Ergodic Theory

Throughout this chapter we assume $(X, \mathfrak{F}, \mu)$ is a probability space, i.e., a measure space with $\mu(X) = 1$. Ergodic theory studies properties of measure-preserving mappings $\varphi : X \to X$. That is, we assume

$$S \in \mathfrak{F} \Longrightarrow \varphi^{-1}(S) \in \mathfrak{F} \quad \text{and} \quad \mu(\varphi^{-1}(S)) = \mu(S). \tag{14.1}$$

The map φ defines a linear map T on functions:

$$Tf(x) = f(\varphi(x)). \tag{14.2}$$

If (14.1) holds, then, given $f \in L^1(X, \mu)$,

$$\int_X f(\varphi(x))\, d\mu = \int_X f(x)\, d\mu. \tag{14.3}$$

Hence $T : L^p(X, \mu) \to L^p(X, \mu)$ is an isometry for each $p \in [1, \infty]$. A central object of study in ergodic theory is the sequence of means:

$$A_m f(x) = \frac{1}{m} \sum_{k=0}^{m-1} T^k f(x). \tag{14.4}$$

In particular, one considers whether $A_m f$ tends to a limit, as $m \to \infty$, and whether that limit is a constant, namely $c = \int_X f\, d\mu$.

The first basic result of this nature, due to J. von Neumann, deals with $f \in L^2(X, \mu)$. Actually it has a Hilbert space setting. Recall that if a linear operator $T : H \to H$ on a Hilbert space H is an isometry, then $T^*T = I$. The abstract result uses the following simple lemma.

Lemma 14.1. *If $T : H \to H$ is a linear isometry on a Hilbert space H, then there is an orthogonal direct sum*

$$H = K \oplus \overline{R}, \tag{14.5}$$

where

$$K = \text{Ker}\,(I - T^*) = \text{Ker}\,(I - T), \quad R = \text{Range}\,(I - T), \tag{14.6}$$

and $\overline{R}$ is the closure of R.

Proof. First, note that

$$\begin{aligned} R^\perp &= \{v \in H : (v, (I-T)w) = 0, \forall\, w \in H\} \\ &= \{v \in H : ((I-T^*)v, w) = 0, \forall\, w \in H\} \\ &= \text{Ker}\,(I - T^*). \end{aligned}$$

Now (14.5) follows by (4.29)–(4.30) and the rest of the paragraph there, which implies $\overline{R} = K^\perp$, with $K = \text{Ker}\,(I - T^*)$.

It remains to show that $\text{Ker}\,(I - T^*) = \text{Ker}\,(I - T)$. Since $T^*T = I, I - T^* = -T^*(I - T)$, so clearly $\text{Ker}\,(I - T) \subset \text{Ker}\,(I - T^*)$. For the reverse inclusion, note that $T^*T = I \Rightarrow (TT^*)^2 = TT^*$, so $Q = TT^*$ is the orthogonal projection of H onto the range of T. (Cf. Exercises 16–17 of Chapter 9.) Now $T^*u = u \Rightarrow Qu = Tu$, but then $\|Qu\| = \|Tu\| = \|u\|$, so $Qu = u$ and hence $Tu = u$, giving the converse.

Here is the abstract Mean Ergodic Theorem.

Proposition 14.2. *In the setting of Lemma 14.1, for each $f \in H$,*

$$A_m f = \frac{1}{m} \sum_{k=0}^{m-1} T^k f \longrightarrow Pf, \tag{14.7}$$

in H-norm, where P is the orthogonal projection of H onto K.

Proof. Clearly $A_m f \equiv f$ if $f \in K$. If $f = (I - T)v \in R$, then

$$\frac{1}{m} \sum_{k=0}^{m-1} T^k f = \frac{1}{m}(v - T^m v) \to 0, \quad \text{as} \;\; m \to \infty, \tag{14.8}$$

and since the operator norm $\|A_m\| \le 1$ for each m, this convergence holds on $\overline{R}$. Now (14.7) follows from (14.5).

Proposition 14.2 immediately applies to (14.4) when $f \in L^2(X, \mu)$. We next establish a more general result.

