2. Shift Maps and Chaos in Conservative Systems

This content has been downloaded from IOPscience. Please scroll down to see the full text.

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 141.211.4.224
This content was downloaded on 09/05/2016 at 03:31

Please note that terms and conditions apply.
CHAPTER 2

Shift Maps and Chaos in Conservative Systems

2.1. Bernoulli shifts, base $\mu$ arithmetic and everywhere dense sets of unstable periodic orbits

The simplest deterministic system yielding statistical mechanical behavior is a one-dimensional (hence non-conservative) map, the Bernoulli shift [1-3]. This map is intimately related to chaos in conservative systems via symbolic dynamics. Consider the map

$$x_{n+1} = D x_n \mod 1$$

(2.1)

with $x_0 \in [0, 1]$, a discrete map on the circle of unit circumference. The solution is

$$x_n = D^n x_0 \mod 1$$

(2.2)

where $x_0$ is the initial condition that must be specified. Since $x_n \to 0$ as $n \to \infty$ for $|D| < 1$, the origin is here a stable fixed point. If $D > 1$, there are no stable fixed points since $\delta x_n = D^n \delta x_0$; all perturbations $\delta x_0$ iterate away from the origin. For $D = 2$, the binary shift, $x_{n+1} = 2x_n \mod 1$ is given by

$$x_{n+1} = \begin{cases} 2x_n & 0 \leq x_n < 1/2 \\ 2x_n - 1 & 1/2 \leq x_n < 1 \end{cases}$$

(2.3)

In order to understand the discrete dynamics and to make contact with “computer arithmetic”, it is useful to do the arithmetic in some definite base $\mu$:

$$x_0 = \sum_{\ell=0}^{\infty} e_{2\ell}/\mu^{2\ell}, \quad e_{2, 0} = 0, 1, 2, \ldots, \mu - 1$$

(2.4)

The proof that such expansions are complete in the unit interval is given in Niven [4]. E.g., $\mu = 10$, with $e_i = 0, 1, 2, \ldots, 9$ yields the everyday decimal expansion where, e.g.,

$$0.10256 = \frac{1}{10} + \frac{2}{10^2} + \frac{5}{10^4} + \frac{6}{10^7} + \frac{0}{10^8} + \ldots$$

(2.5)

whereas $\mu = 2$ yields binary arithmetic, the typical computer language. E.g., in binary

$$3/4 = 1/2 + 1/2^2 + 0/2^4 + \ldots = .11000 \ldots$$

(2.6)

It is easy to find the binary expansion of any rational number $x_0 \in [0, 1]$. Consider $x_0 = 5/7$, e.g. The rule for forming the binary expansion is illustrated as follows:

$$2 \times 5/7 = 10/7 = 1 + 3/7 \quad so \quad e_1 = 1$$

(2.7)

$$2 \times 3/7 = 6/7 \quad e_2 = 0$$

(2.8)

$$2 \times 6/7 = 1 + 5/7 \quad e_3 = 1$$

(2.9)

and from here there is periodicity, so that

$$5/7 = 1/2 + 0/2^2 + 1/2^3 + 1/2^7 + \ldots$$

(2.9b)

**Exercise:** Obtain the binary expansion for $x_0 = 3/13$.

**Solution:**

$$2 \times 3/13 = 6/13 \quad e_1 = 0$$

$$2 \times 10/13 = 1 + 7/13 \quad e_2 = 1$$

$$2 \times 6/13 = 12/13 \quad e_2 = 0$$

$$2 \times 7/13 = 1 + 1/13 \quad e_3 = 1$$

$$2 \times 12/13 = 1 + 11/13 \quad e_3 = 0$$

$$2 \times 1/13 = 2/13 \quad e_4 = 0$$

$$2 \times 11/13 = 1 + 9/13 \quad e_4 = 1$$

$$2 \times 2/13 = 4/13 \quad e_5 = 0$$

$$2 \times 9/13 = 1 + 5/13 \quad e_5 = 1$$

$$2 \times 4/13 = 8/13 \quad e_6 = 0$$

$$2 \times 5/13 = 10/13 \quad e_6 = 0$$

$$2 \times 8/13 = 1 + 3/13 \quad e_7 = 1$$

(Then periodic)

$$3/13 = .0011011001001110110001 \ldots$$

**Basic periodic block**

We will show quite generally that rational numbers have periodic expansions in every integral base $\mu$ and that this periodicity is the source of the dense set of unstable periodic orbits in Bernoulli shifts. It is important to understand this fact because dense sets of unstable periodic orbits are important for chaos theory in general. A rational number is a number $x_0 = p/q$ with $p$ and $q$ both integers. Irrational numbers, e.g., numbers like $\sqrt{2}, e$, or $\pi$ cannot be expressed as the ratio of two integers. Rational numbers are countable, dense in $[0, 1]$, but occupy no space in the unit interval. Irrational numbers occupy all the space between 0 and 1 (cf. Niven).

Given that the rationals are countable, it is easy to show
that they occupy no space in [0, 1]. Label the rationals by integers 1, 2, 3, . . . , and centered upon the nth rational draw a one dimensional “disc” of “radius” $\varepsilon/2^n$:

The total length in [1, 0] occupied by the rationals is then less than

$$L = \sum_{n=1}^{\infty} \varepsilon/2^n = -\varepsilon + \varepsilon(1 - 1/2) = \varepsilon$$

which can be made as small as we like since $\varepsilon$ is arbitrary. Hence, $L = 0$ and all the space is occupied by the irrationals. The irrationals were proven uncountable by G. Cantor. The rationals are dense in [1, 0], which means that between any two points in [1, 0] there exist infinitely many rationals. The rationals are countable, dense, and have Lebesque measure zero in [0, 1].

It is easy to show that a rational number is periodic in every base $\mu = 2, 3, 4, . . .$ and vice-versa. To see this,

let

$$x = \sum_{\varepsilon_i = 0, 1, 2, . . ., \mu - 1 = 0; \varepsilon_i}^{\varepsilon_i} \varepsilon_i/\mu^i,$$

Given $x$ in the form $p/q$, we construct the base-$\mu$ expansion as follows. First we write

$$\mu p/q = [\mu p/q] + p_1/q$$

when $[\mu p/q] = \text{integer part of } p/q$ and $p_1/q$ is the remainder. Then $\varepsilon_i = [\mu p/q]$, and we construct the sequence

$$\mu p_{i+1}/q = [\mu p_i/q] + p_{i+1}/q \varepsilon_i = [\mu p_i/q]$$

Because there are at most $q$ choices for each $p_i$, one of the $p_i$ must eventually repeat for $i \leq q$. Thereafter, the resulting sequence of integers is periodic, so that we can write

$$p/q = \varepsilon_0 \varepsilon_2 \ldots \varepsilon_i \varepsilon_{i+1} \varepsilon_{i+2} \ldots \varepsilon_{N-1} \varepsilon_N \ldots$$

where $\varepsilon_0 \varepsilon_2 \ldots \varepsilon_i$ is the basic periodic string and $\varepsilon_{i+1} \varepsilon_{i+2} \ldots \varepsilon_{N-1} \varepsilon_N$ represents an initial string that is not repeated (this is illustrated below by computing the binary sequence for 1/6). In base $\mu$, a factor $1/\mu^i$ in $p/q$ will not be repeated. So, $N$ is the “period” of $p/q$ in base $\mu$ and the algorithm that was used to obtain the expansion is just the Euclidean algorithm, which forms the basis for ordinary division. It follows that irrational numbers are non-periodic in every base $\mu$, and this fact is essential for distinguishing true chaotic orbits of Bernoulli shifts from unstable periodic ones.

**Exercise:** Obtain the binary representation of 1/3.

**Solution:**

$$2 \times 1/3 = 2/3 \quad \text{so} \quad \varepsilon_1 = 0$$

$$2 \times 2/3 = 1 + 1/3 \quad \varepsilon_2 = 1$$

and periodicity occurs thereafter, yielding

$$1/3 = .01010101 \ldots$$

**Exercise:** Construct the binary expansion for $x = 1/6 = 1/2 \times 3$

**Solution:** $2 \times 1/6 = 1/3$ so $\varepsilon_1 = 0$ and the remaining bits are the same as for the 1/3-sequence:

$$1/6 = .001010101 \ldots$$

(periodicity begins after one binary point shift).

**Exercise:** Construct the ternary representation

$$x = \sum_{n=1}^{\infty} \varepsilon_n/3^n, \quad \varepsilon_n = 0, 1, \text{ or } 2$$

for $x = 1/9$.

