a near-complete reversal of the population with every iteration. Bottom: $N = 25000$ with $N = 1000$, showing long time equilibration and dramatic resurgence of the spin population.

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

$$\big\langle f(\boldsymbol{\varphi})\big\rangle_t = \lim_{T\to\infty}\frac{1}{T}\int\limits_0^T dt\, f\big(\boldsymbol{\varphi}(t)\big) \quad . \tag{3.74}$$

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

$$\big\langle f(\boldsymbol{\varphi})\big\rangle_E = \int d\mu\, f(\boldsymbol{\varphi})\,\delta\big(E-H(\boldsymbol{\varphi})\big)\Big/\int d\mu\,\delta\big(E-H(\boldsymbol{\varphi})\big) \quad , \tag{3.75}$$

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

$$\text{ergodicity} \quad\Longleftrightarrow\quad \big\langle f(\boldsymbol{\varphi})\big\rangle_t = \big\langle f(\boldsymbol{\varphi})\big\rangle_E \quad , \tag{3.76}$$

for all smooth functions $f(\boldsymbol{\varphi})$ for which $\big\langle f(\boldsymbol{\varphi})\big\rangle_E$ exists and is finite. Note that we do not average over *all* of phase space. Rather, we average only over a hypersurface along which $H(\boldsymbol{\varphi}) = E$ 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 $\boldsymbol{\varphi}$ 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 $\chi_{\mathcal{R}}(\boldsymbol{\varphi})$ be the characteristic function of a region $\mathcal{R}$:

$$\chi_{\mathcal{R}}(\boldsymbol{\varphi}) = \begin{cases} 1 & \text{if } \boldsymbol{\varphi}\in\mathcal{R} \\ 0 & \text{otherwise,}\end{cases} \tag{3.77}$$

where $H(\boldsymbol{\varphi}) = E$ for all $\boldsymbol{\varphi}\in\mathcal{R}$, so $\dim\mathcal{R} = 2n-1$, where $n = 2dN$ is the total dimension of phase space for N point particles in d space dimensions. Then

$$\big\langle \chi_{\mathcal{R}}(\boldsymbol{\varphi})\big\rangle_t = \lim_{T\to\infty}\left(\frac{\text{time spent in }\mathcal{R}}{T}\right) \quad . \tag{3.78}$$

If the system is ergodic, then

$$\big\langle \chi_{\mathcal{R}}(\boldsymbol{\varphi})\big\rangle_t = P(\mathcal{R}) = \frac{D_{\mathcal{R}}(E)}{D(E)} \quad , \tag{3.79}$$

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

$$D(E) = \int d\mu\,\delta\big(E-H(\boldsymbol{\varphi})\big) = \int\limits_{\mathcal{S}_E} d\Sigma_E \quad , \tag{3.80}$$

called the *density of states*, is the surface area of phase space at energy E, and $D_{\mathcal{R}}(E)$ is defined below in eqn. 3.83. The hypersurface $\mathcal{S}_E$ is the set of points $\boldsymbol{\varphi}$ satisfying $H(\boldsymbol{\varphi}) = E$, 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[10]

$$d\Sigma_E \equiv \frac{dS_E}{|\boldsymbol{\nabla} H|}\bigg|_{H(\boldsymbol{\varphi})=E} \quad , \tag{3.81}$$

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

$$D(E) = \frac{d}{dE}\int d\mu\, \Theta\big(E - H(\boldsymbol{\varphi})\big) \equiv \frac{d\Omega(E)}{dE} \quad , \tag{3.82}$$

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

$$D_{\mathcal{R}}(E) = \int\limits_{\mathcal{R}} d\Sigma_E \quad . \tag{3.83}$$

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

3.5.2 The microcanonical ensemble

The distribution,

$$\varrho_E(\boldsymbol{\varphi}) = \frac{\delta\big(E - H(\boldsymbol{\varphi})\big)}{D(E)} = \frac{\delta\big(E - H(\boldsymbol{\varphi})\big)}{\int d\mu\, \delta\big(E - H(\boldsymbol{\varphi})\big)} \quad , \tag{3.84}$$