Proposition 14.3. *Let P denote the orthogonal projection of $L^2(X,\mu)$ onto* $\operatorname{Ker}(I-T)$. *Then, for $p \in [1,2]$, P extends to a continuous projection on $L^p(X,\mu)$, and*

$$f \in L^p(X,\mu) \Longrightarrow A_m f \to Pf \tag{14.9}$$

in L^p-norm, as $m \to \infty$.

Proof. Note that the L^p-operator norm $\|A_m\|_{\mathcal{L}(L^p)} \le 1$ for each m, and since $\|g\|_{L^p} \le \|g\|_{L^2}$ for $p \in [1,2]$, we have (14.9) in L^p-norm for each f in the dense subspace $L^2(X,\mu)$ of $L^p(X,\mu)$. Now, given $f \in L^p(X,\mu)$, $\varepsilon > 0$, pick $g \in L^2(X,\mu)$ such that $\|f-g\|_{L^p} < \varepsilon$. Then

$$\|A_n f - A_m f\|_{L^p} \le \|A_n g - A_m g\|_{L^2} + \|A_n(f-g)\|_{L^p} + \|A_m(f-g)\|_{L^p}. \tag{14.10}$$

Hence

$$\limsup_{m,n\to\infty} \|A_n f - A_m f\|_{L^p} \le 2\varepsilon, \quad \forall\, \varepsilon > 0. \tag{14.11}$$

This implies the sequence $(A_n f)$ is Cauchy in $L^p(X,\mu)$, for each $f \in L^p(X,\mu)$. Hence it has a limit; call it Qf. Clearly Qf is linear in f, $\|Qf\|_{L^p} \le \|f\|_{L^p}$, and $Qf = Pf$ for $f \in L^2(X,\mu)$. Hence Q is the unique continuous extension of P from $L^2(X,\mu)$ to $L^p(X,\mu)$ (so we change its name to P). Note that $P^2 = P$ on $L^p(X,\mu)$, since it holds on the dense linear subspace $L^2(X,\mu)$. Proposition 14.3 is proven.

Remark. Note that $P = P^*$. It follows that $P : L^p(X,\mu) \to L^p(X,\mu)$ for all $p \in [1,\infty]$. We will show in Proposition 14.7 that (14.9) holds in L^p-norm for $p < \infty$. The subject of mean ergodic theorems has been considerably extended and abstracted by K. Yosida, S. Kakutani, W. Eberlein, and others. An account can be found in [**Kr**].

Such mean ergodic theorems were complemented by pointwise convergence results on $A_m f(x)$, first by G. Birkhoff. This can be done via estimates of Yosida and Kakutani on the maximal function

$$A^{\#} f(x) = \sup_{m\ge 1} A_m f(x) = \sup_{n \ge 1} A_n^{\#} f(x), \tag{14.12}$$

where

$$A_n^{\#} f(x) = \sup_{1\le m\le n} A_m f(x). \tag{14.13}$$

We follow a clean route to such maximal function estimates given in [**Gar**].

Lemma 14.4. *With A_m given by (14.2)–(14.4) and $f \in L^1(X,\mu)$, set*

$$E_n = \{x \in X : A_n^{\#} f(x) \geq 0\}. \tag{14.14}$$

Then

$$\int_{E_n} f \, d\mu \geq 0. \tag{14.15}$$

Proof. For notational convenience, set

$$S_k f = kA_k f = f + Tf + \cdots + T^{k-1} f, \quad M_k f = kA_k^{\#} f = \sup_{1 \leq \ell \leq k} S_\ell f.$$