**Solution:** $1/9 = .0100 \ldots$ (Trivial periodicity occurs after 3 ternary shifts).

We will prove now that if an expansion in base $\mu$ is periodic, then the expansion necessarily represents a rational number. Assume that (with $x \in [1, 0]$)

$$x = \sum_{n=1}^{\infty} \varepsilon_n/\mu^n = .\varepsilon_1 \varepsilon_2 \ldots \varepsilon_N \ldots$$

is periodic with period $N$, where $\varepsilon_n = 0, 1, 2, \ldots, \mu - 1$. Then

$$\mu^N x = \varepsilon_1 \varepsilon_2 \ldots \varepsilon_N + \varepsilon_1 \varepsilon_2 \ldots \varepsilon_N \ldots = x + x$$

where

$$x = \varepsilon_1 \varepsilon_2 \ldots \varepsilon_N = \sum_{n=0}^{N-1} \varepsilon_{N-n} \mu^n$$

is an integer, so that

$$x = \frac{x}{\mu^N - 1}$$

is the ratio of two integers. It is left for the reader to prove a similar result whenever the rational has an “initial transient”, a block of digits that is never repeated.

**Exercise:** Obtain the general form of rational numbers $x \in [0, 1]$ that constitute the lattice of a binary computer that uses only fixed-point operations [5] (i.e., floating point arithmetic is not used).

**Solution:** All such numbers are of the form

$$x = \sum_{n=0}^{N} \varepsilon_n/2^n, \quad \varepsilon_i = 0 \text{ or } 1.$$
we have $x = 2^a$, e.g., if $x = 7/8$, $x = 1 + 2 + 2^2$, so that in binary, $x = 1.110.0$, and $x = .1110000..$, since multiplication by 1/2 merely shifts the binary point one unit to the left.

For an arbitrary rational number $x = p/q$ in base $\mu = 2$, we have

$$x_0 = \ldots e_n e_{n-1} \ldots e_1 e_0 \ldots$$

where $N$ is the period. With this particular representation, the Bernoulli shift $x_{n+1} = 2x_n \mod 1$ amounts to a binary shift one place to the right, dropping the integral part:

$$x_1 = \ldots e_n e_{n-1} \ldots e_1 e_0 \ldots$$

$$x_2 = \ldots e_n e_{n-1} \ldots e_1 e_0 \ldots$$

and it is clear that the string $e_n e_{n-1} \ldots e_1 e_0$ repeats systematically. Periodic shifts of shifts are (1) dense in the unit interval because rational numbers are dense but (2) unstable because of the exponential magnification of errors. If we know $x_0$ only to within an error $\delta x_0$, then the error after $n$ iterations becomes

$$\delta x_n = 2^n \delta x_0 \mod 1$$

$$= e^{-n/2} \delta x_0 \mod 1$$

where $\lambda = \ln 2 > 0$ is called the Liapunov exponent [1, 2]. If $\delta x_0$ is irrational, then for $\delta x_0 \approx 1$, we have lost all information as to the location in $[0, 1]$ of $x_0$, and the relaxation time for loss of information (or approach to statistical equilibrium) is given by $1 \sim 2^n/\delta x_0$, or

$$n \approx -\ln \delta x_0/\ln 2 = -\ln \delta x_0/\lambda.$$

For $x_0$ irrational, the closure of the orbit as $n \to \infty$, the attractor, is the entire unit interval. The question then arises what is the probability density for the iterates? According to Borel [6], almost all irrationals, the so-called “normal numbers”, yield a uniform distribution of iterates $p(x) = 1$ as $n \to \infty$. This means that the Bernoulli shift is “mixing” in the statistical mechanical sense, for almost all initial conditions. The idea of mixing is illustrated as follows. Let us consider some initial spread of uncertainty in $x_0$ defined by an initial probability density $p_0(x)$, e.g., the simplest case is given by

$$p_0(x) = \begin{cases} 1/\delta x_0, & x \in [x_0, x_0 + \delta x_0] \\ 0, & \text{otherwise} \end{cases}$$

although the following argument does not require such details of $p_0(x)$. If we think of the segment of the $x$-axis with length $\delta x_0$ as a droplet of red dye, and the unit interval as otherwise clear (one-dimensional) body of water, than as $n \to \infty$ the Bernoulli shift colours the water uniformly red. Thus “the” $n = \infty$ limit is the limit of statistical equilibrium. Mathematically, it is the positive Liapunov exponent $\lambda = \ln 2$ that stretches the initially localized droplet of dye into a long, thin filament that wraps round and round the circle of unit circumference.

We can describe the mixing property by the introduction of a “master-equation” for the probability density $p_n(x)$ that evolves after $n$ iterations from some specified initial density $p_0(x)$ [1, 2], Exercise: Show that the above described statistical behavior cannot occur on a binary computer where $x_0$ is represented by a finite binary number $x_n = \sum_{j=1}^{N} e_j/2^j = e_1 e_2 \ldots e_N 00 \ldots$

Solution: $2^n x_0 = 0 \mod 1$, so that all iterations hop in finite

This master equation expresses the fact that both $2x_{n+1}$ and $2x_{n+1} - 1$ equal $x_n$ (the map is deterministic), but there is a 50/50 chance that $x_n$ actually arose from one or the other. This result can be derived formally from the Frobenius-Perron equation [1, 7] as follows, which shows that no extra probabilistic assumptions are needed:

$$p_n(x) = \int_0^1 dp_{n-1}(y) \delta(x - f(y))$$

where $x_n = f(x_{n-1})$ is the Bernoulli shift, and

$$\delta(x - f(y)) = \sum_{j=1}^{n} \frac{1}{f'(x_j)} \delta(y - x_j)$$

where $f'(x_j) = 2$ and, the points $x_j$ are the two points that iterate via the map $f$ to the same point $x$. Note that an invariant density $p(x)$ is a fixed point solution of the Frobenius-Perron equation. A stable fixed point solution corresponds to an irreversible approach of the system to equilibrium. We can prove that an arbitrary smooth initial density $p_0(x)$ evolves irreversibly to $p(x) = 1$ as follows:

$$p_0(x) = \sum_{j=0}^{N-1} \int_{x_{j+1}}^{x_j} dp_{n-1}(x)$$

$$= \int_{x_{N-1}}^{x_0} dp_{n-1}(x) = \frac{1}{2^n}$$

so that $p_n(x)$ is an average of the values of $p_0$ at $2^n$ different points between $x/2^n \sim 0$ and $((x + 2^n - 1)/2^n \sim 1$, as $n \to \infty$. To go further, notice that the sum has the form of a Riemann integral:

$$\int_{x_{N-1}}^{x_0} dp_{n-1}(x) = \frac{1}{2^n}$$

and the range of $x_j$ (as $n \to \infty$) is from 0 to 1. Therefore, for $n \to \infty$, $p_n(x) = \sum_{j=0}^{N-1} \int_{x_{j+1}}^{x_j} p_0(x) \, dx = 1.$

Exercise: Show that the above described statistical behavior cannot occur on a binary computer where $x_0$ is represented by a finite binary number $x_0 = \sum_{j=1}^{N} \epsilon_j/2^j = \epsilon_1 \epsilon_2 \ldots \epsilon_N 00 \ldots$

Solution: $2^n x_0 = 0 \mod 1$, so that all iterations hop in finite

We generalize the above argument for the binary map $f(x) = 2x \mod 1$ to a more general map $f(x) = \mu x \mod 1$ where $\mu$ is any positive number. Consider the map $f(x) = \mu x \mod 1$ where $\mu$ is any positive number. We have $f(x) = (\mu x \mod 1 + 1)/2$. The probability density $p(x)$ that evolves after $n$ iterations from some specified initial density $p_0(x)$ given by

$$p_n(x) = \int_0^1 dp_{n-1}(y) \delta(x - f(y))$$

This master equation expresses the fact that both $2x_{n+1}$ and $2x_{n+1} - 1$ equal $x_n$ (the map is deterministic), but there is a 50/50 chance that $x_n$ actually arose from one or the other. This result can be derived formally from the Frobenius-Perron equation [1, 7] as follows, which shows that no extra probabilistic assumptions are needed:

$$p_n(x) = \int_0^1 dp_{n-1}(y) \delta(x - f(y))$$

where $x_n = f(x_{n-1})$ is the Bernoulli shift, and

$$\delta(x - f(y)) = \sum_{j=1}^{n} \frac{1}{f'(x_j)} \delta(y - x_j)$$

where $f'(x_j) = \mu$ and, the points $x_j$ are the two points that iterate via the map $f$ to the same point $x$. Note that an invariant density $p(x)$ is a fixed point solution of the Frobenius-Perron equation. A stable fixed point solution corresponds to an irreversible approach of the system to equilibrium. We can prove that an arbitrary smooth initial density $p_0(x)$ evolves irreversibly to $p(x) = 1$ as follows:

$$p_0(x) = \sum_{j=0}^{N-1} \int_{x_{j+1}}^{x_j} dp_{n-1}(x)$$

$$= \int_{x_{N-1}}^{x_0} dp_{n-1}(x) = \frac{1}{2^n}$$

so that $p_n(x)$ is an average of the values of $p_0$ at $2^n$ different points between $x/2^n \sim 0$ and $((x + 2^n - 1)/2^n \sim 1$, as $n \to \infty$. To go further, notice that the sum has the form of a Riemann integral:

$$\int_{x_{N-1}}^{x_0} dp_{n-1}(x) = \frac{1}{2^n}$$

and the range of $x_j$ (as $n \to \infty$) is from 0 to 1. Therefore, for $n \to \infty$, $p_n(x) = \sum_{j=0}^{N-1} \int_{x_{j+1}}^{x_j} p_0(x) \, dx = 1.$

