

4

Ergodicity

The theorems presented in the previous chapter fully establish the first part of Boltzmann's ergodic hypothesis: for any measurable set E , the mean sojourn time $\tau(E, x)$ is well defined for almost every point x . The second part of the ergodic hypothesis, that is, the claim that $\tau(E, x)$ should coincide with the measure of E for almost every x , is a statement of a different nature and is the subject of the present chapter.

In this chapter we always take μ to be a probability measure invariant under some measurable transformation $f : M \rightarrow M$. We say that the system (f, μ) is *ergodic* if, given any measurable set E , we have $\tau(E, x) = \mu(E)$ for μ -almost every point $x \in M$. We are going to see that this is equivalent to saying that the system is dynamically indivisible, in the sense that every invariant set has either full measure or zero measure. Other equivalent formulations of the ergodicity property are discussed in Section 4.1. One of them is that time averages coincide with space averages: for every integrable function φ ,

$$\lim_n \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f^j(x)) = \int \varphi d\mu \quad \text{at } \mu\text{-almost every point.}$$

In Section 4.2 we illustrate, by means of examples, several techniques to prove or disprove ergodicity. Most of them will be utilized again later in more complex situations. Next, we take the following viewpoint: we fix the dynamical systems and analyze the properties of ergodic measures within the space of all invariant measures of that dynamical system. As we are going to see in Section 4.3, the ergodic measures are precisely the extremal elements of that space.

In Section 4.4 we give a brief outline of the historical development of ergodic theory in the context of conservative systems. The main highlights are *KAM theory*, thus denominated in homage to Andrey Kolmogorov, Vladimir Arnold and Jürgen Moser, and *hyperbolic dynamics*, which was initiated by Steven Smale, Dmitry Anosov, Yakov Sinai and their collaborators. The two theories deal with distinct types of dynamical behavior, elliptic and hyperbolic,

and they reach opposing conclusions: roughly speaking, hyperbolic systems are ergodic whereas elliptic systems are not.

4.1 Ergodic systems

We use the expressions “the measure μ is ergodic with respect to the transformation f ” or “the transformation f is ergodic with respect to the measure μ ” to mean the same thing, namely, that the system (f, μ) is ergodic. Recall that, by definition, this means that the mean sojourn time in any measurable set of μ -almost every point coincides with the measure of that set. This condition can be rephrased in several equivalent ways, as we are going to see next.

4.1.1 Invariant sets and functions

A measurable function $\varphi : M \rightarrow \mathbb{R}$ is said to be *invariant* if $\varphi = \varphi \circ f$ at μ -almost every point. In other words, φ is invariant if it is constant on every trajectory of f outside a zero measure subset. Moreover, we say that a measurable set $B \subset M$ is *invariant* if its characteristic function χ_B is an invariant function. Equivalently, B is invariant if it differs from its pre-image $f^{-1}(B)$ by a zero measure set:

$$\mu(B \Delta f^{-1}(B)) = 0.$$

Exercise 1.1.4 collects some equivalent formulations of this property. It is easy to check that the family of all invariant sets is a σ -algebra, that is, it is closed under countable unions and intersections and under passage to the complement.

Example 4.1.1. Let $f : [0, 1] \rightarrow [0, 1]$ be the decimal expansion transformation introduced in Section 1.3.1, and μ be the Lebesgue measure. Clearly, the set $A = \mathbb{Q} \cap [0, 1]$ of rational numbers is invariant. Other interesting examples are the sets of points $x = 0.a_1a_2\dots$ in $[0, 1]$ with prescribed proportions of digits a_i with each value $k \in \{0, \dots, 9\}$. More precisely, given any vector $p = (p_0, \dots, p_9)$ such that $p_i \geq 0$ for every i and $\sum_i p_i = 1$, define

$$A_p = \left\{ x : \lim_n \frac{1}{n} \#\{1 \leq i \leq n : a_i = k\} = p_k \text{ for } k = 0, \dots, 9 \right\}.$$

Observe that if $x = 0 \cdot a_1a_2\dots$ then every point $y \in f^{-1}(x)$ may be written as $y = 0 \cdot ba_1a_2\dots$ with $b \in \{0, \dots, 9\}$. It is clear that the extra digit b does not affect the proportion of digits with any of the values $0, \dots, 9$ in the decimal expansion. Thus, $y \in A_p$ if and only if $x \in A_p$. This implies that A_p is indeed invariant under f .

Example 4.1.2. Let $\varphi : [0, 1] \rightarrow \mathbb{R}$ be a function in $L^1(\mu)$. According to the ergodic theorem of Birkhoff (Theorem 3.2.3), the time average $\tilde{\varphi}$ is an invariant function. So, every level set

$$B_c = \{x \in [0, 1]; \tilde{\varphi}(x) = c\}$$

is an invariant set. Observe also that every invariant function is of this form: it is clear that if φ is invariant then it coincides with its time average $\tilde{\varphi}$ at μ -almost every point.

The next proposition collects a few equivalent ways to define ergodicity. We say that a function φ is constant at μ -almost every point if there exists $c \in \mathbb{R}$ such that $\varphi(x) = c$ for μ -almost every $x \in M$.

Proposition 4.1.3. *Let μ be an invariant probability measure of a measurable transformation $f : M \rightarrow M$. The following conditions are all equivalent:*

- (i) *For every measurable set $B \subset M$ one has $\tau(B, x) = \mu(B)$ for μ -almost every point.*
- (ii) *For every measurable set $B \subset M$ the function $\tau(B, \cdot)$ is constant at μ -almost every point.*
- (iii) *For every integrable function $\varphi : M \rightarrow \mathbb{R}$ one has $\tilde{\varphi}(x) = \int \varphi d\mu$ for μ -almost every point.*
- (iv) *For every integrable function $\varphi : M \rightarrow \mathbb{R}$ the time average $\tilde{\varphi} : M \rightarrow \mathbb{R}$ is constant at μ -almost every point.*
- (v) *For every invariant integrable function $\psi : M \rightarrow \mathbb{R}$ one has $\psi(x) = \int \psi d\mu$ for μ -almost every point.*
- (vi) *Every invariant integrable function $\psi : M \rightarrow \mathbb{R}$ is constant at μ -almost every point.*
- (vii) *For every invariant subset A we have either $\mu(A) = 0$ or $\mu(A) = 1$.*

Proof. It is immediate that (i) implies (ii), that (iii) implies (iv) and that (v) implies (vi). It is also clear that (v) implies (iii) and (vi) implies (iv), because the time average is an invariant function (recall Proposition 3.2.4). Analogously, (iii) implies (i) and (iv) implies (ii), because the mean sojourn time is a time average (of the characteristic function of B). We are left to prove the following implications:

(ii) implies (vii): Let A be an invariant set. Then $\tau(A, x) = 1$ for μ -almost every $x \in A$ and $\tau(A, x) = 0$ for μ -almost every $x \in A^c$. Since $\tau(A, \cdot)$ is assumed to be constant at μ -almost every point, it follows that $\mu(A) = 0$ or $\mu(A) = 1$.

(vii) implies (v): Let ψ be an invariant integrable function. Then every level set

$$B_c = \{x \in M : \psi(x) \leq c\}$$

is an invariant set. So, the hypothesis implies that $\mu(B_c) \in \{0, 1\}$ for every $c \in \mathbb{R}$. Since $c \mapsto \mu(B_c)$ is non-decreasing, it follows that there exists $\bar{c} \in \mathbb{R}$

such that $\mu(B_c) = 0$ for every $c < \bar{c}$ and $\mu(B_c) = 1$ for every $c \geq \bar{c}$. Then $\psi = \bar{c}$ at μ -almost every point. Hence, $\int \psi d\mu = \bar{c}$ and so $\psi = \int \psi d\mu$ at μ -almost every point.

4.1.2 Spectral characterization

The next proposition characterizes the ergodicity property in terms of the Koopman operator $U_f(\varphi) = \varphi \circ f$:

Proposition 4.1.4. *Let μ be an invariant probability measure of a measurable transformation $f : M \rightarrow M$. The following conditions are equivalent:*

- (i) (f, μ) is ergodic.
- (ii) For any pair of measurable sets A and B one has

$$\lim_n \frac{1}{n} \sum_{j=0}^{n-1} \mu(f^{-j}(A) \cap B) = \mu(A)\mu(B). \tag{4.1.1}$$

- (iii) For any functions $\varphi \in L^p(\mu)$ and $\psi \in L^q(\mu)$, with $1/p + 1/q = 1$, one has

$$\lim_n \frac{1}{n} \sum_{j=0}^{n-1} \int (U_f^j \varphi) \psi d\mu = \int \varphi d\mu \int \psi d\mu. \tag{4.1.2}$$

Proof. It is clear that (iii) implies (ii): just take $\varphi = \chi_A$ and $\psi = \chi_B$. To show that (ii) implies (i), let A be an invariant set. Taking $A = B$ in hypothesis (ii), we get that

$$\mu(A) = \lim_n \frac{1}{n} \sum_{j=0}^{n-1} \mu(f^{-j}(A) \cap A) = \mu(A)^2.$$

This implies that $\mu(A) = 0$ or $\mu(A) = 1$.

Now it suffices to prove that (i) implies (iii). Consider any $\varphi \in L^p(\mu)$ and $\psi \in L^q(\mu)$. By ergodicity and the ergodic theorem of Birkhoff (Theorem 3.2.3) we have that

$$\frac{1}{n} \sum_{j=0}^{n-1} U_f^j \varphi \rightarrow \int \varphi d\mu \tag{4.1.3}$$

at μ -almost every point. Initially, assume that $|\varphi| \leq k$ for some $k \geq 1$. Then, for every $n \in \mathbb{N}$,

$$\left| \left(\frac{1}{n} \sum_{j=0}^{n-1} U_f^j \varphi \right) \psi \right| \leq k|\psi|.$$

So, since $k|\psi| \in L^1(\mu)$, we may use the dominated convergence theorem (Theorem A.2.11) to conclude that

$$\int \left(\frac{1}{n} \sum_{j=0}^{n-1} U_f^j \varphi \right) \psi d\mu \rightarrow \int \varphi d\mu \int \psi d\mu.$$

This proves the claim (4.1.2) when φ is bounded. All that is left to do is remove this restriction. Given any $\varphi \in L^p(\mu)$ and $k \geq 1$, define

$$\varphi_k(x) = \begin{cases} k & \text{if } \varphi(x) > k \\ \varphi(x) & \text{if } \varphi(x) \in [-k, k] \\ -k & \text{if } \varphi(x) < -k. \end{cases}$$

Fix $\varepsilon > 0$. By the previous argument, for every $k \geq 1$ one has

$$\left| \int \left(\frac{1}{n} \sum_{j=0}^{n-1} U_f^j \varphi_k \right) \psi \, d\mu - \int \varphi_k \, d\mu \int \psi \, d\mu \right| < \varepsilon \tag{4.1.4}$$

if n is large enough (depending on k). Next, observe that $\|\varphi_k - \varphi\|_p \rightarrow 0$ when $k \rightarrow \infty$: this is clear when $p = \infty$, because $\varphi_k = \varphi$ for every $k > \|\varphi\|_\infty$; for $p < \infty$ use the monotone convergence theorem (Theorem A.2.9). Hence, using the Hölder inequality (Theorem A.5.5), we have that

$$\left| \int (\varphi_k - \varphi) \, d\mu \int \psi \, d\mu \right| \leq \|\varphi_k - \varphi\|_p \left| \int \psi \, d\mu \right| < \varepsilon, \tag{4.1.5}$$

for every k sufficiently large. Similarly,

$$\begin{aligned} \left| \int \frac{1}{n} \sum_{j=0}^{n-1} U_f^j (\varphi_k - \varphi) \psi \, d\mu \right| &\leq \frac{1}{n} \sum_{j=0}^{n-1} \left| \int U_f^j (\varphi_k - \varphi) \psi \, d\mu \right| \\ &\leq \frac{1}{n} \sum_{j=0}^{n-1} \|U_f^j (\varphi_k - \varphi)\|_p \|\psi\|_q \, d\mu \\ &= \|\varphi_k - \varphi\|_p \|\psi\|_q < \varepsilon, \end{aligned} \tag{4.1.6}$$

for every n and every k sufficiently large, independent of n . Fix k so that (4.1.5) and (4.1.6) hold and then take n sufficiently large such that (4.1.4) also holds. Summing the three relations (4.1.4) to (4.1.6), we get that

$$\left| \int \left(\frac{1}{n} \sum_{j=0}^{n-1} U_f^j \varphi \right) \psi \, d\mu - \int \varphi \, d\mu \int \psi \, d\mu \right| < 3\varepsilon$$

for every n sufficiently large. This gives condition (iii).

In the case $p = q = 2$, the condition (4.1.2) may be expressed in terms of the inner product \cdot in the space $L^2(\mu)$. In this way we get that (f, μ) is ergodic if and only if:

$$\lim_n \frac{1}{n} \sum_{j=0}^{n-1} [(U_f^j \varphi) - (\varphi \cdot 1)] \cdot \psi = 0 \quad \text{for every } \varphi, \psi \in L^2(\mu). \tag{4.1.7}$$

We will use a few times the following elementary facts: given any measurable sets A and B ,

$$\begin{aligned}
 |\mu(A) - \mu(B)| &= |\mu(A \setminus B) - \mu(B \setminus A)| \\
 &\leq \mu(A \setminus B) + \mu(B \setminus A) = \mu(A \Delta B),
 \end{aligned}
 \tag{4.1.8}$$

and given any sets A_1, A_2, B_1, B_2 ,

$$(A_1 \cap A_2) \Delta (B_1 \cap B_2) \subset (A_1 \Delta B_1) \cup (A_2 \Delta B_2).
 \tag{4.1.9}$$

Corollary 4.1.5. *Assume that the condition (4.1.1) in Proposition 4.1.4 holds for every A and B in some algebra \mathcal{A} that generates the σ -algebra of measurable sets. Then (f, μ) is ergodic.*

Proof. Let A and B be arbitrary measurable sets. By the approximation theorem (Theorem A.1.19), given any $\varepsilon > 0$ there are A_0 and B_0 in \mathcal{A} such that $\mu(A \Delta A_0) < \varepsilon$ and $\mu(B \Delta B_0) < \varepsilon$. Observe that

$$\begin{aligned}
 |\mu(f^{-j}(A) \cap B) - \mu(f^{-j}(A_0) \cap B_0)| &\leq \mu(f^{-j}(A) \Delta f^{-j}(A_0)) + \mu(B \Delta B_0) \\
 &= \mu(A \Delta A_0) + \mu(B \Delta B_0) < 2\varepsilon
 \end{aligned}$$

(the equality uses the fact that μ is an invariant measure) for every j and

$$|\mu(A)\mu(B) - \mu(A_0)\mu(B_0)| \leq \mu(A \Delta A_0) + \mu(B \Delta B_0) < 2\varepsilon.$$

Then, the hypothesis

$$\lim_n \frac{1}{n} \sum_{j=0}^{n-1} \mu(f^{-j}(A_0) \cap B_0) = \mu(A_0)\mu(B_0)$$

implies that

$$\begin{aligned}
 -4\varepsilon &\leq \liminf_n \frac{1}{n} \sum_{j=0}^{n-1} \mu(f^{-j}(A) \cap B) - \mu(A)\mu(B) \\
 &\leq \limsup_n \frac{1}{n} \sum_{j=0}^{n-1} \mu(f^{-j}(A) \cap B) - \mu(A)\mu(B) \leq 4\varepsilon.
 \end{aligned}$$

Since ε is arbitrary, this proves that the condition (4.1.1) holds for all pairs of measurable sets. According to Proposition 4.1.4, it follows that the system is ergodic.

In the same spirit, it suffices to check part (iii) of Proposition 4.1.4 on dense subsets:

Corollary 4.1.6. *Assume that the condition (4.1.2) in Proposition 4.1.4 for every φ and ψ in dense subsets of $L^p(\mu)$ and $L^q(\mu)$, respectively. Then (f, μ) is ergodic.*

We leave the proof of this fact to the reader (see Exercise 4.1.3).

4.1.3 Exercises

- 4.1.1. Let (M, \mathcal{A}) be a measurable space and $f : M \rightarrow M$ be a measurable transformation. Prove that if $p \in M$ is a periodic point of period k , then the measure $\mu_p = \frac{1}{k}(\delta_p + \delta_{f(p)} + \cdots + \delta_{f^{k-1}(p)})$ is ergodic.
- 4.1.2. Let μ be an invariant probability measure, not necessarily ergodic, of a measurable transformation $f : M \rightarrow M$. Show that the following limit exists for any pair of measurable sets A and B :

$$\lim_n \frac{1}{n} \sum_{i=0}^{n-1} \mu(f^{-i}(A) \cap B).$$

- 4.1.3. Show that an invariant probability measure μ is ergodic for a transformation $f : M \rightarrow M$ if and only if any one of the following conditions holds:
- $\mu(\bigcup_{n \geq 0} f^{-n}(A)) = 1$ for every measurable set A with $\mu(A) > 0$;
 - given any measurable sets A, B with $\mu(A)\mu(B) > 0$, there is $n \geq 1$ such that $\mu(f^{-n}(A) \cap B) > 0$;
 - the convergence in condition (iii) of Proposition 4.1.4 holds for some choice of p, q and some dense subset of functions $\varphi \in L^p(\mu)$ and $\psi \in L^q(\mu)$;
 - there is $p \in [1, \infty]$ such that every invariant function $\varphi \in L^p(\mu)$ is constant at μ -almost every point;
 - every integrable function φ satisfying $\varphi \circ f \geq \varphi$ at μ -almost every point (or $\varphi \circ f \leq \varphi$ at μ -almost every point) is constant at μ -almost every point.
- 4.1.4. Take M to be a metric space. Prove that an invariant probability measure μ is ergodic for $f : M \rightarrow M$ if and only if the time average of every bounded uniformly continuous function $\varphi : M \rightarrow \mathbb{R}$ is constant at μ -almost every point.
- 4.1.5. Take M to be a metric space. We call the *basin* of an invariant probability measure μ the set $B(\mu)$ of all points $x \in M$ such that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f^j(x)) = \int \varphi d\mu$$

for every bounded continuous function $\varphi : M \rightarrow \mathbb{R}$. Check that the basin is an invariant set. Moreover, if μ is ergodic then $B(\mu)$ has full μ -measure.