For $k \in \{1, \ldots, n\}$, $(M_n f)^+ \geq S_k f$, and hence (because T is positivity preserving)

$$f + T(M_n f)^+ \geq f + TS_k f = S_{k+1} f.$$

Hence

$$f \geq S_k f - T(M_n f)^+, \quad \text{for} \ \ 1 \leq k \leq n,$$

this holding for $k \geq 2$ by the argument above, and trivially for $k = 1$. Taking the max over $k \in \{1, \ldots, n\}$ yields

$$f \geq M_n f - T(M_n f)^+. \tag{14.16}$$

Integrating (14.16) over E_n yields

$$\begin{aligned}
\int_{E_n} f \, d\mu &\geq \int_{E_n} (M_n f - T(M_n f)^+) \, d\mu \\
&= \int_{E_n} ((M_n f)^+ - T(M_n f)^+) \, d\mu \\
&= \int_X (M_n f)^+ \, d\mu - \int_{E_n} T(M_n f)^+ \, d\mu \\
&\geq \int_X (M_n f)^+ \, d\mu - \int_X T(M_n f)^+ \, d\mu = 0,
\end{aligned} \tag{14.17}$$

the first and second identities on the right because $M_n f \geq 0$ precisely on E_n, the last inequality because $T(M_n f)^+ \geq 0$, and the last identity by (14.3). This proves the lemma.

Lemma 14.4 leads to the following maximal function estimate.

Proposition 14.5. *In the setting of Lemma 14.4, one has, for each $\lambda > 0$,*

$$\mu(\{x \in X : A_n^{\#} f(x) \geq \lambda\}) \leq \frac{1}{\lambda} \|f\|_{L^1}. \tag{14.18}$$

Proof. If we set $E_{n\lambda} = \{x \in X : A_n^{\#} f(x) \geq \lambda\} = \{x \in X : A_n^{\#}(f(x) - \lambda) \geq 0\}$, then Lemma 14.4 yields

$$\int_{E_{n\lambda}} (f - \lambda)\, d\mu \geq 0. \tag{14.19}$$

Thus

$$\|f\|_{L^1} \geq \int_{E_{n\lambda}} f\, d\mu \geq \lambda\, \mu(E_{n\lambda}), \tag{14.20}$$

as asserted in (14.18).

Note that

$$E_{n\lambda} \nearrow \{x \in X : A^{\#} f(x) \geq \lambda\} = E_\lambda, \tag{14.21}$$

so we have $\mu(E_\lambda) \leq \|f\|_{L^1}/\lambda$. Now we introduce the maximal function

$$\mathcal{A}^{\#} f(x) = \sup_{m \geq 1} |A_m f(x)| \leq A^{\#}|f|(x). \tag{14.22}$$

We have

$$\mu(\{x \in X : \mathcal{A}^{\#} f(x) \geq \lambda\}) \leq \frac{1}{\lambda} \|f\|_{L^1}. \tag{14.23}$$

We are now ready for Birkhoff's Pointwise Ergodic Theorem.

Theorem 14.6. *If T and A_m are given by* (14.2)–(14.4), *where φ is a measure-preserving map, then, given $f \in L^1(X, \mu)$,*

$$\lim_{m \to \infty} A_m f(x) = P f(x), \quad \mu\text{-}a.e. \tag{14.24}$$

Proof. Given $f \in L^1(X, \mu)$, $\varepsilon > 0$, let us pick $f_1 \in L^2(X, \mu)$ such that $\|f - f_1\|_{L^1} \leq \varepsilon/2$. Then use Lemma 14.1, with $H = L^2(X, \mu)$, to produce

$$g \in \operatorname{Ker}(I - T), \quad h = (I - T)v, \quad \|f_1 - (g + h)\|_{L^2} \leq \frac{\varepsilon}{2}. \tag{14.25}$$

Here $v \in L^2(X,\mu)$. It follows that

$$\|f-(g+h)\|_{L^1} \le \varepsilon, \tag{14.26}$$

and we have

$$\begin{aligned} A_m f &= A_m g + A_m h + A_m(f-g-h) \\ &= g + \frac{1}{m}(v - T^m v) + A_m(f-g-h). \end{aligned} \tag{14.27}$$

Clearly $v(x)/m \to 0$, μ-a.e., as $m \to \infty$. Also

$$\int\limits_X \sum_{m\ge 1} \Bigl|\frac{1}{m} T^m v(x)\Bigr|^2 \, d\mu = \|v\|_{L^2}^2 \sum_{m\ge 1} \frac{1}{m^2} < \infty, \tag{14.28}$$

which implies $T^m v(x)/m \to 0$, μ-a.e., as $m \to \infty$. We deduce that for each $\lambda > 0$,