Exercise: Show that the above described statistical behavior cannot occur on a binary computer where $x_0$ is represented by a finite binary number $x_0 = \sum_{j=1}^{N} \epsilon_j/2^j = \epsilon_1 \epsilon_2 \ldots \epsilon_N 00 \ldots$

Solution: $2^n x_0 = 0 \mod 1$, so that all iterations hop in finite
time onto the unstable fixed point $x = 0$. This behavior is typical of any Apple, IBM, Commodore, VAX, etc., since all of these computers perform binary arithmetic.

**Exercise:** Study the Bernoulli shift

$$x_{n+1} = 2x_n \mod 1$$

on your pocket calculator. For typical calculators (Hewlett-Packard, T.I., Sharp, etc.) one does not obtain $x_n = 0$ for finite $n$. Explain why this is so.

**Solution:** Typical pocket calculators do base-10 arithmetic, so that one obtains non-trivial unstable periodic orbits with $x_{n+1} = 2x_n \mod 1$. In contrast, the solutions of $x_{n+1} = 10x_n \mod 1$ will iterate to 0 in finite time on such machines. The typical hand-held calculator is more accurate than a typical (single-precision) microcomputer because the former typically carries $N \sim 12$ base 10 digits while the latter typically carries $N \leq 30$ base 2 digits (bits). With $2^N \sim 10^3$, $N \sim 12 \ln 10 / \ln 2 \sim 40$ bits are required for a binary machine if it is to have precision equal to that of typical pocket calculators.

In summary, if we consider the Bernoulli map

$$x_{n+1} = \mu x_n \mod 1,$$  \hspace{1cm} (2.35)

it is called a "shift" because in base $\mu$ arithmetic with $x_0 = .e_1 e_2 \ldots e_N \ldots$ where $e_i = 0, 1, 2, \ldots, \mu - 1$, each iteration of the map corresponds simply to a shift of the base point one place to the right, discarding the integer part:

$$x_0 = .e_1 e_2 \ldots$$

$$x_1 = .e_2 e_3 \ldots$$

$$x_2 = .e_3 e_4 \ldots$$

$$\ldots$$

$$x_N = .e_{N+1} e_{N+2} \ldots .$$  \hspace{1cm} (2.36)

Hence, if $x_0$ is rational then the orbit is periodic and the instability of a nontrivial periodic orbit is reflected in the continued transformation of least significant bits into most significant bits. If $x_0$ is irrational, there is no periodicity and we obtain a true chaotic orbit. The orbit is chaotic because bits that are initially insignificant are transformed continually into most significant bits as $n$ increases, and there is no periodicity. The question whether such motion is in any sense "random" is discussed in Chapter 5.

Since rationals occur with measure 0 and irrationals with measure 1, it would be nice to be able to conclude that the chance of randomly drawing an irrational from the continuum is certain. In fact, the opposite is true: any number $x = \sum_{i=1}^{\infty} e_i / \mu^i$ (2.37)

that is constructed by a finite number of operations is rational, and we are both in practice and in principle confined to the use of a finite number of operations in computation. To write down the general $n$th term of any given rational is possible, because the $e_i$ are periodic. Also, if we toss a fair coin $N$ times, the outcome defines a rational number because we can let $T \rightarrow 0$ and $H \rightarrow 1$ define a binary sequence in a heads-tails game.

If $x_0$ is irrational, there is no periodicity, and the best that one can do is to compute each digit, one after the other, by means of some algorithm. This leads to the interesting question in what sense are computed orbits chaotic, which we have yet to discuss. The connection between irrationals and the game of heads and tails is discussed in Kac [6] and Niven [4].

### 2.2. Mixing and ergodicity

According to Gibbs, who gave us modern statistical mechanics [8], ordinary household mixing is the proper analogy for the approach of a conservative dynamical system to statistical equilibrium [9, 10, 17]. Consider a glass containing two fluids, water and a small droplet of red dye (those who need no colours for the purpose of visualization are free to substitute a jigger of scotch or bourbon). Both fluids are approximately incompressible, so that the volume of each fluid is separately conserved, and both are approximately non-viscous in the sense that we can ignore the effects of molecular diffusion of one fluid across the boundary of another; molecular diffusion alone is too slow to be of any practical use in mixing the two fluids together.

Initially, the dye is localized within a small region $A$. Upon slow stirring, the dye is drawn into a long, thin filament which we denote by $A_1 = U_A$ where $U_A$ is the time-evolution operator representing the stirring. For $t \gg 0$, the stirring causes this filament to spread more or less uniformly throughout the glass of water, so that the entire fluid eventually appears uniformly reddish in color. Here, the human eye performs the "coarse-graining", which is the averaging required to give the appearance of the irreversible tendency toward statistical equilibrium. If one looks at the mixture on a small enough length scale, before molecular diffusion has had time to act,
one will see that in reality there is a very long, thin, branching filament that threads chaotically throughout the water. Gibbs discussed such an example as the explanation for the approach to statistical equilibrium, and we know now that Liapunov exponents provide a definite mechanism for such filamentation. The droplet of dye is analogous to our initial data on the surface of the torus. The Liapunov exponents provide the exact analogy in Liouville's Theorem. A uniformly red fluid corresponds to the state of our maximum ignorance, statistical equilibrium, and the Hamiltonian system is presumably a non-integrable (completely unstable) one.

The mixing condition can be formulated mathematically as follows [10]:

If \( B \) is any region in the fluid with volume \( V_B \), then after a long time \( t \gg 0 \), the fraction of dye that threads throughout \( B \) should, with high probability, equal the fraction of the total space occupied by \( B \):

\[
\lim_{t \to \infty} \frac{V_{A \cap B}}{V_A} = \frac{V_B}{V_T} \tag{2.38}
\]

where \( A \cap B \) denotes the intersection of the two point sets \( A \) and \( B \), \( A = U \cup A \) and \( V_T \) is the total volume of fluid (size of the bounded phase space). So, if the region \( B \) occupies 40% of the total volume (\( V_B/V_T = 0.4 \)), then with high probability, one should find about 40% of the dye threading throughout \( B \) as \( t \to \infty \). If \( \mu(A, \cap B) \) denotes the measure of the intersection, i.e.,

\[
\mu(A, \cap B) = \frac{V_{A \cap B}}{V_T}, \tag{2.39}
\]

then mixing corresponds to

\[
\lim_{t \to \infty} \mu(A, \cap B) = \mu(A) \mu(B) \tag{2.40}
\]

where \( \mu(A) = V_A/V_T \) and \( \mu(B) = V_B/V_T \) represent the fractions of the total space occupied by \( A \) and \( B \) respectively. It is worthwhile to note that 0, Reynolds's test for turbulence in a fluid was to inject a small amount of dye and then look for mixing. If the fluid motion mixes the dye, then the fluid is turbulent [12]. The mixing provided by the Bernoulli shift and the related bakers' transformation represents a laminar filamentation. Turbulence requires vorticity, and chaotic vortex dynamics seem necessary for truly turbulent behavior. The examples of mixing illustrated below by the bakers' transformation do not have any chaotic vortices or eddies in phase space, hence are not-turbulent. In real fluids, turbulent filamentation mixes 2 fluids together rapidly. In the design of internal combustion engines, e.g., care must be taken to design the combustion chamber so as to insure that turbulent mixing occurs for better efficiency.