- 4.1.6. Show that if μ and η are distinct ergodic probability measures of a transformation $f : M \rightarrow M$, then η and μ are mutually singular.
- 4.1.7. Let μ be a probability measure invariant under some transformation $f : M \rightarrow M$. Show that the product measure $\mu_2 = \mu \times \mu$ is invariant under the transformation $f_2 : M \times M \rightarrow M \times M$ defined by $f_2(x, y) = (f(x), f(y))$. Moreover, if (f_2, μ_2) is ergodic then (f, μ) is ergodic. Is the converse true?
- 4.1.8. Let μ be a probability measure invariant under some transformation $f : M \rightarrow M$. Assume that (f^n, μ) is ergodic for every $n \geq 1$. Show that if φ is a non-constant eigenfunction of the Koopman operator U_f then the eigenvalue is *not* a root of unity and any set restricted to which φ is constant has zero measure.

4.2 Examples

In this section we use a number of examples to illustrate different methods for checking whether a system is ergodic or not.

4.2.1 Rotations on tori

Initially, let us consider the case of a rotation $R_\theta : S^1 \rightarrow S^1$ on the circle $S^1 = \mathbb{R}/\mathbb{Z}$. As observed in Section 1.3.3, the Lebesgue measure m is invariant under R_θ . We want to analyze the ergodic behavior of the system (R_θ, m) for different values of θ .

If θ is rational, say $\theta = p/q$ in irreducible form, $R_\theta^q(x) = x$ for every $x \in S^1$. Then, given any segment $I \subset S^1$ with length less than $1/q$, the set

$$A = I \cup R_\theta(I) \cup \dots \cup R_\theta^{q-1}(I)$$

is invariant under R_θ and its Lebesgue measure satisfies $0 < m(A) < 1$. Thus, if θ is rational then the Lebesgue measure is *not* ergodic. The converse is much more interesting:

Proposition 4.2.1. *If θ is irrational then R_θ is ergodic relative to the Lebesgue measure.*

We are going to mention two different proofs of this fact. The first one, which we detail below, uses some simple facts from Fourier analysis. The second one, which we leave as an exercise (Exercise 4.2.6), is based on a density point argument similar to the one we will use in Section 4.2.2 to prove that the decimal expansion map is ergodic relative to the Lebesgue measure.

We denote by $L^2(m)$ the Hilbert space of measurable functions ψ whose square is integrable, that is, such that:

$$\int |\psi|^2 dm < \infty.$$

It is convenient to consider functions with values in \mathbb{C} , and we will do so. We use the well-known fact that the family of functions

$$\phi_k : S^1 \rightarrow \mathbb{C}, x \mapsto e^{2\pi ikx}, \quad k \in \mathbb{Z}$$

is a Hilbert basis of this space: given any $\varphi \in L^2(m)$ there exists a unique sequence $(a_k)_{k \in \mathbb{Z}}$ of complex numbers such that

$$\varphi(x) = \sum_{k \in \mathbb{Z}} a_k e^{2\pi ikx} \quad \text{for almost every } x \in S^1. \quad (4.2.1)$$

This is called the *Fourier series* expansion of $\varphi \in L^2(m)$. Then

$$\varphi(R_\theta(x)) = \sum_{k \in \mathbb{Z}} a_k e^{2\pi ik\theta} e^{2\pi ikx}. \quad (4.2.2)$$

Assume that φ is invariant. Then (4.2.1) and (4.2.2) coincide. By uniqueness of the coefficients in the Fourier series expansion, this happens if and only if

$$a_k e^{2\pi i k \theta} = a_k \quad \text{for every } k \in \mathbb{Z}.$$

The hypothesis that θ is irrational means that $e^{2\pi i k \theta} \neq 1$ for every $k \neq 0$. Hence, the relation that we just obtained implies that $a_k = 0$ for every $k \neq 0$. In other words, $\varphi(z) = a_0$ for m -almost every $z \in S^1$. This proves that every invariant L^2 function is constant m -almost everywhere. In particular, the characteristic function $\varphi = \chi_A$ of any invariant set $A \subset S^1$ is constant at m -almost every point. This is the same as saying that A has either zero measure or full measure. Hence, by Proposition 4.1.3, the measure m is ergodic.

These observations extend in a natural way to the rotation on the d -torus \mathbb{T}^d , for any $d \geq 1$:

Proposition 4.2.2. *If $\theta = (\theta_1, \dots, \theta_d)$ is rationally independent then the rotation $R_\theta : \mathbb{T}^d \rightarrow \mathbb{T}^d$ is ergodic relative to the Lebesgue measure.*

This may be proved by the same argument as in the case $d = 1$, using the fact (see Exercise 4.2.1) that the family of functions

$$\phi_{k_1, \dots, k_d} : \mathbb{T}^d \rightarrow \mathbb{C}, (x_1, \dots, x_d) \mapsto e^{2\pi i(k_1 x_1 + \dots + k_d x_d)}, \quad (k_1, \dots, k_d) \in \mathbb{Z}^d$$

is a Hilbert basis of the space $L^2(m)$.

Corollary 4.2.3. *If $\theta = (\theta_1, \dots, \theta_d)$ is rationally independent then the rotation $R_\theta : \mathbb{T}^d \rightarrow \mathbb{T}^d$ is minimal, that is, every orbit $\mathcal{O}(x) = \{R_\theta^n(x) : n \in \mathbb{N}\}$ is dense in \mathbb{T}^d .*

Proof. Let us consider in \mathbb{T}^d the flat distance, defined by:

$$d([\xi], [\eta]) = \inf\{d(\xi', \eta') : \xi', \eta' \in \mathbb{R}^d, \xi' \sim \xi, \eta' \sim \eta\}.$$

Observe that this distance is preserved by every rotation. Let $\{U_k : k \in \mathbb{N}\}$ be a countable basis of open sets of \mathbb{T}^d and m be the Lebesgue measure on \mathbb{T}^d . By ergodicity, there is $W \subset \mathbb{T}^d$, with total Lebesgue measure, such that $\tau(U_k, x) = m(U_k) > 0$ for every k and every $x \in W$. In particular, the orbit of x is dense in \mathbb{T}^d for every $x \in W$. Now consider an arbitrary point $x \in M$ and consider any $y \in W$. Then, for every $\delta > 0$ there exists $k \geq 1$ such that $d(f^k(y), x) < \delta$. It follows that $d(f^{n+k}(y), f^n(x)) < \delta$ for every $n \geq 1$. Since the orbit of y is dense, this implies that the orbit of x is δ -dense, that is, it intersects the δ -neighborhood of every point. Since δ is arbitrary, this implies that the orbit of x is dense in the ambient torus.

In fact, the irrational rotations on the circle or, more generally, on any torus have a much stronger property than ergodicity: they are *uniquely ergodic*, meaning that they admit a unique invariant probability measure (which is the Lebesgue measure, of course). Uniquely ergodic systems are studied in Chapter 6.

4.2.2 Decimal expansion

Consider the transformation $f : [0, 1] \rightarrow [0, 1]$, $f(x) = 10x - [10x]$ introduced in Section 1.3.1. We have seen that f preserves the Lebesgue measure m .

Proposition 4.2.4. *The transformation f is ergodic relative to the Lebesgue measure m .*

Proof. By Proposition 4.1.3, it suffices to prove that every invariant set A has total measure. The main ingredient is the derivation theorem (Theorem A.2.15), according to which almost every point of A is a density point of A . More precisely (see also Exercise A.2.9), m -almost every point $a \in A$ satisfies

$$\liminf_{\varepsilon \rightarrow 0} \left\{ \frac{m(I \cap A)}{m(I)} : I \text{ an interval such that } a \in I \subset B(a, \varepsilon) \right\} = 1. \quad (4.2.3)$$

Let us fix a density point $a \in A$. Since the set of points of the form $m/10^k$, $k \in \mathbb{N}$, $0 \leq m \leq 10^k$ has zero measure, it is no restriction to suppose that a is not of that form. Let us consider the family of intervals

$$I(k, m) = \left(\frac{m-1}{10^k}, \frac{m}{10^k} \right), \quad k \in \mathbb{N}, \quad m = 1, \dots, 10^k.$$

It is clear that for each $k \in \mathbb{N}$ there exists a unique $m = m_k$ such that $I(k, m_k)$ contains the point a . Denote $I_k = I(k, m_k)$. The property (4.2.3) implies that

$$\frac{m(I_k \cap A)}{m(I_k)} \rightarrow 1 \quad \text{when } k \rightarrow \infty.$$

Observe also that each f^k is an affine bijection from I_k to the interval $(0, 1)$. This has the following immediate consequence, which is crucial for our argument:

Lemma 4.2.5 (Distortion). *For every $k \in \mathbb{N}$, one has*

$$\frac{m(f^k(E_1))}{m(f^k(E_2))} = \frac{m(E_1)}{m(E_2)} \quad (4.2.4)$$

for any measurable subsets E_1 and E_2 of I_k .

Applying this fact to $E_1 = I_k \cap A$ and $E_2 = I_k$ we find that

$$\frac{m(f^k(I_k \cap A))}{m((0, 1))} = \frac{m(I_k \cap A)}{m(I_k)}.$$

Clearly, $m((0, 1)) = 1$. Moreover, as we take A to be invariant, $f^k(I_k \cap A)$ is contained in A . In this way we get that

$$m(A) \geq \frac{m(I_k \cap A)}{m(I_k)} \quad \text{for every } k.$$

Since the sequence on the right-hand side converges to 1 when $k \rightarrow \infty$, it follows that $m(A) = 1$, as we wanted to prove.

The proof of Lemma 4.2.5 relies on the fact that the transformation f is affine on each interval $((m - 1)/10, m/10)$; that may give the impression that the method of proof that we just presented is restricted to a very special class of examples. In fact, this is not so—much to the contrary.

The reason is that there are many situations where one can obtain a slightly weaker version of the statement of Lemma 4.2.5 that is, nevertheless, still sufficient to conclude the proof of ergodicity. In a few words, instead of the claim that the two sides of (4.2.4) are equal, one can often show that the quotient between the two terms is bounded by some uniform constant. That is called the *bounded distortion property*. As an illustration of these ideas, in Section 4.2.4 we prove that the Gauss transformation is ergodic.

Next, we describe an application of Proposition 4.2.4 in the context of number theory. We say that a number $x \in \mathbb{R}$ is *10-normal* if every block of digits $(b_1, \dots, b_l), l \geq 1$ occurs with frequency 10^{-l} in the decimal expansion of x . Rational numbers are never 10-normal, of course, and it is also easy to give irrational examples, such as $x = 0.101001000100001000001 \dots$. Moreover, it is not difficult to construct 10-normal numbers, for example, the Champernowne constant $x = 0.12345678910111213141516171819202122 \dots$, which is obtained by concatenation of the successive natural numbers.

However, it is usually difficult to check whether a given number is 10-normal or not. For example, that remains unknown for the numbers π, e and even $\sqrt{2}$. On the other hand, using the previous proposition one can easily prove that almost every number is 10-normal:

Proposition 4.2.6. *The set of 10-normal numbers $x \in \mathbb{R}$ has full Lebesgue measure in the real line.*

Proof. Since the fact of being 10-normal or not is independent of the integer part of the number, we only need to show that almost every $x \in [0, 1]$ is 10-normal. Consider $f : [0, 1] \rightarrow [0, 1]$ defined by $f(x) = 10x - [10x]$. For each block $(b_1, \dots, b_l) \in \{0, \dots, 9\}^l$, consider the interval

$$I_{b_1, \dots, b_l} = \left[\frac{\kappa}{10^l}, \frac{\kappa + 1}{10^l} \right) \quad \text{with } \kappa = \sum_{i=1}^l b_i 10^{l-i}.$$

Recall that if $x = 0.a_0a_1 \dots a_k a_{k+1} \dots$ then $f^k(x) = 0.a_k a_{k+1} \dots$ for every $k \geq 1$. Hence, $f^k(x) \in I_{b_1, \dots, b_l}$ if and only if $(a_k, \dots, a_{k+l-1}) = (b_1, \dots, b_l)$. So, the mean sojourn time $\tau(I_{b_1, \dots, b_l}, x)$ is equal to the frequency of the block (b_1, \dots, b_l) in the decimal expansion of x . Using the Birkhoff ergodic theorem and the fact that the transformation f is ergodic with respect to the Lebesgue measure m , we conclude that for every (b_1, \dots, b_l) there exists a full Lebesgue measure subset $B(b_1, \dots, b_l)$ of the interval $[0, 1]$ such that

$$\tau(I_{b_1, \dots, b_l}, x) = m(I_{b_1, \dots, b_l}) = \frac{1}{10^l} \quad \text{for every } x \in B(b_1, \dots, b_l).$$

Let B be the intersection of $B(b_1, \dots, b_l)$ over all values of b_1, \dots, b_l in $\{0, \dots, 9\}$ and every $l \geq 1$. Then $m(B) = 1$ and every $x \in B$ is 10-normal.

More generally, for any integer $d \geq 2$, we say that $x \in \mathbb{R}$ is d -normal if every block $(b_1, \dots, b_l) \in \{0, \dots, d-1\}^l$, $l \geq 1$ occurs with frequency d^{-l} in the expansion of x in base d . Finally, we say that x is a *normal number* if it is d -normal for every $d \geq 2$. Everything that was said before for $d = 10$ extends immediately to general d . In particular, the set of d -normal numbers has full Lebesgue measure for every $d \geq 2$. Taking the intersection over all the values of d , we conclude that *Lebesgue-almost every real number is normal* (Borel normal theorem).

4.2.3 Bernoulli shifts

Let (X, \mathcal{C}, ν) be a probability space. In this section we consider the product space $\Sigma = X^{\mathbb{N}}$, endowed with the product σ -algebra $\mathcal{B} = \mathcal{C}^{\mathbb{N}}$ and the product measure $\mu = \nu^{\mathbb{N}}$. As explained in Appendix A.2.3, this means that: Σ is the set of all sequences $(x_n)_{n \in \mathbb{N}}$ with $x_n \in X$ for every n ; \mathcal{B} is the σ -algebra generated by the measurable cylinders

$$[m; A_m, \dots, A_n] = \{(x_i)_{i \in \mathbb{N}} : x_i \in A_i \text{ for } m \leq i \leq n\}$$

with $m \leq n$ and $A_i \in \mathcal{C}$ for each i ; and μ is the probability measure on Σ characterized by

$$\mu([m; A_m, \dots, A_n]) = \prod_{i=m}^n \nu(A_i). \quad (4.2.5)$$

We may think of the elements of Σ as representing the results of a sequence of random experiments with values in X and all subject to the same probability distribution ν : given any measurable set $A \subset X$, the probability of $x_i \in A$ is equal to $\nu(A)$ for every i . Moreover, in this model the results of the successive experiments are independent: indeed, the relation (4.2.5) means that the probability of $x_i \in A_i$ for every $m \leq i \leq n$ is the product of the probabilities of the individual events $x_i \in A_i$.

In this section we introduce a dynamical system $\sigma : \Sigma \rightarrow \Sigma$ on the space Σ , called the *shift map*, which preserves the measure μ . The pair (σ, μ) is called a *Bernoulli shift*. The main result is that every Bernoulli shift is an ergodic system.

It is worth pointing out that \mathbb{N} may be replaced with \mathbb{Z} throughout the construction. That is, we may take Σ to be the space of two-sided sequences $(\dots, x_{-n}, \dots, x_0, \dots, x_n, \dots)$. Up to minor adjustments, which we leave to the reader, all that follows remains valid in that case. In addition, in the two-sided case the shift map is invertible.

The *shift map* $\sigma : \Sigma \rightarrow \Sigma$ is defined by

$$\sigma((x_n)_n) = (x_{n+1})_n.$$

That is, by definition, σ sends each sequence $(x_0, x_1, \dots, x_n, \dots)$ to the sequence (x_1, \dots, x_n, \dots) . Observe that the pre-image of any cylinder is still a cylinder:

$$\sigma^{-1}([m; A_m, \dots, A_n]) = [m+1; A_m, \dots, A_n]. \quad (4.2.6)$$

It follows that the map σ is measurable with respect to the σ -algebra \mathcal{B} . Moreover,

$$\mu(\sigma^{-1}([m; A_m, \dots, A_n])) = \nu(A_m) \cdots \nu(A_n) = \mu([m; A_m, \dots, A_n]),$$

and (using Lemma 1.3.1) that ensures that the measure μ is invariant under σ .

Proposition 4.2.7. *Every Bernoulli shift (σ, μ) is ergodic.*

Proof. Let A be an invariant measurable set. We want to prove that $\mu(A) = 0$ or $\mu(A) = 1$. We use the following fact:

Lemma 4.2.8. *If B and C are finite unions of pairwise disjoint cylinders, then*

$$\mu(B \cap \sigma^{-j}(C)) = \mu(B)\mu(\sigma^{-j}(C)) = \mu(B)\mu(C),$$

for every j sufficiently large.