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

$$\big\langle f(\boldsymbol{\varphi})\big\rangle_E = \frac{1}{D(E)}\int\limits_{\mathcal{S}_E} d\Sigma_E\, f(\boldsymbol{\varphi}) \quad , \tag{3.85}$$

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 $\varrho(\varphi, t) \to \varrho^{\rm eq}(\boldsymbol{\varphi})$, for consider the following motion on the toroidal space $\big\{\boldsymbol{\varphi} = (q,p)\,\big|\, 0 \le q < 1 \quad , \; 0 \le p < 1\big\}$, 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 $p = p_0 + \alpha(q - q_0)$.

[10] 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.

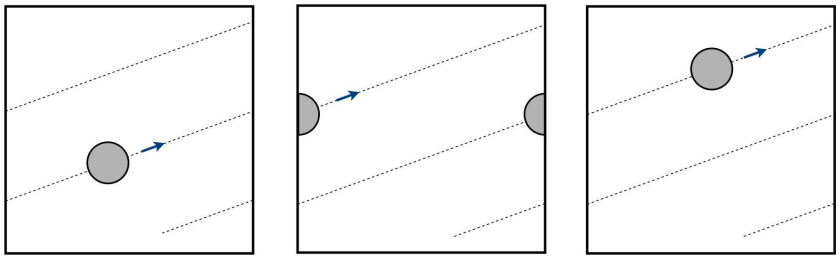

Figure 3.7: Constant phase space velocity at an irrational angle over a toroidal phase space is ergodic, but not mixing. A circle remains a circle, and a blob remains a blob.

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

$$f(q,p) = \sum_{m,n} \hat{f}_{m,n}\, e^{2\pi i(mq+np)} \quad . \tag{3.86}$$

We have, then,

$$f\big(q(t),p(t)\big) = \sum_{m,n} \hat{f}_{m,n}\, e^{2\pi i(mq_0+np_0)}\, e^{2\pi i(m+\alpha n)t} \quad . \tag{3.87}$$

We can now perform the time average of f:

$$\begin{aligned} \big\langle f(q,p)\big\rangle_t &= \hat{f}_{0,0} + \lim_{T\to\infty}\, \frac{1}{T} {\sum_{m,n}}' \hat{f}_{m,n}\, e^{2\pi i(mq_0+np_0)}\, \frac{e^{2\pi i(m+\alpha n)T}-1}{2\pi i(m+\alpha n)} \\ &= \hat{f}_{0,0} \quad \text{if } \alpha \notin \mathbb{Q} \quad . \end{aligned} \tag{3.88}$$

Clearly,

$$\big\langle f(q,p)\big\rangle_E = \int_0^1 dq \int_0^1 dp\, f(q,p) = \hat{f}_{0,0} = \big\langle f(q,p)\big\rangle_t \quad , \tag{3.89}$$

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 $\exp\big\{2\pi i(m+\alpha n)t\big\} = 1$ whenever $(m,n) = (kr,-ks)$ for any $k \in \mathbb{Z}$. Thus,

$$\big\langle f(q,p)\big\rangle_t = \sum_{k=-\infty}^{\infty} \hat{f}_{kr,-ks}\, e^{2\pi ik(rq_0-sp_0)} \quad , \tag{3.90}$$

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

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

$$\varrho(q,p,t=0) = \Theta\big(a^2 - (q-q_0)^2 - (p-p_0)^2\big) \quad , \tag{3.91}$$

then it remains the characteristic function of a disc:

$$\varrho(q,p,t) = \Theta\big(a^2 - (q-q_0-t)^2 - (p-p_0-\alpha t)^2\big) \quad , \tag{3.92}$$