We have previously given an example of a dynamical system which is ergodic but non-mixing: quasi-periodic motion on a torus that arises from irrational frequency ratios. In that case, time-averages equal space averages over the surface of the torus (ergodic motion), but the dynamical equations will not cause any filamentation of a small "droplet" of initial data on the surface of the torus. The Liapunov exponents for that system are all zero, and there is in that case no other mechanism to cause a droplet to undergo filamentation. It is possible in other cases that there may be mixing without such exponents, and this possibility is left open in our formulation of Liapunov exponents for discrete maps in the next section.

Regarding the relation of mixing to ergodicity, Arnol'd and Avez [10] show that a mixing system has no non-trivial invariant sub-space: if we set \( A = B \), then

\[
\lim_{t \to \infty} \mu(A, \cap A) = \mu(A) \mu(A) \quad \text{(mixing condition)} \tag{2.41}
\]

but on the other hand, \( A = B = A \), so that

\[
\mu(A, \cap A) = \mu(A), \tag{2.42}
\]

yielding \( \mu(A)(1 - \mu(A)) = 0 \). Hence, either \( A = \) the entire bounded phase space (\( \mu(A) = 1 \)) or else \( A \) is empty (\( \mu(A) = 0 \)). According to Arnol'd and Avez, this means that mixing implies ergodicity. Let us see how this works, using the language of probabilities in one dimension.

The idea is that uniform mixing implies ergodicity, so we want to understand how time averages can be replaced by phase space averages when mixing holds. Consider a one-dimensional example, for simplicity. A time average is then defined as follows:

\[
\tilde{x}_n = \frac{1}{n} \sum_{i=1}^{n} x_i \quad \text{and} \quad \tilde{x}_n \Delta x \quad \text{yielding} \quad \mu(\Delta x) \approx \mu(x) \tag{2.43}
\]

where (with a 1-D map \( x_{n+1} = f(x_n) \)) is the “time” and \( n \) is the number of iterations falling into the \( i \)-th bin, in some finite subdivision of the unit interval (phase space) into \( N \) bins of equal width.

\( \tilde{x} \) is the approximate value of \( x \) in the \( i \)-th bin. Mixing means that the iterations fall into the \( i \)-th bin in the ratio

\[
\frac{n_i}{n} \approx \mu(\Delta x) \tag{2.44}
\]

where \( \Delta x \) is the size of the bin, and uniform mixing corresponds to \( \mu(\Delta x) = \Delta x \), yielding

\[
\tilde{x}_n \approx \frac{1}{n} \sum_{i=1}^{n} x_i \Delta x \approx \sum_{i=1}^{n} \tilde{x}_i \Delta x = \int_0^1 x \, dx \tag{2.45}
\]

More generally, we would obtain \( \mu(dx) = p(x) \, dx \) where \( p(x) \) is the invariant density of the map. In our example we assumed \( p(x) = 1 \), which is true of all Bernoulli shifts with normal numbers as initial conditions. In order to understand a mechanism for mixing and for deterministic chaos in general, it is useful to introduce the notion of a set of Liapunov exponents.

2.3. Local exponential instability of orbits: Liapunov exponents

In order to find some conceptual ground to stand on, let us return for the moment to the idea of an unstable limit cycle. Such a curve has the property that it repels all nearby trajectories so that they diverge from the limit cycle at an exponential rate. Consider the possibility of a curve that is not a limit cycle but which repels all nearby trajectories in phase space at an exponential rate, and consider the possibility that such repelling trajectories are dense in the system's phase space, such a dense set may be provided by unstable periodic or unstable quasiperiodic orbits. Then for any two repelling trajectories separated initially by the distance \( \delta x(0) \), the
magnitude of the separation at a later time will be
\[ |\delta x(t)| \sim |\delta x(0)| e^{\lambda t} \]  
(2.46)
where \( \lambda \) is called the largest Liapunov exponent and depends
in general upon the location of the trajectory in phase space.
In one dimension, the relation
\[ \lambda = \lim_{t \to \infty} \frac{1}{t} \ln \frac{|\delta x(t)|}{|\delta x(0)|} \]  
(2.47)
defines the Liapunov exponent [1, 2], while a similar relation in
higher dimensions yields only the largest Liapunov exponent. \( \lambda > 0 \)
defines an orbital instability that guarantees
deterministic chaos. The basic idea of deterministic chaos is
that small errors \( \sim |\delta x(0)| \) in initial data are magnified
exponentially fast so that an orbit can be computed only for
small errors \( |\delta x(0)| \rightarrow 0 \)
times \( \lambda \), or \( t \ll \frac{1}{\lambda} \ln (L/|\delta x(0)|) \)  
(2.48)
where \( L \sim \) the extent of the phase space where the (bounded)
chaotic motion occurs.

In general, the equation for \( \lambda \) can be formulated by lin-
erizing the equations of motion about a particular chaotic
orbit, as we will illustrate below for discrete maps in one and
two dimensions.

The presence of a positive Liapunov exponent provides a
mechanism for mixing via filamentation of uncertainty in
initial data. Local exponential instability of trajectories in
phase space was suggested by Max Born as the means by
which a conservative dynamical system might approach
statistical equilibrium; the filamentation mechanism seems
first to have been discussed by Gibbs [8].

In order to formulate analytic expressions for the numeri-
cal computation of Liapunov exponents, let us begin with a
one-dimensional map. If we understand the formulation for
discrete maps, then we can apply it to differential equations
which have been converted into discrete maps for the purpose
of computation. So, we begin with one a dimensional iterated
map
\[ x_{n+1} = f(x_n), \]  
(2.49)
where the iterations \( x_n \) are attracted or confined to a point set
within some bounded interval \( [a, b] \) as \( n \to \infty \) (e.g., if
\( f(x) = 2x \) mod 1, \( [a, b] = [0, 1] \)). The main idea is as
follows: we compute an exact orbit (via a thought experi-
ment, if necessary) \( x_0 \rightarrow x_1 \rightarrow x_2 \rightarrow \ldots \) beginning from
some nearby initial condition \( x_0 + \delta x_0 \), and then linearize
the map about this exact orbit
\[ x_n + \delta x_n = f(x_{n-1} + \delta x_{n-1}) \]
\approx f(x_{n-1}) + f'(x_{n-1}) \delta x_{n-1} \ldots \]  
(2.50)
We then find the error propagation equation
\[ \delta x_n \approx f'(x_{n-1}) \delta x_{n-1} \]  
(2.51)
and since this is a linear equation, we can solve it exactly to obtain
\[ \delta x_n \approx \prod_{i=1}^{n-1} f'(x_{i-1}) \delta x_0 \]  
(2.52)
where \( \delta x_0 \) is the initial separation of the two orbits. With
\[ |\delta x_n| \approx \prod_{i=1}^{n-1} |f'(x_{i-1})| |\delta x_0| \]  
(2.53)
we can define the Liapunov exponent
\[ \lambda = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \ln |f'(x_{i-1})|, \]  
(2.54)
that describes to exponential separation of nearby trajec-
tories, i.e.,
\[ |\delta x_n| \sim |\delta x_0| e^{\lambda n}, \quad n \gg 1. \]  
(2.55)
Naturally, one should take care that, for large \( n, |\delta x_0| \) is small
enough that the neglect of quadratic and higher order terms
is justified. Also, when \( \lambda > 0 \) one will not obtain pointwise
convergence to a definite number \( \lambda \), rather, there will be
fluctuations in \( \lambda \). For an ergodic system with invariant
density \( p(x) \), all time averages can be replaced by phase space
averages and since \( \lambda \) is defined by the “time” average
\[ \lambda = \frac{\sum_{i=1}^{n} \ln |f'(x_{i-1})|}{n}, \quad n \to \infty, \]  
(2.56)
we obtain (as \( n \to \infty \))
\[ \lambda = \int_{a}^{b} dx \, p(x) \ln |f'(x)|, \]  
(2.57)
or \( \lambda = \langle \ln |f'(x)| \rangle \). Given \( p(x) \) one can just as easily com-
pute the root mean square fluctuation
\[ \Delta \lambda \approx \langle (\ln |f'(x)| - \lambda)^2 \rangle^{1/2} \]  
(2.58)
in order to estimate the deterministic noise in \( \lambda \). In order that
the average value \( \lambda \) should be physically meaningful, it is
necessary that \( \Delta \lambda / \lambda \ll 1 \). In this case, one can define a
“relaxation time” to statistical equilibrium as follows. When
\( |\delta x_n| \sim |a - b| \), we have lost all information as to the location
of the orbit; we are in a state of maximum ignorance, which
according to Boltzmann defines the state of statistical equi-
lbrium. The relaxation time for this statistical equilibrium is
estimated as follows:
\[ |a - b| \sim |\delta x_0| e^{\lambda n} \]  
(2.59)
so that
\[ n \sim - (\ln |a - b|/|\delta x_0|)/\lambda \]  
(2.60)
is the required relaxation time. Since we have clearly violated
the linearization condition \( |\delta x_n| \ll |a - b| \), this estimate
must be taken with a small grain of salt. Nevertheless, it
correctly reflects the fact that, the more information you have (the smaller is $|\delta x_0|$, i.e.), the longer it takes before the system relaxes to equilibrium. As we will see in Ch. 5, this feature of chaos defeats any attempt to compute the system’s orbit via fixed-precision arithmetic, but is does not eliminate the possibility that exact chaotic orbits can be computed by more careful means.