(14.29)

$$\begin{aligned} &\mu\bigl(\{x \in X : \limsup A_m f(x) - \liminf A_m f(x) > \lambda\}\bigr) \\ &= \mu\bigl(\{x \in X : \limsup A_m(f-g-h) - \liminf A_m(f-g-h) > \lambda\}\bigr) \\ &\le \mu\Bigl(\Bigl\{x \in X : \mathcal{A}^{\#}(f-g-h) > \frac{\lambda}{2}\Bigr\}\Bigr) \\ &\le \frac{2}{\lambda}\|f-g-h\|_{L^1} \\ &\le \frac{2\varepsilon}{\lambda}. \end{aligned}$$

Since ε can be taken arbitrarily small, this implies that $A_m f(x)$ converges as $m \to \infty$, μ-a.e. We already know it converges to $Pf(x)$ in L^1-norm, so (14.24) follows.

We can use the maximal function estimate (14.23) to extend Proposition 14.3, as follows. First, there is the obvious estimate

$$\|\mathcal{A}^{\#} f\|_{L^\infty} \le \|f\|_{L^\infty}. \tag{14.30}$$

Now the Marcinkiewicz Interpolation Theorem (see Appendix D) applied to (14.23) and (14.30) yields

$$\|\mathcal{A}^{\#} f\|_{L^p} \le C_p \|f\|_{L^p}, \quad 1 < p < \infty. \tag{14.31}$$

Using this, we prove the following.

Proposition 14.7. *In the setting of Proposition 14.3, we have, for all* $p \in [1, \infty)$,

$$f \in L^p(X, \mu) \Longrightarrow A_m f \to Pf, \quad \text{in } L^p\text{-norm}, \tag{14.32}$$

as $m \to \infty$.

Proof. Take $p \in (1, \infty)$. Given $f \in L^p(X, \mu)$, we have

$$|A_m f(x)| \le \mathcal{A}^{\#} f(x), \quad \mathcal{A}^{\#} f \in L^p(X, \mu).$$

Since the convergence (14.24) holds pointwise μ-a.e., (14.32) follows from the Dominated Convergence Theorem. That just leaves $p = 1$, for which we rely on Proposition 14.3.

Remark. Since $P^* = P$, it follows from Proposition 14.7 that

$$f \in L^p(X, \mu) \Longrightarrow A_m^* f \to Pf,$$

weak* in $L^p(X, \mu)$, for $p \in (1, \infty]$. More general ergodic theorems, such as can be found in [**Kr**], imply one has convergence in L^p-norm (and μ-a.e.), for $p \in [1, \infty)$. Of course if φ is invertible, then such a result is a simple application of the results given above, with φ replaced by φ^{-1}.

Having discussed the convergence of $A_m f$, we turn to the second question raised after (14.4), namely whether the limit must be constant. So far we see that the set of limits coincides with $\operatorname{Ker}(I - T)$, i.e., with the set of invariant functions, where we say $f \in L^p(X, \mu)$ is invariant if and only if

$$f(x) = f(\varphi(x)), \quad \mu\text{-a.e.} \tag{14.33}$$

We note that the following conditions are equivalent:

$$\begin{aligned} &\text{(a)} \qquad f \in L^1(X, \mu) \ \text{ invariant } \Rightarrow f \text{ constant } (\mu\text{-a.e.}), \\ &\text{(b)} \qquad f \in L^2(X, \mu) \ \text{ invariant } \Rightarrow f \text{ constant } (\mu\text{-a.e.}), \\ &\text{(c)} \qquad S \in \mathfrak{F} \ \text{ invariant } \Rightarrow \mu(S) = 0 \ \text{ or } \ \mu(S) = 1. \end{aligned} \tag{14.34}$$

