

8

Equivalent systems

This chapter is devoted to the *isomorphism problem* in ergodic theory: under what conditions should two systems (f, μ) and (g, ν) be considered “the same” and how does one decide, for given systems, whether they are in those conditions?

The fundamental notion is called *ergodic equivalence*: two systems are said to be ergodically equivalent if, restricted to subsets with full measure, the corresponding transformations are conjugated by some invertible map that preserves the invariant measures. Through such a map, properties of either system may be translated to corresponding properties of the other system.

Although this is a natural notion of isomorphism in the context of ergodic theory, it is not an easy one to handle. In general, the only way to prove that two given systems are equivalent is by exhibiting the equivalence map more or less explicitly. On the other hand, the most usual way to show that two systems are not equivalent is by finding some property that holds for one but not the other.

Thus, it is useful to consider a weaker notion, called *spectral equivalence*: two systems are spectrally equivalent if their Koopman operators are conjugated by some unitary operator. Two ergodically equivalent systems are always spectrally equivalent, but the converse is not true.

The idea of spectral equivalence leads to a rich family of invariants, related to the spectrum of the Koopman operator, that must have the same value for any two systems that are equivalent and, thus, may be used to exclude that possibility. Other invariants, of non-spectral nature, have an equally crucial role. The most important of all, the entropy, will be treated in Chapter 9.

The notions of ergodic equivalence and spectral equivalence, and the relations between them, are studied in Sections 8.1 and 8.2, respectively. In Sections 8.3 and 8.4 we study two classes of systems with opposite dynamical features: transformations with *discrete spectrum*, that include the ergodic translations on compact abelian groups, and transformations with a *Lebesgue spectrum*, which have the Bernoulli shifts as the main example.

These two classes of systems, as well as others that we introduced previously (ergodicity, strong mixing, weak mixing) are invariants of spectral equivalence

and, hence, also of ergodic equivalence. Finally, in Section 8.5 we discuss a third notion of equivalence, that we call *ergodic isomorphism*, especially in the context of Lebesgue spaces.

8.1 Ergodic equivalence

Let μ and ν be probability measures invariant under measurable transformations $f : M \rightarrow M$ and $g : N \rightarrow N$, respectively. We say that the systems (f, μ) and (g, ν) are *ergodically equivalent* if one can find measurable sets $X \subset M$ and $Y \subset N$ with $\mu(M \setminus X) = 0$ and $\nu(N \setminus Y) = 0$, and a measurable bijection $\phi : X \rightarrow Y$ with measurable inverse, such that

$$\phi_*\mu = \nu \quad \text{and} \quad \phi \circ f = g \circ \phi.$$

We leave it to the reader to check that this is indeed an equivalence relation, that is, reflexive, symmetric and transitive.

Observe also that the sets X and Y in the definition may be chosen to be invariant under f and g , respectively. Indeed, consider $X_0 = \bigcap_{n=0}^{\infty} f^{-n}(X)$. It is clear from the definition that $X_0 \subset X$ and $f(X_0) \subset X_0$. Since $\mu(X) = 1$ and the intersection is countable, we have that $\mu(X_0) = 1$. Analogously, $Y_0 = \bigcap_{m=0}^{\infty} g^{-m}(Y)$ is a measurable subset of Y such that $\nu(Y_0) = 1$ and $g(Y_0) \subset Y_0$. Moreover, by construction, $Y_0 = \phi(X_0)$. Therefore, the restriction of ϕ to X_0 is still a bijection onto Y_0 .

Example 8.1.1. Let $f : [0, 1] \rightarrow [0, 1]$ be defined by $f(x) = 10x - [10x]$. As we saw in Section 1.3.1, this transformation preserves the Lebesgue measure m on $[0, 1]$. If one represents each number $x \in [0, 1]$ by its decimal expansion $x = 0.a_0a_1a_2\dots$, the transformation f corresponds, simply, to shifting all the digits of x one unit to the left. That motivates us to consider:

$$\phi : \{0, 1, \dots, 9\}^{\mathbb{N}} \rightarrow [0, 1], \quad \phi((a_n)_n) = \sum_{n=0}^{\infty} \frac{a_n}{10^{n+1}} = 0.a_0a_1a_2\dots$$

It is clear that ϕ is surjective. On the other hand, it is not injective, since certain real numbers have more than one decimal expansion: for example, $0.1000000\dots = 0.099999\dots$. Actually, this happens if and only the number admits a finite decimal expansion, that is, such that all but finitely many digits are zero. The set of such numbers is countable and, hence, is irrelevant from the point of view of the Lebesgue measure. More precisely, let us consider the set $X \subset \{0, 1, \dots, 9\}^{\mathbb{N}}$ of all sequences with an infinite number of symbols different from zero and the set $Y \subset [0, 1]$ of all numbers whose decimal expansion is infinite (hence, unique). Then the restriction of ϕ to X is a bijection onto Y .

It is easy to check that both $\phi|_X$ and its inverse are measurable: use the fact that the image of the intersection of X with each cylinder $[0; a_0, \dots, a_{m-1}]$ is

the intersection of Y with some interval of length 10^{-m} . This observation also shows that $\phi_*\nu = m$, where ν denotes the Bernoulli measure on $\{0, 1, \dots, 9\}^{\mathbb{N}}$ that assigns equal weights to all the digits. Moreover, denoting by σ the shift map in $\{0, 1, \dots, 9\}^{\mathbb{N}}$, we have that

$$\phi \circ \sigma((a_n)_n) = 0, a_1 a_2 \dots a_n \dots = f \circ \phi((a_n)_n)$$

for every $(a_n)_n \in X$. This proves that (f, m) is ergodically equivalent to the Bernoulli shift (σ, ν) .

Suppose that (f, μ) and (g, ν) are ergodically equivalent. A measurable set $A \subset M$ is invariant under $f : M \rightarrow M$ if and only if $\phi(A)$ is invariant under $g : N \rightarrow N$. Moreover, $\nu(\phi(A)) = \mu(A)$. Therefore, (f, μ) is ergodic if and only if (g, ν) is ergodic. It is just as easy to obtain similar conclusions for the mixing and the weak mixing properties. Indeed, essentially all the properties that we study in this book are *invariants of ergodic equivalence*, that is, if they hold for a given system then they also hold for any system that is ergodically equivalent to it. An exception is unique ergodicity, which is a property of a different nature, since it concerns solely the transformation.

This also means that these properties may be used to try to distinguish systems that are not ergodically equivalent. Still, that is usually a difficult task. For example, nothing of what was said so far is of much help towards answering the following question: are the shift maps

$$\sigma : \{1, 2\}^{\mathbb{Z}} \rightarrow \{1, 2\}^{\mathbb{Z}} \quad \text{and} \quad \zeta : \{1, 2, 3\}^{\mathbb{Z}} \rightarrow \{1, 2, 3\}^{\mathbb{Z}}, \tag{8.1.1}$$

endowed with the corresponding Bernoulli measures giving the same weights to all the symbols, ergodically equivalent? It is easy to see that σ and ζ are not *topologically conjugate* (for example: ζ has three fixed points, whereas σ has only two), but the existence of an ergodic equivalence is a much more delicate issue. In fact, this type of question motivates most of the content of the present chapter and also leads to the notion of entropy, which is the subject of Chapter 9.

Example 8.1.2. Let $\sigma : M \rightarrow M$ be the shift map in $M = X^{\mathbb{N}}$ and let $\mu = \nu^{\mathbb{N}}$ be a Bernoulli measure. Let $\hat{\sigma} : \hat{M} \rightarrow \hat{M}$ be the natural extension of σ and $\hat{\mu}$ be the lift of μ (Section 2.4.2). Moreover, let $\tilde{\sigma} : \tilde{M} \rightarrow \tilde{M}$ be the shift map in $\tilde{M} = X^{\mathbb{Z}}$ and $\tilde{\mu} = \nu^{\mathbb{Z}}$ be the corresponding Bernoulli measure. Then, $(\hat{\sigma}, \hat{\mu})$ is ergodically equivalent to $(\tilde{\sigma}, \tilde{\mu})$. An equivalence may be constructed as follows.

By definition, \hat{M} is the space of pre-orbits of $\sigma : M \rightarrow M$, that is, of all the sequences $\hat{x} = (\dots, x_{-n}, \dots, x_0)$ in M such that $\sigma(x_{-j}) = x_{-j+1}$ for every $j \geq 1$. Moreover, each x_{-j} is a sequence $(x_{-j,i})_{i \in \mathbb{N}}$ in X . So, the previous relation means that

$$x_{-j,i+1} = x_{-j+1,i} \quad \text{for every } i, j \in \mathbb{N}. \tag{8.1.2}$$

Consider the map $\phi : \hat{M} \rightarrow \tilde{M}$, $\hat{x} \mapsto \tilde{x}$ given by

$$\tilde{x}_n = x_{0,n} = x_{-1,n+1} = \dots \quad \text{and} \quad \tilde{x}_{-n} = x_{-n,0} = x_{-n-1,1} = \dots$$

We leave it to the reader to check that ϕ is indeed an ergodic equivalence between the natural extension $(\hat{\sigma}, \hat{\mu})$ and the two-sided shift map $(\tilde{\sigma}, \tilde{\mu})$.

8.1.1 Exercises

- 8.1.1. Let $f : [0, 1] \rightarrow [0, 1]$ be the transformation defined by $f(x) = 2x - [2x]$ and m be the Lebesgue measure on $[0, 1]$. Exhibit a map $g : [0, 1] \rightarrow [0, 1]$ and a probability measure ν invariant under g such that (g, ν) is ergodically equivalent to (f, μ) and the support of ν has empty interior.
- 8.1.2. Let $f : \{1, \dots, k\}^{\mathbb{N}} \rightarrow \{1, \dots, k\}^{\mathbb{N}}$ and $g : \{1, \dots, l\}^{\mathbb{N}} \rightarrow \{1, \dots, l\}^{\mathbb{N}}$ be one-sided shift maps, endowed with Bernoulli measures μ and ν , respectively. Show that, for every set $X \subset \{1, \dots, k\}^{\mathbb{N}}$ with $f^{-1}(X) = X$ and $\mu(X) = 1$, there exists $x \in X$ such that $\#(X \cap f^{-1}(x)) = k$. Conclude that if $k \neq l$ then (f, μ) and (g, ν) cannot be ergodically equivalent.
- 8.1.3. Let $X = \{1, \dots, d\}$ and consider the shift map $\sigma : X^{\mathbb{N}} \rightarrow X^{\mathbb{N}}$ endowed with a Markov measure μ . Given any cylinder $C = [0; c_0, \dots, c_l]$ in $X^{\mathbb{N}}$, let μ_C be the normalized restriction of μ to C . Show that there exists an induced transformation $\sigma_C : C \rightarrow C$ (see Section 1.4.2) preserving μ_C and such that (σ_C, μ_C) is ergodically equivalent to a Bernoulli shift $(\sigma_{\mathbb{N}}, \nu)$ in $\mathbb{N}^{\mathbb{N}}$.

8.2 Spectral equivalence

Let $f : M \rightarrow M$ and $g : N \rightarrow N$ be transformations preserving probability measures μ and ν , respectively. Let $U_f : L^2(\mu) \rightarrow L^2(\mu)$ and $U_g : L^2(\nu) \rightarrow L^2(\nu)$ be the corresponding Koopman operators. We say that (f, μ) and (g, ν) are *spectrally equivalent* if there exists some unitary operator $L : L^2(\mu) \rightarrow L^2(\nu)$ such that

$$U_g \circ L = L \circ U_f. \tag{8.2.1}$$

We leave it to the reader to check that the relation defined in this way is, indeed, an equivalence relation.

It is easy to see that if two systems are ergodically equivalent then they are spectrally equivalent. Indeed, suppose that there exists an invertible map $h : M \rightarrow N$ such that $\phi_*\mu = \nu$ and $\phi \circ f = g \circ \phi$. Then, the Koopman operator

$$U_\phi : L^2(\nu) \rightarrow L^2(\mu), \quad U_\phi(\psi) = \psi \circ \phi$$

is an isometry and is invertible: the inverse is the Koopman operator associated with ϕ^{-1} . In other words, U_ϕ is a unitary operator. Moreover,

$$U_f \circ U_\phi = U_{\phi \circ f} = U_{g \circ \phi} = U_\phi \circ U_g.$$

Therefore, $L = U_\phi$ is a spectral equivalence between the two systems.