If we think of $\delta x_0$ as the error in initial data which is always present in the observations of a physical system, then

$$n \ll - (\ln(a - b/|\delta x_0|))/\lambda$$

(2.61)
is a necessary condition for some crude degree of predictability whenever $\lambda > 0$. Computed orbits that violate this condition are easily generated but cannot be interpreted as orbits of the dynamical system. Such orbits are called pseudo-orbits and are the typical output of floating-point arithmetic computations.

For a two dimensional discrete map

$$x_{n+1} = f(x_n, y_n)$$

$$y_{n+1} = g(x_n, y_n)$$

(2.62)

we can also linearize about some exact orbit $(x_0, y_0) \rightarrow (x_1, y_1) \rightarrow \ldots \rightarrow (x_n, y_n) \rightarrow \ldots$ to obtain the error propagation equation

$$\delta x_{n+1} = A_n \delta x_n$$

(2.63)

where

$$\delta x_n = \begin{pmatrix} \delta x_n \\ \delta y_n \end{pmatrix}$$

(2.64)

and the Jacobi matrix

$$A_n = \begin{pmatrix} \frac{\partial f}{\partial x_n} & \frac{\partial f}{\partial y_n} \\ \frac{\partial g}{\partial x_n} & \frac{\partial g}{\partial y_n} \end{pmatrix}$$

(2.65)
is evaluated for each iteration at the point $(x_n, y_n)$ on the exact trajectory. If one can solve the exact equations for the exact orbit, one can use that information to solve also the linearized equations, we can write

$$\delta x_0 = \sum_{i=1}^n c_i e_i(n)$$

(2.68)

and also obtain

$$\delta x_n = J_n \delta x_0 = \sum_{i=1}^n c_i \mu_i(n) e_i(n).$$

(2.69)

Liapunov exponents $\lambda_1$ and $\lambda_2$ are then defined by assuming

$$\mu_i(n) \sim e^{\lambda_i n}$$

(2.70)

so that

$$\lambda_i = \frac{1}{n} \ln \mu_i(n), \quad n \rightarrow \infty.$$  

(2.71)

$\lambda_i > 0$ corresponds to a local expansion of small areas along $e_i$, while if $\lambda_i < 0$ we have a local contraction of small areas along the direction $e_i(n)$. The assumption that $\mu_i(n) \sim e^{\lambda_i n}$ is made in the spirit of an assumption of scaling behavior in critical phenomena; until one does a definite calculation on a particular model, one does not in general know whether it is true.

Since areas transform as

$$\Delta A_n = |J_n| \Delta A_0,$$  

(2.72)

where $\Delta A_0$ is some initial element of area in phase space and $|J_n| = \det J_n$ is the Jacobian determinant of the transformation $(x_0, y_0) \rightarrow (x_n, y_n)$, we have $\det J_n = \mu_1(n)\mu_2(n)$, and for large $n$ we can write

$$\Delta A_n \approx e^{n(\lambda_1 + \lambda_2)} \Delta A_0.$$  

(2.73)

Hence, $\lambda_1 + \lambda_2 > 0$ corresponds to an area-preserving map while $\lambda_1 + \lambda_2 < 0$ will follow for a contracting map. With $\lambda_1 + \lambda_2 < 0$, deterministic chaos requires at least one positive exponent whereas for regular (“stable”) motion $\lambda_1 \leq 0$ is necessary. The condition $\mu_+ = \mu_2$ is both necessary and sufficient for stable motion in an area-preserving map; this condition defines an elliptic point. As we will see by the following example, the existence of non-zero exponents provide a mechanism whereby filamentation to statistical equilibrium can occur in area-preserving maps. There, since $\lambda_0 = -\lambda_1$, the existence of one exponent guarantees the existence of the other.

2.4. The bakers’ transformation and the approach of a conservative dynamical system to statistical equilibrium

The bakers transformation

$$x_{n+1} = \begin{cases} 2x_n & x_n \in [0, 1/2) \\ y_n/2 & y_n \in [1/2, 1] \end{cases}$$

$$y_{n+1} = \begin{cases} 2x_n - 1 & x_n \in [1/2, 1] \\ (y_n + 1)/2 & y_n \in [1/2, 1] \end{cases}$$

is a mathematical idealization of the process whereby a baker kneads a blob of dough by stretching and folding. This process yields mixing via filamentation, as can be seen graphically by mapping finite regions (cf. Fig. 2.5).

As $n \rightarrow \infty$, the unit square is “painted” by infinitely many thin horizontal coloured filaments. When one can no longer
Fig. 2.5. Mixing via filamentation in the bakers’ transformation; the shaded area represents wet paint and the unshaded area represents an unpainted surface. The baker’s transformation distributes the paint uniformly over the unit square, while conserving the amount of paint.

where
\[ \delta X_n = \begin{pmatrix} \delta x_n \\ \delta y_n \end{pmatrix} \]  
(2.77)

Since \( \delta x_n = e^{n \lambda_1} \delta x_0 \) and \( \delta y_n = e^{-n \lambda_2} \delta y_0 \), we obtain stretching of areas in the x-direction, with Liapunov exponent \( \lambda_1 = \ln 2 > 0 \), and contraction in the y-direction with Liapunov exponent \( \lambda_2 = -\ln 2 < 0 \). The result \( \lambda_1 + \lambda_2 = 0 \), corresponds to the area-preservation property \( J(t) = 1 \), since \( J = (e^{n \lambda_1} e^{-n \lambda_2}) \).

Does the filamentation caused by these Liapunov exponents produce a uniform invariant density \( p(x, y) = 1 \) as \( n \to \infty \), as one might expect? In order to investigate this question from a statistical mechanics viewpoint, it is useful to use the master-equation, which is the equation of motion for \( p_n(x, y) \), the probability density of the iterates. One can derive the equation of motion for \( p_n(x, y) \) from Liouville’s equation, which in this case is given by

\[
P_n(x_n, y_n) \ dx_n \ dy_n = p_0(x_0, y_0) \ dx_0 \ dy_0 
= p_{n-1}(x_{n-1}, y_{n-1}) \ dx_{n-1} \ dy_{n-1} 
\]

(2.79)

Since all the relevant Jacobians equal unity, this means that \( p_n(x_n, y_n) = p_{n-1}(x_{n-1}, y_{n-1}) = \ldots = p_0(x_0, y_0) \).

The inverse bakers’ transformation is given by

\[
x_n = x_{n+1}/2, \quad x_n \in [0, 1/2) \\
y_n = 2y_{n-1} - 1, \quad y_n \in [1/2, 1] 
\]
(2.80)

Fig. 2.6. Binary bakers’ transformation.

Fig. 2.7. Mixing agrees with Poincare recurrence (iteration of a block of initial conditions).
which is the same as
\[
\begin{cases}
  x_{n+1} = \frac{y_{n+1}}{2} & y_{n+1} \in [0, 1/2) \\
  2x_{n+1} + 1 & y_{n+1} \in [1/2, 1]
\end{cases}
\]  
(2.81)

Replacing \((x_{n-1}, y_{n-1})\) by their expressions in terms of \((x_n, y_n)\) on the right hand side of the equation
\[
p_n(x_{n-1}, y_{n-1}) = p_{n-1}(x_{n-1}, y_{n-1})
\]
and setting \((x_n, y_n) = (x, y)\) yields the master equation
\[
p_n(x, y) = \begin{cases}
  p_{n-1}\left(\frac{x}{2}, 2y\right), & y \in [0, 1/2) \\
  p_{n-1}\left(\frac{x+1}{2}, 2y-1\right), & y \in [1/2, 1]
\end{cases}
\]  
(2.82)

Is there an approach to statistical equilibrium, as is suggested by our preceding graphics? The answer is yes, but it is in the “weak” sense that asymptotic averages of well-behaved functions of the dynamical variables \(f(x, y)\) can be computed as if there were convergence of \(p_n(x, y)\) as \(n \to \infty\) to a point-wise uniform invariant density \(p(x, y) = 1:\)
\[
\langle f(x, y) \rangle = \lim_{n \to \infty} \int_0^1 \int_0^1 dx dy \ p_n(x, y) f(x, y)
\]  
(2.83)

The convergence of \(p_n(x, y)\) to unity in the weak sense [13] is adequate for the purpose of computing all the averages of interest.

