

# 1

## Recurrence

Ergodic theory studies the behavior of dynamical systems with respect to measures that remain invariant under time evolution. Indeed, it aims to describe those properties that are valid for the trajectories of almost all initial states of the system, that is, all but a subset that has zero weight for the invariant measure. Our first task, in Section 1.1, will be to explain what we mean by ‘dynamical system’ and ‘invariant measure’.

The roots of the theory date back to the first half of the 19th century. By 1838, the French mathematician Joseph Liouville observed that every energy-preserving system in classical (Newtonian) mechanics admits a natural invariant volume measure in the space of configurations. Just a bit later, in 1845, the great German mathematician Carl Friedrich Gauss pointed out that the transformation

$$(0, 1] \rightarrow \mathbb{R}, \quad x \mapsto \text{fractional part of } \frac{1}{x},$$

which has an important role in number theory, admits an invariant measure equivalent to the Lebesgue measure (in the sense that the two have the same zero measure sets). These are two of the examples of applications of ergodic theory that we discuss in Section 1.3. Many others are introduced throughout this book.

The first important result was found by the great French mathematician Henri Poincaré by the end of the 19th century. Poincaré was particularly interested in the motion of celestial bodies, such as planets and comets, which is described by certain differential equations originating from Newton’s law of universal gravitation. Starting from Liouville’s observation, Poincaré realized that for almost every initial state of the system, that is, almost every value of the initial position and velocity, the solution of the differential equation comes back arbitrarily close to that initial state, unless it goes to infinity. Even more, this *recurrence* property is not specific to (celestial) mechanics: it is shared by any dynamical system that admits a finite invariant measure. That is the theme of Section 1.2.

The same theme reappears in Section 1.5, in a more elaborate context: there, we deal with any finite number of dynamical systems commuting with each other, and we seek *simultaneous* returns of the orbits of all those systems to the neighborhood of the initial state. This kind of result has important applications in combinatorics and number theory, as we will see.

The recurrence phenomenon is also behind the constructions that we present in Section 1.4. The basic idea is to fix some positive measure subset of the domain and to consider the first return to that subset. This first-return transformation is often easier to analyze, and it may be used to shed much light on the behavior of the original transformation.

## 1.1 Invariant measures

Let  $(M, \mathcal{B}, \mu)$  be a measure space and  $f : M \rightarrow M$  be a measurable transformation. We say that the measure  $\mu$  is *invariant* under  $f$  if

$$\mu(E) = \mu(f^{-1}(E)) \quad \text{for every measurable set } E \subset M. \quad (1.1.1)$$

We also say that  $\mu$  is *f-invariant*, or that  $f$  *preserves*  $\mu$ , to mean just the same. Notice that the definition (1.1.1) makes sense, since the pre-image of a measurable set under a measurable transformation is still a measurable set. Heuristically, the definition means that the probability that a point picked “at random” is in a given subset is equal to the probability that its image is in that subset.

It is possible, and convenient, to extend this definition to other types of dynamical systems, beyond transformations. We are especially interested in *flows*, that is, families of transformations  $f^t : M \rightarrow M$ , with  $t \in \mathbb{R}$ , satisfying the following conditions:

$$f^0 = \text{id} \quad \text{and} \quad f^{s+t} = f^s \circ f^t \quad \text{for every } s, t \in \mathbb{R}. \quad (1.1.2)$$

In particular, each transformation  $f^t$  is invertible and the inverse is  $f^{-t}$ . Flows arise naturally in connection with differential equations of the form

$$\frac{d\gamma}{dt}(t) = X(\gamma(t))$$

in the following way: under suitable conditions on the vector field  $X$ , for each point  $x$  in the domain  $M$  there exists exactly one solution  $t \mapsto \gamma_x(t)$  of the differential equation with  $\gamma_x(0) = x$ ; then  $f^t(x) = \gamma_x(t)$  defines a flow in  $M$ .

We say that a measure  $\mu$  is *invariant* under a flow  $(f^t)_t$  if it is invariant under each one of the transformations  $f^t$ , that is, if

$$\mu(E) = \mu(f^{-t}(E)) \quad \text{for every measurable set } E \subset M \text{ and } t \in \mathbb{R}. \quad (1.1.3)$$

**Proposition 1.1.1.** *Let  $f : M \rightarrow M$  be a measurable transformation and  $\mu$  be a measure on  $M$ . Then  $f$  preserves  $\mu$  if and only if*

$$\int \phi d\mu = \int \phi \circ f d\mu \tag{1.1.4}$$

for every  $\mu$ -integrable function  $\phi : M \rightarrow \mathbb{R}$ .

*Proof.* Suppose that the measure  $\mu$  is invariant under  $f$ . We are going to show that the relation (1.1.4) is valid for increasingly broader classes of functions. Let  $\mathcal{X}_B$  denote the characteristic function of any measurable subset  $B$ . Then

$$\mu(B) = \int \mathcal{X}_B d\mu \quad \text{and} \quad \mu(f^{-1}(B)) = \int \mathcal{X}_{f^{-1}(B)} d\mu = \int (\mathcal{X}_B \circ f) d\mu.$$

Thus, the hypothesis  $\mu(B) = \mu(f^{-1}(B))$  means that (1.1.4) is valid for characteristic functions. Then, by linearity of the integral, (1.1.4) is valid for all simple functions. Next, given any integrable  $\phi : M \rightarrow \mathbb{R}$ , consider a sequence  $(s_n)_n$  of simple functions, converging to  $\phi$  and such that  $|s_n| \leq |\phi|$  for every  $n$ . That such a sequence exists is guaranteed by Proposition A.1.33. Then, using the dominated convergence theorem (Theorem A.2.11) twice:

$$\int \phi d\mu = \lim_n \int s_n d\mu = \lim_n \int (s_n \circ f) d\mu = \int (\phi \circ f) d\mu.$$

This shows that (1.1.4) holds for every integrable function if  $\mu$  is invariant. The converse is also contained in the arguments we just presented.

### 1.1.1 Exercises

- 1.1.1. Let  $f : M \rightarrow M$  be a measurable transformation. Show that a Dirac measure  $\delta_p$  is invariant under  $f$  if and only if  $p$  is a fixed point of  $f$ . More generally, a probability measure  $\delta_{p,k} = k^{-1}(\delta_p + \delta_{f(p)} + \dots + \delta_{f^{k-1}(p)})$  is invariant under  $f$  if and only if  $f^k(p) = p$ .
- 1.1.2. Prove the following version of Proposition 1.1.1. Let  $M$  be a metric space,  $f : M \rightarrow M$  be a measurable transformation and  $\mu$  be a measure on  $M$ . Show that  $f$  preserves  $\mu$  if and only if

$$\int \phi d\mu = \int \phi \circ f d\mu$$

for every bounded continuous function  $\phi : M \rightarrow \mathbb{R}$ .

- 1.1.3. Prove that if  $f : M \rightarrow M$  preserves a measure  $\mu$  then, given any  $k \geq 2$ , the iterate  $f^k$  also preserves  $\mu$ . Is the converse true?
- 1.1.4. Suppose that  $f : M \rightarrow M$  preserves a probability measure  $\mu$ . Let  $B \subset M$  be a measurable set satisfying any one of the following conditions:
  - (a)  $\mu(B \setminus f^{-1}(B)) = 0$ ;
  - (b)  $\mu(f^{-1}(B) \setminus B) = 0$ ;
  - (c)  $\mu(B \Delta f^{-1}(B)) = 0$ ;
  - (d)  $f(B) \subset B$ .

Show that there exists  $C \subset M$  such that  $f^{-1}(C) = C$  and  $\mu(B \Delta C) = 0$ .

1.1.5. Let  $f : U \rightarrow U$  be a  $C^1$  diffeomorphism on an open set  $U \subset \mathbb{R}^d$ . Show that the Lebesgue measure  $m$  is invariant under  $f$  if and only if  $|\det Df| \equiv 1$ .

## 1.2 Poincaré recurrence theorem

We are going to study two versions of Poincaré's theorem. The first one (Section 1.2.1) is formulated in the context of (finite) measure spaces. The theorem of Kač, that we state and prove in Section 1.2.2, provides a quantitative complement to that statement. The second version of the recurrence theorem (Section 1.2.3) assumes that the ambient is a topological space with certain additional properties. We will also prove a third version of the recurrence theorem, due to Birkhoff, whose statement is purely topological.

### 1.2.1 Measurable version

Our first result asserts that, given any *finite* invariant measure, almost every point in any positive measure set  $E$  returns to  $E$  an infinite number of times:

**Theorem 1.2.1** (Poincaré recurrence). *Let  $f : M \rightarrow M$  be a measurable transformation and  $\mu$  be a finite measure invariant under  $f$ . Let  $E \subset M$  be any measurable set with  $\mu(E) > 0$ . Then, for  $\mu$ -almost every point  $x \in E$  there exist infinitely many values of  $n$  for which  $f^n(x)$  is also in  $E$ .*

*Proof.* Denote by  $E_0$  the set of points  $x \in E$  that never return to  $E$ . As a first step, let us prove that  $E_0$  has zero measure. To this end, let us observe that the pre-images  $f^{-n}(E_0)$  are pairwise disjoint. Indeed, suppose there exist  $m > n \geq 1$  such that  $f^{-m}(E_0)$  intersects  $f^{-n}(E_0)$ . Let  $x$  be a point in the intersection and  $y = f^n(x)$ . Then  $y \in E_0$  and  $f^{m-n}(y) = f^m(x) \in E_0$ . Since  $E_0 \subset E$ , this means that  $y$  returns to  $E$  at least once, which contradicts the definition of  $E_0$ . This contradiction proves that the pre-images are pairwise disjoint, as claimed.

Since  $\mu$  is invariant, we also have that  $\mu(f^{-n}(E_0)) = \mu(E_0)$  for all  $n \geq 1$ . It follows that

$$\mu\left(\bigcup_{n=1}^{\infty} f^{-n}(E_0)\right) = \sum_{n=1}^{\infty} \mu(f^{-n}(E_0)) = \sum_{n=1}^{\infty} \mu(E_0).$$

The expression on the left-hand side is finite, since the measure  $\mu$  is assumed to be finite. On the right-hand side we have a sum of infinitely many terms that are all equal. The only way such a sum can be finite is if the terms vanish. So,  $\mu(E_0) = 0$  as claimed.

Now let us denote by  $F$  the set of points  $x \in E$  that return to  $E$  a finite number of times. It is clear from the definition that every point  $x \in F$  has some iterate

$f^k(x)$  in  $E_0$ . In other words,

$$F \subset \bigcup_{k=0}^{\infty} f^{-k}(E_0).$$

Since  $\mu(E_0) = 0$  and  $\mu$  is invariant, it follows that

$$\mu(F) \leq \mu\left(\bigcup_{k=0}^{\infty} f^{-k}(E_0)\right) \leq \sum_{k=0}^{\infty} \mu(f^{-k}(E_0)) = \sum_{k=0}^{\infty} \mu(E_0) = 0.$$

Thus,  $\mu(F) = 0$  as we wanted to prove.

Theorem 1.2.1 implies an analogous result for continuous time systems: if  $\mu$  is a finite invariant measure of a flow  $(f^t)$ , then for every measurable set  $E \subset M$  with positive measure and for  $\mu$ -almost every  $x \in E$  there exist times  $t_j \rightarrow +\infty$  such that  $f^{t_j}(x) \in E$ . Indeed, if  $\mu$  is invariant under the flow then, in particular, it is invariant under the so-called *time-1 map*  $f^1$ . So, the statement we just made follows immediately from Theorem 1.2.1 applied to  $f^1$  (the times  $t_j$  one finds in this way are integers). Similar observations apply to the other versions of the recurrence theorem that we present in the sequel.

On the other hand, the theorem in the next section is specific to discrete time systems.

### 1.2.2 Kač theorem

Let  $f : M \rightarrow M$  be a measurable transformation and  $\mu$  be a finite measure invariant under  $f$ . Let  $E \subset M$  be any measurable set with  $\mu(E) > 0$ . Consider the *first-return time function*  $\rho_E : E \rightarrow \mathbb{N} \cup \{\infty\}$ , defined by

$$\rho_E(x) = \min\{n \geq 1 : f^n(x) \in E\} \tag{1.2.1}$$

if the set on the right-hand side is non-empty and  $\rho_E(x) = \infty$  if, on the contrary,  $x$  has no iterate in  $E$ . According to Theorem 1.2.1, the second alternative occurs only on a set with zero measure.

The next result shows that this function is integrable and even provides the value of the integral. For the statement we need the following notation:

$$E_0 = \{x \in E : f^n(x) \notin E \text{ for every } n \geq 1\} \quad \text{and} \\ E_0^* = \{x \in M : f^n(x) \notin E \text{ for every } n \geq 0\}.$$

In other words,  $E_0$  is the set of points in  $E$  that never return to  $E$  and  $E_0^*$  is the set of points in  $M$  that never enter  $E$ . We have seen in Theorem 1.2.1 that  $\mu(E_0) = 0$ .