The converse is false, as will be clear from the sequel. For example, all countably generated two-sided Bernoulli shifts are spectrally equivalent (Corollary 8.4.12); however, not all have the same entropy (Example 9.1.10) and so not all are ergodically equivalent.

8.2.1 Invariants of spectral equivalence

Recall that the spectrum $\text{spec}(A)$ of a bounded linear operator $A : E \rightarrow E$ in a complex Banach space E consists of the complex numbers λ such that $A - \lambda \text{id}$ is not invertible. We say that $\lambda \in \text{spec}(A)$ is an *eigenvalue* if $A - \lambda \text{id}$ is not injective, that is, if there exists $v \neq 0$ such that $Av = \lambda v$. Then, the dimension of the kernel of $A - \lambda \text{id}$ is called the *multiplicity* of the eigenvalue.

By definition, the *spectrum* of a system (f, μ) is the spectrum of the corresponding Koopman operator $U_f : L^2(\mu) \rightarrow L^2(\mu)$. If (f, μ) is spectrally equivalent to (g, ν) then the two systems have the same spectrum: the relation (8.2.1) implies that

$$(U_g - \lambda \text{id}) = L \circ (U_f - \lambda \text{id}) \circ L^{-1} \tag{8.2.2}$$

and, consequently, $U_g - \lambda \text{id}$ is invertible if and only if $U_f - \lambda \text{id}$ is invertible.

In fact, the spectrum itself is a poor invariant: in particular, all the invertible ergodic systems with no atoms have the same spectrum (Exercise 8.2.1). However, the associated spectral measure does provide very useful invariants, as we are going to see. The simplest one is the set of atoms of the spectral measure, that is, the set of eigenvalues of the Koopman operator. Note that (8.2.2) also shows that a given λ is an eigenvalue of U_f if and only if it is an eigenvalue of U_g ; besides, in that case the two multiplicities are equal.

Observe that 1 is always an eigenvalue of the Koopman operator, since $U_f \varphi = \varphi$ for every constant function φ . By Proposition 4.1.3(v), the system (f, μ) is ergodic if and only if this eigenvalue has multiplicity 1 for U_f . Thus, it follows from what we have just said that (f, μ) is ergodic if and only if any system (g, ν) spectrally equivalent to it is ergodic. In other words, ergodicity is an *invariant of spectral equivalence*.

Analogously, suppose that the system (f, μ) is mixing. Then, by Proposition 7.1.12,

$$\lim_n U_f^n \varphi \cdot \psi = \int \varphi d\mu \int \psi d\mu$$

for every $\varphi, \psi \in L^2(\mu)$. Now suppose that (g, ν) is spectrally equivalent to (f, μ) . Let L be the unitary operator in (8.2.1). The inverse L^{-1} maps eigenvectors of U_g associated with the eigenvalue 1 to eigenvectors of U_f associated with the same eigenvalue 1. Since the two systems are ergodic (use the previous paragraph), this means that L^{-1} maps constant functions to constant functions. Since L^{-1} is unitary,

$$U_g^n \varphi \cdot \psi = L^{-1}(U_g^n \varphi) \cdot L^{-1} \psi = U_f^n(L^{-1} \varphi) \cdot L^{-1} \psi$$

and, hence, $\lim_n U_g^n \varphi \cdot \psi = \int L^{-1} \varphi d\mu \int L^{-1} \psi d\mu$ for every $\varphi, \psi \in L^2(v)$. Also,

$$\int L^{-1} \varphi d\mu = L^{-1} \varphi \cdot 1 = L^{-1} \varphi \cdot L^{-1} 1 = \varphi \cdot 1 = \int \varphi dv$$

and, analogously, $\int L^{-1} \psi d\mu = \int \psi d\mu$. In this way, we have shown that

$$\lim_n U_g^n \varphi \cdot \psi = \int \varphi d\mu \int \psi d\mu,$$

for every $\varphi, \psi \in L^2(v)$, that is, (g, v) is also mixing. This shows that the mixing property is an invariant of spectral equivalence.

The same argument may be used for the weak mixing property, though the theorem that we prove in Section 8.2.2 below gives us a more interesting proof of the fact that weak mixing is an invariant of spectral equivalence.

8.2.2 Eigenvalues and weak mixing

As we have seen, the Koopman operator $U_f : L^2(\mu) \rightarrow L^2(\mu)$ of a system (f, μ) is an isometry, that is, it satisfies $U_f^* U_f = \text{id}$. If f is invertible then the Koopman operator is unitary, that is, it satisfies $U_f^* U_f = U_f U_f^* = \text{id}$. In particular, in that case U_f is a normal operator. Then the property of weak mixing admits the following interesting characterization:

Theorem 8.2.1. *An invertible system (f, μ) is weak mixing if and only if the constant functions are the only eigenvectors of the Koopman operator.*

In particular, a system (f, μ) is weak mixing if and only if it is ergodic and 1 is the unique eigenvalue of U_f .

Proof. Suppose that (f, μ) is weak mixing. Let $\varphi \in L^2(\mu)$ be any (non-zero) eigenfunction of U_f and λ be the corresponding eigenvalue. Then

$$\int \varphi d\mu = \int U_f \varphi d\mu = \lambda \int \varphi d\mu,$$

and this implies that $\int \varphi d\mu = 0$ or $\lambda = 1$. In the first case,

$$C_j(\varphi, \bar{\varphi}) = \left| \int (U_f^j \varphi) \bar{\varphi} d\mu \right| = \left| \lambda^j \int \varphi \bar{\varphi} d\mu \right| = \int |\varphi|^2 d\mu$$

for every $j \geq 1$ (recall that $|\lambda| = 1$). But then

$$\lim_n \frac{1}{n} \sum_{j=0}^{n-1} C_j(\varphi, \bar{\varphi}) = \int |\varphi|^2 d\mu > 0,$$

contradicting the hypothesis that the system is weak mixing. In the second case, using that the system is ergodic, we find that φ is constant at μ -almost every point. This shows that if the system is weak mixing then the constant functions are the only eigenvectors.

Now suppose that the only eigenvectors of U_f are the constant functions. To conclude that (f, μ) is weak mixing, we must show that

$$\frac{1}{n} \sum_{j=0}^{n-1} C_j(\varphi, \psi)^2 \rightarrow 0 \quad \text{for any } \varphi, \psi \in L^2(\mu)$$

(recall Exercise 7.1.2). It follows immediately from the definition that

$$C_j(\varphi, \psi) = C_j(\varphi', \psi) \quad \text{where } \varphi' = \varphi - \int \varphi \, d\mu$$

and the integral of φ' vanishes. Hence, it is no restriction to suppose that $\int \varphi \, d\mu = 0$. Then, using the relation (A.7.6) for the unitary operator $L = U_f$, we get:

$$C_j(\varphi, \psi)^2 = \left| \int (U_f^j \varphi) \psi \, d\mu \right|^2 = \left| \int_{\mathbb{C}} z^j \, d\theta(z) \right|^2,$$

where $\theta = E\varphi \cdot \psi$. The expression on the right-hand side may be rewritten as follows:

$$\int_{\mathbb{C}} z^j \, d\theta(z) \int_{\mathbb{C}} \bar{z}^j \, d\bar{\theta}(z) = \int_{\mathbb{C}} \int_{\mathbb{C}} z^j \bar{w}^j \, d\theta(z) \, d\bar{\theta}(w).$$

Therefore, given any $n \geq 1$,

$$\frac{1}{n} \sum_{j=0}^{n-1} C_j(\varphi, \psi)^2 = \int_{\mathbb{C}} \int_{\mathbb{C}} \frac{1}{n} \sum_{j=0}^{n-1} (z\bar{w})^j \, d\theta(z) \, d\bar{\theta}(w). \tag{8.2.3}$$

We claim that the measure $\theta = E\varphi \cdot \psi$ is non-atomic. In fact, suppose that there exists $\lambda \in \mathbb{C}$ such that $\theta(\{\lambda\}) \neq 0$. Then, $E(\{\lambda\}) \neq 0$ and then we may use Proposition A.7.8 to conclude that the function $E(\{\lambda\})\varphi$ is an eigenvector of U_f . By the hypothesis about the operator U_f , this implies that $E(\{\lambda\})\varphi$ is constant at μ -almost every point. Hence,

$$E(\{\lambda\})\varphi \cdot \varphi = E(\{\lambda\})\varphi \int \bar{\varphi} \, d\mu = 0.$$

Lemma A.7.3 also gives that

$$E(\{\lambda\})\varphi \cdot \varphi = E(\{\lambda\})^2 \varphi \cdot \varphi = E(\{\lambda\})\varphi \cdot E(\{\lambda\})\varphi.$$

Putting these two identities together, we conclude that $E(\{\lambda\})\varphi = 0$, which contradicts the hypothesis. Thus, our claim is proved.

The sequence $n^{-1} \sum_{j=0}^{n-1} (z\bar{w})^j$ in (8.2.3) is bounded and (see Exercise 8.2.6) converges to zero on the complement of the diagonal $\Delta = \{(z, w) : z = w\}$. Moreover, the diagonal has measure zero:

$$(\theta \times \bar{\theta})(\Delta) = \int \theta(\{y\}) \, d\bar{\theta}(y) = 0,$$

because θ is non-atomic. Then we may use the monotone convergence theorem to conclude that (8.2.3) converges to zero when $n \rightarrow \infty$. This proves that (f, μ) is weak mixing if U_f has no non-constant eigenvectors.

Suppose that M is a topological space. We say that a continuous map $f : M \rightarrow M$ is *topologically weak mixing* if the Koopman operator U_f has no non-constant *continuous* eigenfunctions. The following fact is an easy consequence of Theorem 8.2.1:

Corollary 8.2.2. *If (f, μ) is weak mixing then the restriction of f to the support of μ is topologically weak mixing.*

Proof. Let φ be a continuous eigenfunction of U_f . By Theorem 8.2.1, the function φ is constant at μ -almost every point. Hence, by continuity, φ is constant (at every point) on the support of μ .

We mentioned in Section 7.3 that almost every interval exchange is weak mixing but not mixing. In the following we describe an explicit construction, based on an extension of ideas that were hinted at in Example 6.3.9. The reader may find this and other variations of those ideas in Section 7.4 of Kalikow and McCutcheon [KM10].

Example 8.2.3 (Chacon). Consider the sequence $(\mathcal{S}_n)_n$ of piles defined as follows. First, $\mathcal{S}_1 = \{[0, 2/3]\}$. Next, for each $n \geq 1$, let \mathcal{S}_n be the pile obtained by dividing \mathcal{S}_{n-1} into three columns, with the same width, and piling those columns up on top of each other, with an additional interval inserted between the second pile and the third one, as illustrated in Figure 8.1.

$$\begin{aligned} \text{For example, } \mathcal{S}_2 &= \{[0, 2/9], [2/9, 4/9], [6/9, 8/9], [4/9, 6/9]\} \text{ and} \\ \mathcal{S}_3 &= \{[0, 2/27], [6/27, 8/27], [18/27, 20/27], [12/27, 14/27], [2/27, 4/27], \\ &\quad [8/27, 10/27], [20/27, 22/27], [14/27, 16/27], [24/27, 26/27], \\ &\quad [4/27, 6/27], [10/27, 12/27], [22/27, 24/27], [16/27, 18/27]\}. \end{aligned}$$

Note that each \mathcal{S}_n is a pile in the interval $J_n = [0, 1 - 3^{-n})$. The sequence $(f_n)_n$ of transformations associated with such piles converges at every point to a transformation $f : [0, 1) \rightarrow [0, 1)$ that preserves the Lebesgue measure m . This system (f, m) is weak mixing but not mixing (Exercise 8.2.7).