Some authors use the coarse-grained density [1, 13]
\[
g_n(x) = \int_0^1 \int_0^1 dx' dy' \ p_n(x, y) f(x, y)
\]  
(2.84)

where
\[
g_n(x) = \left( g_{n-1}(x/2) + g_{n-1}\left(\frac{x+1}{2}\right) \right)/2
\]  
(2.85)

which is the master equation for the binary Bernoulli shift, and state that the result \(g_n(x) \to 1\) as \(n \to \infty\) reflects the relaxation to equilibrium in the bakers’ transformation. This is a little misleading, for in the bakers’ transformation it is \(x\), not \(y\), that is the “fast” variable that one should average over in order to “coarse-grain”. This is clear from an iteration of a block of initial conditions such as we exhibit graphically in Fig. 2.7.

**Exercise:** Derive the master equation from the Frobenius–Peron equation.

**Solution:**
\[
p_n(x, y) = \int_0^1 \int_0^1 dx' dy' \ p_{n-1}(x', y') \delta(x - f(x')) \times \delta(y - g(y'))
\]  
(2.86)

where
\[
f(x) = \begin{cases}
  2x, & 0 \leq x < 1/2 \\
  2x - 1, & 1/2 \leq x \leq 1
\end{cases}
\]  
(2.87)

and
\[
g(y) = \begin{cases}
  y/2, & 0 \leq y < 1/2 \\
  (y + 1)/2, & 1/2 \leq y \leq 1
\end{cases}
\]  
(2.88)

Since
\[
\delta(x - f(x')) = \frac{1}{|f'(x')|} \delta(x' - f^{-1}(x))
\]  
(2.89)

and using the fact that the location of \(x_n\) in the unit interval determines the \(y_n\)-branch that \(y_{n-1}\) came from, we have
\[
p_n(x, y) = \frac{1}{2} \int_0^1 dx' \int_0^1 dy' \ p_n(x', y') \delta(x' - x/2)
\]  
\[
\times \delta(y' - (2y - 1)) + \frac{1}{2} \int_0^1 dx' \int_0^1 dy' \ p_n(x', y') \delta(x' - (x + 1)/2)
\]  
\[
\times \delta(y' - (2y - 1))
\]  
(2.90)

So, if \(0 \leq y < 1/2\) we have
\[
p_n(x, y) = p_{n-1}\left(\frac{x}{2}, 2y\right)
\]  
(2.91)

whereas if \(1/2 \leq y \leq 1\) we obtain
\[
p_n(x, y) = p_{n-1}\left(\frac{x+1}{2}, 2y-1\right).
\]  
(2.91b)

It is interesting to see graphically that the mixing process, i.e., the irreversible approach to statistical equilibrium, is consistent with Poincaré’s recurrence theorem. Consider, e.g., a finite block \(\mathcal{B}\) in the unit square, corresponding to some initial uncertainty in our knowledge of initial data \((x_0, y_0)\); \((x_0, y_0)\) lies within the block with even probability in this case. Then under the action of the map, this initial uncertainty becomes filamented as is shown in Fig. 2.7.

So, for \(n = 2\) there is already a “recurrence” and clearly there will be infinitely many intersections of the filament with the location of the original block \(\mathcal{B}\) as \(n \to \infty\). It follows that a definite orbit \((x_0, y_0) \to (x_1, y_1) \to \cdots \to (x_n, y_n) \to \cdots\) must return arbitrarily closely to \((x_0, y_0)\) infinitely many times as \(n \to \infty\), in agreement with Poincaré’s Theorem. Qualitatively, there seem to be two essential elements in the approach to statistical equilibrium: instability, as is here provided by the two finite Liapunov exponents, and a dense set of unstable (almost-)periodic orbits in phase space. So, far from ruling out mixing-behaviour, the uniform approach to statistical equilibrium in the unit square could not occur if Poincaré’s recurrence theorem did not hold.

In order to understand the unstable almost-periodic nature of the orbits beginning from rational initial conditions, we should study the bakers’ map in binary arithmetic [3]. With our initial conditions expressed as binary expansions
\[
x_0 = \sum_{j=1}^{\infty} e_j/2^j = .e_1 e_2 \ldots
\]  
(2.92)

and
\[
y_0 = \sum_{j=1}^{\infty} \delta_j/2^j = .\delta_1 \delta_2 \ldots
\]  
(2.92b)
The bakers' transformation,  

\[ \begin{align*}  
x_n &= \begin{cases}  
2x_n, & 0 \leq x_n < 1/2 \\
2x_n - 1, & 1/2 \leq x_n \leq 1 
\end{cases} \\
y_n &= \begin{cases}  
y_n/2, & 0 \leq y_n < 1/2 \\
(1 + y_n)/2, & 1/2 \leq y_n \leq 1 
\end{cases}  
\]  

(2.93)

can be understood as follows: \( x_1 = .\delta_1\delta_2 \ldots \) whether \( \delta_1 = 0 \) or 1, but \( y_1 = .0\delta_1\delta_2 \ldots \) if \( \delta_1 = 0 \) and \( y_1 = .1\delta_1\delta_2 \ldots \) if \( \delta_1 = 1 \). I.e., \( x_1 = .\delta_1\delta_2 \ldots \) whether \( \delta_1 = 0 \) or 1, but \( y_1 = .0\delta_1\delta_2 \ldots \) if \( \delta_1 = 0 \) and \( y_1 = .1\delta_1\delta_2 \ldots \) if \( \delta_1 = 1 \). So, \( x_1, y_1, \ldots \). It is by now clear that the initial condition \( y_0 \) is "lost" for large enough \( n \), the bits \( \delta_1\delta_2 \ldots \) are shifted from most significant into least significant bits with each iteration, and in general we find that

\begin{align*}  
x_n &= .\delta_{n+1}\delta_{n+2} \ldots \\
y_n &= .\delta_1\delta_2 \ldots 
\]  

(2.94)

If we then write the two-sided binary sequence representing \( x_0 \) and \( y_0 \)

\[ \ldots \delta_2\delta_1 . \delta_0 \delta_0 . \]  

(2.95)

then after \( n \) iterations the two-sided binary sequence

\[ \ldots \delta_2\delta_1 \delta_0 \ldots \delta_1 \delta_1 \delta_0 \ldots . \]  

(2.96)

represents the bakers' transformation. \( x_n \) is given by the sequence of bits to the right of the binary point whereas \( y_n \) is given by the reverse sequence of bits to the left of the binary point. So, the motion for large \( n \) is determined entirely by the distribution of bits in \( x_n \) and not at all by \( y_n \). If \( x_0 \) is rational, then the \( x \)-motion is unstable periodic while the \( y \)-motion is almost-periodic. The instability corresponds to the fact that a small error in \( x_0 \) will produce an entirely different orbit, typically one with much larger period. If \( x_0 \) is irrational, the motion is truly chaotic and is also ergodic; the neighbourhood of any initial condition \( (x_0, y_0) \) will be visited arbitrarily often by the system's orbit as \( n \) increases without bound. The mixing property and corresponding uniform invariant density for shift maps are deeply connected with the existence of Borel's normal numbers.

The behavior of the baker's transformation on a finite-state binary computer is quite different. Since a binary computer (in fixed-point arithmetic) approximates every \( (x_0, y_0) \) by two numbers of the form

\begin{align*}  
x_0 &= .\delta_1\delta_2 \ldots \delta_N 00 \ldots \\
y_0 &= .\delta_1\delta_2 \ldots \delta_N 00 \ldots 
\]  

(2.97)

we have

\begin{align*}  
x_n &= .\delta_1\delta_2 \ldots \delta_N 00 \ldots \\
y_n &= .\delta_1\delta_2 \ldots \delta_N 00 \ldots 
\]  

(2.97b)

whenever \( n > N \). Notice that the use of floating-point arithmetic does not in any way remedy the fact that all orbits "hop" onto the unstable fixed point at the origin in finite "time" \( N \).

Exercise: Using \( x_{n+1} = 3x_n \mod 1 \), construct area-preserving bakers' transformation that has nontrivial orbits on a finite binary lattice. Show that on such a lattice, the orbits are all unstable-periodic.