**Theorem 1.2.2 (Kač).** *Let  $f : M \rightarrow M$  be a measurable transformation,  $\mu$  be a finite invariant measure and  $E \subset M$  be a positive measure set. Then the function*

$\rho_E$  is integrable and

$$\int_E \rho_E d\mu = \mu(M) - \mu(E_0^*).$$

*Proof.* For each  $n \geq 1$ , define

$$E_n = \{x \in E : f(x) \notin E, \dots, f^{n-1}(x) \notin E, \text{ but } f^n(x) \in E\} \quad \text{and}$$

$$E_n^* = \{x \in M : x \notin E, f(x) \notin E, \dots, f^{n-1}(x) \notin E, \text{ but } f^n(x) \in E\}.$$

That is,  $E_n$  is the set of points of  $E$  that return to  $E$  for the first time exactly at time  $n$ ,

$$E_n = \{x \in E : \rho_E(x) = n\},$$

and  $E_n^*$  is the set points that are not in  $E$  and enter  $E$  for the first time exactly at time  $n$ . It is clear that these sets are measurable and, hence,  $\rho_E$  is a measurable function. Moreover, the sets  $E_n, E_n^*, n \geq 0$  constitute a *partition* of the ambient space: they are pairwise disjoint and their union is the whole of  $M$ . So,

$$\mu(M) = \sum_{n=0}^{\infty} (\mu(E_n) + \mu(E_n^*)) = \mu(E_0^*) + \sum_{n=1}^{\infty} (\mu(E_n) + \mu(E_n^*)). \quad (1.2.2)$$

Now observe that

$$f^{-1}(E_n^*) = E_{n+1}^* \cup E_{n+1} \quad \text{for every } n. \quad (1.2.3)$$

Indeed,  $f(y) \in E_n^*$  means that the first iterate of  $f(y)$  that belongs to  $E$  is  $f^n(f(y)) = f^{n+1}(y)$  and that occurs if and only if  $y \in E_{n+1}^*$  or else  $y \in E_{n+1}$ . This proves the equality (1.2.3). So, given that  $\mu$  is invariant,

$$\mu(E_n^*) = \mu(f^{-1}(E_n^*)) = \mu(E_{n+1}^*) + \mu(E_{n+1}) \quad \text{for every } n.$$

Applying this relation successively, we find that

$$\mu(E_n^*) = \mu(E_m^*) + \sum_{i=n+1}^m \mu(E_i) \quad \text{for every } m > n. \quad (1.2.4)$$

The relation (1.2.2) implies that  $\mu(E_m^*) \rightarrow 0$  when  $m \rightarrow \infty$ . So, taking the limit as  $m \rightarrow \infty$  in the equality (1.2.4), we find that

$$\mu(E_n^*) = \sum_{i=n+1}^{\infty} \mu(E_i). \quad (1.2.5)$$

To complete the proof, replace (1.2.5) in the equality (1.2.2). In this way we find that

$$\mu(M) - \mu(E_0^*) = \sum_{n=1}^{\infty} \left( \sum_{i=n}^{\infty} \mu(E_i) \right) = \sum_{n=1}^{\infty} n\mu(E_n) = \int_E \rho_E d\mu,$$

as we wanted to prove.

In some cases, for example when the system  $(f, \mu)$  is *ergodic* (this property will be defined and studied later, starting from Chapter 4), the set  $E_0^*$  has zero measure. Then the conclusion of the Kač theorem means that

$$\frac{1}{\mu(E)} \int_E \rho_E d\mu = \frac{\mu(M)}{\mu(E)} \tag{1.2.6}$$

for every measurable set  $E$  with positive measure. The left-hand side of this expression is the *mean return time* to  $E$ . So, (1.2.6) asserts that *the mean return time is inversely proportional to the measure of  $E$* .

**Remark 1.2.3.** By definition,  $E_n^* = f^{-n}(E) \setminus \bigcup_{k=0}^{n-1} f^{-k}(E)$ . So, the fact that the sum (1.2.2) is finite implies that the measure of  $E_n^*$  converges to zero when  $n \rightarrow \infty$ . This fact will be useful later.

### 1.2.3 Topological version

Now let us suppose that  $M$  is a topological space, endowed with its Borel  $\sigma$ -algebra  $\mathcal{B}$ . A point  $x \in M$  is *recurrent* for a transformation  $f : M \rightarrow M$  if there exists a sequence  $n_j \rightarrow \infty$  of natural numbers such that  $f^{n_j}(x) \rightarrow x$ . Analogously, we say that  $x \in M$  is recurrent for a flow  $(f^t)_t$  if there exists a sequence  $t_j \rightarrow +\infty$  of real numbers such that  $f^{t_j}(x) \rightarrow x$  when  $j \rightarrow \infty$ .

In the next theorem we assume that the topological space  $M$  admits a countable basis of open sets, that is, there exists a countable family  $\{U_k : k \in \mathbb{N}\}$  of open sets such that every open subset of  $M$  may be written as a union of elements  $U_k$  of this family. This condition holds in most interesting examples.

**Theorem 1.2.4** (Poincaré recurrence). *Suppose that  $M$  admits a countable basis of open sets. Let  $f : M \rightarrow M$  be a measurable transformation and  $\mu$  be a finite measure on  $M$  invariant under  $f$ . Then,  $\mu$ -almost every  $x \in M$  is recurrent for  $f$ .*

*Proof.* For each  $k$ , denote by  $\tilde{U}_k$  the set of points  $x \in U_k$  that never return to  $U_k$ . According to Theorem 1.2.1, every  $\tilde{U}_k$  has zero measure. Consequently, the countable union

$$\tilde{U} = \bigcup_{k \in \mathbb{N}} \tilde{U}_k$$

also has zero measure. Hence, to prove the theorem it suffices to check that every point  $x$  that is not in  $\tilde{U}$  is recurrent. That is easy, as we are going to see. Consider  $x \in M \setminus \tilde{U}$  and let  $U$  be any neighborhood of  $x$ . By definition, there exists some element  $U_k$  of the basis of open sets such that  $x \in U_k$  and  $U_k \subset U$ . Since  $x$  is not in  $\tilde{U}$ , we also have that  $x \notin \tilde{U}_k$ . In other words, there exists  $n \geq 1$  such that  $f^n(x)$  is in  $U_k$ . In particular,  $f^n(x)$  is also in  $U$ . Since the neighborhood  $U$  is arbitrary, this proves that  $x$  is a recurrent point.

Let us point out that the conclusions of Theorems 1.2.1 and 1.2.4 are false, in general, if the measure  $\mu$  is not finite:

**Example 1.2.5.** Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be the translation by 1, that is, the transformation defined by  $f(x) = x + 1$  for every  $x \in \mathbb{R}$ . It is easy to check that  $f$  preserves the Lebesgue measure on  $\mathbb{R}$  (which is infinite). On the other hand, no point  $x \in \mathbb{R}$  is recurrent for  $f$ . According to the recurrence theorem, this last observation implies that  $f$  can not admit any finite invariant measure.

However, it is possible to extend these statements for certain cases of infinite measures: see Exercise 1.2.2.

To conclude, we present a purely topological version of Theorem 1.2.4, called the Birkhoff recurrence theorem, that makes no reference at all to invariant measures:

**Theorem 1.2.6** (Birkhoff recurrence). *If  $f : M \rightarrow M$  is a continuous transformation on a compact metric space  $M$  then there exists some point  $x \in X$  that is recurrent for  $f$ .*

*Proof.* Consider the family  $\mathcal{I}$  of all non-empty closed sets  $X \subset M$  that are invariant under  $f$ , in the sense that  $f(X) \subset X$ . This family is non-empty, since  $M \in \mathcal{I}$ . We claim that an element  $X \in \mathcal{I}$  is minimal for the inclusion relation if and only if the orbit of every  $x \in X$  is dense in  $X$ . Indeed, it is clear that if  $X$  is a closed invariant subset then  $X$  contains the closure of the orbit of each one of its elements. Hence, in order to be minimal,  $X$  must coincide with every one of these closures. Conversely, for the same reason, if  $X$  coincides with the orbit closure of each one of its points then it has no proper subset that is closed and invariant. That is,  $X$  is minimal. This proves our claim. In particular, every point  $x$  in a minimal set is recurrent. Therefore, to prove the theorem it suffices to prove that there exists some minimal set.

We claim that every totally ordered set  $\{X_\alpha\} \subset \mathcal{I}$  admits a lower bound. Indeed, consider  $X = \bigcap_\alpha X_\alpha$ . Observe that  $X$  is non-empty, since the  $X_\alpha$  are compact and they form a totally ordered family. It is clear that  $X$  is closed and invariant under  $f$  and it is equally clear that  $X$  is a lower bound for the set  $\{X_\alpha\}$ . That proves our claim. Now it follows from Zorn's lemma that  $\mathcal{I}$  does contain minimal elements.

Theorem 1.2.6 can also be deduced from Theorem 1.2.4 together with the fact, which we will prove later (in Chapter 2), that every continuous transformation on a compact metric space admits some invariant probability measure.

## 1.2.4 Exercises

- 1.2.1. Show that the following statement is equivalent to Theorem 1.2.1, meaning that each one of them can be obtained from the other. Let  $f : M \rightarrow M$  be a measurable transformation and  $\mu$  be a finite invariant measure. Let  $E \subset M$  be any measurable



set with  $\mu(E) > 0$ . Then there exists  $N \geq 1$  and a positive measure set  $D \subset E$  such that  $f^N(x) \in E$  for every  $x \in D$ .

- 1.2.2. Let  $f : M \rightarrow M$  be an invertible transformation and suppose that  $\mu$  is an invariant measure, not necessarily finite. Let  $B \subset M$  be a set with finite measure. Prove that, given any measurable set  $E \subset M$  with positive measure,  $\mu$ -almost every point  $x \in E$  either returns to  $E$  an infinite number of times or has only a finite number of iterates in  $B$ .
- 1.2.3. Let  $f : M \rightarrow M$  be an invertible transformation and suppose that  $\mu$  is a  $\sigma$ -finite invariant measure: there exists an increasing sequence of measurable subsets  $M_k$  with  $\mu(M_k) < \infty$  for every  $k$  and  $\bigcup_k M_k = M$ . We say that a point  $x$  goes to infinity if, for every  $k$ , there exists only a finite number of iterates of  $x$  that are in  $M_k$ . Show that, given any  $E \subset M$  with positive measure,  $\mu$ -almost every point  $x \in E$  returns to  $E$  an infinite number of times or else goes to infinity.
- 1.2.4. Let  $f : M \rightarrow M$  be a measurable transformation, not necessarily invertible,  $\mu$  be an invariant probability measure and  $D \subset M$  be a set with positive measure. Prove that almost every point of  $D$  spends a positive fraction of time in  $D$ :

$$\limsup_n \frac{1}{n} \#\{0 \leq j \leq n - 1 : f^j(x) \in D\} > 0$$

for  $\mu$ -almost every  $x \in D$ . [Note: One may replace limsup by liminf in the statement, but the proof of that fact will have to wait until Chapter 3.]

- 1.2.5. Let  $f : M \rightarrow M$  be a measurable transformation preserving a finite measure  $\mu$ . Given any measurable set  $A \subset M$  with  $\mu(A) > 0$ , let  $n_1 < n_2 < \dots$  be the sequence of values of  $n$  such that  $\mu(f^{-n}(A) \cap A) > 0$ . The goal of this exercise is to prove that  $V_A = \{n_1, n_2, \dots\}$  is a syndetic, that is, that there exists  $C > 0$  such that  $n_{i+1} - n_i \leq C$  for every  $i$ .
  - (a) Show that for any increasing sequence  $k_1 < k_2 < \dots$  there exist  $j > i \geq 1$  such that  $\mu(A \cap f^{-(k_j - k_i)}(A)) > 0$ .
  - (b) Given any infinite sequence  $\ell = (\ell_j)_j$  of natural numbers, denote by  $S(\ell)$  the set of all finite sums of consecutive elements of  $\ell$ . Show that  $V_A$  intersects  $S(\ell)$  for every  $\ell$ .
  - (c) Deduce that the set  $V_A$  is syndetic.

[Note: Exercise 3.1.2 provides a different proof of this fact.]

- 1.2.6. Show that if  $f : [0, 1] \rightarrow [0, 1]$  is a measurable transformation preserving the Lebesgue measure  $m$  then  $m$ -almost every point  $x \in [0, 1]$  satisfies

$$\liminf_n n|f^n(x) - x| \leq 1.$$

[Note: Boshernitzan [Bos93] proved a much more general result, namely that  $\liminf_n n^{1/d} d(f^n(x), x) < \infty$  for  $\mu$ -almost every point and every probability measure  $\mu$  invariant under  $f : M \rightarrow M$ , assuming  $M$  is a separable metric whose  $d$ -dimensional Hausdorff measure is  $\sigma$ -finite.]

- 1.2.7. Define  $f : [0, 1] \rightarrow [0, 1]$  by  $f(x) = (x + \omega) - [x + \omega]$ , where  $\omega$  represents the golden ratio  $(1 + \sqrt{5})/2$ . Given  $x \in [0, 1]$ , check that  $n|f^n(x) - x| = n^2|\omega - q_n|$  for every  $n$ , where  $(q_n)_n \rightarrow \omega$  is the sequence of rational numbers given by  $q_n = [x + n\omega]/n$ . Using that the roots of the polynomial  $R(z) = z^2 - z - 1$  are precisely  $\omega$  and  $\omega - \sqrt{5}$ , prove that  $\liminf_n n^2|\omega - q_n| \geq 1/\sqrt{5}$ . [Note: This shows that the constant 1 in Exercise 1.2.6 cannot be replaced by any constant smaller than

$1/\sqrt{5}$ . It is not known whether 1 is the smallest constant such that the statement holds for *every* transformation on the interval.]