Here we say $S \in \mathfrak{F}$ is invariant if and only if

$$\mu(\varphi^{-1}(S) \triangle S) = 0, \tag{14.35}$$

where $A \triangle B = (A \setminus B) \cup (B \setminus A)$. Note that if $S \in \mathfrak{F}$ satisfies (14.35), then

$$\widetilde{S} = \bigcap_{j \ge 0} \bigcup_{k \ge j} \varphi^{-k}(S) \Longrightarrow \varphi^{-1}(\widetilde{S}) = \widetilde{S} \ \text{ and } \ \mu(\widetilde{S} \triangle S) = 0. \tag{14.36}$$

To see the equivalence in (14.34), note that if $f \in L^1(X, \mu)$ is invariant, then all the sets $S_\lambda = \{x \in X : f(x) > \lambda\}$ are invariant, so (c)$\Rightarrow$(a). Meanwhile clearly (a)$\Rightarrow$(b)$\Rightarrow$(c). A measure-preserving map $\varphi : X \to X$ satisfying (14.34) is said to be *ergodic*.

Theorem 14.6 and Proposition 14.7 have the following corollary.

Proposition 14.8. *If $\varphi : X \to X$ is ergodic and $f \in L^p(X,\mu)$, $p \in [1,\infty)$, then*

$$A_m f \longrightarrow \int_X f \, d\mu, \quad \text{in } L^p\text{-norm and } \mu\text{-a.e.} \tag{14.37}$$

We now consider some examples of ergodic maps. First take the unit circle, $X = S^1 \approx \mathbb{R}/(2\pi\mathbb{Z})$, with measure $d\mu = d\theta/2\pi$. Take $e^{i\alpha} \in S^1$ and define

$$R_\alpha : S^1 \longrightarrow S^1, \quad R_\alpha(e^{i\theta}) = e^{i(\theta+\alpha)}. \tag{14.38}$$

Proposition 14.9. *The map R_α is ergodic if and only if $\alpha/2\pi$ is irrational.*

Proof. We compare the Fourier coefficients $\hat{f}(k) = \int f(\theta)e^{-ik\theta}\, d\mu = (f, e_k)$ with those of Tf. We have

$$\widehat{Tf}(k) = (Tf, e_k) = (f, T^{-1}e_k) = e^{ik\alpha}\hat{f}(k).$$

Thus

$$Tf = f, \ \hat{f}(k) \neq 0 \Longrightarrow e^{ik\alpha} = 1. \tag{14.39}$$

But $e^{ik\alpha} = 1$ for some nonzero $k \in \mathbb{Z}$ if and only if $\alpha/2\pi$ is rational.

In the next example, let $(X, \mathfrak{F}, \mu)$ be a probability space, and form the two-sided infinite product

$$\Omega = \prod_{k=-\infty}^{\infty} X, \tag{14.40}$$

which comes equipped with a σ-algebra $\mathcal{O}$ and a product measure ω, via the construction given at the end of Chapter 6. There is a map on Ω called the two-sided shift:

$$\Sigma : \Omega \to \Omega, \quad \Sigma(x)_k = x_{k+1}, \quad x = (\dots, x_{-1}, x_0, x_1, \dots). \tag{14.41}$$

Proposition 14.10. *The two-sided shift* (14.41) *is ergodic.*

Proof. We make use of the following orthonormal set. Let $\{u_j : j \in \mathbb{Z}^+\}$ be an orthonormal basis of $L^2(X,\mu)$, with $u_0 = 1$. Let $\mathcal{A}$ be the set of elements of $\prod_{k=-\infty}^{\infty} \mathbb{Z}^+$ of the form $\alpha = (\dots, \alpha_{-1}, \alpha_0, \alpha_1, \dots)$ such that $\alpha_k \neq 0$ for only finitely many k. Set

$$v_\alpha(x) = \prod_{k=-\infty}^{\infty} u_{\alpha_k}(x_k), \quad \alpha \in \mathcal{A}, \tag{14.42}$$

and note that for each $\alpha \in \mathcal{A}$ only finitely many factors in this product are not $\equiv 1$. We have the following:

$$\{v_\alpha : \alpha \in \mathcal{A}\} \text{ is an orthonormal basis of } L^2(\Omega, \omega). \tag{14.43}$$

(Cf. Exercise 13 of Chapter 6.) Note that if $Tf(x) = f(\Sigma(x))$,

$$Tv_\alpha = v_{\sigma(\alpha)}, \quad \sigma(\alpha)_k = \alpha_{k-1}. \tag{14.44}$$