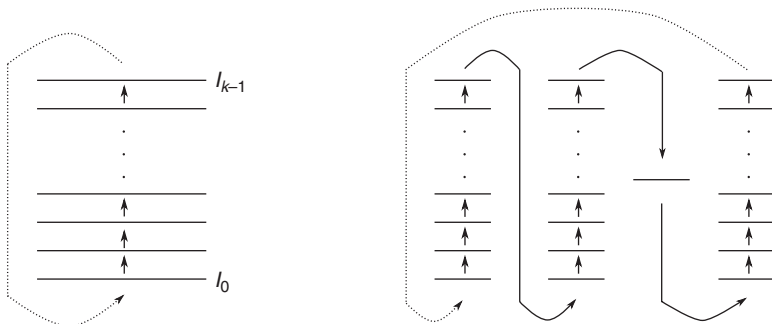


Figure 8.1. Constructing a weak mixing system that is not mixing

8.2.3 Exercises

- 8.2.1. Let (f, μ) be an invertible ergodic system with no atoms. Show that every λ in the unit circle $\{z \in \mathbb{C} : |z| = 1\}$ is an *approximate eigenvalue* of the Koopman operator $U_f : L^2(\mu) \rightarrow L^2(\mu)$: there exists some sequence $(\varphi_n)_n$ such that $\|\varphi_n\| \rightarrow 1$ and $\|U_f \varphi_n - \lambda \varphi_n\| \rightarrow 0$. In particular, the spectrum of U_f coincides with the unit circle.
- 8.2.2. Let m be the Lebesgue measure on the circle and $U_\alpha : L^2(m) \rightarrow L^2(m)$ be the Koopman operator of the irrational rotation $R_\alpha : S^1 \rightarrow S^1$. Calculate the eigenvalues of U_α and deduce that (R_α, m) and (R_β, m) are spectrally equivalent if and only if $\alpha = \pm\beta$. [Observation: Corollary 8.3.6 provides a more complete statement.]
- 8.2.3. Let m be the Lebesgue measure on the circle and, for each integer number $k \geq 2$, let $U_k : L^2(m) \rightarrow L^2(m)$ be the Koopman operator of the transformation $f_k : S^1 \rightarrow S^1$ given by $f_k(x) = kx \pmod{\mathbb{Z}}$. Check that if $p \neq q$ then (f_p, m) and (f_q, m) are not ergodically equivalent. Show that, for any $k \geq 2$,

$$L^2(m) = \{\text{constants}\} \oplus \bigoplus_{j=0}^{\infty} U_k^j(H_k),$$

where $H_k = \{\sum_{n \in \mathbb{Z}} a_n e^{2\pi i n x} : a_n = 0 \text{ if } k \mid n\}$ and the terms in the direct sum are pairwise orthogonal. Conclude that (f_p, m) and (f_q, m) are spectrally equivalent for any p and q .

- 8.2.4. Let $f : S^1 \rightarrow S^1$ be the transformation given by $f(x) = kx \pmod{\mathbb{Z}}$ and μ be the Lebesgue measure. Show that (f, μ) is weak mixing if and only if $|k| \geq 2$.
- 8.2.5. Prove that, for any invertible transformation f , if μ is ergodic for every iterate f^n and there exists $C > 0$ such that

$$\limsup_n \mu(f^{-n}(A) \cap B) \leq C \mu(A) \mu(B),$$

for any measurable sets A and B , then μ is weak mixing. [Observation: This statement is due to Ornstein [Orn72]. In fact, he proved more: under these hypotheses the system is (strongly) mixing.]

- 8.2.6. Let z and w be two complex numbers with absolute value 1. Check that

- (a) $\lim_n \frac{1}{n} \sum_{j=0}^{n-1} |z^j - 1| = 0$ if and only if $z = 1$;
- (b) $\lim_n \frac{1}{n} \sum_{j=0}^{n-1} (z\bar{w})^j = 0$ if $z \neq w$.

- 8.2.7. Consider the system (f, m) in Example 8.2.3. Show that

- (a) the system (f, m) is ergodic;
- (b) the only eigenvalues of the Koopman operator $U_f : L^1(m) \rightarrow L^1(m)$ are the constant functions, and hence (f, m) is weak mixing;
- (c) $\limsup_n m(f^n(A) \cap A) \geq 2/7$ if we take $A = [0, 2/9)$; in particular, (f, m) is not mixing.

8.3 Discrete spectrum

In this section and the next we study two extreme cases, in terms of the type of spectral measure of the Koopman operator: systems with *discrete spectrum*, whose spectral measure is purely atomic, and systems with *Lebesgue spectrum*, whose spectral measure is equivalent to the Lebesgue measure on the unit circle.

We begin by describing some properties of the eigenvalues and eigenvectors of the Koopman operator. It is clear that all the eigenvalues are in the unit circle, since U_f is an isometry.

Proposition 8.3.1. *If $\varphi_1, \varphi_2 \in L^2(\mu)$ satisfy $U_f\varphi_1 = \lambda_1\varphi_1$ and $U_f\varphi_2 = \lambda_2\varphi_2$ with $\lambda_1 \neq \lambda_2$, then $\varphi_1 \cdot \varphi_2 = 0$. Moreover, the eigenvalues of U_f form a subgroup of the unit circle.*

If the system (f, μ) is ergodic then every eigenvalue of U_f is simple and the absolute value of every eigenfunction is constant at μ -almost every point.

Proof. The first claim follows from the identity

$$\varphi_1 \cdot \varphi_2 = U_f\varphi_1 \cdot U_f\varphi_2 = \lambda_1\varphi_1 \cdot \lambda_2\varphi_2 = \lambda_1\bar{\lambda}_2(\varphi_1 \cdot \varphi_2) = \lambda_1\lambda_2^{-1}(\varphi_1 \cdot \varphi_2),$$

since $\lambda_1\lambda_2^{-1} \neq 1$. This identity also shows that the set of all eigenvalues is closed under the operation $(\lambda_1, \lambda_2) \mapsto \lambda_1\lambda_2^{-1}$. Recalling that 1 is always an eigenvalue, it follows that this set is a multiplicative group.

Now assume that (f, μ) is ergodic and suppose that $U_f\varphi = \lambda\varphi$. Then $U_f(|\varphi|) = |U_f\varphi| = |\lambda\varphi| = |\varphi|$ at μ -almost every point. By ergodicity, this implies that $|\varphi|$ is constant at μ -almost every point. Next, suppose that $U_f\varphi_1 = \lambda\varphi_1$, $U_f\varphi_2 = \lambda\varphi_2$ and the functions φ_1 and φ_2 are not identically zero. Since $|\varphi_2|$ is constant at μ -almost every point, $\varphi_2(x) \neq 0$ for μ -almost every x . Then φ_1/φ_2 is well defined. Moreover,

$$U_f\left(\frac{\varphi_1}{\varphi_2}\right) = \frac{U_f(\varphi_1)}{U_f(\varphi_2)} = \frac{\lambda\varphi_1}{\lambda\varphi_2} = \frac{\varphi_1}{\varphi_2}.$$

By ergodicity, it follows that the quotient is constant at μ -almost every point. That is, $\varphi_1 = c\varphi_2$ for some $c \in \mathbb{C}$.

We say that a system (f, μ) has *discrete spectrum* if the eigenvectors of the Koopman operator $U_f : L^2(\mu) \rightarrow L^2(\mu)$ generate the Hilbert space $L^2(\mu)$. Observe that this implies that U_f is invertible and, hence, is a unitary operator. This terminology is justified by the following observation (recall also Theorem A.7.9):

Proposition 8.3.2. *A system (f, μ) has discrete spectrum if and only if its Koopman operator U_f has a spectral representation of the form*

$$T : \bigoplus_j L^2(\sigma_j)^{\chi_j} \rightarrow \bigoplus_j L^2(\sigma_j)^{\chi_j}, \quad (\varphi_{j,l})_{j,l} \mapsto (z \mapsto z\varphi_{j,l}(z))_{j,l}, \quad (8.3.1)$$

where each σ_j is a Dirac measure at a point in the unit circle.

Proof. Suppose that U_f admits a spectral representation of the form (8.3.1) with $\sigma_j = \delta_{\lambda_j}$ for some λ_j in the unit circle. Each $L^2(\sigma_j)^{\chi_j}$ may be canonically identified with a subspace in the direct sum. The restriction of T to that subspace coincides with $\lambda_j \text{id}$, since

$$z\varphi_{j,l}(z) = \lambda_j \varphi_{j,l}(z) \quad \text{at } \sigma_j\text{-almost every point.} \tag{8.3.2}$$

Let $(v_{j,l})_l$ be a Hilbert basis of $L^2(\sigma_j)^{\chi_j}$. Then $(v_{j,l})_{j,l}$ is a Hilbert basis of the direct sum formed by eigenvectors of T . Since T is unitarily conjugate to U_f , it follows that $L^2(\mu)$ admits a Hilbert basis formed by eigenvectors of the Koopman operator.

Now suppose that (f, μ) has discrete spectrum. Let $(\lambda_j)_j$ be the eigenvalues of U_f and, for each j , let $\sigma_j = \delta_{\lambda_j}$ and χ_j be the Hilbert dimension of the eigenspace $\ker(U_f - \lambda_j \text{id})$. Note that the space $L^2(\sigma_j)$ is 1-dimensional, since every function is constant at σ_j -almost every point. Therefore, the Hilbert dimension of $L^2(\delta_{\lambda_j})^{\chi_j}$ is also equal to χ_j . Hence, there exists some unitary isomorphism

$$L_j : \ker(U_f - \lambda_j \text{id}) \rightarrow L^2(\delta_{\lambda_j})^{\chi_j}.$$

It is clear that $L_j \circ U_f \circ L_j^{-1} = \lambda_j \text{id}$. In other words, recalling the observation (8.3.2),

$$L_j \circ U_f \circ L_j^{-1} : (\varphi_{j,l})_l \mapsto (z \mapsto \lambda_j \varphi_{j,l}(z))_l = (z \mapsto z\varphi_{j,l}(z))_l. \tag{8.3.3}$$

The eigenspaces $\ker(U_f - \lambda_j \text{id})$ generate $L^2(\mu)$, by hypothesis, and they are pairwise orthogonal, by Proposition 8.3.1. Hence, we may combine the operators L_j to obtain a unitary isomorphism $L : L^2(\mu) \rightarrow \bigoplus_j L^2(\sigma_j)^{\chi_j}$. The relation (8.3.3) gives that

$$L \circ U_f \circ L^{-1} : (\varphi_{j,l})_{j,l} \mapsto (z \mapsto z\varphi_{j,l}(z))_l$$

is a spectral representation of U_f of the form we are looking for.

Example 8.3.3. Let m be the Lebesgue measure on the torus \mathbb{T}^d . Consider the Fourier basis $\{\phi_k(x) = e^{2\pi i k \cdot x} : k \in \mathbb{Z}^d\}$ of the Hilbert space $L^2(m)$. Let f be the rotation $R_\theta : \mathbb{T}^d \rightarrow \mathbb{T}^d$ corresponding to a given vector $\theta = (\theta_1, \dots, \theta_d)$. Then,

$$U_f \phi_k(x) = \phi_k(x + \theta) = e^{2\pi i k \cdot \theta} \phi_k(x) \quad \text{for every } x \in \mathbb{T}^d.$$

This shows that every ϕ_k is an eigenvector of U_f and, hence, (f, m) has discrete spectrum. Note that the group of eigenvalues is

$$G_\theta = \{e^{2\pi i k \cdot \theta} : k \in \mathbb{Z}^d\}, \tag{8.3.4}$$

that is, the subgroup of the unit circle generated by $\{e^{2\pi i \theta_j} : j = 1, \dots, d\}$.

More generally, every ergodic translation in a compact abelian group has discrete spectrum. Conversely, every ergodic system with discrete spectrum is ergodically isomorphic to a translation in a compact abelian group (the notion of ergodic isomorphism is discussed in Section 8.5). Another interesting result

is that every subgroup of the unit circle is the group of eigenvalues of some ergodic system with discrete spectrum. These facts are proved in Section 3.3 of the book of Peter Walters [Wal82].

Proposition 8.3.4. *Suppose that (f, μ) and (g, ν) are ergodic and have discrete spectrum. Then (f, μ) and (g, ν) are spectrally equivalent if and only if their Koopman operators $U_f : L^2(\mu) \rightarrow L^2(\mu)$ and $U_g : L^2(\nu) \rightarrow L^2(\nu)$ have the same eigenvalues.*

Proof. It is clear that if the Koopman operators are conjugate then they have the same eigenvalues. To prove the converse, let $(\lambda_j)_j$ be the eigenvalues of the two operators. By Proposition 8.3.2, the eigenvalues are simple. For each j , let u_j and v_j be unit vectors in $\ker(U_f - \lambda_j \text{id})$ and $\ker(U_g - \lambda_j \text{id})$, respectively. Then $(u_j)_j$ and $(v_j)_j$ are Hilbert bases of $L^2(\mu)$ and $L^2(\nu)$, respectively. Consider the isomorphism $L : L^2(\mu) \rightarrow L^2(\nu)$ defined by $L(u_j) = v_j$. This operator is unitary, since it maps a Hilbert basis to a Hilbert basis, and it satisfies

$$L \circ U_f(u_j) = L(\lambda_j u_j) = \lambda_j v_j = U_g(v_j) = U_g \circ L(u_j)$$

for every j . By linearity, it follows that $L \circ U_f = U_g \circ L$. Therefore, (f, μ) and (g, ν) are spectrally equivalent.