Proof. First, suppose that B and C are both cylinders: $B = [k; B_k, \dots, B_l]$ and $C = [m; C_m, \dots, C_n]$. Then,

$$\sigma^{-j}(C) = [m+j; C_m, \dots, C_n] \quad \text{for each } j.$$

Consider any j large enough that $m+j > l$. Then,

$$\begin{aligned} B \cap \sigma^{-j}(C) &= \{(x_n)_n : x_k \in B_k, \dots, x_l \in B_l, x_{m+j} \in C_m, \dots, x_{n+j} \in C_n\} \\ &= [k; B_k, \dots, B_l, X, \dots, X, C_m, \dots, C_n], \end{aligned}$$

where X appears exactly $m+j-l-1$ times. By the definition (4.2.5), this gives that

$$\mu(B \cap \sigma^{-j}(C)) = \prod_{i=k}^l \nu(B_i) 1^{m+j-l-1} \prod_{i=m}^n \nu(C_i) = \mu(B)\mu(C).$$

This proves the conclusion of the lemma when both sets are cylinders. The general case follows easily, using the fact that μ is finitely additive.

Proceeding with the proof of Proposition 4.2.7, suppose for a while that the invariant set A belongs to the algebra \mathcal{B}_0 whose elements are the finite unions of pairwise disjoint cylinders. Then, on the one hand, we may apply the previous lemma with $B = C = A$ to conclude that $\mu(A \cap \sigma^{-j}(A)) = \mu(A)^2$ for every large j . On the other hand, since A is invariant, the left-hand side of this identity is equal to $\mu(A)$ for every j . It follows that $\mu(A) = \mu(A)^2$, which means that either $\mu(A) = 0$ or $\mu(A) = 1$.

Now let A be an arbitrary invariant set. By the approximation theorem (Theorem A.1.19), given any $\varepsilon > 0$ there exists $B \in \mathcal{B}_0$ such that $\mu(A \Delta B) < \varepsilon$. By Lemma 4.2.8 we may fix j such that

$$\mu(B \cap \sigma^{-j}(B)) = \mu(B)\mu(\sigma^{-j}(B)) = \mu(B)^2. \quad (4.2.7)$$

Using (4.1.8) and (4.1.9) and the fact that μ is invariant, we get that

$$|\mu(A \cap \sigma^{-j}(A)) - \mu(B \cap \sigma^{-j}(B))| \leq 2\mu(A \Delta B) < 2\varepsilon \quad (4.2.8)$$

(a similar fact was deduced during the proof of Corollary 4.1.5). Moreover,

$$|\mu(A)^2 - \mu(B)^2| \leq 2|\mu(A) - \mu(B)| < 2\varepsilon. \quad (4.2.9)$$

Putting the relations (4.2.7), (4.2.8) and (4.2.9) together, we conclude that $|\mu(A) - \mu(A)^2| < 4\varepsilon$. Since ε is arbitrary, we deduce that $\mu(A) = \mu(A)^2$ and, hence, either $\mu(A) = 0$ or $\mu(A) = 1$.

When X is a topological space and \mathcal{C} is the corresponding Borel σ -algebra, we may endow Σ with the *product topology* which, by definition, is the topology generated by the cylinders $[m; A_m, \dots, A_n]$ where A_m, \dots, A_n are open subsets of X . The property (4.2.6) implies that the shift map $\sigma : \Sigma \rightarrow \Sigma$ is continuous with respect to this topology. The theorem of Tychonoff (see [Dug66]) asserts that Σ is compact if X is compact.

A relevant special case is when X is a finite set endowed with the discrete topology, that is, such that every subset of X is open. A map $f : M \rightarrow M$ in a topological space M is said to be *transitive* if there exists some $x \in M$ whose trajectory $f^n(x)$, $n \geq 1$ is dense in M . We leave it to the reader (Exercise 4.2.2) to prove the following result:

Proposition 4.2.9. *Let X be a finite set and Σ be either $X^{\mathbb{N}}$ or $X^{\mathbb{Z}}$. Then the shift map $\sigma : \Sigma \rightarrow \Sigma$ is transitive. Moreover, the set of all periodic points of σ is dense in Σ .*

The following informal statement, which is one of many versions of the *monkey paradox*, illustrates the meaning of the ergodicity of the Bernoulli measure μ : *A monkey hitting keys at random on a typewriter keyboard for an infinite amount of time will almost surely type the complete text of “Os Lusíadas”.*¹

To “prove” this statement we need to formulate it a bit more precisely. The possible texts typed by the monkey correspond to the sequences $(x_n)_{n \in \mathbb{N}}$ in the (finite) set X of all the characters on the keyboard: letters, digits, space, punctuation signs, and so on. Denote by $\sigma : \Sigma \rightarrow \Sigma$ the shift map in the space $\Sigma = X^{\mathbb{N}}$. It is assumed that each character $* \in X$ has a positive probability p_*

¹ Monumental epic poem by the 16th-century Portuguese poet Luis de Camões.

of being hit at each time. This corresponds to the probability measure

$$\nu = \sum_{* \in X} p_* \delta_*$$

on the set X . Furthermore, it is assumed that the character hit at each time is independent of all the previous ones. This means that the distribution of the sequences of characters $(x_n)_n$ is governed by the Bernoulli probability measure $\mu = \nu^{\mathbb{N}}$.

The text of “Os Lusíadas” corresponds to a certain finite (albeit very long) sequence of characters (l_0, \dots, l_N) . Consider the cylinder $L = [0; l_0, \dots, l_N]$. Then

$$\mu(L) = \prod_{j=1}^N p_{l_j}$$

is positive (although very small). A sequence $(x_n)_n$ contains a complete copy of “Os Lusíadas” precisely if $\sigma^k((x_n)_n) \in L$ for some $k \geq 0$. By the Birkhoff ergodic theorem and the fact that (σ, μ) is ergodic, the set K of values of k for which that happens satisfies

$$\lim_n \frac{1}{n} \#(K \cap [0, n - 1]) = \mu(L) > 0, \tag{4.2.10}$$

with full probability. In particular, for almost all sequences $(x_n)_n$ the set K is infinite, which means that $(x_n)_n$ contains infinitely many copies of “Os Lusíadas”. Actually, (4.2.10) yields an even stronger conclusion: still with full probability, the copies of our poem correspond to a positive (although small) fraction of all the typed characters. In other words, on average, the monkey types a new copy of “Os Lusíadas” every so many (a great many) years.

4.2.4 Gauss map

As we have seen in Section 1.3.2, the gauss map $G(x) = 1/x - [1/x]$ has an invariant probability measure μ equivalent to the Lebesgue measure, namely:

$$\mu(E) = \frac{1}{\log 2} \int_E \frac{dx}{1+x}. \tag{4.2.11}$$

Proposition 4.2.10. *The system (G, μ) is ergodic.*

This can be proved using a more elaborate version of the method introduced in Section 4.2.2. We are going to outline the arguments in the proof, referring to Section 4.2.2 for those parts that are common to both situations and addressing in more detail the main new difficulty.

Let A be an invariant set with positive measure. We want to show that $\mu(A) = 1$. On the one hand, it remains true that for almost every point $a \in [0, 1]$ there exists a sequence of intervals I_k containing a and such that G^k maps I_k

bijectively and differentiably onto $(0, 1)$. Indeed, such intervals can be found as follows. First, consider

$$I(1, m) = \left(\frac{1}{m+1}, \frac{1}{m} \right),$$

for each $m \geq 1$. Next, define, by recurrence,

$$I(k, m_1, \dots, m_k) = I(1, m_1) \cap G^{-k+1}(I(k-1, m_2, \dots, m_k))$$

for $m_1, \dots, m_k \geq 1$. Then, it suffices to take as I_k the interval $I(k, m_1, \dots, m_k)$ that contains a . This is well defined for every $k \geq 1$ and every point a in the complement of a countable set, namely, the set $\bigcup_{k=0}^{\infty} G^{-k}(\{0, 1\})$.

On the other hand, although the restriction of G^k to each I_k is a differentiable bijection, it is not affine. For that reason, the analogue of (4.2.4) cannot hold in the present case. This difficulty is by passed by the result that follows, which is an example of *distortion* control: it is important to note that the constant K is independent of I_k, E_1, E_2 and, most of all, k .

Proposition 4.2.11 (Bounded distortion). *There exists $K > 1$ such that, given any $k \geq 1$ and any interval I_k such that G^k restricted to I_k is a differentiable bijection,*

$$\frac{\mu(G^k(E_1))}{\mu(G^k(E_2))} \leq K \frac{\mu(E_1)}{\mu(E_2)}$$

for any measurable subsets E_1 and E_2 of the interval I_k .

For the proof of this proposition we need the following two auxiliary results:

Lemma 4.2.12. *For every $x \in (0, 1]$ we have*

$$|G'(x)| \geq 1 \quad \text{and} \quad |(G^2)'(x)| \geq 2 \quad \text{and} \quad |G''(x)/G'(x)^2| \leq 2.$$

Proof. Recall that $G(x) = 1/x - m$ on each interval $(1/(m+1), 1/m]$. Therefore,

$$G'(x) = -\frac{1}{x^2} \quad \text{and} \quad G''(x) = \frac{2}{x^3}.$$

The first identity implies that $|G'(x)| \geq 1$ for every $x \in (0, 1]$. Moreover, $|G'(x)| \geq 2$ whenever $x \leq 2/3$. On the other hand, $x \geq 2/3$ implies that $G(x) = 1/x - 1 < 2/3$ and, consequently, $G'(G(x)) \geq 2$. Combining these observations we find that $|(G^2)'(x)| = |G'(x)| |G'(G(x))| \geq 2$ for every $x \in (0, 1]$. Finally, $|G''(x)/G'(x)^2| = 2|x| \leq 2$ also for every $x \in (0, 1]$.

Lemma 4.2.13. *There exists $C > 1$ such that, given any $k \geq 1$ and any interval I_k such that G^k restricted to I_k is a differentiable bijection,*

$$\frac{|(G^k)'(x)|}{|(G^k)'(y)|} \leq C \quad \text{for any } x \text{ and } y \text{ in } I_k.$$

Proof. Let g be a local inverse of G , that is, a differentiable function defined on some interval and such that $G(g(z)) = z$ for every z in the domain of definition. Note that

$$[\log |G' \circ g(z)|]' = \frac{G''(g(z)) g'(z)}{G'(g(z))} = \frac{G''(g(z))}{G'(g(z))^2}.$$

Therefore, the last estimate in Lemma 4.2.12 implies that

$$|[\log |G' \circ g(z)|]'| \leq 2 \quad \text{for every } g \text{ and every } z. \tag{4.2.12}$$

In other words, every function of the form $\log |G' \circ g|$ admits 2 as a Lipschitz constant. Observe also that if $x, y \in I_k$ then

$$\begin{aligned} \log \frac{|(G^k)'(x)|}{|(G^k)'(y)|} &= \sum_{j=0}^{k-1} \log |G'(G^j(x))| - \log |G'(G^j(y))| \\ &= \sum_{j=1}^k \log |G' \circ g_j(G^j(x))| - \log |G' \circ g_j(G^j(y))|, \end{aligned}$$

where g_j denotes a local inverse of G defined on the interval $[G^j(x), G^j(y)]$. Using the estimate (4.2.12), we get that

$$\log \frac{|(G^k)'(x)|}{|(G^k)'(y)|} \leq 2 \sum_{j=1}^k |G^j(x) - G^j(y)| = 2 \sum_{i=0}^{k-1} |G^{k-i}(x) - G^{k-i}(y)|. \tag{4.2.13}$$

Now, the first two estimates in Lemma 4.2.12 imply that

$$|G^k(x) - G^k(y)| \geq 2^{\lfloor i/2 \rfloor} |G^{k-i}(x) - G^{k-i}(y)|$$

for every $i = 0, \dots, k$. Replacing in (4.2.13), we conclude that

$$\log \frac{|(G^k)'(x)|}{|(G^k)'(y)|} \leq 2 \sum_{i=0}^{k-1} 2^{-\lfloor i/2 \rfloor} |G^k(x) - G^k(y)| \leq 8 |G^k(x) - G^k(y)| \leq 8.$$

Now it suffices to take $C = e^8$.

Proof of Proposition 4.2.11. Let m be the Lebesgue measure on $[0, 1]$. It follows from Lemma 4.2.13 that

$$\frac{m(G^k(E_1))}{m(G^k(E_2))} = \frac{\int_{E_1} |(G^k)'| dm}{\int_{E_2} |(G^k)'| dm} \leq C \frac{m(E_1)}{m(E_2)}.$$

On the other hand, the definition (4.2.11) implies that

$$\frac{1}{2 \log 2} m(E) \leq \mu(E) \leq \frac{1}{\log 2} m(E)$$

for every measurable set $E \subset [0, 1]$. Combining these two relations, we find that

$$\frac{\mu(G^k(E_1))}{\mu(G^k(E_2))} \leq 2 \frac{m(G^k(E_1))}{m(G^k(E_2))} \leq 2C \frac{m(E_1)}{m(E_2)} \leq 4C \frac{\mu(E_1)}{\mu(E_2)}.$$

Hence, it suffices to take $K = 4C$.

We are ready to conclude that (G, μ) is ergodic. Let A be an invariant set with $\mu(A) > 0$. Then A also has positive Lebesgue measure, since μ is absolutely continuous with respect to the Lebesgue measure. Let a be a density point of A whose future trajectory is contained in the open interval $(0, 1)$. Consider the sequence $(I_k)_k$ of the intervals $I(k, m_1, \dots, m_k)$ that contain a . It follows from Lemma 4.2.12 that

$$\text{diam } I_k \leq \sup \left\{ \frac{1}{|(G^k)'(x)|} : x \in I_k \right\} \leq 2^{-[k/2]}$$

for every $k \geq 1$. In particular, the diameter of I_k converges to zero and so

$$\frac{\mu(I_k \cap A)}{\mu(I_k)} \rightarrow 1 \quad \text{when } k \rightarrow \infty. \tag{4.2.14}$$

Let us take $E_1 = I_k \cap A^c$ and $E_2 = I_k$. By Proposition 4.2.11,

$$\frac{\mu(G^k(I_k \cap A^c))}{\mu(G^k(I_k))} \leq K \frac{\mu(I_k \cap A^c)}{\mu(I_k)}.$$

Observe that $G^k(I_k \cap A^c) = A^c$ up to a zero measure set, because the set A is assumed to be invariant. Recall also that $G^k(I_k) = (0, 1)$, which has full measure. Therefore, the previous inequality may be written as

$$\mu(A^c) \leq K \frac{\mu(I_k \cap A^c)}{\mu(I_k)}.$$

According to (4.2.14), the expression on the right-hand side converges to zero when $k \rightarrow \infty$. It follows that $\mu(A^c) = 0$, as we wanted to prove.

4.2.5 Linear endomorphisms of the torus

Recall that we call the torus of dimension d (or just d -torus) the quotient space $\mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d$, that is, the space of all equivalence classes of the equivalence relation defined in \mathbb{R}^d by $x \sim y \Leftrightarrow x - y \in \mathbb{Z}^d$. This quotient inherits from \mathbb{R}^d the structure of a differentiable manifold of dimension d . In what follows we assume that \mathbb{T}^d is also endowed with the flat Riemannian metric, which makes it locally isometric to the Euclidean space \mathbb{R}^d . Let m be the volume measure associated with this Riemannian metric (see Appendix A.4.5).

Let A be a d -by- d matrix with integer coefficients and determinant different from zero. Then $A(\mathbb{Z}^d) \subset \mathbb{Z}^d$ and, consequently, A induces a transformation

$$f_A : \mathbb{T}^d \rightarrow \mathbb{T}^d, \quad f_A([x]) = [A(x)],$$

where $[x]$ denotes the equivalence class that contains $x \in \mathbb{R}^d$. These transformations are called *linear endomorphisms* of the torus.

Note that f_A is differentiable and the derivative $Df_A([x])$ at each point is canonically identified with A . In particular, the Jacobian $\det Df_A([x])$ is constant equal to $\det A$. It follows (Exercise 4.2.9) that the degree of f is equal to $|\det A|$.

In particular, f_A is invertible if and only if $|\det A| = 1$. In this case, the inverse

is the transformation $f_{A^{-1}}$ induced by the inverse matrix A^{-1} ; observe that A^{-1} is also a matrix with integer coefficients.

In any case, f_A preserves the Lebesgue measure on \mathbb{T}^d . This may be seen as follows. Since f_A is a local diffeomorphism, the pre-image of any measurable set D with sufficiently small diameter consists of $|\det A|$ (= degree of f_A) pairwise disjoint sets D_i , each of which is mapped diffeomorphically onto D . By the formula of change of variables, $m(D) = |\det A| m(D_i)$ for every i . This proves that $m(D) = m(f_A^{-1}(D))$ for every measurable set D with small diameter. Hence, f_A does preserve the Lebesgue measure m . Next we prove the following fact:

Theorem 4.2.14. *The system (f_A, m) is ergodic if and only if no eigenvalue of the matrix A is a root of unity.*

Proof. Suppose that no eigenvalue of A is a root of unity. Consider any function $\varphi \in L^2(m)$ and let

$$\varphi([x]) = \sum_{k \in \mathbb{Z}^d} c_k e^{2\pi i(k \cdot x)}$$

be its Fourier series expansion (with $k \cdot x = k_1 x_1 + \dots + k_d x_d$). The coefficients $c_k \in \mathbb{C}$ satisfy

$$\sum_{k \in \mathbb{Z}^d} |c_k|^2 = \|\varphi\|_2^2 < \infty. \tag{4.2.15}$$

Then, the Fourier series expansion of $\varphi \circ f_A$ is:

$$\varphi(f_A([x])) = \sum_{k \in \mathbb{Z}^d} c_k e^{2\pi i(k \cdot A(x))} = \sum_{k \in \mathbb{Z}^d} c_k e^{2\pi i(A^*(k) \cdot x)},$$

where A^* denotes the adjoint of A . Suppose that φ is an invariant function, that is, $\varphi \circ f_A = \varphi$ at m -almost every point. Then, since the Fourier series expansion is unique, we must have

$$c_{A^*(k)} = c_k \quad \text{for every } k \in \mathbb{Z}. \tag{4.2.16}$$

We claim that the trajectory of every $k \neq 0$ under the transformation A^* is infinite. Indeed, if the trajectory of some $k \neq 0$ were finite then there would exist $l, r \in \mathbb{Z}$ with $r > 0$ such that $A^{(l+r)*}(k) = A^{l*}(k)$. This could only happen if A^* had some eigenvalue λ such that $\lambda^r = 1$. Since A and A^* have the same eigenvalues, that would mean that A has some eigenvalue which is a root of unity, which is excluded by the hypothesis. Hence, the trajectory of every $k \neq 0$ is infinite, as claimed. Then the identity (4.2.16), together with (4.2.15), implies that $c_k = 0$ for every $k \neq 0$. Thus, $\varphi = c_0$ at m -almost every point. This proves that the system (f_A, m) is ergodic.