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

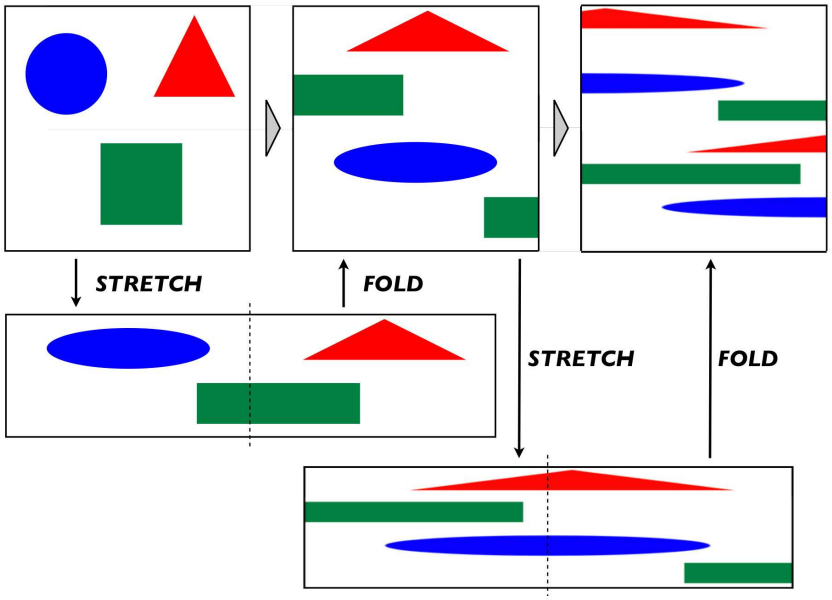

Figure 3.8: The baker's transformation is a successive stretching, cutting, and restacking.

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

$$\nu(A) = \frac{\int d\Sigma_E\, \chi_A(\boldsymbol{\varphi})}{\int d\Sigma_E} = \frac{D_A(E)}{D(E)} \quad , \tag{3.93}$$

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

$$\text{mixing} \qquad \Longleftrightarrow \qquad \lim_{n\to\infty} \nu\Big(g^n A \cap B\Big) = \nu(A)\,\nu(B) \quad . \tag{3.94}$$

In other words, the fraction of B covered by the n^{th} 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

$$\begin{aligned} \big\langle f(\boldsymbol{\varphi})\big\rangle &= \int d\mu\, \varrho(\boldsymbol{\varphi}, t)\, f(\boldsymbol{\varphi}) \xrightarrow[t\to\infty]{} \int d\mu\, f(\boldsymbol{\varphi})\, \delta\big(E - H(\boldsymbol{\varphi})\big) \Big/ \int d\mu\, \delta\big(E - H(\boldsymbol{\varphi})\big) \\ &\equiv \mathsf{Tr}\Big[f(\boldsymbol{\varphi})\, \delta\big(E - H(\boldsymbol{\varphi})\big)\Big] \Big/ \mathsf{Tr}\Big[\delta\big(E - H(\boldsymbol{\varphi})\big)\Big] \quad . \end{aligned} \tag{3.95}$$

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.