Corollary 8.3.5. *If (f, μ) is ergodic, invertible and has discrete spectrum then (f, μ) is spectrally equivalent to (f^{-1}, μ) .*

Proof. It is clear that λ is an eigenvalue of U_f if and only if λ^{-1} is an eigenvalue of $U_{f^{-1}}$; moreover, in that case the eigenvectors are the same. Since the sets of eigenvalues are groups, it follows that the two operators have the same eigenvalues and the same eigenvectors. Apply Proposition 8.3.4.

Let m be the Lebesgue measure on the torus \mathbb{T}^d . Proposition 8.3.4 also allows us to classify the irrational rotations on the torus up to equivalence, ergodic and spectral:

Corollary 8.3.6. *Let $\theta = (\theta_1, \dots, \theta_d)$ and $\tau = (\tau_1, \dots, \tau_d)$ be rationally independent vectors and R_θ and R_τ be the corresponding rotations on the torus \mathbb{T}^d . The following conditions are equivalent:*

- (i) (R_θ, m) and (R_τ, m) are ergodically equivalent;
- (ii) (R_θ, m) and (R_τ, m) are spectrally equivalent;
- (iii) there exists $L \in \text{SL}(d, \mathbb{Z})$ such that $\theta = L\tau \pmod{\mathbb{Z}^d}$.

We leave the proof to the reader (Exercise 8.3.2). In the special case of the circle, we get that two irrational rotations R_θ and R_τ are equivalent if and only if either $R_\theta = R_\tau$ or $R_\theta = R_\tau^{-1}$. See also Exercise 8.3.3.

8.3.1 Exercises

- 8.3.1. Suppose that (f, μ) has discrete spectrum and the Hilbert space $L^2(\mu)$ is separable (this is the case, for instance, if the σ -algebra of measurable sets is countably generated). Show that there exists a sequence $(n_k)_k$ converging to infinity such that $\|U_f^{n_k} \varphi - \varphi\|_2$ converges to zero when $k \rightarrow \infty$, for every $\varphi \in L^2(\mu)$.
- 8.3.2. Prove Corollary 8.3.6.
- 8.3.3. Let m be the Lebesgue measure on S^1 and $\theta = p/q$ and $\tau = r/s$ be two rational numbers, with $\gcd(p, q) = 1 = \gcd(r, s)$. Show that the rotations (R_θ, m) and (R_τ, m) are ergodically equivalent if and only if the denominators q and s are equal.

8.4 Lebesgue spectrum

This section is devoted to the class of systems whose Koopman operator has the following property (the reason for the terminology will become clear in Proposition 8.4.10):

Definition 8.4.1. Let $U : H \rightarrow H$ be an isometry in a Hilbert space. We say that U has *Lebesgue spectrum* if there exists some closed subspace $E \subset H$ such that

- (i) $U(E) \subset E$;
- (ii) $\bigcap_{n \in \mathbb{N}} U^n(E) = \{0\}$;
- (iii) $\sum_{n \in \mathbb{N}} U^{-n}(E) = H$.

Given a probability measure μ , we denote by $L_0^2(\mu) = L_0^2(M, \mathcal{B}, \mu)$ the orthogonal complement, inside the space $L^2(\mu) = L^2(M, \mathcal{B}, \mu)$, of the subspace of constant functions. In other words,

$$L_0^2(\mu) = \{\varphi \in L^2(\mu) : \int \varphi d\mu = 0\}.$$

Note that $L_0^2(\mu)$ is invariant under the Koopman operator: $\varphi \in L_0^2(\mu)$ if and only if $U_f \varphi \in L_0^2(\mu)$. We say that the system (f, μ) has *Lebesgue spectrum* if the restriction of the Koopman operator to $L_0^2(\mu)$ has Lebesgue spectrum.

8.4.1 Examples and properties

We start by observing that all Bernoulli shifts have Lebesgue spectrum. It is convenient to treat one-sided shifts and the two-sided shifts separately.

Example 8.4.2. Consider a one-sided shift map $\sigma : X^{\mathbb{N}} \rightarrow X^{\mathbb{N}}$ and a Bernoulli measure $\mu = \nu^{\mathbb{N}}$ on $X^{\mathbb{N}}$. Let $E = L_0^2(\mu)$. Conditions (i) and (iii) in Definition 8.4.1 are obvious. To prove condition (ii), consider any function $\varphi \in L_0^2(\mu)$ in the intersection, that is, such that for every $n \in \mathbb{N}$ there exists a

function $\psi_n \in L_0^2(\mu)$ satisfying $\varphi = \psi_n \circ \sigma^n$. We want to show that φ is constant at μ -almost every point. For each $c \in \mathbb{R}$, consider

$$A_c = \{x \in X^{\mathbb{N}} : \varphi(x) > c\}.$$

For each $n \in \mathbb{N}$, we may write $A_c = \sigma^{-n}(\{x \in X^{\mathbb{N}} : \psi_n(x) > c\})$. Then A_c belongs to the σ -algebra generated by the cylinders of the form $[n; C_n, \dots, C_m]$ with $m \geq n$. Consequently, $\mu(A_c \cap C) = \mu(A_c)\mu(C)$ for every cylinder C of the form $C = [0; C_0, \dots, C_{n-1}]$. Since n is arbitrary and the cylinders are a generating family, it follows that $\mu(A_c \cap B) = \mu(A_c)\mu(B)$ for every measurable set $B \subset X^{\mathbb{N}}$. Taking $B = A_c$ we conclude that $\mu(A_c) = \mu(A_c)^2$; in other words, $\mu(A_c) \in \{0, 1\}$ for every $c \in \mathbb{R}$. This proves that φ is constant at μ -almost every point, as stated.

Example 8.4.3. Now consider a two-sided shift map $\sigma : X^{\mathbb{Z}} \rightarrow X^{\mathbb{Z}}$ and a Bernoulli measure $\mu = \nu^{\mathbb{Z}}$. Let \mathcal{A} be the σ -algebra generated by the cylinders of the form $[0; C_0, \dots, C_m]$ with $m \geq 0$. Denote by $L_0^2(X^{\mathbb{Z}}, \mathcal{A}, \mu)$ the space of all functions $\varphi \in L_0^2(\mu)$ that are measurable with respect to the σ -algebra \mathcal{A} (in other words, $\varphi(x)$ depends only on the coordinates $x_n, n \geq 0$ of the point). Take $E = L_0^2(X^{\mathbb{Z}}, \mathcal{A}, \mu)$. Condition (i) in Definition 8.4.1 is obvious. Condition (ii) follows from the same arguments that we used in Example 8.4.2. To prove condition (iii), note that $\bigcup_n U_{\sigma^{-n}}(E)$ contains the characteristic functions of all the cylinders. Therefore, it contains all the linear combinations of characteristic functions of sets in the algebra generated by the cylinders. This implies that the union is dense in $L_0^2(\mu)$, as we wanted to prove.

Lemma 8.4.4. *If (f, μ) has Lebesgue spectrum then $\lim_n U_f^n \varphi \cdot \psi = 0$ for every $\varphi \in L_0^2(\mu)$ and every $\psi \in L^2(\mu)$.*

Proof. Observe that the sequence $U_f^n \varphi \cdot \psi$ is bounded. Indeed, by the Cauchy–Schwarz inequality (Theorem A.5.4):

$$|U_f^n \varphi \cdot \psi| \leq \|U_f^n \varphi\|_2 \|\psi\|_2 = \|\varphi\|_2 \|\psi\|_2 \quad \text{for every } n.$$

So, it is enough to prove that every convergent subsequence $U_f^{n_j} \varphi \cdot \psi$ converges to zero. Furthermore, the set $\{U_f^n \varphi : n \in \mathbb{N}\}$ is bounded in $L^2(\mu)$, because U_f is an isometry. By the theorem of Banach–Alaoglu (Theorems A.6.1 and 2.3.1), every sequence in that set admits some weakly convergent subsequence. Hence, it is no restriction to suppose that $U_f^{n_j} \varphi$ converges weakly to some $\hat{\varphi} \in L^2(\mu)$.

Let E be a subspace satisfying the conditions in Definition 8.4.1. Initially, suppose that $\varphi \in U_f^{-k}(E)$ for some k . Then $U_f^{n_j} \varphi \in U_f^{n_j-k}(E)$. Hence, given any $l \in \mathbb{N}$, we have that $U_f^{n_j} \varphi \in U_f^l(E)$ for every j sufficiently large. It follows (see Exercise A.6.8) that $\hat{\varphi} \in U_f^l(E)$ for every $l \in \mathbb{N}$. By condition (ii) in the definition, this implies that $\hat{\varphi} = 0$ at μ -almost every point. In particular, $\lim_j U_f^{n_j} \varphi \cdot \psi = \hat{\varphi} \cdot \psi = 0$.

Now consider any $\varphi \in L_0^2(\mu)$. By condition (iii) in the definition, for every $\varepsilon > 0$ there exist $k \in \mathbb{N}$ and $\varphi_k \in U_f^{-k}(E)$ such that $\|\varphi - \varphi_k\|_2 \leq \varepsilon$. Using the Cauchy–Schwarz once more inequality:

$$|U_f^n \varphi \cdot \psi - U_f^n \varphi_k \cdot \psi| \leq \|\varphi - \varphi_k\|_2 \|\psi\|_2 \leq \varepsilon \|\psi\|_2$$

for every n . Recalling that $\lim_n U_f^n U_f^k \cdot \psi = 0$ (by the previous paragraph), we find that

$$-\varepsilon \|\psi\|_2 \leq \liminf_n U_f^n \varphi \cdot \psi \leq \limsup_n U_f^n \varphi \cdot \psi \leq \varepsilon \|\psi\|_2.$$

Making $\varepsilon \rightarrow 0$, it follows that $\lim_n U_f^n \varphi \cdot \psi = 0$, as we wanted to prove.

Corollary 8.4.5. *If (f, μ) has Lebesgue spectrum then (f, μ) is mixing.*

Proof. It suffices to observe that

$$C_n(\varphi, \psi) = |U_f^n \varphi \cdot \psi - \left(\int \varphi d\mu \right) \cdot \psi| = |U_f^n \left(\varphi - \int \varphi d\mu \right) \cdot \psi|$$

and the function $\varphi' = \varphi - \int \varphi d\mu$ is in $L_0^2(\mu)$.

The converse to Corollary 8.4.5 is false, in general: in Example 8.4.13 we present certain mixing systems that do not have Lebesgue spectrum.

The class of systems with Lebesgue spectrum is invariant under spectral equivalence. Indeed, suppose that (f, μ) has Lebesgue spectrum and (g, ν) is spectrally equivalent to (f, μ) . Let $L : L^2(\mu) \rightarrow L^2(\nu)$ be a unitary operator conjugating the Koopman operators U_f and U_g . It follows from the hypothesis and Corollary 8.4.5 that (f, μ) is weak mixing. Hence, by Theorem 8.2.1, the constant functions are the only eigenvectors of U_f . Then the same holds for U_g and so the conjugacy L maps constant functions to constant functions. Then, as L is unitary, its restriction to the orthogonal complement $L_0^2(\mu)$ is a unitary operator onto $L_0^2(\nu)$. Now, given any subspace $E \subset L_0^2(\mu)$ satisfying the conditions (i), (ii), (iii) in Definition 8.4.1 for U_f , it is clear that the subspace $L(E) \subset L_0^2(\nu)$ satisfies the corresponding conditions for U_g . Hence, (g, ν) has Lebesgue spectrum.