To prove the converse, assume that A admits some eigenvalue which is a root of unity. Then the same holds for A^* and, hence, there exists $r > 0$ such that 1 is an eigenvalue of A^{r*} . Since A^{r*} has integer coefficients, it follows (see

Exercise 4.2.8) that there exists some $k \in \mathbb{Z}^d \setminus \{0\}$ such that $A^{r^*}(k) = k$. Fix k and consider the function $\varphi \in L^2(m)$ defined by

$$\varphi([x]) = \sum_{i=0}^{r-1} e^{2\pi i(A^{i^*}(k) \cdot x)} = \sum_{i=0}^{r-1} e^{2\pi i(k \cdot A^i(x))}.$$

Then φ is an invariant function for f_A and it is not constant at m -almost every point. Hence, (f_A, m) is not ergodic.

4.2.6 Hopf argument

In this section we present an alternative, more geometric, method to prove the ergodicity of certain linear endomorphisms of the torus. This is based on an argument introduced by Eberhard F. Hopf in his pioneering work [Hop39] on the ergodicity of geodesic flows on surfaces with negative Gaussian curvature.

In the present linear context, the Hopf argument may be used whenever $|\det A| = 1$ and the matrix A is hyperbolic, that is, A has no eigenvalues in the unit circle. But its strongest point is that it may be extended to much more general differentiable systems, not necessarily linear. Some of these extensions are mentioned in Section 4.4.

The hypothesis that the matrix A is hyperbolic means that the space \mathbb{R}^d may be written as a direct sum $\mathbb{R}^d = E^s \oplus E^u$ such that:

1. $A(E^s) = E^s$ and all the eigenvalues of $A|_{E^s}$ have absolute value smaller than 1;
2. $A(E^u) = E^u$ and all the eigenvalues of $A|_{E^u}$ have absolute value bigger than 1.

Then there exist constants $C > 0$ and $\lambda < 1$ such that

$$\begin{aligned} \|A^n(v^s)\| &\leq C\lambda^n \|v^s\| \quad \text{for every } v^s \in E^s \text{ and every } n \geq 0, \\ \|A^{-n}(v^u)\| &\leq C\lambda^n \|v^u\| \quad \text{for every } v^u \in E^u \text{ and every } n \geq 0. \end{aligned} \tag{4.2.17}$$

Example 4.2.15. Consider $A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$. The eigenvalues of A are

$$\lambda_u = \frac{3 + \sqrt{5}}{2} > 1 > \lambda_s = \frac{3 - \sqrt{5}}{2} > 0$$

and the corresponding eigenspaces are:

$$E^u = \left\{ (x, y) \in \mathbb{R}^2 : y = \frac{\sqrt{5} - 1}{2}x \right\} \quad \text{and} \quad E^s = \left\{ (x, y) \in \mathbb{R}^2 : y = -\frac{\sqrt{5} + 1}{2}x \right\}.$$

The family of all affine subspaces of \mathbb{R}^d of the form $v + E^s$, with $v \in \mathbb{R}^d$, defines a partition \mathcal{F}^s of \mathbb{R}^d that we call *stable foliation* and whose elements we call *stable leaves* of A . This partition is invariant under A , meaning that

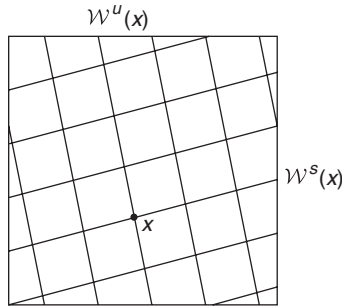


Figure 4.1. Stable foliation and unstable foliation in the torus

the image of any stable leaf is still a stable leaf. Moreover, by (4.2.17), the transformation A contracts distances uniformly inside each stable leaf. Analogously, the family of all affine subspaces of \mathbb{R}^d of the form $v + E^u$ with $v \in \mathbb{R}^d$ defines the *unstable foliation* \mathcal{F}^u of \mathbb{R}^d , whose elements are called *unstable leaves*. The unstable foliation is also invariant and the transformation A expands distances uniformly inside unstable leaves.

Mapping \mathcal{F}^s and \mathcal{F}^u by the canonical projection $\pi : \mathbb{R}^d \rightarrow \mathbb{T}^d$, we obtain foliations \mathcal{W}^s and \mathcal{W}^u of the torus that we call *stable foliation* and *unstable foliation* of the transformation f_A . See Figure 4.1. The previous observations show that these foliations are invariant under f_A . Moreover:

- (i) $d(f_A^j(x), f_A^j(y)) \rightarrow 0$ when $j \rightarrow +\infty$, for any points x and y in the same stable leaf;
- (ii) $d(f_A^j(y), f_A^j(z)) \rightarrow 0$ when $j \rightarrow -\infty$, for any points y and z in the same unstable leaf.

We are going to use this geometric information to prove that (f_A, m) is ergodic. To that end, let $\varphi : \mathbb{T}^d \rightarrow \mathbb{R}$ be any continuous function and consider the time averages

$$\varphi^+(x) = \lim_n \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f_A^j(x)) \quad \text{and} \quad \varphi^-(x) = \lim_n \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f_A^{-j}(x)),$$

which are defined for m -almost every $x \in \mathbb{T}^d$. By Corollary 3.2.8, there exists a full measure set $X \subset \mathbb{T}^d$ such that

$$\varphi^+(x) = \varphi^-(x) \quad \text{for every } x \in X. \tag{4.2.18}$$

Let us denote by $\mathcal{W}^s(x)$ and $\mathcal{W}^u(x)$, respectively, the stable leaf and the unstable leaf of f_A through each point $x \in \mathbb{T}^d$.

Lemma 4.2.16. *The function φ^+ is constant on each leaf of \mathcal{W}^s : if $\varphi^+(x)$ exists and $y \in \mathcal{W}^s(x)$ then $\varphi^+(y)$ exists and it is equal to $\varphi^+(x)$. Analogously, φ^- is constant on each leaf of \mathcal{W}^u .*

Proof. According to property (i) above, $d(f_A^j(x), f_A^j(y))$ converges to zero when $j \rightarrow \infty$. Noting that φ is uniformly continuous (because its domain is compact), it follows that

$$\varphi(f_A^j(x)) - \varphi(f_A^j(y)) \rightarrow 0 \quad \text{when } j \rightarrow \infty.$$

In particular, the Cesàro limit

$$\lim_n \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f_A^j(x)) - \varphi(f_A^j(y))$$

is also zero. That implies that $\varphi^+(y)$ exists and is equal to $\varphi^+(x)$. The argument for φ^- is entirely analogous.

Given any open subset R of the torus and any $x \in R$, denote by $\mathcal{W}^s(x, R)$ the connected component of $\mathcal{W}^s(x) \cap R$ that contains x and by $\mathcal{W}^u(x, R)$ the connected component of $\mathcal{W}^u(x) \cap R$ that contains x . We call R a *rectangle* if $\mathcal{W}^s(x, R)$ intersects $\mathcal{W}^u(y, R)$ at a unique point, for every x and y in R . See Figure 4.2.

Lemma 4.2.17. *Given any rectangle $R \subset \mathbb{T}^d$, there exists a measurable set $Y_R \subset X \cap R$ such that $m(R \setminus Y_R) = 0$ and, given any x and y in Y_R , there exist points x' and y' in $X \cap R$ such that $x' \in \mathcal{W}^s(x, R)$ and $y' \in \mathcal{W}^s(y, R)$ and $y' \in \mathcal{W}^u(x')$.*

Proof. Let us denote by m_x^s the Lebesgue measure on the stable leaf $\mathcal{W}^s(x)$ of each point $x \in \mathbb{T}^d$. Note that $m(R \setminus X) = 0$, since X has full measure in \mathbb{T}^d . Then, by the theorem of Fubini,

$$m_x^s(\mathcal{W}^s(x, R) \setminus X) = 0 \quad \text{for } m\text{-almost every } x \in R.$$

Define $Y_R = \{x \in X \cap R : m_x^s(\mathcal{W}^s(x, R) \setminus X) = 0\}$. Then Y_R has full measure in R . Given $x, y \in R$, consider the map $\pi : \mathcal{W}^s(x, R) \rightarrow \mathcal{W}^s(y, R)$ defined by

$$\pi(x') = \text{intersection between } \mathcal{W}^u(x', R) \text{ and } \mathcal{W}^s(y, R).$$

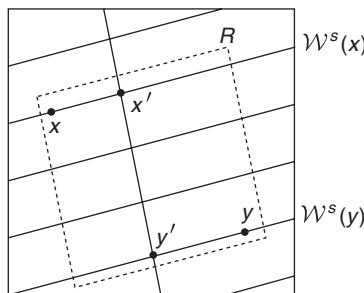


Figure 4.2. Rectangle in \mathbb{T}^d

This map is affine and, consequently, it has the following property, called *absolute continuity*:

$$m_x^s(E) = 0 \iff m_y^s(\pi(E)) = 0.$$

In particular, the image of $\mathcal{W}^s(x,R) \cap X$ has full measure in $\mathcal{W}^s(y,R)$ and, consequently, it intersects $\mathcal{W}^s(y,R) \cap X$. So, there exists $x' \in \mathcal{W}^s(x,R) \cap X$ whose image $y' = \pi(x')$ is in $\mathcal{W}^s(y,R) \cap X$. Observing that x' and y' are in the same unstable leaf, by the definition of π , we see that these points satisfy the conditions in the conclusion of the lemma.

Consider any rectangle R . Given any x, y in Y_R , consider the points x', y' in X given by Lemma 4.2.17. Using Lemma 4.2.16 as well, we obtain

$$\varphi^-(x) = \varphi^+(x) = \varphi^+(x') = \varphi^-(x') = \varphi^-(y') = \varphi^+(y') = \varphi^+(y) = \varphi^-(y).$$

This shows that the functions φ^+ and φ^- coincide with one another and are constant in Y_R .

Now let R_1, \dots, R_N be a finite cover of the torus by rectangles. Consider the set

$$Y = \bigcup_{j=1}^N Y_j, \quad \text{where } Y_j = Y_{R_j}.$$

Observe that $m(Y) = 1$, since $Y \cap R_j \supset Y_j$ has full measure in R_j for every j . We claim that $\varphi^+ = \varphi^-$ is constant on the whole Y . Indeed, given any $k, l \in \{1, \dots, N\}$ we may find $j_0 = k, j_1, \dots, j_{n-1}, j_n = l$ such that each R_{j_i} intersects $R_{j_{i-1}}$ (that is just because the torus is path-connected). Recalling that R_j is an open set and Y_j is a full measure subset, we get that each Y_{j_i} intersects $Y_{j_{i-1}}$. Then, $\varphi^+ = \varphi^-$ is constant on the union of all the Y_{j_i} . This proves our claim.

In this way, we have shown that the time averages φ^\pm of any continuous function φ are constant at m -almost every point. Consequently (see Exercise 4.1.4), the system (f_A, m) is ergodic.

4.2.7 Exercises

- 4.2.1. Prove Proposition 4.2.2.
- 4.2.2. Prove Proposition 4.2.9.
- 4.2.3. Let $I = [0, 1]$ and $f : I \rightarrow I$ be the function defined by

$$f(x) = \begin{cases} 2x & \text{if } 0 \leq x < 1/3 \\ 2x - 2/3 & \text{if } 1/3 \leq x < 1/2 \\ 2x - 1/3 & \text{if } 1/2 \leq x < 2/3 \\ 2x - 1 & \text{if } 2/3 \leq x \leq 1. \end{cases}$$

Show that f is ergodic with respect to the Lebesgue measure m .

- 4.2.4. Let X be a finite set and $\Sigma = X^{\mathbb{N}}$. Prove that every infinite compact subset of Σ invariant under the shift map $\sigma : \Sigma \rightarrow \Sigma$ contains some non-periodic point.

- 4.2.5. Let X be a topological space, endowed with the corresponding Borel σ -algebra \mathcal{C} , and let $\Sigma = X^{\mathbb{N}}$. Show that if X has a countable basis of open sets then the Borel σ -algebra of Σ (for the product topology) coincides with the product σ -algebra $\mathcal{B} = \mathcal{C}^{\mathbb{N}}$. The same is true for $\Sigma = X^{\mathbb{Z}}$ and $\mathcal{B} = \mathcal{C}^{\mathbb{Z}}$.
- 4.2.6. In this exercise we propose an alternative proof of Proposition 4.2.1. Assume that θ is irrational. Let A be an invariant set with positive measure. Recalling that the orbit $\{R_\theta^n(a) : n \in \mathbb{Z}\}$ of every $a \in S^1$ is dense in S^1 , show that no point of S^1 is a density point of A^c . Conclude that $\mu(A) = 1$.
- 4.2.7. Assume that θ is irrational. Let $\varphi : S^1 \rightarrow \mathbb{R}$ be any continuous function. Show that

$$\tilde{\varphi}(x) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi(R_\theta^j(x)) \tag{4.2.19}$$

exists at every point and, in fact, the limit is uniform. Deduce that $\tilde{\varphi}$ is constant at every point. Conclude that R_θ has a unique invariant probability measure.

- 4.2.8. Let A be a square matrix of dimension d with rational coefficients and let λ be a rational eigenvalue of A . Show that there exists some eigenvector with integer coefficients, that is, some $k \in \mathbb{Z}^d \setminus \{0\}$ such that $Ak = \lambda k$.
- 4.2.9. Show that if $f : M \rightarrow M$ is a local diffeomorphism on a compact Riemannian manifold then

$$\text{degree } f = \int |\det Df| \, dm,$$

where m denotes the volume measure induced by the Riemannian metric of M , normalized in such a way that $m(M) = 1$. In particular, for any square matrix A of dimension d with integer coefficients, the degree of the linear endomorphism $f_A : \mathbb{T}^d \rightarrow \mathbb{T}^d$ is equal to $|\det A|$.

- 4.2.10. A number $x \in (0, 1)$ has continued fraction expansion of bounded type if the sequence $(a_n)_n$ constructed in Section 1.3.2 is bounded. Prove that the set $\mathcal{L} \subset (0, 1)$ of points with continued fraction expansion of bounded type has Lebesgue measure zero.
- 4.2.11. Let $f : M \rightarrow M$ be a measurable transformation, μ be an ergodic invariant measure and $\varphi : M \rightarrow \mathbb{R}$ be a measurable function with $\int \varphi \, d\mu = +\infty$. Prove that $\lim_n (1/n) \sum_{j=0}^{n-1} \varphi(f^j(x)) = +\infty$ for μ -almost every $x \in M$.
- 4.2.12. Observe that the number b in Exercise 3.2.4 is independent of x in a set with full Lebesgue measure. Prove that the arithmetic mean of the numbers a_1, \dots, a_n, \dots goes to infinity: $\lim_n (1/n)(a_1 + \dots + a_n) = +\infty$.

4.3 Properties of ergodic measures

In this section we take the transformation $f : M \rightarrow M$ to be fixed and we analyze the set $\mathcal{M}_e(f)$ of probability measures that are ergodic with respect to f as a subset of the space $\mathcal{M}_1(f)$ of all probability measures invariant under f .

Recall that a measure ν is said to be *absolutely continuous* with respect to another measure μ if $\mu(E) = 0$ implies $\nu(E) = 0$. Then we write $\nu \ll \mu$. This

relation is transitive: if $\nu \ll \mu$ and $\mu \ll \lambda$ then $\nu \ll \lambda$. The first result asserts that the ergodic probability measures are minimal for this order relation:

Lemma 4.3.1. *If μ and ν are invariant probability measures such that μ is ergodic and ν is absolutely continuous with respect to μ then $\mu = \nu$.*

Proof. Let $\varphi : M \rightarrow \mathbb{R}$ be any bounded measurable function. Since μ is invariant and ergodic, the time average

$$\tilde{\varphi}(x) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} \varphi(f^j(x))$$

is constant: $\tilde{\varphi}(x) = \int \varphi d\mu$ at μ -almost every point. Since $\nu \ll \mu$, it follows that the equality also holds at ν -almost every point. In particular,

$$\int \varphi d\nu = \int \tilde{\varphi} d\nu = \int \varphi d\mu$$

(the first equality is part of the Birkhoff ergodic theorem). Therefore, the integrals of each bounded measurable function φ with respect to μ and with respect to ν coincide. In particular, considering characteristic functions, we conclude that $\mu = \nu$.