### 1.3 Examples

Next, we describe some simple examples of invariant measures for transformations and flows that help us interpret the significance of the Poincaré recurrence theorems and also lead to some interesting conclusions.

#### 1.3.1 Decimal expansion

Our first example is the transformation defined on the interval  $[0, 1]$  in the following way:

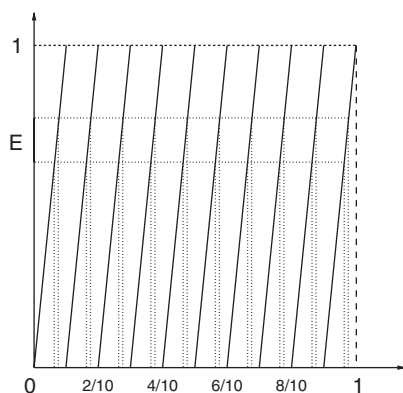
$$f : [0, 1] \rightarrow [0, 1], \quad f(x) = 10x - [10x].$$

Here and in what follows, we use  $[y]$  as the *integer part* of a real number  $y$ , that is, the largest integer smaller than or equal  $y$ . So,  $f$  is the map sending each  $x \in [0, 1]$  to the *fractional part* of  $10x$ . Figure 1.1 represents the graph of  $f$ .

We claim that the Lebesgue measure  $\mu$  on the interval is invariant under the transformation  $f$ , that is, it satisfies

$$\mu(E) = \mu(f^{-1}(E)) \quad \text{for every measurable set } E \subset M. \quad (1.3.1)$$

This can be checked as follows. Let us begin by supposing that  $E$  is an interval. Then, as illustrated in Figure 1.1, its pre-image  $f^{-1}(E)$  consists of ten intervals, each of which is ten times shorter than  $E$ . Hence, the Lebesgue measure of  $f^{-1}(E)$  is equal to the Lebesgue measure of  $E$ . This proves that (1.3.1) does hold in the case of intervals. As a consequence, it also holds when  $E$  is a finite



**Figure 1.1.** Fractional part of  $10x$

union of intervals. Now, the family of all finite unions of intervals is an algebra that generates the Borel  $\sigma$ -algebra of  $[0, 1]$ . Hence, to conclude the proof it is enough to use the following general fact:

**Lemma 1.3.1.** *Let  $f : M \rightarrow M$  be a measurable transformation and  $\mu$  be a finite measure on  $M$ . Suppose that there exists some algebra  $\mathcal{A}$  of measurable subsets of  $M$  such that  $\mathcal{A}$  generates the  $\sigma$ -algebra  $\mathcal{B}$  of  $M$  and  $\mu(E) = \mu(f^{-1}(E))$  for every  $E \in \mathcal{A}$ . Then the latter remains true for every set  $E \in \mathcal{B}$ , that is, the measure  $\mu$  is invariant under  $f$ .*

*Proof.* We start by proving that  $\mathcal{C} = \{E \in \mathcal{B} : \mu(E) = \mu(f^{-1}(E))\}$  is a monotone class. Let  $E_1 \subset E_2 \subset \dots$  be any increasing sequence of elements of  $\mathcal{C}$  and let  $E = \bigcup_{i=1}^{\infty} E_i$ . By Theorem A.1.14 (see Exercise A.1.9),

$$\mu(E) = \lim_i \mu(E_i) \quad \text{and} \quad \mu(f^{-1}(E)) = \lim_i \mu(f^{-1}(E_i)).$$

So, using the fact that  $E_i \in \mathcal{C}$ ,

$$\mu(E) = \lim_i \mu(E_i) = \lim_i \mu(f^{-1}(E_i)) = \mu(f^{-1}(E)).$$

Hence,  $E \in \mathcal{C}$ . In precisely the same way, one gets that the intersection of any decreasing sequence of elements of  $\mathcal{C}$  is in  $\mathcal{C}$ . This proves that  $\mathcal{C}$  is indeed a monotone class.

Now it is easy to deduce the conclusion of the lemma. Indeed, since  $\mathcal{C}$  is assumed to contain  $\mathcal{A}$ , we may use the monotone class theorem (Theorem A.1.18), to conclude that  $\mathcal{C}$  contains the  $\sigma$ -algebra  $\mathcal{B}$  generated by  $\mathcal{A}$ . That is precisely what we wanted to prove.

Now we explain how one may use the fact that the Lebesgue measure is invariant under  $f$ , together with the Poincaré recurrence theorem, to reach some interesting conclusions. The transformation  $f$  is directly related to the usual decimal expansion of a real number: if  $x$  is given by

$$x = 0.a_0a_1a_2a_3 \dots$$

with  $a_i \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$  and  $a_i \neq 9$  for infinitely many values of  $i$ , then its image under  $f$  is given by

$$f(x) = 0.a_1a_2a_3 \dots$$

Thus, more generally, the  $n$ -th iterate of  $f$  can be expressed in the following way, for every  $n \geq 1$ :

$$f^n(x) = 0.a_n a_{n+1} a_{n+2} \dots \tag{1.3.2}$$

Let  $E$  be the subset of points  $x \in [0, 1]$  whose decimal expansion starts with the digit 7, that is, such that  $a_0 = 7$ . According to Theorem 1.2.1, almost every element in  $E$  has infinitely many iterates that are also in  $E$ . By the expression (1.3.2), this means that there are infinitely many values of  $n$  such that  $a_n = 7$ .

So, we have shown that *almost every number  $x$  whose decimal expansion starts with 7 has infinitely many digits equal to 7.*

Of course, instead of 7 we may consider any other digit. Even more, there is a similar result (see Exercise 1.3.2) when, instead of a single digit, one considers a block of  $k \geq 1$  consecutive digits. Later on, in Chapter 3, we will prove a much stronger fact: for almost every number  $x \in [0, 1]$ , every digit occurs with frequency  $1/10$  (more generally, every block of  $k \geq 1$  digits occurs with frequency  $1/10^k$ ) in the decimal expansion of  $x$ .

### 1.3.2 Gauss map

The system we present in this section is related to another important algorithm in number theory, the *continued fraction expansion*, which plays a central role in the problem of finding the best rational approximation to any real number. Let us start with a brief presentation of this algorithm.

Given any number  $x_0 \in (0, 1)$ , let

$$a_1 = \left[ \frac{1}{x_0} \right] \quad \text{and} \quad x_1 = \frac{1}{x_0} - a_1.$$

Note that  $a_1$  is a natural number,  $x_1 \in [0, 1)$  and

$$x_0 = \frac{1}{a_1 + x_1}.$$

Supposing that  $x_1$  is different from zero, we may repeat this procedure, defining

$$a_2 = \left[ \frac{1}{x_1} \right] \quad \text{and} \quad x_2 = \frac{1}{x_1} - a_2.$$

Then

$$x_1 = \frac{1}{a_1 + x_2} \quad \text{and so} \quad x_0 = \frac{1}{a_1 + \frac{1}{a_2 + x_2}}.$$

Now we may proceed by induction: for each  $n \geq 1$  such that  $x_{n-1} \in (0, 1)$ , define

$$a_n = \left[ \frac{1}{x_{n-1}} \right] \quad \text{and} \quad x_n = \frac{1}{x_{n-1}} - a_n = G(x_{n-1}),$$

and observe that

$$x_0 = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\dots + \frac{1}{a_n + x_n}}}}. \quad (1.3.3)$$

It can be shown that the sequence

$$z_n = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\dots + \frac{1}{a_n}}}} \tag{1.3.4}$$

converges to  $x_0$  when  $n \rightarrow \infty$ . This is usually expressed through the expression

$$x_0 = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\dots + \frac{1}{a_n + \dots}}}}, \tag{1.3.5}$$

which is called *continued fraction expansion* of  $x_0$ .

Note that the sequence  $(z_n)_n$  defined by the relation (1.3.4) consists of rational numbers. Indeed, one can show that these are the *best rational approximations* of the number  $x_0$ , in the sense that each  $z_n$  is closer to  $x_0$  than any other rational number whose denominator is smaller than or equal to the denominator of  $z_n$  (written in irreducible form). Observe also that to obtain (1.3.5) we had to assume that  $x_n \in (0, 1)$  for every  $n \in \mathbb{N}$ . If in the course of the process one encounters some  $x_n = 0$ , then the algorithm halts and we consider (1.3.3) to be the continued fraction expansion of  $x_0$ . It is clear that this can happen only if  $x_0$  itself is a rational number.

This continued fraction algorithm is intimately related to a certain dynamical system on the interval  $[0, 1]$  that we describe in the following. The *Gauss map*  $G : [0, 1] \rightarrow [0, 1]$  is defined by

$$G(x) = \frac{1}{x} - \left[ \frac{1}{x} \right] = \text{fractional part of } 1/x,$$

if  $x \in (0, 1]$  and  $G(0) = 0$ . The graph of  $G$  can be easily sketched (see Figure 1.2), starting from the following observation: for every  $x$  in each interval  $I_k = (1/(k+1), 1/k]$ , the integer part of  $1/x$  is equal to  $k$  and so  $G(x) = 1/x - k$ .

The continued fraction expansion of any number  $x_0 \in (0, 1)$  can be obtained from the Gauss map in the following way: for each  $n \geq 1$ , the natural number  $a_n$  is determined by

$$G^{n-1}(x_0) \in I_{a_n},$$

and the real number  $x_n$  is simply the  $n$ -th iterate  $G^n(x_0)$  of the point  $x_0$ . This process halts whenever we encounter some  $x_n = 0$ ; as we explained previously, this can only happen if  $x_0$  is a rational number (see Exercise 1.3.4). In particular, there exists a full Lebesgue measure subset of  $(0, 1)$  such that all the iterates of  $G$  are defined for all the points in that subset.

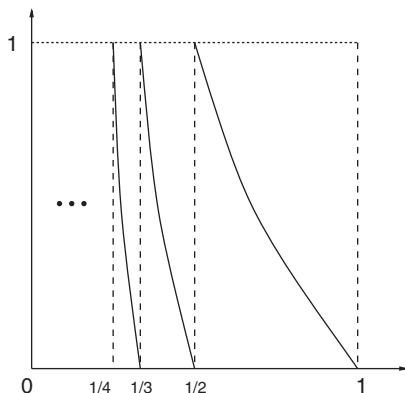


Figure 1.2. Gauss map

A remarkable fact that makes this transformation interesting from the point of view of ergodic theory is that  $G$  admits an invariant probability measure that, in addition, is equivalent to the Lebesgue measure on the interval. Indeed, consider the measure defined by

$$\mu(E) = \int_E \frac{c}{1+x} dx \quad \text{for every measurable set } E \subset [0, 1], \quad (1.3.6)$$

where  $c$  is a positive constant. The integral is well defined, since the function in the integral is continuous on the interval  $[0, 1]$ . Moreover, this function takes values inside the interval  $[c/2, c]$  and that implies

$$\frac{c}{2} m(E) \leq \mu(E) \leq cm(E) \quad \text{for every measurable set } E \subset [0, 1]. \quad (1.3.7)$$

In particular,  $\mu$  is indeed equivalent to the Lebesgue measure  $m$ .

**Proposition 1.3.2.** *The measure  $\mu$  is invariant under  $G$ . Moreover, if we choose  $c = 1/\log 2$  then  $\mu$  is a probability measure.*

*Proof.* We are going to use the following lemma:

**Lemma 1.3.3.** *Let  $f : [0, 1] \rightarrow [0, 1]$  be a transformation such that there exist pairwise disjoint open intervals  $I_1, I_2, \dots$  satisfying*

1. *the union  $\bigcup_k I_k$  has full Lebesgue measure in  $[0, 1]$  and*
2. *the restriction  $f_k = f | I_k$  to each  $I_k$  is a diffeomorphism onto  $(0, 1)$ .*

*Let  $\rho : [0, 1] \rightarrow [0, \infty)$  be an integrable function (relative to the Lebesgue measure) satisfying*

$$\rho(y) = \sum_{x \in f^{-1}(y)} \frac{\rho(x)}{|f'(x)|} \quad (1.3.8)$$

*for almost every  $y \in [0, 1]$ . Then the measure  $\mu = \rho dx$  is invariant under  $f$ .*

*Proof.* Let  $\phi = \chi_E$  be the characteristic function of an arbitrary measurable set  $E \subset [0, 1]$ . Changing variables in the integral,

$$\int_{I_k} \phi(f(x))\rho(x) dx = \int_0^1 \phi(y)\rho(f_k^{-1}(y))|(f_k^{-1})'(y)| dy.$$

Note that  $(f_k^{-1})'(y) = 1/f'(f_k^{-1}(y))$ . So, the previous relation implies that

$$\begin{aligned} \int_0^1 \phi(f(x))\rho(x) dx &= \sum_{k=1}^{\infty} \int_{I_k} \phi(f(x))\rho(x) dx \\ &= \sum_{k=1}^{\infty} \int_0^1 \phi(y) \frac{\rho(f_k^{-1}(y))}{|f'(f_k^{-1}(y))|} dy. \end{aligned} \tag{1.3.9}$$

Using the monotone convergence theorem (Theorem A.2.9) and the hypothesis (1.3.8), we see that the last expression in (1.3.9) is equal to

$$\int_0^1 \phi(y) \sum_{k=1}^{\infty} \frac{\rho(f_k^{-1}(y))}{|f'(f_k^{-1}(y))|} dy = \int_0^1 \phi(y)\rho(y) dy.$$

In this way we find that  $\int_0^1 \phi(f(x))\rho(x) dx = \int_0^1 \phi(y)\rho(y) dy$ . Since  $\mu = \rho dx$  and  $\phi = \chi_E$ , this means that  $\mu(f^{-1}(E)) = \mu(E)$  for every measurable set  $E \subset [0, 1]$ . In other words,  $\mu$  is invariant under  $f$ .