Given closed subspaces $V \subset W$ of a Hilbert space H , we denote by $W \ominus V$ the orthogonal complement of V inside W , that is,

$$W \ominus V = W \cap V^\perp = \{w \in W : v \cdot w = 0 \text{ for every } v \in V\}.$$

The proof of the following fact is discussed in the next section:

Proposition 8.4.6. *If $U : H \rightarrow H$ is an isometry and E_1 and E_2 are subspaces satisfying the conditions in Definition 8.4.1, then the orthogonal complements $E_1 \ominus U(E_1)$ and $E_2 \ominus U(E_2)$ have the same Hilbert dimension.*

This leads to the following definition: the *rank* of an operator $U : H \rightarrow H$ with Lebesgue spectrum is the Hilbert dimension of $E \ominus U(E)$ for any subspace E satisfying the conditions in Definition 8.4.1.

Then we define the *rank* of a system (f, μ) with Lebesgue spectrum to be the rank of the associated Koopman operator restricted to $L^2_0(\mu)$. It is clear that the rank is less than or equal to the Hilbert dimension of $L^2_0(\mu)$. In particular, if $L^2(\mu)$ is separable then the rank is countable, possibly finite. The majority of interesting examples fall into this category:

Example 8.4.7. Suppose that the probability space (M, \mathcal{B}, μ) is *countably generated*, that is, there exists a countable family G of measurable subsets such that every element of \mathcal{B} coincides, up to measure zero, with some element of the σ -algebra generated by G . Then $L^2(\mu)$ is separable: the algebra \mathcal{A} generated by G is countable and the linear combinations with rational coefficients of characteristic functions of elements of \mathcal{A} form a countable dense subset of $L^2(\mu)$.

It is interesting to point out that no examples are known of systems with Lebesgue spectrum of finite rank. For Bernoulli shifts, the rank coincides with the dimension of the corresponding $L^2(\mu)$:

Example 8.4.8. Let (σ, μ) be a one-sided Bernoulli shift (similar considerations apply in the two-sided case). As we have seen in Example 8.4.2, we may take $E = L^2_0(\mu)$. Then, denoting $x = (x_1, \dots, x_n, \dots)$ and recalling that $\mu = \nu^{\mathbb{N}}$,

$$\begin{aligned} \varphi \in E \ominus U_\sigma(E) &\Leftrightarrow \int \varphi(x_0, x) \psi(x) d\mu(x_0, x) = 0 && \forall \psi \in L^2_0(\mu) \\ &\Leftrightarrow \int \left(\int \varphi(x_0, x) d\nu(x_0) \right) \psi(x) d\mu(x) = 0 && \forall \psi \in L^2_0(\mu). \end{aligned}$$

Hence, $E \ominus U_\sigma(E) = \{ \varphi \in L^2(\mu) : \int \varphi(x_0, x) d\nu(x_0) = 0 \text{ for } \mu\text{-almost every } x \}$. We claim that $\dim(E \ominus U_\sigma(E)) = \dim L^2(\mu)$. The inequality \leq is obvious. To prove the other inequality, fix any measurable function $\phi : X \rightarrow \mathbb{R}$ with $\int \phi d\nu = 0$ and $\int \phi^2 d\nu = 1$. Consider the linear map $I : L^2(\mu) \rightarrow L^2(\mu)$ associating with each $\psi \in L^2(\mu)$ the function $I\psi(x_0, x) = \phi(x_0)\psi(x)$. The assumptions on ϕ imply that

$$I\psi \in E \ominus U_\sigma(E) \quad \text{and} \quad \|I\psi\|_2 = \|\psi\|_2 \quad \text{for every } \psi \in L^2(\mu).$$

This shows that $E \ominus U_\sigma(E)$ contains a subspace isometric to $L^2(\mu)$ and, hence, $\dim E \ominus U_\sigma(E) \geq \dim L^2(\mu)$. This concludes the argument.

We say that the shift is *of countable type* if the probability space X is countably generated. This is automatic, for example, if X is finite, or even countable. In that case, the space $\Sigma = X^{\mathbb{N}}$ (or $\Sigma = X^{\mathbb{Z}}$) is also countably generated: if G is a countable generator of X then the cylinders $[m; C_m, \dots, C_n]$ with $C_j \in G$ form a countable generator of Σ . Then, as observed in Example 8.4.7, the space $L^2(\mu)$ is separable. Therefore, it follows from Example 8.4.8 that *every Bernoulli shift of countable type has Lebesgue spectrum with countable rank*.

8.4.2 The invertible case

In this section we take the system (f, μ) to be invertible. In this context, the notion of Lebesgue spectrum may be formulated in a more transparent way:

Proposition 8.4.9. *Let $U : H \rightarrow H$ be a unitary operator in a Hilbert space H . Then U has Lebesgue spectrum if and only if there exists a closed subspace $F \subset H$ such that the iterates $U^k(F)$, $k \in \mathbb{Z}$ are pairwise orthogonal and satisfy*

$$H = \bigoplus_{k \in \mathbb{Z}} U^k(F).$$

Proof. Suppose that there exists some subspace F as in the statement. Take $E = \bigoplus_{k=0}^{\infty} U^k(F)$. Condition (i) in Definition 8.4.1 is immediate:

$$U(E) = \bigoplus_{k=1}^{\infty} U^k(F) \subset E.$$

As for condition (ii), note that $\varphi \in \bigcap_{n=0}^{\infty} U^n(E)$ means that $\varphi \in \bigoplus_{k=n}^{\infty} U^k(F)$ for every $n \geq 0$. This implies that φ is orthogonal to $U^k(F)$ for every $k \in \mathbb{Z}$. Hence, $\varphi = 0$. Finally, by hypothesis, we may write any $\varphi \in H$ as an orthogonal sum $\varphi = \sum_{k \in \mathbb{Z}} \varphi_k$ with $\varphi_k \in U^k(F)$ for every k . Then

$$\sum_{k=-n}^{\infty} \varphi_k \in \bigoplus_{k=-n}^{\infty} U^k(F) = U^{-n}(E)$$

for every n and the sequence on the left-hand side converges to φ when $n \rightarrow \infty$. This gives condition (iii) in the definition.

Now we prove the converse. Given E satisfying the conditions (i), (ii) and (iii) in the definition, take $F = E \ominus U(E)$. It is easy to see that the iterates of F are pairwise orthogonal. We claim that

$$\bigoplus_{k=0}^{\infty} U^k(F) = E. \tag{8.4.1}$$

Indeed, consider any $v \in E$. It follows immediately from the definition of F that there exist sequences $v_n \in U^n(F)$ and $w_n \in U^n(E)$ such that $v = v_0 + \dots + v_{n-1} + w_n$ for each $n \geq 1$. We want to show that $(w_n)_n$ converges to zero, to conclude that $v = \sum_{j=0}^{\infty} v_j$. For that, observe that

$$\|v\|^2 = \sum_{j=0}^{n-1} \|v_j\|^2 + \|w_n\|^2 \quad \text{for every } n$$

and, thus, the series $\sum_{j=0}^{\infty} \|v_j\|^2$ is summable. Given $\varepsilon > 0$, fix $m \geq 1$ such that the sum of the terms with $j \geq m$ is less than ε . For every $n \geq m$,

$$\|w_m - w_n\|^2 = \|v_m + \dots + v_{n-1}\|^2 = \|v_m\|^2 + \dots + \|v_{n-1}\|^2 < \varepsilon.$$

This proves that $(w_n)_n$ is a Cauchy sequence in H . Let w be its limit. Since $w_n \in U^n(E) \subset U^m(E)$ for every $m \leq n$, taking the limit we get that $w \in U^m(E)$

for every m . By condition (ii) in the hypothesis, this implies that $w = 0$. Therefore, the proof of the claim (8.4.1) is complete. To conclude the proof of the proposition it suffices to observe that

$$\bigoplus_{k \in \mathbb{Z}} U^k(F) = \sum_{n=0}^{\infty} \bigoplus_{k=-n}^{\infty} U^k(F) = \sum_{n=0}^{\infty} U^{-n}(E).$$

Condition (iii) in the hypothesis implies that this subspace coincides with H .

In particular, an invertible system (f, μ) has Lebesgue spectrum if and only if there exists a closed subspace $F \subset L_0^2(\mu)$ such that

$$L_0^2(\mu) = \bigoplus_{k \in \mathbb{Z}} U_f^k(F). \tag{8.4.2}$$

The next result is the reason why systems with Lebesgue spectrum are denominated in this way, and it also leads naturally to the notion of rank:

Proposition 8.4.10. *Let $U : H \rightarrow H$ be a unitary operator in a Hilbert space. Let λ denote the Lebesgue measure on the unit circle. Then U has Lebesgue spectrum if and only if it admits a spectral representation*

$$T : L^2(\lambda)^\chi \rightarrow L^2(\lambda)^\chi \quad (\varphi_\alpha)_\alpha \mapsto (z \mapsto z\varphi_\alpha(z))_\alpha$$

for some cardinal χ . Moreover, χ is uniquely determined by U .

Proof. Let us start by proving the “if” claim. As we know, the Fourier family $\{z^n : n \in \mathbb{Z}\}$ is a Hilbert basis of the space $L^2(\lambda)$. Let V_n be the one-dimensional subspace generated by $\varphi(z) = z^n$. Then, $L^2(\lambda) = \bigoplus_{n \in \mathbb{Z}} V_n$ and, consequently,

$$L^2(\lambda)^\chi = \left(\bigoplus_{n \in \mathbb{Z}} V_n \right)^\chi = \bigoplus_{n \in \mathbb{Z}} V_n^\chi \tag{8.4.3}$$

(W^χ denotes the orthogonal direct sum of χ copies of a space W). Moreover, the restriction of T to each V_n^χ is a unitary operator onto V_{n+1}^χ . Take $F' = V_0^\chi$. The relation (8.4.3) means that the iterates $T^n(F') = V_n^\chi$ are pairwise orthogonal and their orthogonal direct sum is the space $L^2(\lambda)^\chi$. Using the conjugacy of T to the Koopman operator in $L_0^2(\mu)$, we conclude that there exists a subspace F in the conditions of Proposition 8.4.9.

Conversely, suppose that there exists F in the conditions of Proposition 8.4.9. Let $\{v_q : q \in Q\}$ be a Hilbert basis of F . Then $\{U^n(v_q) : n \in \mathbb{Z}, q \in Q\}$ is a Hilbert basis of H . Given $q \in Q$, denote by δ_q the element of the space $L^2(\lambda)^Q$ that is equal to 1 in the coordinate q and identically zero in all the other coordinates. Define

$$L : H \rightarrow L^2(\lambda)^Q, \quad L(U^n(v_q)) = z^n \delta_q \quad \text{for each } n \in \mathbb{Z} \text{ and } q \in Q.$$

Observe that L is a unitary operator, since $\{z^n \delta_q\}$ is a Hilbert basis of $L^2(\lambda)^Q$. Observe also that $LU = TL$. This provides the spectral representation in the

statement of the proposition, with χ equal to the cardinal of the set Q , that is, equal to the Hilbert dimension of the subspace F .

Let $E \subset H$ be any subspace satisfying the conditions in Definition 8.4.1. Then the orthogonal difference $F = E \ominus U(E)$ satisfies the conclusion of Proposition 8.4.9, as we saw during the proof of that proposition. Moreover, according to the proof of Proposition 8.4.10, we may take the cardinal χ equal to the Hilbert dimension of F . Since χ is uniquely determined, the same holds for the Hilbert dimension of $E \ominus U(E)$. This proves Proposition 8.4.6 in the invertible case. In Exercise 8.4.3 we invite the reader to prove the general case.

We have just shown that the rank of a system with Lebesgue spectrum is well defined. Next, we are going to see that for invertible systems the rank is a complete invariant of spectral equivalence:

Corollary 8.4.11. *Two invertible systems with Lebesgue spectrum are spectrally equivalent if and only if they have the same rank.*

Proof. It is clear that two invertible systems are spectrally equivalent if and only if they admit the same spectral representation. By Proposition 8.4.10, this happens if and only if the value of the cardinal χ is the same, that is, if the rank is the same.

Corollary 8.4.12. *All two-sided Bernoulli shifts of countable type are spectrally equivalent.*

Proof. As we saw in the previous section, all Bernoulli shifts of countable type have countable rank.