Solution: With \( x_{n+1} = 3^x x_0 \mod 1 \) we can use

\begin{align*}  
y_{n+1} &= \begin{cases}  
1/3y_n, & 0 \leq y_n < 1/3 \\
1/3y_n + 1/3, & 1/3 \leq y_n < 2/3 \\
1/3y_n + 2/3, & 2/3 \leq y_n < 1 
\end{cases}  
\]  

(2.98)

and since

\[ J = \left| \frac{\partial(x_{n+1}, y_{n+1})}{\partial(x_n, y_n)} \right| = \begin{pmatrix} 3 & 0 \\
0 & 1/3 \end{pmatrix}, \]

(2.99)

the map is area-preserving.

Furthermore, since

\begin{align*}  
3x_0 &= x_0 + 2x_0, \\
3^2x_0 &= x_0 + 2x_0 + 2x_0 + 2^2x_0, \\
&= x_0 + 2^1x_0, \\
3^3x_0 &= x_0 + 2x_0 + 2^2x_0 + 2^3x_0, \\
&= x_0 + 2^1x_0 + 2^3x_0, \\
3^4x_0 &= x_0 + 2^2x_0 + 2^3x_0, \\
&= x_0 + 2^1x_0 + 2^3x_0, \\
&\cdots 
\]  

(2.100)

it is clear the fractional part of \( 3^x x_0 - x_0 \) is "annihilated" (mod 1) by powers of 3 if \( n \) is large enough. Hence, every \( N \)-bit binary initial condition \( x_0 \) yields a periodic orbit. These orbits are all unstable because \( \lambda = \log 3 > 0 \).

The shift maps discussed in this chapter are completely non-integrable: there are no regions of finite size in phase space where the motion is stable in the sense of a null Liapunov exponent. In the bakers' transformation, the resulting filamentation of small errors in initial data produces an irreversible approach to statistical equilibrium that is consistent with Poincare's recurrence theorem [14]. The non-integrability encountered here is characterized by the presence of a dense set of unstable (almost-)periodic orbits that begin from rational initial conditions. We expect that a similar qualitative picture will be true of conservative chaos quite generally in completely non-integrable Hamiltonian systems. Poincare discovered complex behavior in dynamical systems having a homoclinic point [1, 18] and Smale showed that the motion near such points is describable in a certain sense by a shift map [17].
Finally, although the qualitative picture presented here is very attractive and shows how simple two degree of freedom conservative systems can have an irreversible approach to statistical equilibrium, it is nontrivial to prove whether or not a given system is mixing. [9–11, 15] For example, we do not know the behavior of the Henon-Heiles system with $E \to 1/6$, a system which shows a transition from integrable to near-integrable behavior at $E \sim 1/12$. However, the system of $N$ point vortices in an unbounded ideal fluid in two dimensions is Hamiltonian [19] and is known to exhibit chaotic behavior when $N \geq 4$ [20]; for large $N$ ordinary statistical thermodynamics is expected to be valid [21].

### 2.5. Symbolic dynamics

At the most detailed level of description of a dynamical system, one knows an exact orbit for some length of time. Except for periodic orbits, it is impossible to know the orbit of a chaotic system for all times, even though the system is deterministic. At the most coarse-grained level of description, one may know an equilibrium probability density, $p(x)$. If the phase space is broken up into definite number of bins, then for one dimensional maps of the unit interval, the number of iterations of the map

$$x_n = f(x_{n-1})$$

(2.101)

can be predicted to be

$$N_k \approx \int_{x_k}^{x_{k+1}} p(x) \, dx.$$  

(2.102)

In the long run and for almost all initial data one expects this prediction to be correct. Whether it can be realized computationally on a finite-state machine is a problem that we shall discuss in Ch. 5.

From this point of view, a conceptually useful “coarse-graining of dynamics that retains the detail of the time-sequence of the iterations is the method of symbolic dynamics [17]. Consider first “2-bin” symbolic dynamics, where we map an entire orbital section $x_0 \to x_1 \to x_2 \to \ldots \to x_n$ onto a single binary fraction

$$\phi_n(x_0) = \sum_{i=0}^{n-1} e_i/2^i = e_0e_1e_2 \ldots e_n$$

(2.103)

which is constructed as follows: if $0 \leq x_1 < 1/2$ then $e_1 = 0$ while if $1/2 \leq x_1 \leq 1$ then $e_1 = 1$, and in general

$$e_i = \begin{cases} 0, & 0 \leq x_i < 1/2 \\ 1, & 1/2 \leq x_i \leq 1 \end{cases}$$

(2.104)

Here, the details of the dynamics are replaced by a hopping back and forth between two bins, and for Bernoulli shifts, one is led to the question how random can a deterministic system be, in analogy with coin-tossing ($e_i = 0 \leftrightarrow T$, $e_i = 1 \leftrightarrow H$, e.g., in the tosses of a fair coin). Here, long-period rational fractions yield “pseudo-chaos” in the form of unstable periodic-orbits, while irrational sequences are required for true chaos. If $n \ll N$ where $N$ is the period of a rational binary fraction, then the symbolic dynamics may appear to be chaotic. Symbolic dynamics can be generalized to any number $\mu$ of bins by mapping the orbit onto a base $\mu$ fraction:

$$e_i = K \text{ if } K/\mu \leq x_i < (K + 1)/\mu \text{ where } K = 0, 1, 2, \ldots, \mu - 1,$$

$$x_i = f(x_{i-1})$$

(2.105)

and

$$\phi_n(x_0) = \sum_{i=0}^{n-1} e_i/\mu.$$  

(2.106)

Again, if there is an invariant density $p(x)$ then the long-time expected frequency of visitation of each bin in principle is predicted. Conversely one can use symbolic dynamics to build up a histogram that represents the probability distribution for some observed orbit. However, a careful look shows that in order to actually compute these symbol sequences, the orbit must be known to high precision at long times (cf. Ch. 5). That is, the conceptual simplicity of symbolic dynamics does not extend to the actual prediction of the symbol sequences. For the simplest chaotic maps, the linear shift maps, the symbol sequences can be generated directly from algorithms for irrational numbers.

### 2.6. Perturbed twist maps and near integrable conservative systems

If we return now to the nature of the chaos in the Henon-Heiles system, we must admit from the start that we face a more difficult problem than that encountered in the bakers' transformation. In the former problem there is a sequence of bifurcations as $E$ is increased, with motion ranging from integrable to more and more chaotic whereby the total size of regular regions in phase space decreases as chaos gains more and more ground. The bakers' transformation gives qualitative information about the dynamics in the chaotic regions via symbolic dynamics. In order to attempt to describe such systems such as Henon–Heiles near the transition to chaos, it

---

**Fig. 2.9.** In Computation, the phase space is coarse-grained into a finite number of bins.

**Fig. 2.10.** Integrable systems yield motion on tori in phase space.
is useful to start from a point of view near that of an integral system, that of the perturbed twist map [1].

If we begin with a two degree of freedom integrable system in the action-angle variable formulation, then \( H(J_1, J_2) = E \), \( \theta_1 = \omega_1 \theta \) and \( \theta_2 = \omega_2 \theta \), and as we have discussed in Section 1.6, the motion is either periodic or quasiperiodic on a two-dimensional torus. The radii of the torus are \( J_1 \) and \( J_2 \), as is shown in Fig. 2.10. We can obtain a “twist map” by taking a Poincare section of the motion: consider a plane transverse to the torus and consider the intersection of the orbit with that plane whenever \( \theta_1 \) advances through \( 2\pi \), starting from its initial value \( \theta_{0,1} \). If we denote \( J_i \) by \( J_{\theta_i} \) and \( \theta_i \) by \( \theta_{\theta_i} \) in the \( n \)th intersection of the orbit with the plane, then the twist map is given by

\[
J_{\theta_{i+1}} = J_i + \epsilon f(J_{\theta_i}, \theta_{\theta_i}) + \epsilon g(J_{\theta_i}, \theta_{\theta_i})
\]

(2.107)

where \( \epsilon \) is the frequency ratio \( \omega_1 / \omega_2 \). If \( \epsilon \) is rational then the \( \theta_\epsilon \)-motion is periodic, consisting of a finite set of points. If \( \epsilon \) is irrational, then the \( \theta_\epsilon \) motion is quasiperiodic and yields a dense set of points, a circle as \( n \to \infty \). This is the nature of the return map for a two degree of freedom integrable system.