Now assume $f \in L^2(\Omega, \omega)$ is invariant. Then

$$\hat{f}(\alpha) = (f, v_\alpha) = (Tf, Tv_\alpha) = \hat{f}(\sigma(\alpha)), \tag{14.45}$$

for each $\alpha \in \mathcal{A}$. Iterating this gives $\hat{f}(\alpha) = \hat{f}(\sigma^\ell(\alpha))$ for each $\ell \in \mathbb{Z}^+$. Since

$$\|f\|_{L^2}^2 = \sum_{\alpha \in \mathcal{A}} |\hat{f}(\alpha)|^2 < \infty, \tag{14.46}$$

and $\{\sigma^\ell(\alpha) : \ell \in \mathbb{Z}^+\}$ is an infinite set except for $\alpha = 0 = (\dots, 0, 0, 0, \dots)$, we deduce that $\hat{f}(\alpha) = 0$ for nonzero $\alpha \in \mathcal{A}$, and hence f must be constant.

A variant of the construction above yields the one-sided shift, on

$$\Omega_0 = \prod_{k=0}^{\infty} X, \tag{14.47}$$

with σ-algebra $\mathcal{O}_0$ and product measure ω_0 constructed in the same fashion. As in (14.41), one sets

$$\Sigma_0 : \Omega_0 \to \Omega_0, \quad \Sigma_0(x)_k = x_{k+1}, \quad x = (x_0, x_1, x_2, \dots). \tag{14.48}$$

The following result has essentially the same proof as Proposition 14.10.

Proposition 14.11. *The one-sided shift* (14.48) *is ergodic.*

Another proof of Proposition 14.11 goes as follows. Suppose that $f \in L^2(\Omega_0, \omega_0)$ is invariant, so $f(x_1, x_2, x_3, \dots) = f(x_{k+1}, x_{k+2}, x_{k+3}, \dots)$. Multiplying both sides by $g(x_1, \dots, x_k)$ and integrating, we have

$$(f, g)_{L^2} = (f, 1)_{L^2}(1, g)_{L^2}$$

for each $g \in L^2(\Omega_0, \omega_0)$ of the form $g = g(x_1, \dots, x_k)$, for any $k < \infty$. Since the set of such g is dense in $L^2(\Omega_0, \omega_0)$, we have this identity for all $g \in L^2(\Omega_0, \omega_0)$, and this implies that f is constant.

The concept of ergodicity defined above extends to a semigroup of measure-preserving transformations, i.e., a collection $\mathcal{S}$ of maps on X satisfying (14.1) for each $\varphi \in \mathcal{S}$ and such that

$$\varphi, \psi \in \mathcal{S} \Longrightarrow \varphi \circ \psi \in \mathcal{S}. \tag{14.49}$$

In such a case, one says a function $f \in L^p(X, \mu)$ is invariant provided (14.33) holds for each $\varphi \in \mathcal{S}$, one says $S \in \mathfrak{F}$ is invariant provided (14.35) holds for all $\varphi \in \mathcal{S}$, and one says the action of $\mathcal{S}$ on $(X, \mathfrak{F}, \mu)$ is ergodic provided the (equivalent) conditions in (14.34) hold. The study so far in this chapter has dealt with $\mathcal{S} = \{\varphi^k : k \in \mathbb{Z}^+\}$. Now we will consider one example of the action of a semigroup (actually a group) not isomorphic to $\mathbb{Z}^+$ (nor to $\mathbb{Z}$). This will lead to a result complementary to Proposition 14.10.

Let $\mathcal{S}_\infty$ denote the group of bijective maps $\sigma : \mathbb{Z} \to \mathbb{Z}$ with the property that $\sigma(k) = k$ for all but finitely many k. Let $(X, \mathfrak{F}, \mu)$ be a probability space and let $\Omega = \prod_{k=-\infty}^{\infty} X$, as in (14.40), with the product measure ω. The group $\mathcal{S}_\infty$ acts on Ω by

$$\varphi_\sigma : \Omega \longrightarrow \Omega, \quad \varphi_\sigma(x)_k = x_{\sigma(k)}, \tag{14.50}$$

where $x = (\dots, x_{-1}, x_0, x_1, \dots) \in \Omega$, $\sigma \in \mathcal{S}_\infty$. The following result is called the Hewitt-Savage 01 Law.