Proofs of the facts that are quoted in the following may be found in Mañé [Mañ87, Section II.10]:

Example 8.4.13 (Gaussian shifts). Let $A = (a_{i,j})_{i,j \in \mathbb{Z}}$ be an infinite real matrix. We say that A is *positive definite* if every finite restriction $A_{m,n} = (a_{i,j})_{m \leq i,j < n}$ is positive definite, for any $m < n$. We say that A is *symmetric* if $a_{i,j} = a_{j,i}$ for any $i, j \in \mathbb{Z}$. Let μ be a Borel probability measure on $\Sigma = \mathbb{R}^{\mathbb{Z}}$ (similar considerations hold for $\Sigma = \mathbb{R}^{\mathbb{N}}$). We say that μ is a *Gaussian measure* if there exists some symmetric positive definite matrix A such that $\mu([m; B_m, \dots, B_{n-1}])$ is equal to

$$\frac{1}{(\det A_{m,n})^{1/2}} \frac{1}{(2\pi)^{(n-m)/2}} \int_{B_m \times \dots \times B_{n-1}} \exp\left(-\frac{1}{2}(A_{m,n}^{-1}z \cdot z)\right) dz$$

for any $m < n$ and any measurable sets $B_m, \dots, B_{n-1} \subset \mathbb{R}$. The reason for the factor on the left-hand side is explained in Exercise 8.4.4. A is called the *covariance matrix* of μ . It is uniquely determined by

$$a_{i,j} = \int x_i x_j d\mu(x) \quad \text{for each } i, j \in \mathbb{Z}.$$

For each symmetric positive definite matrix A there exists a unique Gaussian probability measure μ that has A as its covariance matrix. Moreover, μ is invariant under the shift map $\sigma : \Sigma \rightarrow \Sigma$ if and only if $a_{ij} = a_{i+1,j+1}$ for any $i, j \in \mathbb{Z}$. In that case, the properties of the system (σ, μ) are directly related to the behavior of the *covariance sequence*

$$a_n = a_{n,0} = U_\sigma^n x_0 \cdot x_0 \quad \text{for each } n \geq 0.$$

In particular, (f, μ) is mixing if and only if the covariance sequence converges to zero.

Now, Exercise 8.4.5 shows that if (f, μ) has Lebesgue spectrum then the covariance sequence is generated by some absolutely continuous probability measure ν on the unit circle, in the following sense:

$$a_n = \int z^n d\nu(z) \quad \text{for each } n \geq 0.$$

(The Riemann–Lebesgue lemma asserts that if ν is a probability measure absolutely continuous with respect to the Lebesgue measure λ on the unit circle then the sequence $\int z^n d\nu(z)$ converges to zero when $n \rightarrow \infty$.) But Exercise 8.4.6 shows that not every sequence that converges to zero is of this form. Therefore, there exist Gaussian shifts (σ, μ) that are mixing but do not have Lebesgue spectrum.

8.4.3 Exercises

- 8.4.1. Show that every mixing Markov shift has Lebesgue spectrum with countable rank. [Observation: In Section 9.5.3 we mention stronger results.]
- 8.4.2. Let μ be the Haar measure on \mathbb{T}^d and $f_A : \mathbb{T}^d \rightarrow \mathbb{T}^d$ be a surjective endomorphism. Assume that no eigenvalue of the matrix A is a root of unity. Check that every orbit of A^t in the set $\mathbb{Z}^d \setminus \{0\}$ is infinite and use this fact to conclude that (f_A, μ) has Lebesgue spectrum. Conversely, if (f_A, μ) has Lebesgue spectrum then no eigenvalue of A is a root of unity.
- 8.4.3. Complete the proof of Proposition 8.4.6, using Exercise 2.3.6 to reduce the general case to the invertible one.
- 8.4.4. Check that $\int_{\mathbb{R}} e^{-x^2/2} dx = \sqrt{2\pi}$. Use this fact to show that if A is a symmetric positive definite matrix of dimension $d \geq 1$ then

$$\int_{\mathbb{R}^d} \exp(- (A^{-1}z \cdot z)/2) dz = (\det A)^{1/2} (2\pi)^{d/2}.$$

- 8.4.5. Let (f, μ) be an invertible system with Lebesgue spectrum. Show that for every $\varphi \in L^2_0(\mu)$ there exists a probability measure ν absolutely continuous with respect to the Lebesgue measure λ on the unit circle $\{z \in \mathbb{C} : |z| = 1\}$ and such that $U_f^n \varphi \cdot \varphi = \int z^n d\nu(z)$ for every $n \in \mathbb{Z}$.
- 8.4.6. Let λ be the Lebesgue measure in the unit circle. Consider the linear operator $F : L^1(\lambda) \rightarrow c_0$ defined by

$$F(\varphi) = \left(\int z^n \varphi(z) d\lambda(z) \right)_n.$$

Show that F is continuous and injective but not surjective. Therefore, not every sequence of complex numbers $(\alpha_n)_n$ converging to zero may be written as $\alpha_n = \int z^n d\nu(z)$ for $n \geq 0$, for some probability measure ν absolutely continuous with respect to λ .

8.5 Lebesgue spaces and ergodic isomorphism

The main subject this section are the *Lebesgue spaces* (also called *standard probability spaces*), a class of probability spaces introduced by the Russian mathematician Vladimir A. Rokhlin [Rok62]. These spaces have a distinguished role in measure theory, for two reasons: on the one hand, they exhibit much better properties than a general probability space; on the other hand, they include most interesting examples. In particular, every complete separable metric space endowed with a Borel probability measure is a Lebesgue space.

Initially, we discuss yet another notion of equivalence, intermediate to ergodic equivalence and spectral equivalence, that we call *ergodic isomorphism*. One of the highlights is that for transformations in Lebesgue spaces the notions of ergodic equivalence and ergodic isomorphism turn out to coincide.

8.5.1 Ergodic isomorphism

Let (M, \mathcal{B}, μ) be a probability space. We denote by $\tilde{\mathcal{B}}$ the quotient of the σ -algebra by the equivalence relation $A \sim B \Leftrightarrow \mu(A \Delta B) = 0$. Observe that if $A_k \sim B_k$ for every $k \in \mathbb{N}$ then $\bigcup_k A_k \sim \bigcup_k B_k$, $\bigcap_k A_k \sim \bigcap_k B_k$ and $A_k^c \sim B_k^c$ for every $k \in \mathbb{N}$. Therefore, the basic operations of set theory are well defined in the quotient $\tilde{\mathcal{B}}$. Moreover, the measure μ induces a measure $\tilde{\mu}$ on $\tilde{\mathcal{B}}$. The pair $(\tilde{\mathcal{B}}, \tilde{\mu})$ is called the *measure algebra* of the probability space.

Now let (M, \mathcal{B}, μ) and (N, \mathcal{C}, ν) be two probability spaces, and $(\tilde{\mathcal{B}}, \tilde{\mu})$ and $(\tilde{\mathcal{C}}, \tilde{\nu})$ be their measure algebras. A *homomorphism* of measure algebras is a map $H : \tilde{\mathcal{B}} \rightarrow \tilde{\mathcal{C}}$ that preserves the operations of union, intersection and complement and also preserves the measures: $\mu(B) = \nu(H(B))$ for every $B \in \tilde{\mathcal{B}}$. If H is a bijection, we call it an *isomorphism* of measure algebras. In that case the inverse H^{-1} is also an isomorphism of measure algebras.

Every measurable map $h : M \rightarrow N$ satisfying $h_*\mu = \nu$ defines a homomorphism $\tilde{h} : \tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{B}}$, through $B \mapsto h^{-1}(B)$. Moreover, if h is invertible then \tilde{h} is an isomorphism. In the same way, transformations $f : M \rightarrow M$ and $g : N \rightarrow N$ preserving the measures in the corresponding probability spaces define homomorphisms $\tilde{f} : \tilde{\mathcal{B}} \rightarrow \tilde{\mathcal{B}}$ and $\tilde{g} : \tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{C}}$, respectively. We say that the systems (f, μ) and (g, ν) are *ergodically isomorphic* if these homomorphisms are conjugate, that is, if $\tilde{f} \circ H = H \circ \tilde{g}$ for some isomorphism $H : \tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{B}}$.

Ergodically equivalent systems are always ergodically isomorphic: given any ergodic equivalence h , it suffices to take $H = \tilde{h}$. We also have the following relation between ergodic isomorphism and spectral equivalence:

Proposition 8.5.1. *If two systems (f, μ) and (g, ν) are ergodically isomorphic then they are spectrally equivalent.*

Proof. Let $H : \tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{B}}$ be an isomorphism such that $\tilde{f} \circ H = H \circ \tilde{g}$. Consider the linear operator $L : L^2(\nu) \rightarrow L^2(\mu)$ constructed as follows. Initially, $L(\mathcal{X}_C) = \mathcal{X}_{H(C)}$ for every $B \in \tilde{\mathcal{C}}$. Note that $\|L(\mathcal{X}_C)\| = \|\mathcal{X}_C\|$. Extend the definition to the set of simple functions, preserving linearity:

$$L\left(\sum_{j=1}^k c_j \mathcal{X}_{C_j}\right) = \sum_{j=1}^k c_j \mathcal{X}_{H(C_j)} \quad \text{for any } k \geq 1, c_j \in \mathbb{R} \text{ and } C_j \in \tilde{\mathcal{C}}.$$

The definition does not depend on the representation of the simple function as a linear combination of characteristic functions (Exercise 8.5.1). Moreover, $\|L(\varphi)\| = \|\varphi\|$ for every simple function. Recall that the set of simple functions is dense in $L^2(\nu)$. Then, by continuity, L extends uniquely to a linear isometry defined on the whole of $L^2(\nu)$. Observe that this isometry is invertible: the inverse is constructed in the same way, starting from the inverse of H . Finally,

$$U_f \circ L(\mathcal{X}_C) = U_f(\mathcal{X}_{H(C)}) = \mathcal{X}_{\tilde{f}(H(C))} = \mathcal{X}_{H(\tilde{g}(C))} = L(\mathcal{X}_{\tilde{g}(C)}) = L \circ U_g(\mathcal{X}_C)$$

for every $C \in \tilde{\mathcal{C}}$. By linearity, it follows that $U_f \circ L(\varphi) = L \circ U_g(\varphi)$ for every simple function; then, by continuity, the same holds for every $\varphi \in L^2(\nu)$.

Summarizing these observations, we have the following relation between the three equivalence relations:

$$\text{ergodic equivalence} \Rightarrow \text{ergodic isomorphism} \Rightarrow \text{spectral equivalence}.$$

In what follows we discuss some partial converses, starting with the relation between ergodic isomorphism and spectral equivalence.

The following result of Paul Halmos and John von Neumann [HvN42] broadens Proposition 8.3.4 and shows that for systems with discrete spectrum the notions of ergodic isomorphism and spectral equivalence coincide. The reader may find a proof in Section 3.2 of Walters [Wal75].

Theorem 8.5.2 (Discrete spectrum). *If (f, μ) and (g, ν) are ergodic systems with discrete spectrum then the following conditions are equivalent:*

1. (f, μ) and (g, ν) are spectrally equivalent;
2. the Koopman operators of (f, μ) and (g, ν) have the same eigenvalues;
3. (f, μ) and (g, ν) are ergodically isomorphic.

In particular, every invertible ergodic system with discrete spectrum is ergodically isomorphic to its inverse.

8.5.2 Lebesgue spaces

Let (M, \mathcal{B}, μ) be any probability space. Initially, suppose that the measure μ is non-atomic, that is, that $\mu(\{x\}) = 0$ for every $x \in M$. Let $\mathcal{P}_1 \prec \dots \prec \mathcal{P}_n \prec \dots$ be an increasing sequence of finite partitions of M into measurable sets. We call the sequence *separating* if, given any two different points $x, y \in M$, there exists $n \geq 1$ such that $\mathcal{P}_n(x) \neq \mathcal{P}_n(y)$. In other words, the non-empty elements of the partition $\bigvee_{n=1}^{\infty} \mathcal{P}_n$ contain a unique point.

Let $M_{\mathcal{P}}$ be the subset one obtains by removing from M all the $P \in \bigcup_n \mathcal{P}_n$ with measure zero. Observe that $M_{\mathcal{P}}$ has full measure. We denote by $\mathcal{B}_{\mathcal{P}}$ and $\mu_{\mathcal{P}}$ the restrictions of \mathcal{B} and μ , respectively, to $M_{\mathcal{P}}$. Let m be the Lebesgue measure on \mathbb{R} . The next proposition means that the separating sequence allows one to represent the probability space $(M_{\mathcal{P}}, \mathcal{B}_{\mathcal{P}}, \mu_{\mathcal{P}})$ as a kind of subspace of the real line. We say “kind of” because, in general, the image $\iota(M_{\mathcal{P}})$ is *not* a measurable subset of \mathbb{R} .