It is clear that if μ_1 and μ_2 are probability measures invariant under the transformation f then so is $(1 - t)\mu_1 + t\mu_2$, for any $t \in (0, 1)$. This means that the space $\mathcal{M}_1(f)$ of all probability measures invariant under f is *convex*. The next proposition asserts that the ergodic probability measures are the *extremal elements* of this convex set:

Proposition 4.3.2. *An invariant probability measure μ is ergodic if and only if it is not possible to write it as $\mu = (1 - t)\mu_1 + t\mu_2$ with $t \in (0, 1)$ and $\mu_1, \mu_2 \in \mathcal{M}_1(f)$ with $\mu_1 \neq \mu_2$.*

Proof. To prove the “if” claim, assume that μ is not ergodic. Then there exists some invariant set A with $0 < \mu(A) < 1$. Define μ_1 and μ_2 to be the normalized restriction of μ to the set A and to its complement A^c , respectively:

$$\mu_1(E) = \frac{\mu(E \cap A)}{\mu(A)} \quad \text{and} \quad \mu_2(E) = \frac{\mu(E \cap A^c)}{\mu(A^c)}.$$

Since A and A^c are invariant sets and μ is an invariant measure, both μ_1 and μ_2 are still invariant probability measures. Moreover,

$$\mu = \mu(A)\mu_1 + \mu(A^c)\mu_2$$

and, consequently, μ is not extremal.

To prove the converse, assume that μ is ergodic and $\mu = (1 - t)\mu_1 + t\mu_2$ for some $t \in (0, 1)$. It is clear that $\mu(E) = 0$ implies $\mu_1(E) = \mu_2(E) = 0$, that is, μ_1 and μ_2 are absolutely continuous with respect to μ . Hence, by Lemma 4.3.1, $\mu_1 = \mu = \mu_2$. This shows that μ is extremal.

Let us also point out that distinct ergodic measures “live” in disjoint subsets of the space M (see also Exercise 4.3.6):

Lemma 4.3.3. *Assume that the σ -algebra of M admits some countable generating subset Γ . Let $\{\mu_i : i \in \mathcal{I}\}$ be an arbitrary family of ergodic probability measures, all distinct. Then these measures μ_i are mutually singular: there exist pairwise disjoint measurable subsets $\{P_i : i \in \mathcal{I}\}$ invariant under f and such that $\mu_i(P_i) = 1$ for every $i \in \mathcal{I}$.*

Proof. Let \mathcal{A} be the algebra generated by Γ . Note that \mathcal{A} is countable, since it coincides with the union of the (finite) algebras generated by the finite subsets of Γ . For each $i \in \mathcal{I}$, define

$$P_i = \bigcap_{A \in \mathcal{A}} \{x \in M : \tau(A, x) = \mu_i(A)\}.$$

Since μ_i is ergodic, $\{x \in M : \tau(A, x) = \mu_i(A)\}$ has full measure for each A . Using that \mathcal{A} is countable, it follows that $\mu_i(P_i) = 1$ for every $i \in \mathcal{I}$. Moreover, if there exists $x \in P_i \cap P_j$ then $\mu_i(A) = \tau(A, x) = \mu_j(A)$ for every $A \in \mathcal{A}$. In other words, $\mu_i = \mu_j$. This proves that the P_i are pairwise disjoint.

Now assume that $f : M \rightarrow M$ is a continuous transformation in a topological space M . We say that f is *transitive* if there exists some $x \in M$ such that $\{f^n(x) : n \in \mathbb{N}\}$ is dense in M . The next lemma provides a useful characterization of transitivity. Recall that a topological space M is called a *Baire space* if the intersection of any countable family of open dense subsets is dense in M . Every complete metric space is a Baire space and the same is true for every locally compact topological space (see [Dug66]).

Lemma 4.3.4. *Let M be a Baire space with a countable basis of open sets. Then $f : M \rightarrow M$ is transitive if and only if for every pair of open sets U and V there exists $k \geq 1$ such that $f^{-k}(U)$ intersects V .*

Proof. Assume that f is transitive and let $x \in M$ be a point whose orbit $\{f^n(x) : n \in \mathbb{N}\}$ is dense. Then there exists $m \geq 1$ such that $f^m(x) \in V$ and (using the fact that $\{f^n(x) : n > m\}$ is also dense) there exists $n > m$ such that $f^n(x) \in U$. Take $k = n - m$. Then $f^m(x) \in f^{-k}(U) \cap V$. This proves the “only if” part of the statement.

To prove the converse, let $\{U_j : j \in \mathbb{N}\}$ be a countable basis of open subsets of M . The hypothesis ensures that the open set $\bigcup_{k=1}^{\infty} f^{-k}(U_j)$ is dense in M for every $j \in \mathbb{N}$. Then the intersection

$$X = \bigcap_{j=1}^{\infty} \bigcup_{k=1}^{\infty} f^{-k}(U_j)$$

is a dense subset of M . In particular, it is non-empty. On the other hand, by definition, if $x \in X$ then for every $j \in \mathbb{N}$ there exists $k \geq 1$ such that $f^k(x) \in U_j$.

Since the U_j constitute a basis of open subsets of M , this means that $\{f^k(x) : k \in \mathbb{N}\}$ is dense in M .

Proposition 4.3.5. *Let M be a Baire space with a countable basis of open sets. If μ is an ergodic probability measure then the restriction of f to the support of μ is transitive.*

Proof. Start by noting that $\text{supp } \mu$ has a countable basis of open sets, because it is a subspace of M , and it is a Baire space, since it is closed in M . Let U and V be open subsets of $\text{supp } \mu$. By the definition of support, $\mu(U) > 0$ and $\mu(V) > 0$. Define $B = \bigcup_{k=1}^{\infty} f^{-k}(U)$. Then $\mu(B) > 0$, because $B \supset U$, and $f^{-1}(B) \subset B$. By ergodicity (see Exercise 1.1.4) it follows that $\mu(B) = 1$. Then B must intersect V . This proves that there exists $k \geq 1$ such that $f^{-k}(U)$ intersects V . By Lemma 4.3.4, it follows that the restriction $f : \text{supp } \mu \rightarrow \text{supp } \mu$ is transitive.

4.3.1 Exercises

- 4.3.1. Let M be a topological space M with a countable basis of open sets, $f : M \rightarrow M$ be a measurable transformation and μ be an ergodic probability measure. Show that the orbit $\{f^n(x) : n \geq 0\}$ of μ -almost every point $x \in M$ is dense in the support of μ .
- 4.3.2. Let $f : M \rightarrow M$ be a continuous transformation in a compact metric space. Given a function $\varphi : M \rightarrow \mathbb{R}$, prove that there exists an invariant probability measure μ_φ such that

$$\int \varphi d\mu_\varphi = \sup_{\eta \in \mathcal{M}_1(f)} \int \varphi d\eta.$$

- 4.3.3. Let $g : E \rightarrow E$ be a transformation induced by $f : M \rightarrow M$, that is, a transformation of the form $g(x) = f^{\rho(x)}(x)$ with $\rho : E \rightarrow \mathbb{N}$ (see Section 1.4.2). Let ν be an invariant probability measure of g and ν_g be the invariant measure of f defined by (1.4.5). Assume that $\nu_g(M) < \infty$ and denote $\mu = \nu_g/\nu_g(M)$. Show that (f, μ) is ergodic if and only if (g, ν) is ergodic.
- 4.3.4. Let $f : M \rightarrow M$ be a continuous transformation in a separable complete metric space. Given any invariant probability measure μ , let $\hat{\mu}$ be its lift to the natural extension $\hat{f} : \hat{M} \rightarrow \hat{M}$ (see Section 2.4.2). Show that $(\hat{f}, \hat{\mu})$ is ergodic if and only if (f, μ) is ergodic.
- 4.3.5. Let $f : M \rightarrow M$ be a measurable transformation and μ be an invariant measure. Let $g^t : N \rightarrow N$, $t \in \mathbb{R}$ be a suspension flow of f and ν be the corresponding suspension of the measure μ (see Section 3.4.1). Assume that $\nu(N) < \infty$ and denote $\hat{\nu} = \nu/\nu(N)$. Show that $\hat{\nu}$ is ergodic for the flow $(g^t)_t$ if and only if μ is ergodic for f .
- 4.3.6. Show that for finite or countable families of ergodic measures the conclusion of Lemma 4.3.3 holds even if the σ -algebra is not countably generated.
- 4.3.7. Give an example of a metric space M and a transformation $f : M \rightarrow M$ such that there exists a sequence of ergodic Borel measures μ_n converging, in the weak* topology, to a non-ergodic invariant measure μ .

- 4.3.8. Let M be a metric space, $f : M \rightarrow M$ be a continuous transformation and μ be an ergodic probability measure. Show that $\frac{1}{n} \sum_{j=0}^{n-1} f_*^j \nu$ converges to μ in the weak* topology for any probability measure ν on M absolutely continuous with respect to μ , but not necessarily invariant.
- 4.3.9. Let $X = \{1, \dots, d\}$ and $\sigma : \Sigma \rightarrow \Sigma$ be the shift map in $\Sigma = X^{\mathbb{N}}$ or $\Sigma = X^{\mathbb{Z}}$.
- (1) Show that for every $\delta > 0$ there exists $k \geq 1$ such that, given $x^1, \dots, x^s \in \Sigma$ and $m_1, \dots, m_s \geq 1$, there exists a periodic point $y \in \Sigma$ with period n_s and such that $d(f^{j+n_i}(y), f^j(x^i)) < \delta$ for every $0 \leq j < m_i$, where $n_1 = 0$ and $n_i = (m_1 + k) + \dots + (m_{i-1} + k)$ for $1 < i \leq s$.
 - (2) Let $\varphi : \Sigma \rightarrow \mathbb{R}$ be a continuous function and $\tilde{\varphi}$ be its Birkhoff average. Show that, given $\varepsilon > 0$, points $x^1, \dots, x^s \in \Sigma$ where the Birkhoff average of φ is well defined, and numbers $\alpha^1, \dots, \alpha^s > 0$ such that $\sum_i \alpha^i = 1$, there exists a periodic point $y \in \Sigma$ satisfying $|\tilde{\varphi}(y) - \sum_i \alpha^i \tilde{\varphi}(x^i)| < \varepsilon$.
 - (3) Conclude that the set $\mathcal{M}_e(\sigma)$ of ergodic probability measures is dense in the space $\mathcal{M}_1(\sigma)$ of all invariant probability measures.

4.4 Comments in conservative dynamics

The ergodic theorem of Birkhoff, proven in the 1930's, provided a solid mathematical foundation to the statement of the Boltzmann ergodic hypothesis, but left entirely open the question of its *veracity*. In this section we briefly survey the main results obtained since then, in the context of *conservative* systems, that is, dynamical systems that preserve a volume measure on a manifold.

Let us start by mentioning that, in a certain abstract sense, the majority of conservative systems are ergodic. That is the sense of the theorem that we state next, which was proven in the early 1940's by John Oxtoby and Stanislaw Ulam [OU41]. Recall that a subset of a Baire space is called *residual* if it may be written as a countable intersection of open and dense subsets. By the definition of Baire space, every residual subset is dense.

Theorem 4.4.1 (Oxtoby, Ulam). *For every compact Riemannian manifold M there exists a residual subset \mathcal{R} of the space $\text{Homeo}_{\text{vol}}(M)$ of all conservative homeomorphisms of M such that every element of \mathcal{R} is ergodic.*

The results presented below imply that the conclusion of this theorem is no longer true when one replaces $\text{Homeo}_{\text{vol}}(M)$ by the space $\text{Diffeo}_{\text{vol}}^k(M)$ of conservative diffeomorphisms of class C^k , at least for $k > 3$. Essentially nothing is known in this regard in the cases $k = 2$ and $k = 3$. On the other hand, Artur Avila, Sylvain Crovisier and Amie Wilkinson have recently announced a C^1 version of the previous theorem: *for every compact Riemannian manifold M , there exists a residual subset \mathcal{R} of the space $\text{Diffeo}_{\text{vol}}^1(M)$ of conservative diffeomorphisms of class C^1 such that every $f \in \mathcal{R}$ with positive entropy $h_{\text{vol}}(f)$ is ergodic.* The notion of entropy will be studied in Chapter 9.

4.4.1 Hamiltonian systems

The systems that interested Boltzmann, relative to the motion of gas molecules, may, in principle, be described by the laws of Newtonian classical mechanics. In the so-called Hamiltonian formalism of classical mechanics, the states of the system are represented by “generalized coordinates” q_1, \dots, q_d and “generalized momenta” p_1, \dots, p_d , and the system’s evolution is described by the solutions of the Hamilton–Jacobi equations:

$$\frac{dq_j}{dt} = \frac{\partial H}{\partial p_j} \quad \text{and} \quad \frac{dp_j}{dt} = -\frac{\partial H}{\partial q_j}, \quad j = 1, \dots, d, \tag{4.4.1}$$

where H (the total energy of the system) is a C^2 function of the variables $q = (q_1, \dots, q_d)$ and $p = (p_1, \dots, p_d)$; the integer $d \geq 1$ is the number of degrees of freedom.

Example 4.4.2 (Harmonic pendulum). Let $d = 1$ and $H(q, p) = p^2/2 - g \cos q$, where g is a positive constant and $(q, p) \in \mathbb{R}^2$. The Hamilton–Jacobi equations

$$\frac{dq}{dt} = p \quad \text{and} \quad \frac{dp}{dt} = -g \sin q$$

describe the motion of a pendulum subject to a constant gravitational field: the coordinate q measures the angle with respect to the position of (stable) equilibrium and p measures the angular momentum. Then $p^2/2$ is the kinetic energy and $-g \cos q$ is the potential energy. Thus, the *Hamiltonian* H is the total energy.

Note that H is always a *first integral* of the system, that is, it is constant along the flow trajectories:

$$\frac{dH}{dt} = \sum_{j=1}^d \frac{\partial H}{\partial q_j} \frac{dq_j}{dt} + \frac{\partial H}{\partial p_j} \frac{dp_j}{dt} \equiv 0.$$

Thus, we may consider the restriction of the flow to each *energy hypersurface* $H_c = \{(q, p) : H(q, p) = c\}$. The volume measure $dq_1 \cdots dq_d dp_1 \cdots dp_d$ is called the *Liouville measure*. Observe that the divergence of the vector field

$$F = \left(-\frac{\partial H}{\partial p_1}, \dots, -\frac{\partial H}{\partial p_d}, \frac{\partial H}{\partial q_1}, \dots, \frac{\partial H}{\partial q_d} \right)$$

is identically zero. Thus (recall Section 1.3.6) the Liouville measure is invariant under the Hamiltonian flow. It follows (see Exercise 1.3.12) that the restriction of the flow to each energy hypersurface H_c admits an invariant measure μ_c that is given by

$$\mu_c(E) = \int_E \frac{ds}{\|\text{grad}H\|} \quad \text{for every measurable set } E \subset H_c,$$

where ds denotes the volume element on the hypersurface. Then, the ergodic hypothesis may be viewed as claiming that, in general, Hamiltonian systems

are ergodic with respect to this invariant measure μ_c on (almost) every energy hypersurface.

The first important result in this context was announced by Andrey Kolmogorov at the International Congress of Mathematicians ICM 1954 and was substantiated, soon afterwards, by the works of Vladimir Arnold and Jürgen Moser. This led to the deep theory of so-called *almost integrable systems* that is known as KAM theory, in homage to its founders, and to which several other mathematicians contributed in a decisive manner, including Helmut Rüssmann, Michael Herman, Eduard Zehnder, Jean-Christophe Yoccoz and Jürgen Pöschel, among others. Let us explain what is meant by “almost integrable”.

A Hamiltonian system with d degrees of freedom is said to be *integrable* (in the sense of Liouville) if it admits d first integrals I_1, \dots, I_d :

- *independent*: that is, such that the gradients

$$\text{grad} I_j = \left(\frac{\partial I_j}{\partial q_1}, \frac{\partial I_j}{\partial p_1}, \dots, \frac{\partial I_j}{\partial q_d}, \frac{\partial I_j}{\partial p_d} \right), \quad 1 \leq j \leq d,$$

are linearly independent at every point on an open and dense subset of the domain;

- *in involution*: that is, such that the Poisson brackets

$$\{I_j, I_k\} = \sum_{i=1}^d \left[\frac{\partial I_j}{\partial q_i} \frac{\partial I_k}{\partial p_i} - \frac{\partial I_j}{\partial p_i} \frac{\partial I_k}{\partial q_i} \right]$$

are all identically zero.

It follows from the previous remarks that every system with $d = 1$ degree of freedom is integrable: the Hamiltonian H itself is a first integral. Another important example:

Example 4.4.3. For any number $d \geq 1$ of degrees of freedom, assume that the Hamiltonian H depends only on the variables $p = (p_1, \dots, p_d)$. Then the Hamilton–Jacobi equations (4.4.1) reduce to

$$\frac{dq_j}{dt} = \frac{\partial H}{\partial p_j}(p) \quad \text{and} \quad \frac{dp_j}{dt} = -\frac{\partial H}{\partial q_j}(p) = 0.$$

The second equation means that each p_j is a first integral; it is easy to see that the first integrals are independent and in involution. Then the expression on the right-hand side of the first equation is independent of time. Hence, the solution is given by

$$q_j(t) = q_j(0) + \frac{\partial H}{\partial p_j}(p(0))t.$$

As we are going to comment in the following, this example is totally typical of integrable systems.