To conclude the proof of Proposition 1.3.2 we must show that the condition (1.3.8) holds for  $\rho(x) = c/(1+x)$  and  $f = G$ . Let  $G_k$  denote the restriction of  $G$  to the interval  $I_k = (1/(k+1), 1/k)$ , for  $k \geq 1$ . Note that  $G_k^{-1}(y) = 1/(y+k)$  for every  $k$ . Note also that  $G'(x) = (1/x)' = -1/x^2$  for every  $x \neq 0$ . Therefore,

$$\sum_{k=1}^{\infty} \frac{\rho(G_k^{-1}(y))}{|G'(G_k^{-1}(y))|} = \sum_{k=1}^{\infty} \frac{c(y+k)}{y+k+1} \left(\frac{1}{y+k}\right)^2 = \sum_{k=1}^{\infty} \frac{c}{(y+k)(y+k+1)}. \tag{1.3.10}$$

Observing that

$$\frac{1}{(y+k)(y+k+1)} = \frac{1}{y+k} - \frac{1}{y+k+1},$$

we see that the last sum in (1.3.10) has a telescopic structure: except for the first one, all the terms occur twice, with opposite signs, and so they cancel out. This means that the sum is equal to the first term:

$$\sum_{k=1}^{\infty} \frac{c}{(y+k)(y+k+1)} = \frac{c}{y+1} = \rho(y).$$

This proves that the equality (1.3.8) is indeed satisfied and, hence, we may use Lemma 1.3.1 to conclude that  $\mu$  is invariant under  $f$ .

Finally, observing that  $c \log(1+x)$  is a primitive of the function  $\rho(x)$ , we find that

$$\mu([0, 1]) = \int_0^1 \frac{c}{1+x} dx = c \log 2.$$

So, picking  $c = 1/\log 2$  ensures that  $\mu$  is a probability measure.

This proposition allows us to use ideas from ergodic theory, applied to the Gauss map, to obtain interesting conclusions in number theory. For example (see Exercise 1.3.3), the natural number 7 occurs infinitely many times in the continued fraction expansion of almost every number  $x_0 \in (1/8, 1/7)$ , that is, one has  $a_n = 7$  for infinitely many values of  $n \in \mathbb{N}$ . Later on, in Chapter 3, we will prove a much more precise statement, that contains the following conclusion: for almost every  $x_0 \in (0, 1)$  the number 7 occurs with frequency

$$\frac{1}{\log 2} \log \frac{64}{63}$$

in the continued fraction expansion of  $x_0$ . Try to guess right away where this number comes from!

### 1.3.3 Circle rotations

Let us consider on the real line  $\mathbb{R}$  the equivalence relation  $\sim$  that identifies any numbers whose difference is an integer number:

$$x \sim y \iff x - y \in \mathbb{Z}.$$

We represent by  $[x] \in \mathbb{R}/\mathbb{Z}$  the equivalence class of each  $x \in \mathbb{R}$  and denote by  $\mathbb{R}/\mathbb{Z}$  the space of all equivalence classes. This space is called the *circle* and is also denoted by  $S^1$ . The reason for this terminology is that  $\mathbb{R}/\mathbb{Z}$  can be identified in a natural way with the unit circle  $\{z \in \mathbb{C} : |z| = 1\}$  on the complex plane, by means of the map

$$\phi : \mathbb{R}/\mathbb{Z} \rightarrow \{z \in \mathbb{C} : |z| = 1\}, \quad [x] \mapsto e^{2\pi xi}. \quad (1.3.11)$$

Note that  $\phi$  is well defined: since the function  $x \mapsto e^{2\pi xi}$  is periodic of period 1, the expression  $e^{2\pi xi}$  does not depend on the choice of a representative  $x$  for the class  $[x]$ . Moreover,  $\phi$  is a bijection.

The circle  $\mathbb{R}/\mathbb{Z}$  inherits from the real line  $\mathbb{R}$  the structure of an abelian group, given by the operation

$$[x] + [y] = [x + y].$$

Observe that this is well defined: the equivalence class on the right-hand side does not depend on the choice of representatives  $x$  and  $y$  for the classes on the left-hand side. Given  $\theta \in \mathbb{R}$ , we call *rotation* of angle  $\theta$  the transformation

$$R_\theta : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}, \quad [x] \mapsto [x + \theta] = [x] + [\theta].$$



Note that  $R_\theta$  corresponds, via the identification (1.3.11), to the transformation  $z \mapsto e^{2\pi\theta i}z$  on  $\{z \in \mathbb{C} : |z| = 1\}$ . The latter is just the restriction to the unit circle of the rotation of angle  $2\pi\theta$  around the origin in the complex plane. It is clear from the definition that  $R_0$  is the identity map and  $R_\theta \circ R_\tau = R_{\theta+\tau}$  for every  $\theta$  and  $\tau$ . In particular, every  $R_\theta$  is invertible and the inverse is  $R_{-\theta}$ .

We can also endow  $S^1$  with a natural structure of a probability space, as follows. Let  $\pi : \mathbb{R} \rightarrow S^1$  be the canonical projection, that assigns to each  $x \in \mathbb{R}$  its equivalence class  $[x]$ . We say that a set  $E \subset S^1$  is measurable if  $\pi^{-1}(E)$  is a measurable subset of the real line. Next, let  $m$  be the Lebesgue measure on the real line. We define the *Lebesgue measure*  $\mu$  on the circle to be given by

$$\mu(E) = m(\pi^{-1}(E) \cap [k, k + 1)) \quad \text{for every } k \in \mathbb{Z}.$$

Note that the left-hand side of this equality does not depend on  $k$ , since, by definition,  $\pi^{-1}(E) \cap [k, k + 1) = (\pi^{-1}(E) \cap [0, 1)) + k$  and the measure  $m$  is invariant under translations.

It is clear from the definition that  $\mu$  is a probability. Moreover,  $\mu$  is invariant under every rotation  $R_\theta$  (according to Exercise 1.3.8, it is the only probability measure with this property). This can be shown as follows. By definition,  $\pi^{-1}(R_\theta^{-1}(E)) = \pi^{-1}(E) - \theta$  for every measurable set  $E \subset S^1$ . Let  $k$  be the integer part of  $\theta$ . Since  $m$  is invariant under all the translations,

$$\begin{aligned} m((\pi^{-1}(E) - \theta) \cap [0, 1)) &= m(\pi^{-1}(E) \cap [\theta, \theta + 1)) \\ &= m(\pi^{-1}(E) \cap [\theta, k + 1)) + m(\pi^{-1}(E) \cap [k + 1, \theta + 1)). \end{aligned}$$

Note that  $\pi^{-1}(E) \cap [k + 1, \theta + 1) = (\pi^{-1}(E) \cap [k, \theta)) + 1$ . So, the expression on the right-hand side of the previous equality may be written as

$$m(\pi^{-1}(E) \cap [\theta, k + 1)) + m(\pi^{-1}(E) \cap [k, \theta)) = m(\pi^{-1}(E) \cap [k, k + 1)).$$

Combining these two relations we find that

$$\mu(R_\theta^{-1}(E)) = m(\pi^{-1}(R_\theta^{-1}(E) \cap [0, 1))) = m(\pi^{-1}(E) \cap [k, k + 1)) = \mu(E)$$

for every measurable set  $E \subset S^1$ .

The rotations  $R_\theta : S^1 \rightarrow S^1$  exhibit two very different types of dynamical behavior, depending on the value of  $\theta$ . If  $\theta$  is rational, say  $\theta = p/q$  with  $p \in \mathbb{Z}$  and  $q \in \mathbb{N}$ , then

$$R_\theta^q([x]) = [x + q\theta] = [x] \quad \text{for every } [x].$$

Consequently, in this case every point  $x \in S^1$  is periodic with period  $q$ . In the opposite case we have:

**Proposition 1.3.4.** *If  $\theta$  is irrational then  $\mathcal{O}([x]) = \{R_\theta^n([x]) : n \in \mathbb{N}\}$  is a dense subset of the circle  $\mathbb{R}/\mathbb{Z}$  for every  $[x]$ .*

*Proof.* We claim that the set  $\mathcal{D} = \{m + n\theta : m \in \mathbb{Z}, n \in \mathbb{N}\}$  is dense in  $\mathbb{R}$ . Indeed, consider any number  $r \in \mathbb{R}$ . Given any  $\varepsilon > 0$ , we may choose  $p \in \mathbb{Z}$  and  $q \in \mathbb{N}$

such that  $|q\theta - p| < \varepsilon$ . Note that the number  $a = q\theta - p$  is necessarily different from zero, since  $\theta$  is irrational. Let us suppose that  $a$  is positive (the case when  $a$  is negative is analogous). Subdividing the real line into intervals of length  $a$ , we see that there exists an integer  $l$  such that  $0 \leq r - la < a$ . This implies that

$$|r - (lq\theta - lp)| = |r - la| < a < \varepsilon.$$

As  $m = lq$  and  $n = -lp$  are integers and  $\varepsilon$  is arbitrary, this proves that  $r$  is in the closure of the set  $\mathcal{D}$ , for every  $r \in \mathbb{R}$ .

Now, given  $y \in \mathbb{R}$  and  $\varepsilon > 0$ , we may take  $r = y - x$  and, using the previous paragraph, we may find  $m, n \in \mathbb{Z}$  such that  $|m + n\theta - (y - x)| < \varepsilon$ . This is equivalent to saying that the distance from  $[y]$  to the iterate  $R_\theta^n([x])$  is less than  $\varepsilon$ . Since  $x, y$  and  $\varepsilon$  are arbitrary, this shows that every orbit  $\mathcal{O}([x])$  is dense in  $S^1$ .

In particular, it follows that every point on the circle is recurrent for  $R_\theta$  (this is also true when  $\theta$  is rational). The previous proposition also leads to some interesting conclusions in the study of the invariant measures of  $R_\theta$ . Among other things, we will learn later (in Chapter 6) that if  $\theta$  is irrational then the Lebesgue measure is the unique probability measure that is preserved by  $R_\theta$ . Related to this, we will see that the orbits of  $R_\theta$  are uniformly distributed subsets of  $S^1$ .

### 1.3.4 Rotations on tori

The notions we just presented can be generalized to arbitrary dimension, as we are going to explain. For each  $d \geq 1$ , consider the equivalence relation on  $\mathbb{R}^d$  that identifies any two vectors whose difference is an integer vector:

$$(x_1, \dots, x_d) \sim (y_1, \dots, y_d) \iff (x_1 - y_1, \dots, x_d - y_d) \in \mathbb{Z}^d.$$

We denote by  $[x]$  or  $[(x_1, \dots, x_d)]$  the equivalence class of any  $x = (x_1, \dots, x_d)$ . Then we call the  $d$ -dimensional torus, or simply the  $d$ -torus, the space

$$\mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d = (\mathbb{R} / \mathbb{Z})^d$$

formed by those equivalence classes. Let  $m$  be the Lebesgue measure on  $\mathbb{R}^d$ . The operation

$$[(x_1, \dots, x_d)] + [(y_1, \dots, y_d)] = [(x_1 + y_1, \dots, x_d + y_d)]$$

is well defined and turns  $\mathbb{T}^d$  into an abelian group. Given  $\theta = (\theta_1, \dots, \theta_d) \in \mathbb{R}^d$ , we call

$$R_\theta : \mathbb{T}^d \rightarrow \mathbb{T}^d, \quad R_\theta([x]) = [x] + [\theta]$$

the rotation by  $\theta$  (sometimes,  $R_\theta$  is also called the translation by  $\theta$ ). The map

$$\phi : [0, 1]^d \rightarrow \mathbb{T}^d, \quad (x_1, \dots, x_d) \mapsto [(x_1, \dots, x_d)]$$

is surjective and allows us to define a Lebesgue probability measure  $\mu$  on the  $d$ -torus, through the following formula:

$$\mu(B) = m(\phi^{-1}(B)) \quad \text{for every } B \subset \mathbb{T}^d \text{ such that } \phi^{-1}(B) \text{ is measurable.}$$

This measure  $\mu$  is invariant under  $R_\theta$  for every  $\theta$ .

We say that a vector  $\theta = (\theta_1, \dots, \theta_d) \in \mathbb{R}^d$  is *rationally independent* if, for any integer numbers  $n_0, n_1, \dots, n_d$ , we have that

$$n_0 + n_1\theta_1 + \dots + n_d\theta_d = 0 \quad \Rightarrow \quad n_0 = n_1 = \dots = n_d = 0.$$

Otherwise, we say that  $\theta$  is rationally dependent. One can show that  $\theta$  is rationally independent if and only if the rotation  $R_\theta$  is *minimal*, meaning that the orbit  $\mathcal{O}([x]) = \{R_\theta^n([x]) : n \in \mathbb{N}\}$  of every  $[x] \in \mathbb{T}^d$  is a dense subset of  $\mathbb{T}^d$ . In this regard, see Exercises 1.3.9–1.3.10 and also Corollary 4.2.3.