Proposition 8.5.3. *Given any separating sequence $(\mathcal{P}_n)_n$, there exists a compact totally disconnected set $K \subset \mathbb{R}$ and there exists a measurable injective map $\iota : M_{\mathcal{P}} \rightarrow K$ such that the closure of the image $\overline{\iota(P)}$ of every $P \in \bigcup_n \mathcal{P}_n$ is an open and closed subset of K with $m(\overline{\iota(P)}) = \mu(P)$. In particular, $\iota_*\mu$ coincides with the restriction of the Lebesgue measure m to the set K .*

Proof. Let $\alpha_n = 1 + 1/n$ for $n \geq 1$. We are going to construct a sequence of bijective maps $\psi_n : \mathcal{P}_n \rightarrow \mathcal{I}_n$, $n \geq 1$ satisfying:

- (i) each \mathcal{I}_n is a finite family of compact pairwise disjoint intervals;
- (ii) each element of \mathcal{I}_n , $n > 1$ is contained in some element of \mathcal{I}_{n-1} ;
- (iii) $m(\psi_n(P)) = \alpha_n \mu(P)$ for every $P \in \mathcal{P}_n$ and every $n \geq 1$.

To do this, we start by writing $\mathcal{P}_1 = \{P_1, \dots, P_N\}$. Consider any family $\mathcal{I}_1 = \{I_1, \dots, I_N\}$ of compact pairwise disjoint intervals such that $m(I_j) = \alpha_1 \mu(P_j)$ for every j . Let $\psi_1 : \mathcal{P}_1 \rightarrow \mathcal{I}_1$ be the map associating with each P_j the corresponding I_j . Now suppose that, for a given $n \geq 1$, we have already constructed maps ψ_1, \dots, ψ_n satisfying (i), (ii), (iii). For each $P \in \mathcal{P}_n$, let $I = \psi_n(P)$ and let P_1, \dots, P_N be the elements of \mathcal{P}_{n+1} contained in P . Take compact pairwise disjoint intervals $I_1, \dots, I_N \subset I$ satisfying $m(I_j) = \alpha_{n+1} \mu(P_j)$ for each $j = 1, \dots, N$. This is possible because, by the induction hypothesis,

$$m(I) = \alpha_n \mu(P) = \alpha_n \sum_{j=1}^N \mu(P_j) > \alpha_{n+1} \sum_{j=1}^N \mu(P_j).$$

Then, define $\psi_{n+1}(P_j) = I_j$ for each $j = 1, \dots, N$. Repeating this procedure for each $P \in \mathcal{P}_n$, we complete the definition of ψ_{n+1} and \mathcal{I}_{n+1} . It is clear that the conditions (i), (ii), (iii) are preserved. This finishes the construction.

Now, let $K = \bigcap_n \bigcup_{I \in \mathcal{I}_n} I$. It is clear that K is compact and its intersection with any $I \in \mathcal{I}_n$ is an open and closed subset of K . Moreover,

$$\max\{m(I) : I \in \mathcal{I}_n\} = \alpha_n \max\{\mu(P) : P \in \mathcal{P}_n\} \rightarrow 0 \quad \text{when } n \rightarrow \infty \quad (8.5.1)$$

because the sequence $(\mathcal{P}_n)_n$ is separating and the measure μ is non-atomic. Hence, K is totally disconnected. For each $x \in M_{\mathcal{P}}$, the intervals $\psi_n(\mathcal{P}_n(x))$ form a decreasing sequence of compact sets whose lengths decrease to zero. Define $\iota(x)$ to be the unique point in $\bigcap_n \psi_n(\mathcal{P}_n(x))$. The hypothesis that the sequence is separating ensures that ι is injective: if $x \neq y$ then there exists $n \geq 1$ such that $\mathcal{P}_n(x) \cap \mathcal{P}_n(y) = \emptyset$ and, thus, $\iota(x) \neq \iota(y)$. By construction, the pre-image of $K \cap I$ is in $\bigcup_n \mathcal{P}_n$ for every $I \in \bigcup_n \mathcal{I}_n$. Consider the algebra \mathcal{A} formed by the finite disjoint unions of sets $K \cap I$ of this form. This algebra is generating and we have just checked that $\iota^{-1}(A)$ is a measurable set for every $A \in \mathcal{A}$. Therefore, the transformation ι is measurable.

To check the other properties in the statement of the proposition, begin by noting that, for every $n \geq 1$ and $P \in \mathcal{P}_n$,

$$\overline{\iota(P)} = \bigcap_{k=n}^{\infty} \bigcup_Q \psi_k(Q), \quad (8.5.2)$$

where the union is over all the $Q \in \mathcal{P}_k$ that are contained in P . To get the inclusion \subset it suffices to note that $\iota(P) = \bigcup_Q \iota(Q)$ and $\iota(Q) \subset \psi(Q)$ for every $Q \in \mathcal{P}_k$ and every k . The converse follows from the fact that $\iota(P)$ intersects every $\psi_k(Q)$ (the intersection contains $\iota(Q)$) and the length of the $\psi_k(Q)$ converges to zero when $k \rightarrow \infty$. In this way, (8.5.2) is proven. It follows that

$$m(\overline{\iota(P)}) = \lim_k m\left(\bigcup_Q \psi_k(Q)\right) = \lim_k \sum_Q \alpha_k \mu(Q) = \lim_k \alpha_k \mu(P) = \mu(P).$$

Moreover, (8.5.2) means that $\overline{\iota(P)} = \bigcap_{k=n}^{\infty} \bigcup_{I \in \mathcal{I}_k} I$, where the union is over all the $I \in \mathcal{I}_k$ that are contained in $\psi_n(P)$. The right-hand side of this equality coincides with $K \cap \psi_n(P)$ and, hence, is an open and closed subset of K . It also follows from the construction that $\iota^{-1}(\overline{\iota(P)}) = P$. Consequently, $\iota_* \mu(\overline{\iota(P)}) = \mu(P) = m(\overline{\iota(P)})$ for every $P \in \bigcup_n \mathcal{P}_n$. Since the algebra of finite pairwise disjoint unions of sets $\overline{\iota(P)}$ generates the measurable structure of K , we conclude that $\iota_* \mu = m \upharpoonright K$.

We say that a probability space without atoms (M, \mathcal{B}, μ) is a *Lebesgue space* if, for some separating sequence, the image $\iota(M_{\mathcal{P}})$ is a Lebesgue measurable set. Actually, this property does not depend on the choice of the generating sequence (nor on the families \mathcal{I}_n in the proof of Proposition 8.5.3), but we do not prove this fact here: the reader may find a proof in [Rok62, §2.2]. Exercise 8.5.6 shows that it is possible to define Lebesgue space in a more direct way, without using Proposition 8.5.3.

Note that if $\iota(M_{\mathcal{P}})$ is measurable then $\iota(P) = \iota(M_{\mathcal{P}}) \cap \psi_n(P)$ is measurable for every $P \in \mathcal{P}_n$ and every n . Hence, the inverse ι^{-1} is also a measurable transformation. Moreover, $m(\iota(M_{\mathcal{P}})) = \mu(M_{\mathcal{P}}) = 1 = m(K)$. Therefore, every Lebesgue space (M, \mathcal{B}, μ) is isomorphic, as a measure space, to a measurable subset of a compact totally disconnected subset of the real line.

Observe that if the cardinal of M is strictly larger than the cardinal of the continuum then (M, \mathcal{B}, μ) admits no separating sequence and, thus, cannot be a Lebesgue space. In Exercise 8.5.8 we propose another construction of probability spaces that are not Lebesgue spaces. Despite examples such as these, practically all the probability spaces we deal with are Lebesgue spaces:

Theorem 8.5.4. *If M is a complete separable metric space and μ is a Borel probability measure with no atoms then (M, \mathcal{B}, μ) is a Lebesgue space.*

Proof. Let $X \subset M$ be a countable dense subset and $\{B_n : n \in \mathbb{N}\}$ be an enumeration of the set of balls $B(x, 1/k)$ with $x \in X$ and $k \geq 1$. We are going to construct an increasing sequence $(\mathcal{P}_n)_n$ of finite partitions such that

- (i) \mathcal{P}_n is finer than $\{B_1, B_1^c\} \vee \dots \vee \{B_n, B_n^c\}$, and
- (ii) $E_n = \{x \in M : \mathcal{P}_n(x) \text{ is not compact}\}$ satisfies $\mu(E_n) \leq 2^{-n}$.

We start by considering $\mathcal{Q}_1 = \{B_1, B_1^c\}$. By Proposition A.3.7, there exist compact sets $K_1 \subset B_1$ and $K_2 \subset B_1^c$ such that $\mu(B_1 \setminus K_1) \leq 2^{-1}\mu(B_1)$ and $\mu(B_1^c \setminus K_2) \leq 2^{-1}\mu(B_1^c)$. Then take $\mathcal{P}_1 = \{K_1, B_1 \setminus K_1, K_2, B_1^c \setminus K_2\}$. Now, for each $n \geq 1$, assume that one has already constructed partitions $\mathcal{P}_1 < \dots < \mathcal{P}_n$ satisfying (i) and (ii). Consider the partition $\mathcal{Q}_{n+1} = \mathcal{P}_n \vee \{B_{n+1}, B_{n+1}^c\}$ and let Q_1, \dots, Q_m be its elements. By Proposition A.3.7, there exist compact sets $K_j \subset Q_j$ such that $\mu(Q_j \setminus K_j) \leq 2^{-(n+1)}\mu(Q_j)$ for every $j = 1, \dots, m$. Take

$$\mathcal{P}_{n+1} = \{K_1, Q_1 \setminus K_1, \dots, K_m, Q_m \setminus K_m\}.$$

It is clear that \mathcal{P}_{n+1} satisfies (i) and (ii). Therefore, our construction is complete.

All that is left is to show that the existence of such a sequence $(\mathcal{P}_n)_n$ implies the conclusion of the theorem. Property (i) ensures that the sequence is separating. Let $\iota : M_{\mathcal{P}} \rightarrow K$ be a map as in Proposition 8.5.3. Fix any $N \geq 1$ and consider any point $y \in K \setminus \iota(M_{\mathcal{P}})$. For each $n > N$, let I_n be the interval in the family \mathcal{I}_n that contains y and let P_n be the element of \mathcal{P}_n such that $\psi_n(P_n) = I_n$. Note that $(P_n)_n$ is a decreasing sequence. If they were all compact, there would be $x \in \bigcap_{n > N} P_n$ and, by definition, $\iota(x)$ would be equal to y . Since we are assuming that y is not in the image of ι , this proves that there exists $l > N$ such that P_l is not compact. Take $l > N$ minimum and let $I_l = \psi_l(P_l)$. Recall that $m(I_l) = \alpha_l \mu(P_l) \leq 2\mu(P_l)$. Let \tilde{I}_N and \tilde{P}_N be the unions of all these I_l and P_l , respectively, when we vary y on the whole $K \setminus \iota(M_{\mathcal{P}})$. On the one hand, \tilde{I}_N contains $K \setminus \iota(M_{\mathcal{P}})$; on the other hand, \tilde{P}_N is contained in $\bigcup_{l > N} E_l$. Moreover,

the I_l are pairwise disjoint (because we took l minimum) and the same holds for the P_l . Hence,

$$m(\tilde{I}_N) \leq 2\mu(\tilde{P}_N) \leq 2\mu\left(\bigcup_{l>N} E_l\right) \leq 2^{-N+1}.$$

Then the intersection $\bigcap_N \tilde{I}_N$ has Lebesgue measure zero and contains $K \setminus \iota(M_{\mathcal{P}})$. Since K is a Borel set, this shows that $\iota(M_{\mathcal{P}})$ is a Lebesgue measurable set.