A classical theorem of Liouville asserts that if the system is integrable then the Hamilton–Jacobi equations may be solved completely by quadratures. In the proof (see Arnold [Arn78]) one constructs certain functions $\varphi = (\varphi_1, \dots, \varphi_d)$ with values in \mathbb{T}^d which, together with the first integrals $I = (I_1, \dots, I_d) \in \mathbb{R}^d$, constitute *canonical coordinates* of the system (they are called *action-angle coordinates*). What we mean by “canonical” is that the coordinate change

$$\Psi : (q, p) \mapsto (\varphi, I)$$

preserves the form of the Hamilton–Jacobi equations: (4.4.1) becomes

$$\frac{d\varphi_j}{dt} = \frac{\partial H'}{\partial I_j} \quad \text{and} \quad \frac{dI_j}{dt} = -\frac{\partial H'}{\partial \varphi_j}, \tag{4.4.2}$$

where $H' = H \circ \Psi^{-1}$ is the expression of the Hamiltonian in the new coordinates. Since the I_j are first integrals, the second equation yields

$$0 = \frac{dI_j}{dt} = -\frac{\partial H'}{\partial \varphi_j}.$$

This means that H does not depend on the variables φ_j and so we are in the type of situation described in Example 4.4.3. Each trajectory of the Hamiltonian flow is constrained inside a torus $\{I = \text{const}\}$ and, according to the first equation in (4.4.2), it is linear in the coordinate φ :

$$\varphi_j(t) = \varphi_j(0) + \omega_j(I)t, \quad \text{where } \omega_j(I) = \frac{\partial H'}{\partial I_j}(I).$$

In terms of the original coordinates (q, p) , we conclude that the trajectories of the Hamiltonian flow are given by

$$t \mapsto \Psi^{-1}(\varphi(0) + \omega(I)t, I) = \Phi_{\varphi(0), I}(\omega(I)t), \tag{4.4.3}$$

where $\Phi_{\varphi(0), I} : \mathbb{R}^d \rightarrow M$ is a \mathbb{Z}^d -periodic function and $\omega(I) = (\omega_1(I), \dots, \omega_d(I))$ is called a *frequency vector*. We say that the trajectory is *quasi-periodic*.

4.4.2 Kolmogorov–Arnold–Moser theory

It is clear that integrable systems are never ergodic. However, since integrability is a very rare property, this alone would not be an obstruction to most Hamiltonian systems being ergodic. Nevertheless, the fundamental result that we state next asserts that generic integrable systems are *robustly* non-ergodic: every nearby Hamiltonian flow is also non-ergodic.

Let H_0 be an integrable Hamiltonian, written in action-angle coordinates (φ, I) . More precisely, let B^d be a ball in \mathbb{R}^d and assume that $H_0(\varphi, I)$ is defined for every $(\varphi, I) \in \mathbb{T}^d \times B^d$ but depends only on the coordinate I . We call H_0 *non-degenerate* if its Hessian matrix is invertible:

$$\det \left(\frac{\partial^2 H_0}{\partial I_i \partial I_j} \right)_{i,j} \neq 0 \quad \text{at every point.} \tag{4.4.4}$$

Observe that the Hessian matrix of H_0 coincides with the Jacobian matrix of the function $I \mapsto \omega(I)$. Therefore, the *twist condition* (4.4.4) means that the map assigning to each value of I the corresponding frequency vector $\omega(I)$ is a local diffeomorphism.

The next theorem means that, under this condition, most of the invariant tori of the Hamiltonian flow of H_0 persist for any nearby system:

Theorem 4.4.4. *Let H_0 be an integrable non-degenerate Hamiltonian of class C^∞ . Then there exists a neighborhood \mathcal{V} of H_0 in the space $C^\infty(\mathbb{T}^d \times B^d, \mathbb{R})$ such that for every $H \in \mathcal{V}$ there exists a compact set $K \subset \mathbb{T}^d \times B^d$ satisfying:*

- (i) *K is a union of differentiable tori of the form $\{(\varphi, u(\varphi)) : \varphi \in \mathbb{T}^d\}$ each of which is invariant under the Hamiltonian flow of H ;*
- (ii) *the restriction of the Hamiltonian flow of H to each of these tori is conjugate to a linear flow on \mathbb{T}^d ;*
- (iii) *the set K has positive volume and, in fact, the volume of its complement goes to zero when $H \rightarrow H_0$.*

In particular, the Hamiltonian flow of H cannot be ergodic.

The latter is because the set K may be decomposed into positive volume subsets that are also unions of invariant tori and, thus, are invariant. The proof of the theorem shows that the persistence or not of a given invariant torus of H_0 is intimately related to the arithmetic properties of the corresponding frequency vector. Let us explain this.

Given $c > 0$ and $\tau > 0$, we say that a vector $\omega_0 \in \mathbb{R}^d$ is (c, τ) -Diophantine if

$$\|k \cdot \omega_0\| \geq \frac{c}{\|k\|^\tau} \quad \text{for every } k \in \mathbb{Z}^d, \quad (4.4.5)$$

where $\|k\| = |k_1| + \dots + |k_d|$. Diophantine vectors are rationally independent; in fact, the condition (4.4.5) means that ω_0 is badly approximated by rationally dependent vectors. We say that ω_0 is τ -Diophantine if it is (c, τ) -Diophantine for some $c > 0$. The set of τ -Diophantine vectors is non-empty if and only if $\tau \geq d - 1$; moreover, it has full measure in \mathbb{R}^d if τ is strictly larger than $d - 1$ (see Exercise 4.4.1).

While proving Theorem 4.4.4, it is shown that, given $c > 0$, $\tau \geq d - 1$ and any compact set $\Omega \subset \omega(B^d)$, one can find a neighborhood \mathcal{V} of H_0 such that, for every $H \in \mathcal{V}$ and every (c, τ) -Diophantine vector $\omega_0 \in \Omega$, the Hamiltonian flow of H admits a differentiable invariant torus restricted to which the flow is conjugate to the linear flow $t \mapsto \varphi(t) = \varphi(0) + t\omega_0$.

Next, we discuss a version of Theorem 4.4.4 for discrete time systems or, more precisely, symplectic transformations. We call a *symplectic manifold* (see Arnold [Arn78, Chapter 8]) any differentiable manifold M endowed with a *symplectic form*, that is, a non-degenerate differential 2-form θ . By “non-degenerate” we mean that for every $x \in M$ and every $u \neq 0$ there exists

v such that $\theta_x(u, v) \neq 0$. Existence of a symplectic form on M implies that the dimension is even: write $\dim M = 2d$. Moreover, the d -th power $\theta^d = \theta \wedge \dots \wedge \theta$ is a volume form on M .

A differentiable transformation $f : M \rightarrow M$ is said to be *symplectic* if it preserves the symplectic form, meaning that $\theta_x(u, v) = \theta_{f(x)}(Df(x)u, Df(x)v)$ for every $x \in M$ and any $u, v \in T_xM$. Then, in particular, f preserves the volume form θ^d .

Example 4.4.5. Let $M = \mathbb{R}^{2d}$, with coordinates $(q_1, \dots, q_d, p_1, \dots, p_d)$, and let θ be the differential 2-form defined by

$$\theta_x = dq_1 \wedge dp_1 + \dots + dq_d \wedge dp_d \tag{4.4.6}$$

for every x . Then θ is a symplectic form on M . Actually, a classical theorem of Darboux states that for every symplectic form there exists some atlas of the manifold such that the expression of the symplectic form in any local chart is of the type (4.4.6). Consider any transformation of the form

$$f_0(q_1, \dots, q_d, p_1, \dots, p_d) = (q_1 + \omega_1(p), \dots, q_d + \omega_d(p), p_1, \dots, p_d).$$

Using

$$Df_0 \cdot \frac{\partial}{\partial q_j} = \frac{\partial}{\partial q_j} \quad \text{and} \quad Df_0 \cdot \frac{\partial}{\partial p_j} = \frac{\partial}{\partial p_j} + \sum_i \frac{\partial \omega_i}{\partial p_j} \frac{\partial}{\partial q_i},$$

we see that f_0 is symplectic with respect to the form θ .

Example 4.4.6 (Cotangent bundle). Let M be a manifold of class C^r with $r \geq 3$. By definition, the *cotangent space* T_q^*M at each point $q \in M$ is the dual of the tangent space T_qM , and the *cotangent bundle* of M is the disjoint union $T^*M = \bigcup_{q \in M} T_q^*M$ of all cotangent spaces. See Appendix A.4.3. The cotangent bundle is a manifold of class C^{r-1} and the canonical projection $\pi^* : T^*M \rightarrow M$ mapping each T_q^*M to the corresponding base point q is a map of class C^{r-1} .

A very important feature of the cotangent bundle is that it always admits a canonical symplectic form, that is, one that depends only on the manifold M . That can be seen as follows. Let α be the differential 1-form on T^*M defined by

$$\alpha_{(q,p)} : T_{(q,p)}(T^*M) \rightarrow \mathbb{R}, \quad \alpha_{(q,p)} = p \circ D\pi^*(q, p)$$

for each $(q, p) \in T^*M$. It is clear that α is well defined and of class C^{r-2} . Consider the exterior derivative $\theta^* = d\alpha$. One can check (for instance, using local coordinates) that θ^* is non-degenerate at every point and, thus, is a symplectic form in T^*M .

There is no corresponding statement for the tangent bundle TM . However, once we fix a Riemannian metric on M it is possible to endow the tangent bundle with a (non-canonical) symplectic form:

Example 4.4.7 (Tangent bundle). Let M be a Riemannian manifold of class C^r with $r \geq 3$. Then we may identify the tangent bundle TM with the cotangent bundle T^*M through the map $\Xi : TM \rightarrow T^*M$ that maps each point (q, v) with $v \in T_qM$ to the point (q, p) with $p \in T_q^*M$ defined by

$$p : T_qM \rightarrow \mathbb{R}, \quad p(w) = v \cdot_q w.$$

Indeed, Ξ is a diffeomorphism and it maps fibers of TM to fibers of T^*M , preserving the base point. In particular, we may use Ξ to transport the symplectic form θ^* in Example 4.4.6 to a symplectic form θ in TM :

$$\theta_{(q,v)}(w_1, w_2) = \theta_{\Xi(q,v)}^*(D\Xi(q, v)w_1, D\Xi(q, v)w_2)$$

for any $w_1, w_2 \in T_{(q,v)}(TM)$. It is clear from the construction that, unlike θ^* , this form θ depends on the Riemannian metric in M .

By analogy with the case of flows, we call a transformation f_0 *integrable* if there exist coordinates $(q, p) \in \mathbb{T}^d \times B^d$ such that $f_0(q, p) = (q + \omega(p), p)$ for every (q, p) . Moreover, we say that f_0 is *non-degenerate* if

$$\text{the map } p \mapsto \omega(p) \text{ is a local diffeomorphism.} \tag{4.4.7}$$

Theorem 4.4.8. *Let f_0 be a non-degenerate integrable transformation of class C^∞ . Then there exists a neighborhood \mathcal{V} of f_0 in the space $C^\infty(\mathbb{T}^d \times B^d, \mathbb{R}^d)$ such that for every symplectic transformation² $f \in \mathcal{V}$ there exists a compact set $K \subset \mathbb{T}^d \times B^d$ satisfying:*

- (i) K is a union of differentiable tori of the form $\{(q, u(q)) : q \in \mathbb{T}^d\}$, each of which is invariant under f ;
- (ii) the restriction of the transformation f to each of these tori is conjugate to a translation on \mathbb{T}^d ;
- (iii) the set K has positive volume and, in fact, the volume of the complement converges to zero when $f \rightarrow f_0$.

In particular, the transformation f cannot be ergodic.

Just as in the previous (continuous time) situation, the set K is formed by tori restricted to which the dynamics is conjugate to a Diophantine rotation.

Theorems 4.4.4 and 4.4.8 extend to systems of class C^r with r finite but sufficiently large, depending on the dimension. For example, the version of Theorem 4.4.8 for $d = 1$ is true for $r > 3$ and false for $r < 3$; in the boundary case $r = 3$, parts (i) and (ii) of the theorem remain valid but part (iii) does not.

The notion of Hamiltonian flow extends to any symplectic manifold (M, θ) , as follows. Let $H : M \rightarrow \mathbb{R}$ be a function of class C^2 and $dH(z) : T_zM \rightarrow \mathbb{R}$ denote its derivative at each point $z \in M$. By the definition of symplectic form,

² Relative to the canonical symplectic form (4.4.6).

each $\theta_z : T_zM \times T_zM \rightarrow \mathbb{R}$ is a non-degenerate alternate 2-form. Hence, there exists exactly one vector $X_H(z) \in T_zM$ such that

$$\theta_z(X_H(z), v) = dH(z)v \quad \text{for every } v \in T_zM.$$

The map $z \mapsto X_H(z)$ is a vector field of class C^1 on the manifold M . This is called the *Hamiltonian vector field* associated with H . The corresponding flow, given by the differential equation

$$\frac{dz}{dt} = X_H(z), \tag{4.4.8}$$

is the *Hamiltonian flow* associated with H . We leave it to the reader to check that (4.4.8) corresponds precisely to the Hamilton–Jacobi equations (4.4.1) when $M = \mathbb{R}^{2d}$ and θ is the symplectic form in Example 4.4.5.

4.4.3 Elliptic periodic points

The ideas behind the results stated in the previous section may be used to describe the behavior of conservative systems in the neighborhood of elliptic periodic points. Let us explain this briefly, starting with the symplectic case in dimension 2.

When M is a surface, the notions of symplectic form and area form coincide. Thus, a differentiable transformation $f : M \rightarrow M$ is symplectic if and only if it preserves area. Let $\zeta \in M$ be an *elliptic* fixed point, that is, such that the eigenvalues of $Df(\zeta)$ are in the unit circle. Let λ and $\bar{\lambda}$ be the eigenvalues. We say that the fixed point ζ is *non-degenerate* if $\lambda^k \neq 1$ for every $1 \leq k \leq 4$. Then, by the Birkhoff normal form theorem (see Arnold [Arn78, Appendix 7]), there exist canonical coordinates $(x, y) \in \mathbb{R}^2$ in the neighborhood of the fixed point, with $\zeta = (0, 0)$, such that the transformation f has the form:

$$f(\theta, \rho) = (\theta + \omega_0 + \omega_1\rho, \rho) + R(\theta, \rho) \quad \text{with } |R(\theta, \rho)| \leq C|\rho|^2, \tag{4.4.9}$$

where $(\theta, \rho) \in S^1 \times \mathbb{R}$ are the “polar” coordinates defined by

$$x = \sqrt{\rho} \cos 2\pi\theta \quad \text{and} \quad y = \sqrt{\rho} \sin 2\pi\theta.$$

Observe that the *normal form* $f_0 : (\theta, \rho) \mapsto (\theta + \omega_0 + \omega_1\rho, \rho)$ is integrable. Moreover, f_0 satisfies the twist condition (4.4.7) as long as $\omega_1 \neq 0$ (this condition does not depend on the choice of the canonical coordinates, just on the transformation f). Then one may apply the methods of Theorem 4.4.8 to conclude that there exists a set K with positive area that is formed by invariant circles with *Diophantine rotation numbers*, that is, such that the restriction of f to each of these circles is conjugate to a Diophantine rotation. Even more, the fixed point ζ is a density point of this set:

$$\lim_{r \rightarrow 0} \frac{m(B(\zeta, r) \setminus K)}{m(B(\zeta, r))} = 0,$$

where $B(\zeta, r)$ represents the ball of radius $r > 0$ around ζ .

We will refer to points ζ as in the previous paragraph as *generic elliptic fixed points*. An important consequence of what we just said is that *generic elliptic fixed points of area-preserving transformations are stable*: the trajectory of any point close to ζ remains close to ζ for all times, as it is “trapped” on the inside of some small invariant circle. This feature does not extend to higher dimensions, as we will explain shortly.

Still in dimension two, we want to mention other important dynamical phenomena that take place in the neighborhood of generic elliptic fixed points. Let us start by presenting a very useful tool, known as the *Poincaré–Birkhoff fixed point theorem* or *Poincaré last theorem*. The statement was proposed by Poincaré, who also presented some special cases, a few months before his death; the general case was proved by Birkhoff [Bir13] in the following year.

Let $A = S^1 \times [a, b]$, with $0 < a < b$, and let $f : A \rightarrow A$ be a homeomorphism that preserves each of the boundary components of the annulus A . We say that f is a *twist homeomorphism* if it rotates the two boundary components in opposite senses or, more precisely, if there exists some lift $F : \mathbb{R} \times [a, b] \rightarrow \mathbb{R} \times [a, b]$, $F(\theta, \rho) = (\Theta(\theta, \rho), R(\theta, \rho))$, of the map f to the universal cover of the annulus, such that

$$[\Theta(\theta, a) - \theta][\Theta(\theta, b) - \theta] < 0 \quad \text{for every } \theta \in \mathbb{R}. \quad (4.4.10)$$

Theorem 4.4.9 (Poincaré–Birkhoff fixed point). *If $f : A \rightarrow A$ is a twist homeomorphism that preserves area then f admits at least two fixed points in the interior of A .*

As mentioned previously, every generic elliptic fixed point ζ is accumulated by invariant circles with Diophantine rotation numbers. Any two such disks bound an annulus around ζ . Applying Theorem 4.4.9 (or, more precisely, its corollary in Exercise 4.4.6) one gets that any such annulus contains, at least, a pair of periodic orbits with the same period.

In a sense, these pairs of periodic orbits are what is left of the invariant circles of the normal form f_0 with *rational* rotation numbers, which are usually destroyed by the addition of the term R in (4.4.9). Their periods go to infinity when one approaches ζ . Generically, one of these periodic orbits is hyperbolic (saddle points) and the other one is elliptic. An example is sketched in Figure 4.3: the elliptic fixed point ζ is surrounded by a hyperbolic periodic orbit and an elliptic periodic orbit, marked with the letters p and q , respectively, both with period 4. Two invariant circles around ζ are also represented.