Liouville’s theorem from classical mechanics guarantees an area-preserving map in the \((J_\theta, \theta)\) plane because

\[
\frac{\partial (J_{\theta_{i+1}}, \theta_{\theta_{i+1}})}{\partial (J_\theta, \theta)} = 1.
\]

(2.108)

In order to consider the perturbation of this integrable system, we follow Lichtenberg and Lieberman [1] who expect the twist map to take the form

\[
J_{\theta_{i+1}} = J_i + \epsilon f(J_{\theta_i}, \theta_{\theta_i}) + \epsilon g(J_{\theta_i}, \theta_{\theta_i}),
\]

(2.109)

Qualitatively, the extra terms must arise from a perturbation that makes the Hamiltonian non-integrable, such as the non-quadratic part of the potential energy in the Henon-Heiles problem (which problem has been proven to be non-integrable). Our perturbed twist map must be area-preserving, so that

\[
J = \frac{\partial (J_{\theta_{i+1}}, \theta_{\theta_{i+1}})}{\partial (J_\theta, \theta)} = \{J_{\theta_{i+1}}, \theta_{\theta_{i+1}}\} = 1
\]

(2.110)

and we will satisfy this condition to \( O(\epsilon) \):

\[
J = \{J_{\theta_{i+1}}, \theta_\epsilon + 2\pi \epsilon \} = \{J_\theta, \theta_\epsilon \} + \epsilon J_{\theta_{i+1}}, \theta_{\theta_{i+1}}
\]

(2.111)

so that we must require

\[
\{\theta, f\} + \{g, J_{\theta_{i+1}}\} = 0
\]

(2.112)

to \( O(\epsilon^2) = 0 \).

Since

\[
\{\theta, f\} = \frac{\partial f}{\partial J_\theta} = \frac{\partial f}{\partial J_{\theta_{i+1}}} \frac{\partial J_{\theta_{i+1}}}{\partial J_\theta} \approx \frac{\partial f}{\partial J_{\theta_{i+1}}} + O(\epsilon)
\]

(2.113)

and since

\[
\{g, J_{\theta_{i+1}}\} = \{g, J_\theta\} + \epsilon \{g, f\} = \{g, J_\theta\} + O(\epsilon)
\]

(2.114)

Then the area-preserving property requires that

\[
\frac{\partial f}{\partial J_{\theta_{i+1}}} + \frac{\partial g}{\partial \theta_{\theta_{i+1}}} = 0.
\]

(2.115)

So, we are free to attempt to model perturbed twist maps by different choices of \( f \) and \( g \) if we satisfy this constraint. A map that has been much discussed in the literature is the radial twist map, which follows from assuming that

\[
J_{\theta_{i+1}} = 0 \quad \text{and} \quad g = 0:
\]

(2.116)

\[
\theta_{\theta_{i+1}} = \theta_\epsilon + 2\pi \phi(J_{\theta_{i+1}}).
\]

We can now develop a perturbation theory as follows. If we linearize about a period one fixed point of the map \( J_{\theta_{i+1}} = J_\theta = J_0 \), so that \( \phi(J_0) = m \) is an integer, then we set \( J_{\theta_{i+1}} = J_0 + \Delta J_{\theta_{i+1}} \) and obtain

\[
\Delta J_{\theta_{i+1}} = \Delta J_\theta + \epsilon \phi(J_0) \Delta J_{\theta_{i+1}}.
\]

(2.117)

If then we set \( L_\theta = 2\pi \epsilon \Delta J_\theta \) we find that

\[
L_{\theta_{i+1}} = L_\theta + 2\pi \epsilon \phi(J_0),
\]

(2.118)

\[
\theta_{\theta_{i+1}} = \theta_\epsilon + 2\pi m + L_{\theta_{i+1}} = \theta_\epsilon + L_{\theta_{i+1}}, \mod 2\pi
\]

(2.119)

With \( \kappa = 2\pi \epsilon \phi_{\max} \) and \( f^* = f/f_{\max} \) we obtain the radial twist map

\[
L_{\theta_{i+1}} = L_\theta + \kappa f^*(\theta_\epsilon),
\]

(2.119)

\[
\theta_{\theta_{i+1}} = \theta_\epsilon + L_{\theta_{i+1}}, \mod 2\pi
\]

The Chirikov-Taylor model, also called the standard map, follows by setting \( f^*(\theta_\epsilon) = \sin \theta_\epsilon \). Analytic and numerical results for this model are discussed extensively in Lichtenberg and Lieberman’s excellent text. We turn now to a brief qualitative description of the Kolomogorov Arnol’d Moser (or KAM) Theorem.

We begin with the perturbed twist map in the form [18]

\[
J_{\theta_{i+1}} = J_\theta + \epsilon f(J_\theta, \theta_\theta),
\]

(2.120)

\[
\theta_{\theta_{i+1}} = \theta_\theta + 2\pi \phi(J_\theta) + \epsilon \phi(J_\theta, \theta_\theta)
\]

which follows from eq. (2.118) by iteration. When \( \phi = 0 \), the “bare” winding number \( \epsilon \) characterizes an invariant torus. When does this motion persist under weak perturbation: i.e., when is the motion with winding number \( \epsilon \) stable against perturbation? KAM gives a sufficient condition for the existence of perturbed invariant tori. The condition is one of “sufficient incommensurability” of the winding number with respect to rational approximations. Consider the perturbed system, and consider the systematic approximation of an irrational winding number \( \epsilon \) by a sequence of rational approximations \( p_\theta/q_\theta, p_\theta/q_\theta, \ldots, p_\theta/q_\theta, \ldots \). If as \( \theta \) we find that

\[
|\epsilon - p_\theta/q_\theta| > K(\epsilon)/q_\theta^2
\]

(2.121)

as \( \epsilon \to 0 \), then the orbit with irrational winding no. \( \epsilon \) survives as a stable quasiperiodic orbit under the perturbation, and \( K(\epsilon) \) is a positive constant dependent upon the perturbation strength \( \epsilon \). Useful sequences of rational approximants are
generated by continued fractions [4]. The details of this approach are discussed in Ch. 3 under the heading of circle maps, as is a definition of the winding number. Circle maps can be derived from standard maps by making the latter dissipative. The theory of both systems has been discussed by Arnol’d [10, 16].

We can easily see that according to the condition (2.121) almost all irrational tori should survive under weak perturbation, and that therefore almost all of phase space should remain permeated by tori with quasiperiodic orbits. This can be understood by calculating the relative measure of the intervals where the condition fails, for example, what is the fraction of the unit interval where the condition

\[ |x - p_n/q_n| \leq K(e)/q_n^{5/2} \]  

(2.122)

holds as \( n \to \infty \) when \( \varepsilon \ll 1 \).

If we denote the measure of that point set by \( L(e) \), we have

\[ L(e) < \kappa(e) \sum_{q=1}^{\infty} q^{-5/2} \sim \kappa(e) \]  

(2.123)

since the \( q^{-5/2} \) series converges. Hence, the size of the gaps where the KAM condition (2.121) is violated is \( O(e) \) as \( \varepsilon \to 0 \). What happens to the periodic and quasiperiodic orbits that don’t survive? Of the periodic orbits, some remain stable while others become unstable periodic. Discussions of these are to be found under the heading of the Poincare–Birkhoff Theorem [1]. We know from the numerical work on the Henon–Heiles problem that new pairs of stable periodic orbits emerge as older stable orbits are destroyed via pitchfork bifurcations (period doubling bifurcations). Universal critical exponents for period doubling bifurcations of area preserving maps have been discussed by MacKay [22].

An example of a winding number which is destroyed under perturbation is given by \( \pi \), which has a rapidly convergent continued fraction expansion. On the other hand, the Golden Mean [1, 2] or any winding number whose continued fraction expansion is populated asymptotically by ones,

\[ \alpha = \ldots + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \ldots}}} \]  

(2.124)

will survive for small enough \( \varepsilon \) since these winding numbers are sufficiently difficult to approximate by rationals. However, to state that the Golden Mean is the most irrational winding number ignores the fact that almost all winding numbers cannot be approximated by continued fractions or by any other algorithm; the Golden mean winding number belongs to the class of computable winding numbers that is most difficult to destroy via weak perturbation.

Our discussion of KAM for twist maps follows the nice article by Henon [18] and reflects Moser’s proof of KAM. Arnol’d generalized the theorem to include Hamiltonian systems with an arbitrary but finite number of degrees of freedom. Hamiltonian chaos is discussed thoroughly in the book by Lichtenberg and Lieberman [1] and modern results based upon the renormalization group method can be found in the research papers by Kadanoff et al. [23].

References to Chapter 2

18. Henon, M. in Ref. [17].