The next result implies that all the Lebesgue spaces with no atoms are isomorphic:

Proposition 8.5.5. *If (M, \mathcal{B}, μ) is a Lebesgue space with no atoms, there exists an invertible measurable map $h : M \rightarrow [0, 1]$ (defined between subsets of full measure) such that $h_*\mu$ coincides with the Lebesgue measure on $[0, 1]$.*

Proof. Let $\iota : M_{\mathcal{P}} \rightarrow K$ be a map as in Proposition 8.5.3. Consider the map $g : K \rightarrow [0, 1]$ defined by $g(x) = m([a, x] \cap K)$, where $a = \min K$. It follows immediately from the definition that g is non-decreasing and Lipschitz:

$$g(x_2) - g(x_1) = m([x_1, x_2] \cap K) \leq x_2 - x_1,$$

for any $x_1 < x_2$ in K . In particular, g is measurable. By monotonicity, the pre-image of any interval $[y_1, y_2] \subset [0, 1]$ is a set of the form $[x_1, x_2] \cap K$ with $x_1, x_2 \in K$ and $g(x_1) = y_1$ and $g(x_2) = y_2$. In particular,

$$m([x_1, x_2] \cap K) = g(x_2) - g(x_1) = y_2 - y_1 = m([y_1, y_2]).$$

This shows that $g_*(m \upharpoonright K) = m \upharpoonright [0, 1]$. Let Y be the set of points $y \in [0, 1]$ such that $g^{-1}(\{y\}) = [x_1, x_2] \cap K$ with $x_1, x_2 \in K$ and $x_1 < x_2$. Let $X = g^{-1}(Y)$. Then $m(X) = m(Y) = 0$ because Y is countable. Moreover, the restriction $g : K \setminus X \rightarrow [0, 1] \setminus Y$ is bijective. Its inverse is non-decreasing and, consequently, measurable. Now, take $h = g \circ \iota$. It follows from the previous observations that

$$h : M_{\mathcal{P}} \setminus \iota^{-1}(X) \rightarrow g(\iota(M_{\mathcal{P}})) \setminus Y$$

is a measurable bijection with measurable inverse such that $h_*\mu = m \upharpoonright [0, 1]$.

Now we extend this discussion to general probability spaces (M, \mathcal{B}, μ) , possibly with atoms. Let $A \subset M$ be the set of all the atoms; note that A is at most countable, possibly finite. If the space is purely atomic, that is, if $\mu(A) = 1$, then, by definition, it is a Lebesgue space. More generally, let $M' = M \setminus A$, let \mathcal{B}' be the restriction of \mathcal{B} to M' and let μ' be the normalized restriction of μ to \mathcal{B}' . By definition, (M, \mathcal{B}, μ) is a Lebesgue space if (M', \mathcal{B}', μ') is a Lebesgue space.

It is clear that Theorem 8.5.4 remains valid in the general case: every complete separable metric space endowed with a Borel probability measure, possibly with atoms, is a Lebesgue space. Moreover, Proposition 8.5.5 has the following extension to the atomic case: if (M, \mathcal{B}, μ) is a Lebesgue space and

$A \subset M$ denotes the set of atoms of the measure μ , then there exists an invertible measurable map $h : M \rightarrow [0, 1 - \mu(A)] \cup A$ such that $h_*\mu$ coincides with m on the interval $[0, 1 - \mu(A)]$ and coincides with μ on A .

Proposition 8.5.6. *Let (M, \mathcal{B}, μ) and (N, \mathcal{C}, ν) be two Lebesgue spaces and $H : \tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{B}}$ be an isomorphism between the corresponding measure algebras. Then there exists an invertible measurable map $h : M \rightarrow N$ such that $h_*\mu = \nu$ and $H = \tilde{h}$ for every $C \in \tilde{\mathcal{C}}$. Moreover, h is essentially unique: any two maps satisfying these conditions coincide at μ -almost every point.*

We are going to sketch the proof of this proposition in the non-atomic case. The arguments are based on the ideas and use the notations in the proof of Proposition 8.5.3.

Let us start with the uniqueness claim. Let $h_1, h_2 : M \rightarrow N$ be any two maps such that $(h_1)_*\mu = (h_2)_*\mu = \nu$. Suppose that $h_1(x) \neq h_2(x)$ for every x in a set $E \subset M$ with $\mu(E) > 0$. Let $(Q_n)_n$ be a separating sequence in (N, \mathcal{C}, ν) . Then $Q_n(h_1(x)) \neq Q_n(h_2(x))$ for every $x \in E$ and every n sufficiently large. Hence, we may fix n (large) and $E' \subset E$ with $\mu(E') > 0$ such that $Q_n(h_1(x)) \neq Q_n(h_2(x))$ for every $x \in E'$. Consequently, there exist $Q \in Q_n$ and $E'' \subset E'$ with $\mu(E'') > 0$ such that $Q_n(h_1(x)) = Q$ and $Q_n(h_2(x)) \neq Q$ for every $x \in E''$. Therefore, $E'' \subset h_1^{-1}(Q) \setminus h_2^{-1}(Q)$. This implies that $\tilde{h}_1(Q) \neq \tilde{h}_2(Q)$ and, hence, $\tilde{h}_1 \neq \tilde{h}_2$.

Next we comment on the existence claim. Let $(P'_n)_n$ and $(Q'_n)_n$ be separating sequences in (M, \mathcal{B}, μ) and (N, \mathcal{C}, ν) , respectively. Define $\mathcal{P}_n = P'_n \vee H(Q'_n)$ and $\mathcal{Q}_n = Q'_n \vee H^{-1}(P'_n)$. Then $(\mathcal{P}_n)_n$ and $(\mathcal{Q}_n)_n$ are also separating sequences and $\mathcal{P}_n = H(\mathcal{Q}_n)$ for each n . Let $\iota : M_{\mathcal{P}} \rightarrow K$ be a map as in Proposition 8.5.3 and $\psi_n : \mathcal{P}_n \rightarrow \mathcal{I}_n, n \geq 1$ be the family of bijections used in its construction. Let $J : N_{\mathcal{Q}} \rightarrow L$ and $\varphi_n : \mathcal{Q}_n \rightarrow \mathcal{J}_n$ be corresponding objects for (N, \mathcal{C}, ν) . Since we are assuming that (M, \mathcal{B}, μ) and (N, \mathcal{C}, ν) are Lebesgue spaces, ι and J are invertible maps over subsets with full measure. Recall also that $m(\psi_n(P)) = \alpha_n\mu(P)$ for each $P \in \mathcal{P}_n$ and, analogously, $m(\varphi_n(Q)) = \alpha_n\nu(Q)$ for each $Q \in \mathcal{Q}_n$. Hence, $m(\psi_n(P)) = m(\varphi_n(Q))$ if $P = H(Q)$. Then, for each n ,

$$\psi_n \circ H \circ \varphi_n^{-1} : \mathcal{J}_n \rightarrow \mathcal{I}_n \tag{8.5.3}$$

is a bijection that preserves length. Given $z \in K$ and $n \geq 1$, let I_n be the element of \mathcal{I}_n that contains z and let J_n be the corresponding element of \mathcal{J}_n , via (8.5.3). By construction, $(J_n)_n$ is a nested sequence of compact intervals whose length converges to zero. Let $\phi(z)$ be the unique point in the intersection. In this way, one has defined a measurable map $\phi : K \rightarrow L$ that preserves the Lebesgue measure. It is clear from the construction that ϕ is invertible and the inverse is also measurable. Now it suffices to take $h = J^{-1} \circ \phi \circ \iota$.

All that is left is to check that h is invertible. Applying the construction in the previous paragraph to the inverse H^{-1} we find $h' : N \rightarrow M$ such that $h'_*\nu = \mu$

and $H^{-1} = \tilde{h}'$. Then, $\widetilde{h' \circ h} = \tilde{h} \circ \tilde{h}' = \text{id}$ and $\widetilde{h \circ h'} = \tilde{h}' \circ \tilde{h} = \text{id}$. By uniqueness, it follows that $h' \circ h = \text{id}$ and $h \circ h' = \text{id}$ at almost every point.

Corollary 8.5.7. *Let (M, \mathcal{B}, μ) and (N, \mathcal{C}, ν) be two Lebesgue spaces and let $f : M \rightarrow M$ and $g : N \rightarrow N$ be measurable transformations preserving the measures in their corresponding domains. Then (f, μ) and (g, ν) are ergodically equivalent if and only if they are ergodically isomorphic.*

Proof. We only need to show that if the systems are ergodically isomorphic then they are ergodically equivalent. Let $H : \tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{B}}$ be an ergodic isomorphism. By Proposition 8.5.6, there exists an invertible measurable map $h : M \rightarrow N$ such that $h_*\mu = \nu$ and $H = \tilde{h}$. Then,

$$\widetilde{h \circ f} = \tilde{f} \circ \tilde{h} = \tilde{f} \circ H = H \circ \tilde{g} = \tilde{h} \circ \tilde{g} = \widetilde{g \circ h}.$$

By the uniqueness part of Proposition 8.5.6, it follows that $h \circ f = g \circ h$ at μ -almost every point. This shows that h is an ergodic equivalence.

8.5.3 Exercises

8.5.1. Let $H : \tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{B}}$ be a homomorphism of measure algebras. Show that

$$\sum_{i=1}^l b_i \mathcal{X}_{B_i} = \sum_{j=1}^k c_j \mathcal{X}_{C_j} \Rightarrow \sum_{i=1}^l b_i \mathcal{X}_{H(B_i)} = \sum_{j=1}^k c_j \mathcal{X}_{H(C_j)}.$$

8.5.2. Check that the homomorphism of measure algebras $\tilde{g} : \mathcal{C} \rightarrow \mathcal{B}$ induced by a measure-preserving map $g : M \rightarrow N$ is injective. Suppose that N is a Lebesgue space. Show that, given another measure-preserving map $h : M \rightarrow N$, the corresponding homomorphisms \tilde{g} and \tilde{h} coincide if and only if $g = h$ at almost every point.

8.5.3. Let $f : M \rightarrow M$ be a measurable transformation in a Lebesgue space (M, \mathcal{B}, μ) , preserving the measure μ . Show that (f, μ) is invertible at almost every point (that is, there exists an invariant full measure subset restricted to which f is a measurable bijection with measurable inverse) if and only if the corresponding homomorphism of measure algebras $\tilde{f} : \tilde{\mathcal{B}} \rightarrow \tilde{\mathcal{B}}$ is surjective.

8.5.4. Show that the Koopman operator of a system (f, μ) is surjective if and only if the corresponding homomorphism of measure algebras $\tilde{f} : \tilde{\mathcal{B}} \rightarrow \tilde{\mathcal{B}}$ is surjective. In Lebesgue spaces this happens if and only if the system is invertible at almost every point.

8.5.5. Show that every system (f, μ) with discrete spectrum in a Lebesgue space is invertible at almost every point.

8.5.6. Given a separating sequence $\mathcal{P}_1 \prec \dots \prec \mathcal{P}_n \prec \dots$, we call a *chain* any sequence $(P_n)_n$ with $P_n \in \mathcal{P}_n$ and $P_{n+1} \subset P_n$ for every n . We say that a chain is *empty* if $\bigcap_n P_n = \emptyset$. Consider the map $\iota : M_{\mathcal{P}} \rightarrow K$ constructed in Proposition 8.5.3. Show that the image $\iota(M_{\mathcal{P}})$ is a Lebesgue measurable set and $m(K \setminus \iota(M_{\mathcal{P}})) = 0$ if and only if the empty chains have zero measure in the following sense: for every $\delta > 0$ there exists $B \subset M$ such that B is a union of elements of $\bigcup_n \mathcal{P}_n$ with $\mu(B) < \delta$ and every empty chain $(P_n)_n$ has $P_n \subset B$ for every n sufficiently large.

- 8.5.7. Prove the following extension of Proposition 2.4.4: If $f : M \rightarrow M$ preserves a probability measure μ and (M, μ) is a Lebesgue space then μ admits a (unique) lift $\hat{\mu}$ to the natural extension $\hat{f} : \hat{M} \rightarrow \hat{M}$.
- 8.5.8. Let M be a subset of $[0, 1]$ with exterior measure $m^*(M) = 1$ but which is not a Lebesgue measurable set. Consider the σ -algebra \mathcal{M} of all sets of the form $M \cap B$, where B is a Lebesgue measurable subset of \mathbb{R} . Check that $\mu(M \cap B) = m(B)$ defines a probability measure on (M, \mathcal{M}) such that (M, \mathcal{M}, μ) is not a Lebesgue space.