The Swiss mathematician Eduard Zehnder proved that, generically, the hyperbolic periodic orbits exhibit transverse homoclinic points, that is, their stable manifolds and unstable manifolds intersect transversely, as depicted in Figure 4.3. This implies that the geometry of the stable manifolds and unstable manifolds is extremely complex. Moreover, the elliptic periodic orbits satisfy the genericity conditions mentioned previously. This means that all

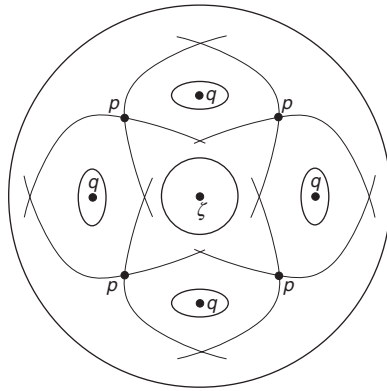


Figure 4.3. Invariant circles, periodic orbits and homoclinic intersections in the neighborhood of a generic elliptic fixed point

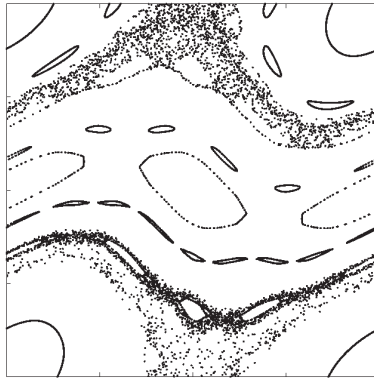


Figure 4.4. Computational evidence for the presence of invariant circles, elliptic islands and transverse homoclinic intersections

the dynamical complexity that we are describing in the neighborhood of ζ is reproduced in the neighborhood of each one of these “satellite” elliptic orbits (which have their own “satellites”, etc.).

Moreover, a theory developed by the French physicist Serge Aubry and the American mathematician John Mather shows that ζ is also accumulated by certain infinite, totally disconnected invariant sets, restricted to which the transformation f is minimal (all the orbits are dense). In a sense, these *Aubry–Mather sets* are a souvenir of the invariant circles of the normal form f_0 with irrational *non-Diophantine* rotation numbers that are also typically destroyed by the addition of the perturbation term R in (4.4.9).

Figure 4.4 illustrates a good part of what we have been saying. It depicts several computer-calculated trajectories of an area-preserving transformation. The behavior of these trajectories suggests the presence of invariant circles, elliptic satellites with their own invariant circles, and even hyperbolic orbits

with associated transverse homoclinic intersections. One can also observe the presence of certain trajectories with “chaotic” behavior, apparently related to those homoclinic intersections.

More generally, let $f : M \rightarrow M$ be a symplectic diffeomorphism on a symplectic manifold M of any (even) dimension $2d \geq 2$. We say that a fixed point $\zeta \in M$ is *elliptic* if all the eigenvalues of the derivative $Df(\zeta)$ are in the unit circle. Let $\lambda_1, \bar{\lambda}_1, \dots, \lambda_d, \bar{\lambda}_d$ be those eigenvalues. We say that ζ is *non-degenerate* if $\lambda_1^{k_1} \dots \lambda_d^{k_d} \neq 1$ for every $(k_1, \dots, k_d) \in \mathbb{Z}^d$ with $|k_1| + \dots + |k_d| \leq 4$ (in particular, the eigenvalues are all distinct). Then, by the Birkhoff normal form theorem (see Arnold [Arn78, Appendix 7]), there exist canonical coordinates $(x_1, \dots, x_d, y_1, \dots, y_d) \in \mathbb{R}^{2d}$ in a neighborhood of ζ such that $\zeta = (0, \dots, 0, 0, \dots, 0)$ and the transformation f has the form

$$f(\theta, \rho) = (\theta + \omega_0 + \omega_1(\rho), \rho) + R(\theta, \rho) \quad \text{with } \|R(\theta, \rho)\| \leq \text{const} \|\rho\|^2,$$

where $\omega_0 \in \mathbb{R}^d$, $\omega_1 : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is a linear map and $(\theta, \rho) \in \mathbb{T}^d \times \mathbb{R}^d$ are the “polar” coordinates defined by

$$x_j = \sqrt{\rho_j} \cos 2\pi \theta_j \quad \text{and} \quad y_j = \sqrt{\rho_j} \sin 2\pi \theta_j, \quad j = 1, \dots, d.$$

Assuming that ω_1 is an isomorphism (this is yet another generic condition on the transformation f), we have that the normal form

$$f_0 : (\theta, \rho) \mapsto (\theta + \omega_0 + \omega_1(\rho), \rho)$$

is integrable and satisfies the twist condition (4.4.7). Applying the ideas of Theorem 4.4.8, one concludes that ζ is a density point of a set K formed by invariant tori of dimension d , restricted to which the transformation f is conjugate to a Diophantine rotation.

In particular, symplectic transformations with generic elliptic fixed (or periodic) points are never ergodic. Observe, on the other hand, that for $d > 1$ a torus of dimension d does not separate the ambient space M into two connected components. Therefore, the argument we used before to conclude that generic elliptic fixed points on surfaces are stable does not extend to higher dimensions. In fact, it is known that when $d > 1$ elliptic fixed points are usually *unstable*: trajectories starting arbitrarily close to the fixed point may escape from a fixed neighborhood of it. This is related to the phenomenon known as *Arnold diffusion*, which is a very active research topic in this area.

Finally, let us mention that this theory also applies to continuous time conservative systems. We say that a stationary point ζ of a Hamiltonian flow is *elliptic* if all the eigenvalues of the derivative of the vector field at the point ζ are pure imaginary numbers. Arguments similar to those in the discrete time case show that, under generic hypothesis, ζ is a density point of a set formed by invariant tori of dimension d restricted to each of which the Hamiltonian flow is conjugate to a linear flow.

Moreover, there are corresponding results for periodic trajectories of Hamiltonian flows. One way to obtain such results is by considering a cross-section to the flow at some point of the periodic trajectory and applying the previous ideas to the corresponding Poincaré map. In this way one finds that, under generic conditions, elliptic periodic trajectories of Hamiltonian flows are accumulated by sets with positive volume consisting of invariant tori of the flow.

The theory of Kolmogorov–Arnold–Moser has many other applications, in a wide variety of situations in mathematics that go beyond the scope of this book. The reader may find more complete information in the following references: Arnold [Arn78], Bost [Bos86], Yoccoz [Yoc92], de la Llave [dLL93] and Arnold, Kozlov and Neishtadt [AKN06], among others.

4.4.4 Geodesic flows

Let M be a compact Riemannian manifold. Some of the notions that are used here are recalled in Appendix A.4.

It follows from the theory of ordinary differential equations that for each $(x, v) \in TM$ there exists a unique geodesic $\gamma_{x,v} : \mathbb{R} \rightarrow M$ of the manifold M such that $\gamma_{x,v}(0) = x$ and $\dot{\gamma}_{x,v}(0) = v$. Moreover, the family of transformations defined by

$$f^t : (x, v) \mapsto (\gamma_{x,v}(t), \dot{\gamma}_{x,v}(t))$$

is a flow on the tangent bundle TM , which is called the *geodesic flow* of M . We denote by T^1M the unit tangent bundle, formed by the pairs $(x, v) \in TM$ with $\|v\| = 1$. The unit tangent bundle is invariant under the geodesic flow.

Equivalently, the geodesic flow may be defined as the Hamiltonian flow in the tangent bundle TM (with the symplectic structure defined in Example 4.4.7) associated with the Hamiltonian function $H(x, v) = \|v\|^2$. So, $(f^t)_t$ preserves the Liouville measure of the tangent bundle.

In this context, the Liouville measure may be described as follows. Every inner product in a finite-dimensional vector space induces a *volume element*³ in that space, relative to which the cube spanned by any orthonormal basis has volume 1. In particular, the Riemannian metric induces a volume element dv on each tangent space T_xM . Integrating this volume element along M , we get a volume measure dx on the manifold itself. The Liouville measure of TM is given, locally, by the product $dx dv$. Moreover, its restriction m to the unit tangent bundle is given, locally, by the product $dx da$, where da is the measure of angle on the unit sphere of T_xM .

The fact that H is a first integral means that the norm $\|v\|$ is constant along trajectories of the flow. In particular, $(f^t)_t$ leaves the unit tangent bundle

³ That is, a volume form defined up to sign: the sign is not determined because the inner product does not detect the orientation of the vector space.

invariant. Furthermore, the geodesic flow preserves the restriction m of the Liouville measure to T^1M . However, the behavior of geodesic flows is, usually, very different from the dynamics of the almost integrable systems that we described in Section 4.4.2.

For example, the Austrian mathematician Eberhard F. Hopf [Hop39] proved in 1939 that if M is a compact surface with *negative Gaussian curvature* at every point then its geodesic flow is ergodic. Almost three decades later, his theorem was extended to manifolds in any dimension, through the following remarkable result of the Russian mathematician Dmitry Anosov [Ano67]:

Theorem 4.4.10 (Anosov). *Let M be a compact manifold with negative sectional curvature. Then the geodesic flow on the unit tangent bundle is ergodic with respect to the Liouville measure on T^1M .*

Thus, the geodesic flows of manifolds with negative curvature were the first important class of Hamiltonian systems for which the ergodic hypothesis could be validated rigorously.

4.4.5 Anosov systems

There are two fundamental steps in the proof of Theorem 4.4.10. The first one is to show that every geodesic flow on a manifold with negative curvature is *uniformly hyperbolic*. This means that every trajectory γ of the flow is contained in invariant submanifolds $W^s(\gamma)$ and $W^u(\gamma)$ that intersect each other transversely along γ and satisfy:

- every trajectory in $W^s(\gamma)$ is exponentially asymptotic to γ in the future;
- every trajectory in $W^u(\gamma)$ is exponentially asymptotic to γ in the past;

(see Figure 4.5), with exponential convergence rates that are uniform, that is, independent of γ . Moreover, the geodesic flow is transitive. The second main step in the proof of Theorem 4.4.10 consists of showing that every transitive, uniformly hyperbolic flow (or transitive *Anosov flow*) of class C^2 that preserves volume is ergodic. We will comment on this last issue in a little while.

There exists a corresponding notion for discrete time systems: we say that a diffeomorphism $f : N \rightarrow N$ on a compact Riemannian manifold is *uniformly*

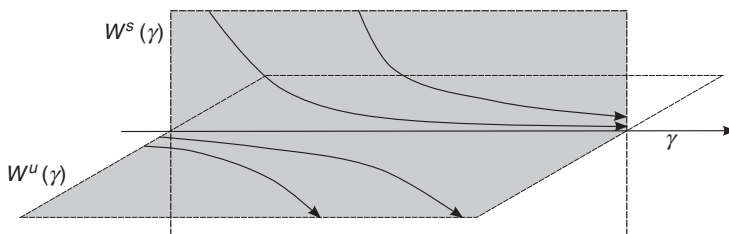


Figure 4.5. Hyperbolic behavior

hyperbolic (or an Anosov diffeomorphism) if the tangent space to the manifold at every point $z \in N$ admits a direct sum decomposition $T_z N = E_z^s \oplus E_z^u$ such that the decomposition is invariant under the derivative of f :

$$Df(z)E_z^s = E_{f(z)}^s \quad \text{and} \quad Df(z)E_z^u = E_{f(z)}^u \quad \text{for every } z \in N, \tag{4.4.11}$$

and the derivative contracts E_z^s and expands E_z^u , uniformly:

$$\sup_{z \in N} \|Df(z) | E_z^s\| < 1 \quad \text{and} \quad \sup_{z \in N} \|Df(z)^{-1} | E_z^u\| < 1 \tag{4.4.12}$$

(for some choice of a norm compatible with the Riemannian metric on M).

One can prove that for each $z \in N$ the set $W^s(z)$ of points whose forward trajectory is asymptotic to the trajectory of z is a differentiable (immersed) submanifold of N tangent to E_z^s at the point z ; analogously, the set $W^u(z)$ of points whose backward trajectory is asymptotic to the trajectory of z is a differentiable submanifold tangent to E_z^u at the point z . These submanifolds form foliations (that is, decompositions of N into differentiable submanifolds) that are invariant under the diffeomorphism:

$$f(W^s(z)) = W^s(f(z)) \quad \text{and} \quad f(W^u(z)) = W^u(f(z)) \quad \text{for every } z \in N.$$

We call $W^s(z)$ the *stable manifold* (or *stable leaf*) and $W^u(z)$ the *unstable manifold* (or *unstable leaf*) of the point $z \in M$.

Concerning the second part of the proof of Theorem 4.4.10, the crucial technical tool to prove that every transitive, uniformly hyperbolic diffeomorphism of class C^2 that preserves volume is ergodic is the following theorem of Anosov and Sinai [AS67]:

Theorem 4.4.11 (Absolute continuity). *The stable and unstable foliations of any Anosov diffeomorphism (or flow) of class C^2 are absolutely continuous:*

1. if $X \subset N$ has zero volume then $X \cap W^s(x)$ has volume zero inside $W^s(x)$ for almost every $x \in N$;
2. if $Y \subset \Sigma$ is a zero volume subset of some submanifold Σ transverse to the stable foliation, then the union of the stable manifolds through the points of Y has zero volume in N ;

and analogously for the unstable foliation.

Ergodicity of the system may then be deduced using the Hopf argument, which we introduced in a special case in Section 4.2.6. Let us explain this. Given any continuous function $\varphi : N \rightarrow \mathbb{R}$, let E_φ be the set of all points $z \in N$ for which the forward and backward time averages, $\varphi^+(z)$ and $\varphi^-(z)$, are well defined and coincide. This set E_φ has full volume, as we have seen in Corollary 3.2.8. Observe also that φ^+ is constant on each stable manifold and φ^- is constant on each unstable manifold. So, by the first part of Theorem 4.4.11, the intersection $Y_z = W^u(z) \cap E_\varphi$ has full volume in $W^u(z)$ for almost every $z \in N$. Moreover, $\varphi^- = \varphi^+$ is constant on each Y_z . Fix any

such z . The transitivity hypothesis implies that the union of all stable manifolds through the points of $W^u(z)$ is the whole ambient manifold N . Hence, using the second part of Theorem 4.4.11, the union of the stable manifolds through the points of Y_z has full volume in N . Clearly, φ^+ is constant on this union. This shows that the time average of every continuous function φ is constant on a full measure set. Hence, f is ergodic.

We close this section by observing that all the known examples of Anosov diffeomorphisms are transitive. The corresponding statement for Anosov flows is false (see Verjovsky [Ver99]). Another open problem in this setting is whether ergodicity still holds when the Anosov system is only of class C^1 . It is known (see [Bow75b, RY80]) that in this case the absolute continuity theorem (Theorem 4.4.11) is false, in general.

4.4.6 Billiards

As we have seen in Sections 4.4.2 and 4.4.3, non-ergodic systems are quite common in the realm of Hamiltonian flows and symplectic transformations. However, this fact alone is not sufficient to invalidate the ergodic hypothesis of Boltzmann in the context where it was formulated. Indeed, ideal gases are a special class of systems and it is conceivable that ergodicity could be typical in this more restricted setting, even it is not typical for general Hamiltonian systems.

In the 1960's, the Russian mathematician and theoretical physicist Yakov Sinai [Sin63] conjectured that Hamiltonian systems formed by spherical hard balls that hit each other elastically are ergodic. Hard ball systems (see Example 4.4.13 for a precise definition) had been proposed as a model for the behavior of ideal gases by the American scientist Josiah Willard Gibbs who, together with Boltzmann and Scottish mathematician and theoretical physicist James Clark Maxwell, created the area of statistical mechanics. The *ergodic hypothesis of Boltzmann–Sinai*, as Sinai's conjecture is often referred to, is the main topic in the present section.

In fact, we are going to discuss the problem of ergodicity for somewhat more general systems, called *billiards*, whose formal definition was first given by Birkhoff in the 1930's.

In its simplest form, a billiard is given by a bounded connected domain $\Omega \subset \mathbb{R}^2$, called the *billiard table*, whose boundary $\partial\Omega$ is formed by a finite number of differentiable curves. We call the *corners* those points of the boundary where it fails to be differentiable; by hypothesis, they constitute a finite set $\mathcal{C} \subset \partial\Omega$. One considers a point particle moving uniformly along straight lines inside Ω , with elastic reflections on the boundary. That is, whenever the particle hits $\partial\Omega \setminus \mathcal{C}$ it is reflected in such a way that the angle of incidence equals the angle of reflection. When the particle hits some corner it is absorbed: its trajectory is not defined from then on.

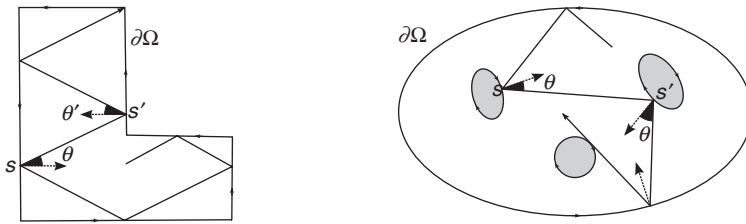


Figure 4.6. Dynamics of billiards

Let us denote by \mathbf{n} the unit vector field orthogonal to the boundary $\partial\Omega$ and pointing to the inside of Ω . It defines an orientation in $\partial\Omega \setminus \mathcal{C}$: a vector t tangent to the boundary is *positive* if the basis $\{t, \mathbf{n}\}$ of \mathbb{R}^2 is positive. It is clear that the motion of the particle is characterized completely by the sequence of collisions with the boundary. Moreover, each such collision may be described by the position $s \in \partial\Omega$ and the angle of reflection $\theta \in (-\pi/2, \pi/2)$. Therefore, the evolution of the billiard is governed by the transformation

$$f : (\partial\Omega \setminus \mathcal{C}) \times (-\pi/2, \pi/2) \rightarrow \partial\Omega \times (-\pi/2, \pi/2), \tag{4.4.13}$$

that associates with each collision (s, θ) the subsequent one (s', θ') . See Figure 4.6.

In the example on the left-hand side of Figure 4.6 the billiard table is a polygon, that is, the boundary consists of a finite number of straight line segments. The one trajectory represented in the figure hits one of the corners. Nearby trajectories, to either side, collide with distinct boundary segments, with very different angles of incidence. In particular, it is clear that the billiard transformation (4.4.13) cannot be continuous. Discontinuities may occur even in the absence of corners. For example, on the right-hand side of Figure 4.6 the boundary has four connected components, all of which are differentiable curves. Consider the trajectory represented in the figure, tangent to one of the boundary components. Nearby trajectories, to either side, hit with different boundary components. Consequently, the billiard map is discontinuous in this case also.