### 1.3.5 Conservative maps

Let  $M$  be an open subset of the Euclidian space  $\mathbb{R}^d$  and  $f : M \rightarrow M$  be a  $C^1$  diffeomorphism. This means that  $f$  is a bijection, both  $f$  and its inverse  $f^{-1}$  are differentiable and the two derivatives are continuous. Denote by  $\text{vol}$  the restriction to  $M$  of the Lebesgue measure (volume measure) on  $\mathbb{R}^d$ . The formula of change of variables asserts that, for any measurable set  $B \subset M$ ,

$$\text{vol}(f(B)) = \int_B |\det Df| \, dx. \tag{1.3.12}$$

One can easily deduce the following consequence:

**Lemma 1.3.5.** *A  $C^1$  diffeomorphism  $f : M \rightarrow M$  preserves the volume measure  $\text{vol}$  if and only if the absolute value  $|\det Df|$  of its Jacobian is equal to 1 at every point.*

*Proof.* Suppose that the absolute value  $|\det Df|$  of its Jacobian is equal to 1 at every point. Let  $E$  be any measurable set  $E$  and  $B = f^{-1}(E)$ . The formula (1.3.12) yields

$$\text{vol}(E) = \int_B 1 \, dx = \text{vol}(B) = \text{vol}(f^{-1}(E)).$$

This means that  $f$  preserves the measure  $\text{vol}$  and so we proved the “if” part of the statement.

To prove the “only if,” suppose that  $|\det Df(x)| > 1$  for some point  $x \in M$ . Then, since the Jacobian is continuous, there exists a neighborhood  $U$  of  $x$  and some number  $\sigma > 1$  such that

$$|\det Df(y)| \geq \sigma \quad \text{for all } y \in U.$$

Then, applying (1.3.12) to  $B = U$ , we get that

$$\text{vol}(f(U)) \geq \int_U \sigma \, dx \geq \sigma \text{vol}(U).$$

Denote  $E = f(U)$ . Since  $\text{vol}(U) > 0$ , the previous inequality implies that  $\text{vol}(E) > \text{vol}(f^{-1}(E))$ . Hence,  $f$  does not leave  $\text{vol}$  invariant. In precisely the same way, one shows that if  $|\det Df(x)| < 1$  for some point  $x \in M$  then  $f$  does not leave the measure  $\text{vol}$  invariant.

### 1.3.6 Conservative flows

Now we discuss the invariance of the volume measure in the setting of flows  $f^t : M \rightarrow M, t \in \mathbb{R}$ . As before, take  $M$  to be an open subset of the Euclidean space  $\mathbb{R}^d$ . Let us suppose that the flow is  $C^1$ , in the sense that the map  $(t, x) \mapsto f^t(x)$  is differentiable and all the derivatives are continuous. Then, in particular, every flow transformation  $f^t : M \rightarrow M$  is a  $C^1$  diffeomorphism: the inverse is  $f^{-t}$ . Since  $f^0$  is the identity map and the Jacobian varies continuously, we have that  $\det Df^t(x) > 0$  at every point.

Applying Lemma 1.3.5 in this context, we find that the flow preserves the volume measure if and only if

$$\det Df^t(x) = 1 \quad \text{for every } x \in U \text{ and every } t \in \mathbb{R}. \tag{1.3.13}$$

However, this is not very useful in practice because most of the time we do not have an explicit expression for  $f^t$  and, hence, it is not clear how to check the condition (1.3.13). Fortunately, there is a reasonably explicit expression for the Jacobian of the flow that can be used in some interesting situations. Let us explain this.

Let us suppose that the flow  $f^t : M \rightarrow M$  corresponds to the trajectories of a  $C^1$  vector field  $F : M \rightarrow \mathbb{R}^d$ . In other words, each  $t \mapsto f^t(x)$  is the solution of the differential equation

$$\frac{dy}{dt} = F(y) \tag{1.3.14}$$

that has  $x$  as the initial condition (when dealing with differential equations we always assume that their solutions are defined for all times).

The *Liouville formula* relates the Jacobian of  $f^t$  to the *divergence*  $\text{div} F$  of the vector field:

$$\det Df^t(x) = \exp \left( \int_0^t \text{div} F(f^s(x)) ds \right) \quad \text{for every } x \text{ and every } t.$$

Recall that the divergence of a vector field  $F$  is the trace of its Jacobian matrix, that is

$$\text{div} F = \frac{\partial F_1}{\partial x_1} + \dots + \frac{\partial F_d}{\partial x_d}. \tag{1.3.15}$$

Combining the Liouville formula with (1.3.13), we obtain:

**Lemma 1.3.6** (Liouville). *The flow  $(f^t)_t$  associated with a  $C^1$  vector field  $F$  preserves the volume measure if and only if the divergence of  $F$  is identically zero.*

We can extend this discussion to the case when  $M$  is any Riemannian manifold of dimension  $d \geq 2$ . The reader who is unfamiliar with this notion may wish to check Appendix A.4.5 before proceeding.

For simplicity, let us suppose that the manifold is orientable. Then the volume measure on  $M$  is given by a differentiable  $d$ -form  $\omega$ , called the *volume form* (this remains true in the non-orientable case, except that the form  $\omega$  is defined up to sign only). What this means is that the volume of any measurable set  $B$  contained in the domain of local coordinates  $(x_1, \dots, x_d)$  is given by

$$\text{vol}(B) = \int_B \rho(x_1, \dots, x_d) dx_1 \cdots dx_d,$$

where  $\omega = \rho dx_1 \cdots dx_d$  is the expression of the volume form in those local coordinates. Let  $F$  be a  $C^1$  vector field on  $M$ . Writing

$$F(x_1, \dots, x_d) = (F_1(x_1, \dots, x_d), \dots, F_d(x_1, \dots, x_d)),$$

we may express the divergence as

$$\text{div } F = \frac{\partial(\rho F)}{\partial x_1} + \cdots + \frac{\partial(\rho F)}{\partial x_d}$$

(it can be shown that the right-hand side does not depend on the choice of the local coordinates). A proof of the following generalization of Lemma 1.3.6 can be found in Sternberg [Ste58]:

**Theorem 1.3.7 (Liouville).** *The flow ( $f^t$ ), associated with a  $C^1$  vector field  $F$  on a Riemannian manifold preserves the volume measure on the manifold if and only if  $\text{div } F = 0$  at every point.*

Then, it follows from the recurrence theorem for flows that, assuming that the manifold has finite volume (for example, if  $M$  is compact) and  $\text{div } F = 0$ , then almost every point is recurrent for the flow of  $F$ .

### 1.3.7 Exercises

- 1.3.1. Use Lemma 1.3.3 to give another proof of the fact that the decimal expansion transformation  $f(x) = 10x - [10x]$  preserves the Lebesgue measure on the interval.
- 1.3.2. Prove that, for any number  $x \in [0, 1]$  whose decimal expansion contains the block 617 (for instance,  $x = 0.3375\mathbf{617}264\dots$ ), that block occurs infinitely many times in the decimal expansion of  $x$ . Even more, the block 617 occurs infinitely many times in the decimal expansion of almost every  $x \in [0, 1]$ .
- 1.3.3. Prove that the number 617 appears infinitely many times in the continued fraction expression of almost every number  $x_0 \in (1/618, 1/617)$ , that is, one has  $a_n = 617$  for infinitely many values of  $n \in \mathbb{N}$ .

- 1.3.4. Let  $G$  be the Gauss map. Show that a number  $x \in (0, 1)$  is rational if and only if there exists  $n \geq 1$  such that  $G^n(x) = 0$ .
- 1.3.5. Consider the sequence  $1, 2, 4, 8, \dots, a_n = 2^n, \dots$  of all the powers of 2. Prove that, given any digit  $i \in \{1, \dots, 9\}$ , there exist infinitely many values of  $n$  for which  $a_n$  starts with that digit.
- 1.3.6. Prove the following extension of Lemma 1.3.3. Let  $f : M \rightarrow M$  be a  $C^1$  local diffeomorphism on a compact Riemannian manifold  $M$ . Let  $\text{vol}$  be the volume measure on  $M$  and  $\rho : M \rightarrow [0, \infty)$  be a continuous function. Then  $f$  preserves the measure  $\mu = \rho \text{vol}$  if and only if

$$\sum_{x \in f^{-1}(y)} \frac{\rho(x)}{|\det Df(x)|} = \rho(y) \quad \text{for every } y \in M.$$

When  $f$  is invertible this means that  $f$  preserves the measure  $\mu$  if and only if  $\rho(x) = \rho(f(x))|\det Df(x)|$  for every  $x \in M$ .

- 1.3.7. Check that if  $A$  is a  $d \times d$  matrix with integer coefficients and determinant different from zero then the transformation  $f_A : \mathbb{T}^d \rightarrow \mathbb{T}^d$  defined on the torus by  $f_A([x]) = [A(x)]$  preserves the Lebesgue measure on  $\mathbb{T}^d$ .
- 1.3.8. Show that the Lebesgue measure on  $S^1$  is the only probability measure invariant under all the rotations of  $S^1$ , even if we restrict to *rational* rotations. [Note: We will see in Chapter 6 that, for any *irrational*  $\theta$ , the Lebesgue measure is the unique probability measure invariant under  $R_\theta$ .]
- 1.3.9. Suppose that  $\theta = (\theta_1, \dots, \theta_d)$  is rationally dependent. Show that there exists a continuous non-constant function  $\varphi : \mathbb{T}^d \rightarrow \mathbb{C}$  such that  $\varphi \circ R_\theta = \varphi$ . Conclude that there exist non-empty open subsets  $U$  and  $V$  of  $\mathbb{T}^d$  that are disjoint and invariant under  $R_\theta$ , in the sense that  $R_\theta(U) = U$  and  $R_\theta(V) = V$ . Deduce that no orbit  $\mathcal{O}([x])$  of the rotation  $R_\theta$  is dense in  $\mathbb{T}^d$ .
- 1.3.10. Suppose that  $\theta = (\theta_1, \dots, \theta_d)$  is rationally independent. Prove that if  $V$  is a non-empty open subset of  $\mathbb{T}^d$  invariant under  $R_\theta$ , then  $V$  is dense in  $\mathbb{T}^d$ . Conclude that  $\bigcup_{n \in \mathbb{Z}} R_\theta^n(U)$  is dense in the torus, for every non-empty open subset  $U$ . Deduce that there exists  $[x]$  whose orbit  $\mathcal{O}([x])$  under the rotation  $R_\theta$  is dense in  $\mathbb{T}^d$ . Conclude that  $\mathcal{O}([y])$  is dense in  $\mathbb{T}^d$  for every  $[y]$ .
- 1.3.11. Let  $U$  be an open subset of  $\mathbb{R}^{2d}$  and  $H : U \rightarrow \mathbb{R}$  be a  $C^2$  function. Denote by  $(p_1, \dots, p_d, q_1, \dots, q_d)$  the coordinate variables in  $\mathbb{R}^{2d}$ . The *Hamiltonian vector field* associated with  $H$  is defined by

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

Check that the flow defined by  $F$  preserves the volume measure.

- 1.3.12. Let  $f : U \rightarrow U$  be a  $C^1$  diffeomorphism preserving the volume measure on an open subset  $U$  of  $\mathbb{R}^d$ . Let  $H : U \rightarrow \mathbb{R}$  be a first integral of  $f$ , that is, a  $C^1$  function such that  $H \circ f = H$ . Let  $c$  be a regular value of  $H$  and  $ds$  be the volume measure defined on the hypersurface  $H_c = H^{-1}(c)$  by the restriction of the Riemannian metric of  $\mathbb{R}^d$ . Prove that the restriction of  $f$  to the hypersurface  $H_c$  preserves the measure  $ds / \|\text{grad} H\|$ .

### 1.4 Induction

In this section we describe a general method, based on the Poincaré recurrence theorem, to construct from a given system  $(f, \mu)$  other systems, that we refer to as *systems induced* by  $(f, \mu)$ . The reason this is interesting is the following. On the one hand, it is often the case that an induced system is easier to analyze, because it has better global properties than the original one. On the other hand, interesting conclusions about the original system can often be obtained from analyzing the induced one. Examples will appear in a while.

#### 1.4.1 First-return map

Let  $f : M \rightarrow M$  be a measurable transformation and  $\mu$  be an invariant probability measure. Let  $E \subset M$  be a measurable set with  $\mu(E) > 0$  and  $\rho(x) = \rho_E(x)$  be the first-return time of  $x$  to  $E$ , as given by (1.2.1). The *first-return map* to the domain  $E$  is the map  $g$  given by

$$g(x) = f^{\rho(x)}(x)$$

whenever  $\rho(x)$  is finite. The Poincaré recurrence theorem ensures that this is the case for  $\mu$ -almost every  $x \in E$  and so  $g$  is defined on a full measure subset of  $E$ . We also denote by  $\mu_E$  the restriction of  $\mu$  to the measurable subsets  $E$ .