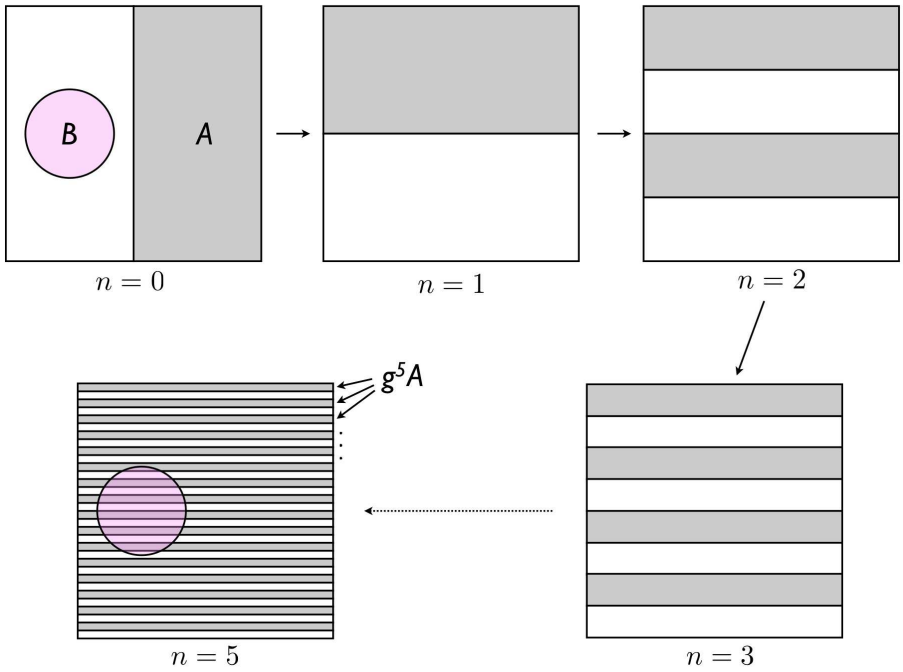

Figure 3.9: The multiply iterated baker's transformation. The set A covers half the phase space and its area is preserved under the map. Initially, the fraction of B covered by A is zero. After many iterations, the fraction of B covered by g^nA approaches $\frac{1}{2}$.

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

$$g(q,p) = \begin{cases} \left(2q\,,\,\frac{1}{2}p\right) & \text{if} \quad 0 \le q < \frac{1}{2} \\ \left(2q-1\,,\,\frac{1}{2}p+\frac{1}{2}\right) & \text{if} \quad \frac{1}{2} \le q < 1 \end{cases} \quad . \tag{3.96}$$

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,

$$g^{-1}(q,p) = \begin{cases} \left(\frac{1}{2}q\,,\,2p\right) & \text{if} \quad 0 \le p < \frac{1}{2} \\ \left(\frac{1}{2}q+\frac{1}{2}\,,\,2p-1\right) & \text{if} \quad \frac{1}{2} \le p < 1 \end{cases} \quad . \tag{3.97}$$

Another example of a mixing system is Arnold's 'cat map'[11]

$$g(q,p) = \big(\,[q+p]\,,\,[q+2p]\,\big) \quad , \tag{3.98}$$

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

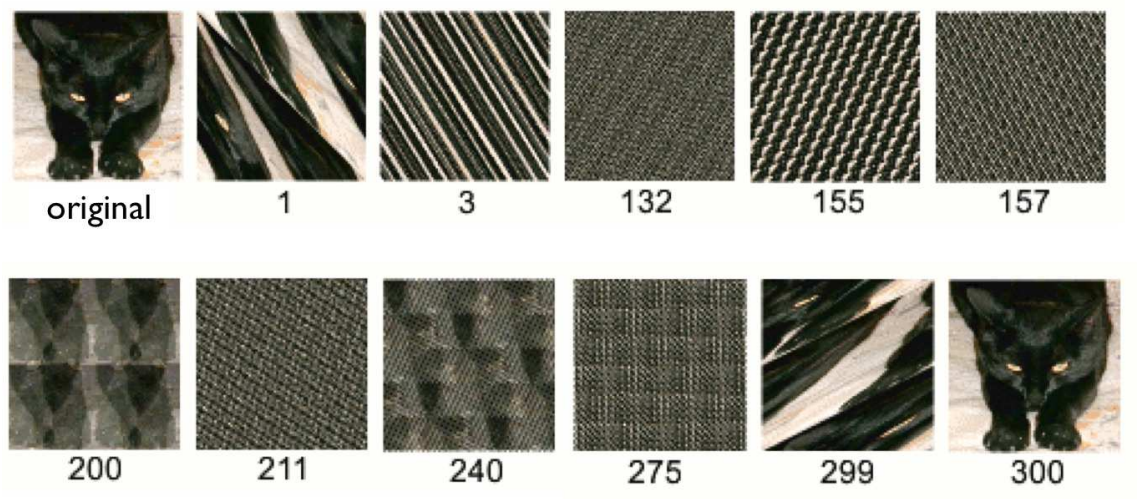

Figure 3.10: The Arnold cat map applied to an image of 150×150 pixels. After 300 iterations, the image repeats itself. (*Source: Wikipedia*)