Example 4.4.12 (Circular billiard table). On the left-hand side of Figure 4.7 we represent a billiard in the unit ball $\Omega \subset \mathbb{R}^2$. The corresponding billiard transformation is given by

$$f : (s, \theta) \mapsto (s - (\pi - 2\theta), \theta).$$

The behavior of this transformation is described geometrically on the right-hand side of Figure 4.7. Observe that f preserves the area measure $ds d\theta$ and satisfies the twist condition (4.4.4). Note also that f is integrable (in the sense of Section 4.4.2) and, in particular, the area measure is not ergodic. We will see in a while (Theorem 4.4.14) that every planar billiard preserves a natural measure equivalent to the area measure on $\partial\Omega \times (-\pi/2, \pi/2)$.

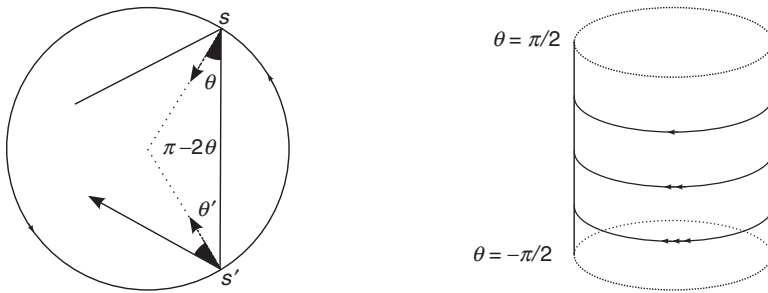


Figure 4.7. Billiard on a circular table

Then, using the previous observations, the KAM theory allows us to prove that billiards with almost circular tables are not ergodic with respect to that invariant measure.

The definition of billiard extends immediately to bounded connected domains Ω in any Euclidean space \mathbb{R}^d , $d \geq 1$, whose boundary consists of a finite number of differentiable hypersurfaces intersecting each other along submanifolds with codimension larger than 1. We denote by \mathcal{C} the union of the submanifolds. As before, we endow $\partial\Omega$ with the orientation induced by the unit vector \mathbf{n} orthogonal to the boundary and pointing to the “inside” of Ω . Elastic reflections on the boundary are defined by the following two conditions: (i) the incident trajectory segment, the reflected trajectory segment and the orthogonal vector \mathbf{n} are co-planar and (ii) the angle of incidence equals the angle of reflection. The billiard transformation is defined as in (4.4.13), having as domain

$$\{(s, v) \in (\partial\Omega \setminus \mathcal{C}) \times S^{d-1} : v \cdot \mathbf{n}(s) > 0\}.$$

Even more generally, we may take as a billiard table any bounded connected domain in a Riemannian surface, whose boundary is formed by a finite number of differentiable hypersurfaces intersecting along higher codimension submanifolds. The definitions are analogous, except that the trajectories between consecutive reflections on the boundary are given by segments of geodesics and angles are measured according to the Riemannian metric on the manifolds.

Example 4.4.13 (Ideal gases and billiards). Ideally, a gas is formed by a large number N of molecules ($N \approx 10^{27}$) that move uniformly along straight lines, between collisions, and collide with each other elastically. Check the right-hand side of Figure 4.8. For simplicity, let us assume that the molecules are identical spheres and that they are contained in the torus⁴ of dimension

⁴ One may replace the torus \mathbb{T}^d by a more plausible container, such as the d -dimensional cube $[0, 1]^d$, for example. However, the analysis is a bit more complicated in that case, because we must take into account the collisions of the balls with the container’s walls.

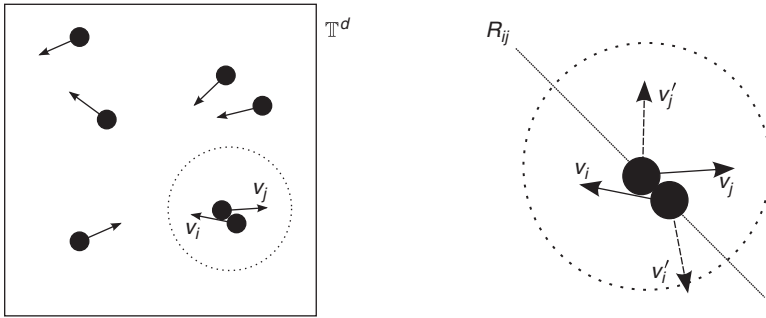


Figure 4.8. Model for an ideal gas

$d \geq 2$. Let us also assume that all the molecules move with constant unit speed. This system can be modelled by a billiard, as follows.

For $1 \leq i \leq N$, denote by $p_i \in \mathbb{T}^d$ the position of the center of the i -th molecule M_i . Let $\rho > 0$ be the radius of each molecule. Then, each state of the system is entirely described by a value of $p = (p_1, \dots, p_N)$ in the set

$$\Omega = \{p = (p_1, \dots, p_N) \in \mathbb{T}^{Nd} : \|p_i - p_j\| \geq 2\rho \text{ for every } i \neq j\}$$

(this set is connected, as long as the radius ρ is sufficiently small).

In the absence of collisions, the point p moves along a straight line inside Ω , with constant speed. When two molecules M_i and M_j collide, $\|p_i - p_j\| = 2\rho$ and the velocity vectors change in the following way. Let v_i and v_j be the velocity vectors of the two molecules immediately before the collision and let R_{ij} be the straight line through p_i and p_j . The elasticity hypothesis means that the velocity vectors v'_i and v'_j immediately after the collision are given by (check the right-hand side of Figure 4.8):

- (i) the components of v_i and v'_i in the direction of R_{ij} are symmetric and the same is true for v_j and v'_j ;
- (ii) the components of v_i and v'_i in the direction orthogonal to R_{ij} are equal and the same is true for v_j and v'_j .

This means, precisely, that the point p undergoes elastic reflection on the hypersurface $\{p \in \partial\Omega : \|p_i - p_j\| = 2\rho\}$ of the boundary of Ω (see Exercise 4.4.4). Therefore, the motion of the point p corresponds exactly to the evolution of the billiard in the table Ω .

The next result places billiards well inside the domain of interest of ergodic theory. Let ds be the volume measure induced on the boundary $\partial\Omega$ by the Riemannian metric of the ambient manifold; in the planar case (that is, when $\Omega \subset \mathbb{R}^2$), ds is just the arc-length. Denote by $d\theta$ the angle measure on each hemisphere $\{v \in S^{d-1} : v \cdot \mathbf{n}(s) > 0\}$.

Theorem 4.4.14. *The transformation f preserves the measure $\nu = \cos\theta ds d\theta$ on the domain $\{(s, v) \in \partial\Omega \times S^{d-1} : v \cdot \mathbf{n}(s) > 0\}$.*

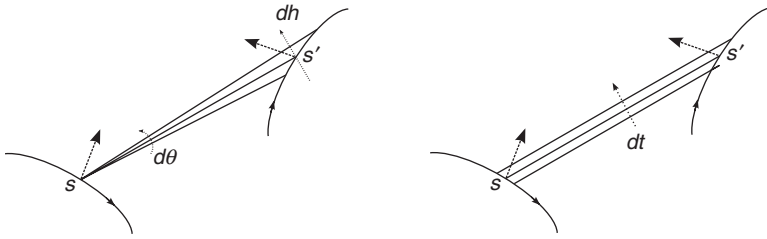


Figure 4.9. Calculating the derivative of the billiard map

In what follows we sketch the proof for planar billiards. The reader should have no trouble checking that all the arguments extend naturally to arbitrary dimension.

Consider any family of trajectories starting from a given boundary point (this means that s is fixed), as represented on the left-hand side of Figure 4.9. Let this family be parameterized by the angle of reflection θ . Denote by $\ell(s, s')$ the length of the line segment connecting s to s' . Then $\ell(s, s')d\theta = dh = \cos\theta' ds'$ and, thus,

$$\frac{\partial s'}{\partial \theta} = \frac{\ell(s, s')}{\cos\theta'}$$

To calculate the derivative of θ' with respect to θ , observe that the variation of θ' is the sum of two components: the first one corresponds to the variation of θ , whereas the second one arises from the variation of the normal vector $\mathbf{n}(s')$ as the collision point s' varies. By the definition of curvature, this second component is equal to $\kappa(s')ds'$. It follows that $d\theta' = d\theta + \kappa(s')ds'$ and, consequently,

$$\frac{\partial \theta'}{\partial \theta} = 1 + \kappa(s') \frac{\partial s'}{\partial \theta} = 1 + \kappa(s') \frac{\ell(s, s')}{\cos\theta'}$$

This can be summarized as follows:

$$Df(s, \theta) \frac{\partial}{\partial \theta} = \frac{\ell(s, s')}{\cos\theta'} \frac{\partial}{\partial s'} + \left(1 + \kappa(s') \frac{\ell(s, s')}{\cos\theta'} \right) \frac{\partial}{\partial \theta'} \tag{4.4.14}$$

Next, consider any family of parallel trajectories, as represented on the right-hand side of Figure 4.9. Let this family be parameterized by the arc-length t in the direction orthogonal to the trajectories. The variations of s and s' along this family are given by $-\cos\theta ds = dt = \cos\theta' ds'$. Since the trajectories all have the same direction, the variations of the angles θ and θ' arise, solely, from the variations of the normal vectors $\mathbf{n}(s)$ and $\mathbf{n}(s')$ as s and s' vary. That is, $d\theta = \kappa(s)ds$ and $d\theta' = \kappa(s')ds'$. Therefore,

$$Df(s, \theta) \left(-\frac{1}{\cos\theta} \frac{\partial}{\partial s} - \frac{\kappa(s)}{\cos\theta} \frac{\partial}{\partial \theta} \right) = \frac{1}{\cos\theta'} \frac{\partial}{\partial s'} + \frac{\kappa(s')}{\cos\theta'} \frac{\partial}{\partial \theta'} \tag{4.4.15}$$

Let $J(s, \theta)$ be the matrix of the derivative $Df(s, \theta)$ with respect to the bases $\{\partial/\partial s, \partial/\partial \theta\}$ and $\{\partial/\partial s', \partial/\partial \theta'\}$. The relations (4.4.14) and (4.4.15) imply that

$$\det J(s, \theta) = \frac{\begin{vmatrix} \frac{\ell(s, s')}{\cos \theta'} & \frac{1}{\cos \theta'} \\ 1 + \kappa(s') \frac{\ell(s, s')}{\cos \theta'} & \frac{\kappa(s')}{\cos \theta'} \end{vmatrix}}{\begin{vmatrix} 0 & -\frac{1}{\cos \theta} \\ 1 & -\frac{\kappa(s)}{\cos \theta} \end{vmatrix}} = \frac{\cos \theta}{\cos \theta'}. \tag{4.4.16}$$

So, by change of variables,

$$\begin{aligned} \int \varphi \, dv &= \int \varphi(s', \theta') \cos \theta' \, ds' \, d\theta' = \int \varphi(f(s, \theta)) \cos \theta' \frac{\cos \theta}{\cos \theta'} \, ds \, d\theta \\ &= \int \varphi(f(s, \theta)) \cos \theta \, ds \, d\theta = \int (\varphi \circ f) \, dv \end{aligned}$$

for every bounded measurable function φ . This proves that f preserves the measure $\nu = \cos \theta \, ds \, d\theta$, as we stated.

We call a billiard *dispersing* if the boundary of the billiard table is strictly convex at every point, when viewed from the inside. In the planar case, with the orientation conventions that we adopted, this means that the curvature κ is negative at every point. Figure 4.10 presents two examples. In the first one, $\Omega \subset \mathbb{R}^2$ and the boundary is a connected set formed by the union of five differentiable curves. In the second example, $\Omega \subset \mathbb{T}^2$ and the boundary has three connected components, all of which are differentiable and convex.

The class of dispersing billiards was introduced by Sinai in his 1970 article [Sin70]. The denomination “dispersing” refers to the fact that in such billiards any (thin) beam of parallel trajectories becomes divergent upon reflection on the boundary, as illustrated on the left-hand side of Figure 4.10. Sinai observed that dispersing billiards are hyperbolic systems, in a non-uniform sense: invariant sub-bundles E_z^s and E_z^u as in (4.4.11) exist at *almost every* point and, instead of (4.4.12), we have that the derivative is contracting along E_z^s and expanding along E_z^u *asymptotically*, that is, for sufficiently large iterates (depending on the point z).

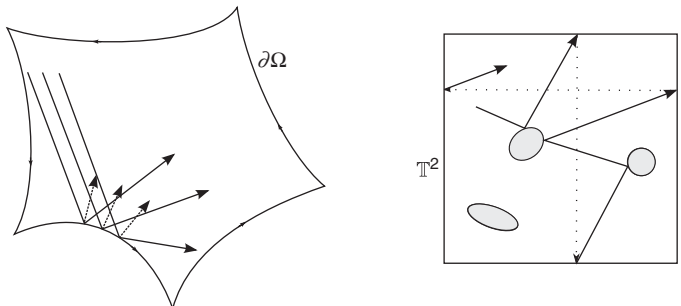


Figure 4.10. Dispersive billiards

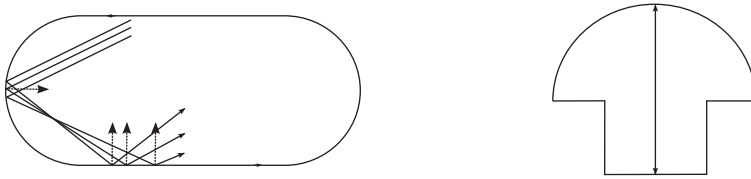


Figure 4.11. Bunimovich stadium and mushroom

The billiards associated with ideal gases (Example 4.4.13) with $N = 2$ molecules are dispersing: it is easy to see that $\{(p_1, p_2) \in \mathbb{R}^{2d} : \|p_1 - p_2\| = 2\rho\}$ is a convex hypersurface. Consequently, these billiards are hyperbolic, in the sense of the previous paragraph. Using a subtle version of the Hopf argument, Sinai proved in [Sin70] that such billiards are ergodic, at least when $d = 2$. This was later extended to arbitrary dimension $d \geq 2$ by Sinai and his student Nikolai Chernov [SC87], still in the case $N = 2$. Thus, dispersing billiards were the first class of billiards for which ergodicity was proven rigorously.

The case $N \geq 3$ of the Boltzmann–Sinai ergodic hypothesis is a lot more difficult because the corresponding billiards are *not* dispersing: the hypersurface

$$\{(p_1, p_2, \dots, p_N) \in \mathbb{R}^{Nd} : \|p_1 - p_2\| = 2\rho\}$$

has cylinder geometry, with zero curvature along the direction of the variables p_i , $i > 2$. Such billiards are called *semi-dispersing*. Most results in this setting are due to the Hungarian mathematicians András Krámlı, Nándor Simányi and Domoko Szász. In [KSS91, KSS92] they proved hyperbolicity and ergodicity for $N = 3$ and also for $N = 4$ assuming that $d \geq 3$. Later, Simányi [Sim02] proved hyperbolicity for the general case: any number of spheres, in any dimension. The problem of ergodicity remains open, in general, although there are many other partial results.

There are now several known examples of ergodic billiards that are not dispersing. This even includes some billiards whose boundary curvature is non-negative at every point. The best-known example is the *Bunimovich stadium*, whose boundary is formed by two semi-circles and two straight line segments. See Figure 4.11. This billiard is hyperbolic, but this property arises from a different mechanism, called *defocusing*: a beam of parallel trajectories reflecting on a concave segment of the billiard table wall starts by focusing, but then gets dispersed. Another interesting example is the *Bunimovich mushroom*: hyperbolic behavior and elliptic behavior coexist on disjoint invariant sets both with positive measure.

4.4.7 Exercises

- 4.4.1. We say that $\omega \in \mathbb{R}^d$ is τ -Diophantine if it is (c, τ) -Diophantine, that is, if it satisfies (4.4.5), for some $c > 0$. Prove that the set of τ -Diophantine vectors is

non-empty if and only if $\tau \geq d - 1$. Moreover, show that the set has full Lebesgue measure in \mathbb{R}^d whenever τ is strictly larger than $d - 1$.

- 4.4.2. Consider a billiard on a rectangular table. Check that every trajectory that does not hit any corner either is periodic or is dense in the billiard table.
- 4.4.3. Show that every billiard on an acute triangle exhibits some periodic trajectory. [Observation: the same is true for right triangles, but the problem is open for obtuse triangles.]
- 4.4.4. Consider the billiard model for ideal gases in Example 4.4.13. Check that elastic collisions between any two molecules correspond to the elastic reflections of the billiard point particle on the boundary of Ω .
- 4.4.5. Prove Theorem 4.4.9 under the additional hypothesis that the function $\rho \mapsto \Theta(\theta, \rho)$ is monotone (increasing or decreasing) for every $\theta \in \mathbb{R}$.
- 4.4.6. Consider the context of Theorem 4.4.9 but, instead of (4.4.10), assume that f rotates the two boundary components of A with different velocities: there exists some lift $F : \mathbb{R} \times [a, b] \rightarrow \mathbb{R} \times [a, b]$ and there exist $p, q \in \mathbb{Z}$ with $q \geq 1$, such that, denoting $F^q = (\Theta^q, R^q)$,

$$[\Theta^q(\theta, a) - p - \theta][\Theta^q(\theta, b) - p - \theta] < 0 \quad \text{for every } \theta \in \mathbb{R}. \quad (4.4.17)$$

Show that f has two periodic orbits with period q in the interior of A , at least.

- 4.4.7. Let Ω be a convex domain in the plane whose boundary $\partial\Omega$ is a differentiable curve. Show that the billiard on Ω has infinitely many periodic orbits.