**Proposition 1.4.1.** *The measure  $\mu_E$  is invariant under the map  $g : E \rightarrow E$ .*

*Proof.* For every  $k \geq 1$ , denote by  $E_k$  the subset of points  $x \in E$  such that  $\rho(x) = k$ . By definition,  $g(x) = f^k(x)$  for every  $x \in E_k$ . Let  $B$  be any measurable subset of  $E$ . Then

$$\mu(g^{-1}(B)) = \sum_{k=1}^{\infty} \mu\left(f^{-k}(B) \cap E_k\right). \tag{1.4.1}$$

On the other hand, since  $\mu$  is  $f$ -invariant,

$$\mu(B) = \mu(f^{-1}(B)) = \mu(f^{-1}(B) \cap E_1) + \mu(f^{-1}(B) \setminus E). \tag{1.4.2}$$

Analogously,

$$\begin{aligned} \mu(f^{-1}(B) \setminus E) &= \mu(f^{-2}(B) \setminus f^{-1}(E)) \\ &= \mu(f^{-2}(B) \cap E_2) + \mu(f^{-2}(B) \setminus (E \cup f^{-1}(E))). \end{aligned}$$

Replacing this expression in (1.4.2), we find that

$$\mu(B) = \sum_{k=1}^2 \mu(f^{-k}(B) \cap E_k) + \mu\left(f^{-2}(B) \setminus \bigcup_{k=0}^1 f^{-k}(E)\right).$$

Repeating this argument successively, we obtain

$$\mu(B) = \sum_{k=1}^n \mu(f^{-k}(B) \cap E_k) + \mu\left(f^{-n}(B) \setminus \bigcup_{k=0}^{n-1} f^{-k}(E)\right). \tag{1.4.3}$$

Now let us go to the limit when  $n \rightarrow \infty$ . It is clear that the last term is bounded above by  $\mu(f^{-n}(E) \setminus \bigcup_{k=0}^{n-1} f^{-k}(E))$ . So, using Remark 1.2.3, that term converges to zero when  $n \rightarrow \infty$ . In this way we conclude that

$$\mu(B) = \sum_{k=1}^{\infty} \mu(f^{-k}(B) \cap E_k).$$

Together with (1.4.1), this shows that  $\mu(g^{-1}(B)) = \mu(B)$  for every measurable subset  $B$  of  $E$ . That is to say, the measure  $\mu_E$  is invariant under  $g$ .

**Example 1.4.2.** Consider the transformation  $f : [0, \infty) \rightarrow [0, \infty)$  defined by

$$f(0) = 0 \quad \text{and} \quad f(x) = 1/x \text{ if } x \in (0, 1) \quad \text{and} \quad f(x) = x - 1 \text{ if } x \geq 1.$$

Let  $E = [0, 1]$ . The time  $\rho$  of first return to  $E$  is given by

$$\rho(0) = 1 \quad \text{and} \quad \rho(x) = k + 1 \text{ if } x \in (1/(k + 1), 1/k] \text{ with } k \geq 1.$$

So, the first-return map to  $E$  is given by

$$g(0) = 0 \quad \text{and} \quad g(x) = 1/x - k \text{ if } x \in (1/(k + 1), 1/k] \text{ with } k \geq 1.$$

In other words,  $g$  is the Gauss map. We saw in Section 1.3.2 that the Gauss map admits an invariant probability measure equivalent to the Lebesgue measure on  $[0, 1)$ . From this, one can draw some interesting conclusions about the original map  $f$ . For instance, using the ideas in the next section one finds that  $f$  admits an (infinite) invariant measure equivalent to the Lebesgue measure on  $[0, \infty)$ .

### 1.4.2 Induced transformations

In an opposite direction, given any measure  $\nu$  invariant under  $g : E \rightarrow E$ , we may construct a certain related measure  $\nu_\rho$  that is invariant under  $f : M \rightarrow M$ . For this,  $g$  does not even have to be a first-return map: the construction that we present below is valid for any map *induced* from  $f$ , that is, any map of the form

$$g : E \rightarrow E, \quad g(x) = f^{\rho(x)}(x), \tag{1.4.4}$$

where  $\rho : E \rightarrow \mathbb{N}$  is a measurable function (it suffices that  $\rho$  is defined on some full measure subset of  $E$ ). As before, we denote by  $E_k$  the subset of points  $x \in E$  such that  $\rho(x) = k$ . Then we define

$$\nu_\rho(B) = \sum_{n=0}^{\infty} \sum_{k>n} \nu(f^{-n}(B) \cap E_k), \tag{1.4.5}$$

for every measurable set  $B \subset M$ .

**Proposition 1.4.3.** *The measure  $\nu_\rho$  defined in (1.4.5) is invariant under  $f$  and satisfies  $\nu_\rho(M) = \int_E \rho \, d\nu$ . In particular,  $\nu_\rho$  is finite if and only if the function  $\rho$  is integrable with respect to  $\nu$ .*



*Proof.* First, let us prove that  $\nu_\rho$  is invariant. By the definition (1.4.5),

$$\nu_\rho(f^{-1}(B)) = \sum_{n=0}^\infty \sum_{k>n} \nu(f^{-(n+1)}(B) \cap E_k) = \sum_{n=1}^\infty \sum_{k \geq n} \nu(f^{-n}(B) \cap E_k).$$

We may rewrite this expression as follows:

$$\nu_\rho(f^{-1}(B)) = \sum_{n=1}^\infty \sum_{k>n} \nu(f^{-n}(B) \cap E_k) + \sum_{k=1}^\infty \nu(f^{-k}(B) \cap E_k). \tag{1.4.6}$$

Concerning the last term, observe that

$$\sum_{k=1}^\infty \nu(f^{-k}(B) \cap E_k) = \nu(g^{-1}(B)) = \nu(B) = \sum_{k=1}^\infty \nu(B \cap E_k),$$

since  $\nu$  is invariant under  $g$ . Replacing this in (1.4.6), we see that

$$\nu_\rho(f^{-1}(B)) = \sum_{n=0}^\infty \sum_{k>n} \nu(f^{-n}(B) \cap E_k) + \sum_{k=1}^\infty \nu(B \cap E_k) = \nu_\rho(B)$$

for every measurable set  $B \subset E$ . The second claim is a direct consequence of the definitions:

$$\nu_\rho(M) = \sum_{n=0}^\infty \sum_{k>n} \nu(f^{-n}(M) \cap E_k) = \sum_{n=0}^\infty \sum_{k>n} \nu(E_k) = \sum_{k=1}^\infty k\nu(E_k) = \int_E \rho \, d\nu.$$

This completes the proof.

It is interesting to analyze how this construction relates to the one in the previous section when  $g$  is a first-return map of  $f$  and the measure  $\nu$  is the restriction  $\mu \mid E$  of some invariant measure  $\mu$  of  $f$ :

**Corollary 1.4.4.** *If  $g$  is the first-return map of  $f$  to a measurable subset  $E$  and  $\nu = \mu \mid E$ , then*

1.  $\nu_\rho(B) = \nu(B) = \mu(B)$  for every measurable set  $B \subset E$ .
2.  $\nu_\rho(B) \leq \mu(B)$  for every measurable set  $B \subset M$ .

*Proof.* By definition,  $f^{-n}(E) \cap E_k = \emptyset$  for every  $0 < n < k$ . This implies that, given any measurable set  $B \subset E$ , all the terms with  $n > 0$  in the definition (1.4.5) are zero. Hence,  $\nu_\rho(B) = \sum_{k>0} \nu(B \cap E_k) = \nu(B)$  as claimed in the first part of the statement.

Consider any measurable set  $B \subset M$ . Then,

$$\begin{aligned} \mu(B) &= \mu(B \cap E) + \mu(B \cap E^c) = \nu(B \cap E) + \mu(B \cap E^c) \\ &= \sum_{k=1}^\infty \nu(B \cap E_k) + \mu(B \cap E^c). \end{aligned} \tag{1.4.7}$$

Since  $\mu$  is invariant,  $\mu(B \cap E^c) = \mu(f^{-1}(B) \cap f^{-1}(E^c))$ . Then, as in the previous equality,

$$\begin{aligned} \mu(B \cap E^c) &= \mu(f^{-1}(B) \cap E \cap f^{-1}(E^c)) + \mu(f^{-1}(B) \cap E^c \cap f^{-1}(E^c)) \\ &= \sum_{k=2}^{\infty} \nu(f^{-1}(B) \cap E_k) + \mu(f^{-1}(B) \cap E^c \cap f^{-1}(E^c)). \end{aligned}$$

Replacing this in (1.4.7), we find that

$$\mu(B) = \sum_{n=0}^1 \sum_{k>n} \nu(f^{-n}(B) \cap E_k) + \mu\left(f^{-1}(B) \cap \bigcap_{n=0}^1 f^{-n}(E^c)\right).$$

Repeating this argument successively, we get that

$$\begin{aligned} \mu(B) &= \sum_{n=0}^N \sum_{k>n} \nu(f^{-n}(B) \cap E_k) + \mu\left(f^{-N}(B) \cap \bigcap_{k=0}^N f^{-n}(E^c)\right) \\ &\geq \sum_{n=0}^N \sum_{k>n} \nu(f^{-n}(B) \cap E_k) \quad \text{for every } N \geq 1. \end{aligned}$$

Taking the limit as  $N \rightarrow \infty$ , we conclude that  $\mu(B) \geq \nu_\rho(B)$ .

We also have from the Kač theorem (Theorem 1.2.2) that

$$\nu_\rho(M) = \int_E \rho \, d\nu = \int_E \rho \, d\mu = \mu(M) - \mu(E_0^*).$$

So, it follows from Corollary 1.4.4 that  $\nu_\rho = \mu$  if and only if  $\mu(E_0^*) = 0$ .

**Example 1.4.5** (Manneville–Pomeau). Given  $d > 0$ , let  $a$  be the only number in  $(0, 1)$  such that  $a(1 + a^d) = 1$ . Then define  $f : [0, 1] \rightarrow [0, 1]$  as follows:

$$f(x) = x(1 + x^d) \quad \text{if } x \in [0, a] \quad \text{and} \quad f(x) = \frac{x - a}{1 - a} \quad \text{if } x \in (a, 1].$$

The graph of  $f$  is depicted on the left-hand side of Figure 1.3. Observe that  $|f'(x)| \geq 1$  at every point, and the inequality is strict at every  $x > 0$ . Let  $(a_n)_n$  be the sequence on the interval  $[0, a]$  defined by  $a_1 = a$  and  $f(a_{n+1}) = a_n$  for  $n \geq 1$ . We also write  $a_0 = 1$ . Some properties of this sequence are studied in Exercise 1.4.2.

Now consider the map  $g(x) = f^{\rho(x)}(x)$ , where

$$\rho : [0, 1] \rightarrow \mathbb{N}, \quad \rho(x) = 1 + \min\{n \geq 0 : f^n(x) \in (a, 1]\}.$$

In other words,  $\rho(x) = k$  and so  $g(x) = f^k(x)$  for every  $x \in (a_k, a_{k-1}]$ . The graph of  $g$  is represented on the right-hand side of Figure 1.3. Note that the restriction to each interval  $(a_k, a_{k-1}]$  is a bijection onto  $(0, 1]$ . A key point is that the induced map  $g$  is *expanding*:

$$|g'(x)| \geq \frac{1}{1 - a} > 1 \quad \text{for every } x \in [0, 1].$$

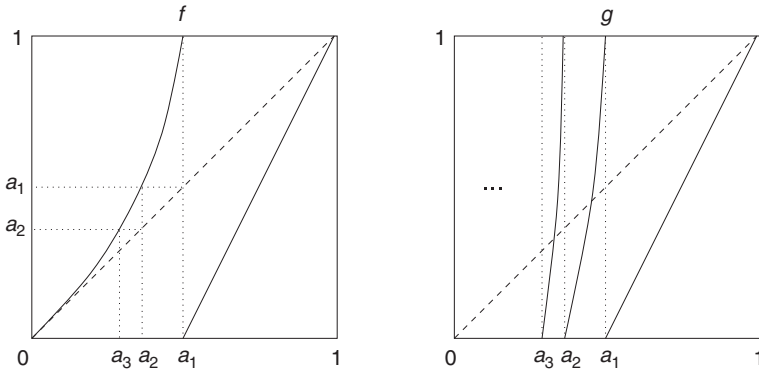


Figure 1.3. Construction of an induced transformation

Using the ideas that will be developed in Chapter 11, one can show that  $g$  admits a unique invariant probability measure  $\nu$  equivalent to the Lebesgue measure on  $(0, 1]$ . In fact, the density (Radon–Nikodym derivative) of  $\nu$  with respect to the Lebesgue measure is bounded from zero and infinity. Then, the  $f$ -invariant measure  $\nu_\rho$  in (1.4.5) is equivalent to Lebesgue measure. It follows (see Exercise 1.4.2) that this measure is finite if and only if  $d \in (0, 1)$ .

### 1.4.3 Kakutani–Rokhlin towers

It is possible, and useful, to generalize the previous constructions even further, by omitting the initial transformation  $f : M \rightarrow M$  altogether. More precisely, given a transformation  $g : E \rightarrow E$ , a measure  $\nu$  on  $E$  invariant under  $g$  and a measurable function  $\rho : E \rightarrow \mathbb{N}$ , we are going to *construct* a transformation  $f : M \rightarrow M$  and a measure  $\nu_\rho$  invariant under  $f$  such that  $E$  can be identified with a subset of  $M$ ,  $g$  is the first-return map of  $f$  to  $E$ , with first-return time given by  $\rho$ , and the restriction of  $\nu_\rho$  to  $E$  coincides with  $\nu$ .