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

$$\begin{pmatrix} q' \\ p' \end{pmatrix} = \overbrace{\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}}^{M} \begin{pmatrix} q \\ p \end{pmatrix} \mod \mathbb{Z}^2 \quad . \tag{3.99}$$

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

$$\begin{pmatrix} q \\ p \end{pmatrix} = \overbrace{\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}}^{M^{-1}} \begin{pmatrix} q' \\ p' \end{pmatrix} \mod \mathbb{Z}^2 \quad . \tag{3.100}$$

Now for something cool. Suppose that our image consists of a set of discrete points 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. **??**, understanding the characteristic features of each successive refinement[12].

3.6 Appendix : 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

$$\frac{dP_n}{dt} = \sum_m \big(W_{nm}\, P_m - W_{mn}\, P_n \big) \quad . \tag{3.101}$$

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

3.6.1 Poisson process

Consider the *Poisson process*, for which

$$W_{nm} = \begin{cases} \lambda & \text{if } n = m+1 \\ 0 & \text{if } n \neq m+1 \; . \end{cases} \tag{3.102}$$

We then have

$$\frac{dP_n}{dt} = \lambda \big(P_{n-1} - P_n \big) \; . \tag{3.103}$$

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

$$\frac{\partial P}{\partial t} = \lambda (z-1)\, P \quad \Rightarrow \quad P(z,t) = e^{(z-1)\lambda t}\, P(z,0) \; . \tag{3.104}$$

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

$$P_n(t) = \frac{(\lambda t)^n}{n!}\, e^{-\lambda t} \; , \tag{3.105}$$

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

$$\langle n^k \rangle = e^{-\alpha} \left(\alpha \frac{\partial}{\partial \alpha} \right)^k e^\alpha \; . \tag{3.106}$$

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

[12] 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.

3.6.2 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

$$W_{nm} = \begin{cases} 0 & \text{if } n \geq m \\ m\gamma & \text{if } n = m-1 \\ 0 & \text{if } n < m-1 \end{cases} \quad . \tag{3.107}$$

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

$$\frac{dP_n}{dt} = (n+1)\,\gamma\,P_{n+1} - n\,\gamma\,P_n \quad . \tag{3.108}$$

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

$$\begin{aligned} \frac{dN}{dt} &= \sum_{n=0}^\infty n\Big[(n+1)\,\gamma\,P_{n+1} - n\,\gamma\,P_n\Big] \\ &= \gamma\sum_{n=0}^\infty \Big[n(n-1)\,P_n - n^2 P_n\Big] = -\gamma\sum_{n=0}^\infty n\,P_n = -\gamma\,N \quad . \end{aligned} \tag{3.109}$$

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

We can go a bit farther here. Let us define

$$P(z,t) \equiv \sum_{n=0}^\infty z^n\,P_n(t) \quad . \tag{3.110}$$

This is sometimes called a *generating function*. Then

$$\begin{aligned} \frac{\partial P}{\partial t} &= \gamma\sum_{n=0}^\infty z^n\Big[(n+1)\,P_{n+1} - n\,P_n\Big] \\ &= \gamma\,\frac{\partial P}{\partial z} - \gamma z\,\frac{\partial P}{\partial z} \quad . \end{aligned} \tag{3.111}$$

Thus,

$$\frac{1}{\gamma}\,\frac{\partial P}{\partial t} - (1-z)\,\frac{\partial P}{\partial z} = 0 \quad . \tag{3.112}$$

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

$$P(z,t) = f\big(\gamma t - \ln(1-z)\big) \quad . \tag{3.113}$$

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

$$P(z, t) = P\big(1 + (z - 1)\, e^{-\gamma t}\,,\, 0\big) \quad . \tag{3.114}$$

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

$$N(t) = \sum_{n=0}^{\infty} n\, P_n(t) = \left.\frac{\partial P}{\partial z}\right|_{z=1} = \; e^{-\gamma t}\, P(1, 0) = N(0)\, e^{-\gamma t} \quad . \tag{3.115}$$