This transformation  $f$  is called the *Kakutani–Rokhlin tower* of  $g$  with time  $\rho$ . The measure  $\nu_\rho$  is finite if and only if  $\rho$  is integrable with respect to  $\nu$ . They are constructed as follows. Begin by defining

$$M = \{(x, n) : x \in E \text{ and } 0 \leq n < \rho(x)\}$$

$$= \bigcup_{k=1}^{\infty} \bigcup_{n=0}^{k-1} E_k \times \{n\}.$$

In other words,  $M$  consists of  $k$  copies of each set  $E_k = \{x \in E : \rho(x) = k\}$ , “piled up” on top of each other. We call each  $\bigcup_{k>n} E_k \times \{n\}$  the  $n$ -th floor of  $M$ . See Figure 1.4.

Next, define  $f : M \rightarrow M$  as follows:

$$f(x, n) = \begin{cases} (x, n + 1) & \text{if } n < \rho(x) - 1 \\ (g(x), 0) & \text{if } n = \rho(x) - 1 \end{cases}.$$

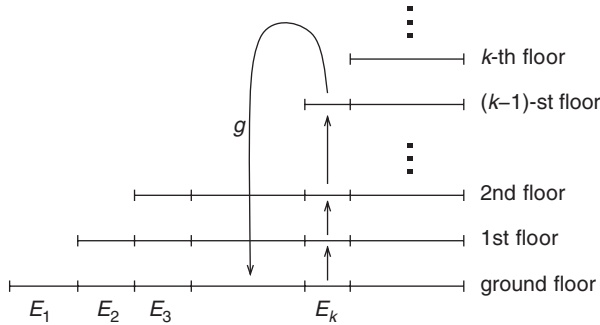


Figure 1.4. Kakutani–Rokhlin tower of  $g$  with time  $\rho$

In other words, each point  $(x, n)$  is “lifted” one floor at a time, until reaching the floor  $\rho(x) - 1$ ; at that stage, the point “falls” directly to  $(g(x), 0)$  on the ground (zero-th) floor. The ground floor  $E \times \{0\}$  is naturally identified with the set  $E$ . Besides, the first-return map to  $E \times \{0\}$  corresponds precisely to  $g : E \rightarrow E$ .

Finally, the measure  $\nu_\rho$  is defined by

$$\nu_\rho \mid (E_k \times \{n\}) = \nu \mid E_k$$

for every  $0 \leq n < k$ . It is clear that the restriction of  $\nu_\rho$  to the ground floor coincides with  $\nu$ . Moreover,  $\nu_\rho$  is invariant under  $f$  and

$$\nu_\rho(M) = \sum_{k=1}^{\infty} k\nu(E_k) = \int_E \rho \, d\nu.$$

This completes the construction of the Kakutani–Rokhlin tower.

### 1.4.4 Exercises

- 1.4.1. Let  $f : S^1 \rightarrow S^1$  be the transformation  $f(x) = 2x \pmod{\mathbb{Z}}$ . Show that the function  $\tau(x) = \min\{k \geq 0 : f^k(x) \in (1/2, 1)\}$  is integrable with respect to the Lebesgue measure. State and prove a corresponding result for any  $C^1$  transformation  $g : S^1 \rightarrow S^1$  that is close to  $f$ , in the sense that  $\sup_x \{\|g(x) - f(x)\|, \|g'(x) - f'(x)\|\}$  is sufficiently small.
- 1.4.2. Consider the measure  $\nu_\rho$  and the sequence  $(a_n)_n$  defined in Example 1.4.5. Check that  $\nu_\rho$  is always  $\sigma$ -finite. Show that  $(a_n)_n$  is decreasing and converges to zero. Moreover, there exist  $c_1, c_2, c_3, c_4 > 0$  such that

$$c_1 \leq a_j^{1/d} \leq c_2 \quad \text{and} \quad c_3 \leq (a_j - a_{j+1})^{1+1/d} \leq c_4 \quad \text{for every } j. \quad (1.4.8)$$

Deduce that the  $g$ -invariant measure  $\nu_\rho$  is finite if and only if  $d \in (0, 1)$ .

- 1.4.3. Let  $\sigma : \Sigma \rightarrow \Sigma$  be the map defined on the space  $\Sigma = \{1, \dots, d\}^{\mathbb{Z}}$  by  $\sigma((x_n)_n) = (x_{n+1})_n$ . Describe the first-return map  $g$  to the subset  $\{(x_n)_n \in \Sigma : x_0 = 1\}$ .
- 1.4.4. [Kakutani–Rokhlin lemma] Let  $f : M \rightarrow M$  be an invertible transformation and  $\mu$  be an invariant probability measure without atoms and such that  $\mu(\bigcup_{n \in \mathbb{N}} f^n(E)) = 1$  for every  $E \subset M$  with  $\mu(E) > 0$ . Show that for every

$n \geq 1$  and  $\varepsilon > 0$  there exists a measurable set  $B \subset M$  such that the iterates  $B, f(B), \dots, f^{n-1}(B)$  are pairwise disjoint and the complement of their union has measure less than  $\varepsilon$ . In particular, this holds for every invertible system that is *aperiodic*, that is, whose periodic points form a zero measure set.

- 1.4.5. Let  $f : M \rightarrow M$  be a transformation and  $(H_j)_{j \geq 1}$  be a collection of subsets of  $M$  such that if  $x \in H_n$  then  $f^j(x) \in H_{n-j}$  for every  $0 \leq j < n$ . Let  $H$  be the set of points that belong to  $H_j$  for infinitely many values of  $j$ , that is,  $H = \bigcap_{k=1}^{\infty} \bigcup_{j=k}^{\infty} H_j$ . For  $y \in H$ , define  $\tau(y) = \min\{j \geq 1 : y \in H_j\}$  and  $T(y) = f^{\tau(y)}(y)$ . Observe that  $T$  maps  $H$  inside  $H$ . Moreover, show that

$$\limsup_n \frac{1}{n} \#\{1 \leq j \leq n : x \in H_j\} \geq \theta > 0 \implies \liminf_k \frac{1}{k} \sum_{i=0}^{k-1} \tau(T^i(x)) \leq \frac{1}{\theta}.$$

- 1.4.6. Let  $f : M \rightarrow M$  be a transformation preserving a measure  $\mu$ . Let  $(H_j)_{j \geq 1}$  and  $\tau : M \rightarrow \mathbb{N}$  be as in Exercise 1.4.5. Consider the sequence of functions  $(\tau_n)_n$  defined by  $\tau_1(x) = \tau(x)$  and  $\tau_n(x) = \tau(f^{\tau_{n-1}(x)}(x)) + \tau_{n-1}(x)$  for  $n > 1$ . Suppose that

$$\limsup_n \frac{1}{n} \#\{1 \leq j \leq n : x \in H_j\} \geq \theta > 0 \quad \text{for } \mu\text{-almost every } x \in M.$$

Show that  $\tau_{n+1}(x)/\tau_n(x) \rightarrow 1$  for  $\mu$ -almost every  $x \in M$ . [Note: Sequences with this property are called *non-lacunary*.]

### 1.5 Multiple recurrence theorems

Now we consider finite families of *commuting maps*  $f_i : M \rightarrow M, i = 1, \dots, q$ , that is, such that

$$f_i \circ f_j = f_j \circ f_i \quad \text{for every } i, j \in \{1, \dots, q\}.$$

Our goal is to explain that the results in Section 1.2 extend to this setting: we find points that are *simultaneously recurrent* for these transformations.

The first result in this direction generalizes the Birkhoff recurrence theorem (Theorem 1.2.6):

**Theorem 1.5.1** (Birkhoff multiple recurrence). *Let  $M$  be a compact metric space and  $f_1, \dots, f_q : M \rightarrow M$  be continuous commuting maps. Then there exists  $a \in M$  and a sequence  $(n_k)_k \rightarrow \infty$  such that*

$$\lim_k f_i^{n_k}(a) = a \quad \text{for every } i = 1, \dots, q. \tag{1.5.1}$$

The key point here is that the sequence  $(n_k)_k$  does not depend on  $i$ : we say that the point  $a$  is *simultaneously recurrent* for all the maps  $f_i, i = 1, \dots, q$ . A proof of Theorem 1.5.1 is given in Section 1.5.1. Next, we discuss the following generalization of the Poincaré recurrence theorem (Theorem 1.2.1):

**Theorem 1.5.2** (Poincaré multiple recurrence). *Let  $(M, \mathcal{B}, \mu)$  be a probability space and  $f_i : M \rightarrow M, i = 1, \dots, q$  be measurable commuting maps that*

preserve the measure  $\mu$ . Then, given any set  $E \subset M$  with positive measure, there exists  $n \geq 1$  such that

$$\mu(E \cap f_1^{-n}(E) \cap \cdots \cap f_q^{-n}(E)) > 0.$$

In other words, for a positive measure subset of points  $x \in E$ , their orbits under all the maps  $f_i$ ,  $i = 1, \dots, q$  return to  $E$  *simultaneously* at time  $n$  (we say that  $n$  is a *simultaneous return* of  $x$  to  $E$ ): once more, the crucial point with the statement is that  $n$  does not depend on  $i$ .

The proof of Theorem 1.5.2 will not be presented here; we refer the interested reader to the book of Furstenberg [Fur77]. We are just going to mention some direct consequences and, in Chapter 2, we will use this theorem to prove the Szemerédi theorem on the existence of arithmetic progressions inside “dense” subsets of integer numbers.

To begin with, observe that the set of simultaneous returns is always infinite. Indeed, let  $n$  be as in the statement of Theorem 1.5.2. Applying the theorem to the set  $F = E \cap f_1^{-n}(E) \cap \cdots \cap f_q^{-n}(E)$ , we find  $m \geq 1$  such that

$$\begin{aligned} \mu(E \cap f_1^{-(m+n)}(E) \cap \cdots \cap f_q^{-(m+n)}(E)) \\ \geq \mu(F \cap f_1^{-m}(F) \cap \cdots \cap f_q^{-m}(F)) > 0. \end{aligned}$$

Thus,  $m+n$  is also a simultaneous return to  $E$ , for all the points in some subset of  $E$  with positive measure.

It follows that, for any set  $E \subset M$  with  $\mu(E) > 0$  and for  $\mu$ -almost every point  $x \in E$ , there exist infinitely many *simultaneous returns* of  $x$  to  $E$ . Indeed, suppose there is a positive measure set  $F \subset E$  such that every point of  $F$  has a finite number of simultaneous returns to  $E$ . On the one hand, up to replacing  $F$  by a suitable subset, we may suppose that the simultaneous returns to  $E$  of all the points of  $F$  are bounded by some  $k \geq 1$ . On the other hand, using the previous paragraph, there exists  $n > k$  such that  $G = F \cap f_1^{-n}(F) \cap \cdots \cap f_q^{-n}(F)$  has positive measure. Now, it is clear from the definition that  $n$  is a simultaneous return to  $E$  of every  $x \in G$ . This contradicts the choice of  $F$ , thus proving our claim.

Another direct corollary is the Birkhoff multiple recurrence theorem (Theorem 1.5.1). Indeed, if  $f_i : M \rightarrow M$ ,  $i = 1, \dots, q$  are continuous commuting transformations on a compact metric space then there exists some probability measure  $\mu$  that is invariant under all these transformations (this fact will be checked in the next chapter, see Exercise 2.2.2). From this point on, we may argue exactly as in the proof of Theorem 1.2.4. More precisely, consider any countable basis  $\{U_k\}$  for the topology of  $M$ . According to the previous paragraph, for every  $k$  there exists a set  $\tilde{U}_k \subset U_k$  with zero measure such that every point in  $U_k \setminus \tilde{U}_k$  has infinitely many simultaneous returns to  $U_k$ . Then  $\tilde{U} = \bigcup_k \tilde{U}_k$  has measure zero and every point in its complement is simultaneously recurrent, in the sense of Theorem 1.5.1.

### 1.5.1 Birkhoff multiple recurrence theorem

In this section we prove Theorem 1.5.1 in the case when the transformations  $f_1, \dots, f_q$  are homeomorphisms of  $M$ , which suffices for all our purposes in the present chapter. The general case may be deduced easily (see Exercise 2.4.7) using the concept of natural extension, which we will present in the next chapter.

The theorem may be reformulated in the following useful way. Consider the transformation  $F : M^q \rightarrow M^q$  defined on the product space  $M^q = M \times \dots \times M$  by  $F(x_1, \dots, x_q) = (f_1(x_1), \dots, f_q(x_q))$ . Denote by  $\Delta_q$  the *diagonal* of  $M^q$ , that is, the subset of points of the form  $\tilde{x} = (x, \dots, x)$ . Theorem 1.5.1 claims, precisely, that there exist  $\tilde{a} \in \Delta_q$  and  $(n_k)_k \rightarrow \infty$  such that

$$\lim_k F^{n_k}(\tilde{a}) = \tilde{a}. \tag{1.5.2}$$

The proof of Theorem 1.5.1 is by induction on the number  $q$  of transformations. The case  $q = 1$  is contained in Theorem 1.2.6. Consider any  $q \geq 2$  and suppose that the statement is true for every family of  $q - 1$  commuting homeomorphisms. We are going to prove that it is true for the family  $f_1, \dots, f_q$ .

Let  $\mathcal{G}$  be the (abelian) group generated by the homeomorphisms  $f_1, \dots, f_q$ . We say that a set  $X \subset M$  is  $\mathcal{G}$ -invariant if  $g(X) \subset X$  for every  $g \in \mathcal{G}$ . Observing that the inverse  $g^{-1}$  is also in  $\mathcal{G}$ , we see that this implies  $g(X) = X$  for every  $g \in \mathcal{G}$ . Just as we did in Theorem 1.2.6, we may use Zorn’s lemma to conclude that there exists some minimal, non-empty, closed,  $\mathcal{G}$ -invariant set  $X \subset M$  (this is Exercise 1.5.2). The statement of the theorem is not affected if we replace  $M$  by  $X$ . Thus, it is no restriction to assume that the ambient space  $M$  is minimal. This assumption is used as follows:

**Lemma 1.5.3.** *If  $M$  is minimal then for every non-empty open set  $U \subset M$  there exists a finite subset  $\mathcal{H} \subset \mathcal{G}$  such that*

$$\bigcup_{h \in \mathcal{H}} h^{-1}(U) = M.$$

*Proof.* For any  $x \in M$ , the closure of the orbit  $\mathcal{G}(x) = \{g(x) : g \in \mathcal{G}\}$  is a non-empty, closed,  $\mathcal{G}$ -invariant subset of  $M$ . So, the hypothesis that  $M$  is minimal implies that every orbit  $\mathcal{G}(x)$  is dense in  $M$ . In particular, there is  $g \in \mathcal{G}$  such that  $g(x) \in U$ . This proves that  $\{g^{-1}(U) : g \in \mathcal{G}\}$  is an open cover of  $M$ . By compactness, it follows that there exists a finite subcover, as claimed.

Consider the product  $M^q$  endowed with the distance function

$$d((x_1, \dots, x_q), (y_1, \dots, y_q)) = \max\{d(x_i, y_i) : 1 \leq i \leq q\}.$$

Note that the map  $M \rightarrow \Delta_q, x \mapsto \tilde{x} = (x, \dots, x)$  is a homeomorphism, and even an isometry for this choice of a distance. Every open set  $U \subset M$  corresponds to an open set  $\tilde{U} \subset \Delta_q$  through this homeomorphism. Given any  $g \in \mathcal{G}$ ,

we denote by  $\tilde{g} : M^q \rightarrow M^q$  the homeomorphism defined by  $\tilde{g}(x_1, \dots, x_q) = (g(x_1), \dots, g(x_q))$ . The fact that the group  $\mathcal{G}$  is abelian implies that  $\tilde{g}$  commutes with  $F$ ; note also that every  $\tilde{g}$  preserves the diagonal  $\Delta_q$ . Then the conclusion of Lemma 1.5.3 may be rewritten in the following form:

$$\bigcup_{h \in \mathcal{H}} \tilde{h}^{-1}(\tilde{U}) = \Delta_q. \tag{1.5.3}$$

**Lemma 1.5.4.** *Given  $\varepsilon > 0$  there exist  $\tilde{x} \in \Delta_q$ ,  $\tilde{y} \in \Delta_q$  and  $n \geq 1$  such that  $d(F^n(\tilde{x}), \tilde{y}) < \varepsilon$ .*

*Proof.* Define  $g_i = f_i \circ f_q^{-1}$  for each  $i = 1, \dots, q - 1$ . Since the maps  $f_i$  commute with each other, so do the maps  $g_i$ . Then, by induction, there exist  $y \in M$  and  $(n_k)_k \rightarrow \infty$  such that

$$\lim_k g_i^{n_k}(y) = y \quad \text{for every } i = 1, \dots, q - 1.$$

Denote  $x_k = f_q^{-n_k}(y)$  and consider  $\tilde{x}_k = (x_k, \dots, x_k) \in \Delta_q$ . Then,

$$\begin{aligned} F^{n_k}(\tilde{x}_k) &= (f_1^{n_k} f_q^{-n_k}(y), \dots, f_{q-1}^{n_k} f_q^{-n_k}(y), f_q^{n_k} f_q^{-n_k}(y)) \\ &= (g_1^{n_k}(y), \dots, g_{q-1}^{n_k}(y), y) \end{aligned}$$

converges to  $(y, \dots, y, y)$  when  $k \rightarrow \infty$ . This proves the lemma, with  $\tilde{x} = \tilde{x}_k$ ,  $\tilde{y} = (y, \dots, y, y)$  and  $n = n_k$  for every  $k$  sufficiently large.

The next step is to show that the point  $\tilde{y}$  in Lemma 1.5.4 is arbitrary:

**Lemma 1.5.5.** *Given  $\varepsilon > 0$  and  $\tilde{z} \in \Delta_q$  there exist  $\tilde{w} \in \Delta_q$  and  $m \geq 1$  such that  $d(F^m(\tilde{w}), \tilde{z}) < \varepsilon$ .*

*Proof.* Given  $\varepsilon > 0$  and  $\tilde{z} \in \Delta_q$ , consider  $\tilde{U} =$  open ball of center  $\tilde{z}$  and radius  $\varepsilon/2$ . By Lemma 1.5.3 and the observation (1.5.3), we may find a finite set  $\mathcal{H} \subset \mathcal{G}$  such that the sets  $\tilde{h}^{-1}(\tilde{U})$ ,  $h \in \mathcal{H}$  cover  $\Delta_q$ . Since the elements of  $\mathcal{G}$  are (uniformly) continuous functions, there exists  $\delta > 0$  such that

$$d(\tilde{x}_1, \tilde{x}_2) < \delta \implies d(\tilde{h}(\tilde{x}_1), \tilde{h}(\tilde{x}_2)) < \varepsilon/2 \quad \text{for every } h \in \mathcal{H}.$$

By Lemma 1.5.4 there exist  $\tilde{x}, \tilde{y} \in \Delta_q$  and  $n \geq 1$  such that  $d(F^n(\tilde{x}), \tilde{y}) < \delta$ . Fix  $h \in \mathcal{H}$  such that  $\tilde{y} \in \tilde{h}^{-1}(\tilde{U})$ . Then,

$$d(\tilde{h}(F^n(\tilde{x})), \tilde{z}) \leq d(\tilde{h}(F^n(\tilde{x})), \tilde{h}(\tilde{y})) + d(\tilde{h}(\tilde{y}), \tilde{z}) < \varepsilon/2 + \varepsilon/2.$$

Take  $\tilde{w} = \tilde{h}(\tilde{x})$ . Since  $\tilde{h}$  commutes with  $F^n$ , the previous inequality implies that  $d(F^n(\tilde{w}), \tilde{z}) < \varepsilon$ , as we wanted to prove.

Next, we prove that one may take  $\tilde{x} = \tilde{y}$  in Lemma 1.5.4:

**Lemma 1.5.6** (Bowen). *Given  $\varepsilon > 0$  there exist  $\tilde{v} \in \Delta_q$  and  $k \geq 1$  with  $d(F^k(\tilde{v}), \tilde{v}) < \varepsilon$ .*



*Proof.* Given  $\varepsilon > 0$  and  $\tilde{z}_0 \in \Delta_q$ , consider the sequences  $\varepsilon_j, m_j$  and  $\tilde{z}_j, j \geq 1$  defined by recurrence as follows. Initially, take  $\varepsilon_1 = \varepsilon/2$ .

- By Lemma 1.5.5 there are  $\tilde{z}_1 \in \Delta_q$  and  $m_1 \geq 1$  with  $d(F^{m_1}(\tilde{z}_1), \tilde{z}_0) < \varepsilon_1$ .
- By the continuity of  $F^{m_1}$ , there exists  $\varepsilon_2 < \varepsilon_1$  such that  $d(\tilde{z}, \tilde{z}_1) < \varepsilon_2$  implies  $d(F^{m_1}(\tilde{z}), \tilde{z}_0) < \varepsilon_1$ .

Next, given any  $j \geq 2$ :

- By Lemma 1.5.5 there are  $\tilde{z}_j \in \Delta_q$  and  $m_j \geq 1$  with  $d(F^{m_j}(\tilde{z}_j), \tilde{z}_{j-1}) < \varepsilon_j$ .
- By the continuity of  $F^{m_j}$ , there exists  $\varepsilon_{j+1} < \varepsilon_j$  such that  $d(\tilde{z}, \tilde{z}_j) < \varepsilon_{j+1}$  implies  $d(F^{m_j}(\tilde{z}), \tilde{z}_{j-1}) < \varepsilon_j$ .

In particular, for any  $i < j$ ,

$$d(F^{m_{i+1}+\dots+m_j}(\tilde{z}_j), \tilde{z}_i) < \varepsilon_{i+1} \leq \frac{\varepsilon}{2}.$$

Since  $\Delta_q$  is compact, we can find  $i, j$  with  $i < j$  such that  $d(\tilde{z}_i, \tilde{z}_j) < \varepsilon/2$ . Take  $k = m_{i+1} + \dots + m_j$ . Then,

$$d(F^k(\tilde{z}_j), \tilde{z}_j) \leq d(F^k(\tilde{z}_j), \tilde{z}_i) + d(\tilde{z}_i, \tilde{z}_j) < \varepsilon.$$

This completes the proof of the lemma.

Now we are ready to conclude the proof of Theorem 1.5.1. For that, let us consider the function

$$\phi : \Delta_q \rightarrow [0, \infty), \quad \phi(\tilde{x}) = \inf\{d(F^n(\tilde{x}), \tilde{x}) : n \geq 1\}.$$

Observe that  $\phi$  is upper semi-continuous: given any  $\varepsilon > 0$ , every point  $\tilde{x}$  admits some neighborhood  $V$  such that  $\phi(\tilde{y}) < \phi(\tilde{x}) + \varepsilon$  for every  $\tilde{y} \in V$ . This is an immediate consequence of the fact that  $\phi$  is given by the infimum of a family of continuous functions. Then (Exercise 1.5.4),  $\phi$  admits some continuity point  $\tilde{a}$ . We are going to show that this point satisfies the conclusion of Theorem 1.5.1.

Let us begin by observing that  $\phi(\tilde{a}) = 0$ . Indeed, suppose that  $\phi(\tilde{a})$  is positive. Then, by continuity, there exist  $\beta > 0$  and a neighborhood  $V$  of  $\tilde{a}$  such that  $\phi(\tilde{y}) \geq \beta > 0$  for every  $\tilde{y} \in V$ . Then,

$$d(F^n(\tilde{y}), \tilde{y}) \geq \beta \quad \text{for every } \tilde{y} \in V \text{ and } n \geq 1. \tag{1.5.4}$$

On the other hand, according to (1.5.3), for every  $\tilde{x} \in \Delta_q$  there exists  $h \in \mathcal{H}$  such that  $\tilde{h}(\tilde{x}) \in V$ . Since the transformations  $h$  are uniformly continuous, we may fix  $\alpha > 0$  such that

$$d(\tilde{z}, \tilde{w}) < \alpha \implies d(\tilde{h}(\tilde{z}), \tilde{h}(\tilde{w})) < \beta \quad \text{for every } \tilde{h} \in \mathcal{H}. \tag{1.5.5}$$

By Lemma 1.5.6, there exists  $n \geq 1$  such that  $d(\tilde{x}, F^n(\tilde{x})) < \alpha$ . Then, using (1.5.5) and recalling that  $F$  commutes with every  $\tilde{h}$ ,

$$d(\tilde{h}(\tilde{x}), F^n(\tilde{h}(\tilde{x}))) < \beta.$$

This contradicts (1.5.4). This contradiction proves that  $\phi(\tilde{a}) = 0$ , as claimed.

In other words, there exists  $(n_k)_k \rightarrow \infty$  such that  $d(F^{n_k}(\tilde{a}), \tilde{a}) \rightarrow 0$  when  $k \rightarrow \infty$ . This means that (1.5.2) is satisfied and, hence, the proof of Theorem 1.5.1 is complete.

### 1.5.2 Exercises

- 1.5.1. Show, by means of examples, that the conclusion of Theorem 1.5.1 is generally false if the transformations  $f_i$  do not commute with each other.
- 1.5.2. Let  $\mathcal{G}$  be the abelian group generated by commuting homeomorphisms  $f_1, \dots, f_q : M \rightarrow M$  on a compact metric space. Prove that there exists some minimal element  $X \subset M$  for the inclusion relation in the family of non-empty, closed,  $\mathcal{G}$ -invariant subsets of  $M$ .
- 1.5.3. Show that if  $\varphi : M \rightarrow \mathbb{R}$  is an upper semi-continuous function on a compact metric space then  $\varphi$  attains its maximum, that is, there exists  $p \in M$  such that  $\varphi(p) \geq \varphi(x)$  for every  $x \in M$ .
- 1.5.4. Show that if  $\varphi : M \rightarrow \mathbb{R}$  is an (upper or lower) semi-continuous function on a compact metric space then the set of continuity points of  $\varphi$  contains a countable intersection of open and dense subsets of  $M$ . In particular, the set of continuity points is dense in  $M$ .
- 1.5.5. Let  $f : M \rightarrow M$  be a measurable transformation preserving a finite measure  $\mu$ . Given  $k \geq 1$  and a positive measure set  $A \subset M$ , show that for almost every  $x \in A$  there exists  $n \geq 1$  such that  $f^{jn}(x) \in A$  for every  $1 \leq j \leq k$ .
- 1.5.6. Let  $f_1, \dots, f_q : M \rightarrow M$  be commuting homeomorphisms on a compact metric space. A point  $x \in M$  is called *non-wandering* if for every neighborhood  $U$  of  $x$  there exist  $n_1, \dots, n_q \geq 1$  such that  $f_1^{n_1} \dots f_q^{n_q}(U)$  intersects  $U$ . The *non-wandering set* is the set  $\Omega(f_1, \dots, f_q)$  of all non-wandering points. Prove that  $\Omega(f_1, \dots, f_q)$  is non-empty and compact.