Proposition 14.12. *The action of $\mathcal{S}_\infty$ on Ω defined by* (14.50) *is ergodic.*

Proof. Let $\{v_\alpha : \alpha \in \mathcal{A}\}$ be the orthonormal basis of $L^2(\Omega, \omega)$ given by (14.42)–(14.43). Note that if $T_\sigma f(x) = f(\varphi_\sigma(x))$, then

$$T_\sigma v_\alpha = v_{\sigma^\# \alpha}, \quad (\sigma^\# \alpha)_k = \alpha_{\sigma^{-1}(k)}. \tag{14.51}$$

Now if $f \in L^2(\Omega, \omega)$ is invariant under the action of $\mathcal{S}_\infty$, then, parallel to (14.45), we have

$$\hat{f}(\alpha) = (f, v_\alpha) = (T_\sigma f, T_\sigma v_\alpha) = \hat{f}(\sigma^\# \alpha), \quad \forall\, \alpha \in \mathcal{A},\ \sigma \in \mathcal{S}_\infty. \tag{14.52}$$

Since $\|f\|_{L^2}^2 = \sum_\alpha |\hat{f}(\alpha)|^2 < \infty$ and $\{\sigma^\# \alpha : \sigma \in \mathcal{S}_\infty\}$ is an infinite set, for each nonzero $\alpha \in \mathcal{A}$, it follows that $\hat{f}(\alpha) = 0$ for nonzero $\alpha \in \mathcal{A}$, and hence f must be constant.

The same proof establishes the following result, which contains both Proposition 14.10 and Proposition 14.12. As above, $\mathcal{A}$ is the set defined in the beginning of the proof of Proposition 14.10.

Proposition 14.13. *Let $\mathcal{G}$ be a group of bijective maps on $\mathbb{Z}$ with the property that*

$$\{\sigma^{\#}\alpha : \sigma \in \mathcal{G}\} \quad \text{is an infinite set, for each nonzero} \quad \alpha \in \mathcal{A}, \tag{14.53}$$

where $\sigma^{\#}\alpha$ is given by (14.51)*. Then the action of $\mathcal{G}$ on Ω, given by* (14.50)*, is ergodic.*

See Exercises 10–14 for a Mean Ergodic Theorem that applies in the setting of Proposition 14.12. Other ergodic theorems that apply to semigroups of transformations can be found in [**Kr**].

Exercises

1. Let $\mathbb{T}^n = S^1 \times \cdots \times S^1 \subset \mathbb{C}^n$, where $S^1 = \{z \in \mathbb{C} : |z| = 1\}$. Given $(e^{-\alpha_1}, \dots, e^{i\alpha_n}) \in \mathbb{T}^n$, define

$$R_\alpha : \mathbb{T}^n \to \mathbb{T}^n, \quad R_\alpha(e^{i\theta_1}, \dots, e^{i\theta_n}) = (e^{i(\theta_1+\alpha_1)}, \dots, e^{i(\theta_n+\alpha_n)}).$$

   Give necessary and sufficient conditions that R_α be ergodic.
   Hint. Adapt the argument used to prove Proposition 14.9.

2. Define $\varphi : S^1 \to S^1$ by $\varphi(z) = z^2$. Show that φ is ergodic.
   Hint. Examine the Fourier series of an invariant function.

3. A measure-preserving map φ on $(X, \mathfrak{F}, \mu)$ is said to be *mixing* provided

$$\mu\big(\varphi^{-k}(E) \cap F\big) \to \mu(E)\mu(F), \quad \text{as} \quad k \to \infty, \tag{14.54}$$

   for each $E, F \in \mathfrak{F}$. Show that φ is mixing if and only if $Tf(x) = f(\varphi(x))$ has the property

$$(T^k f, g)_{L^2} \longrightarrow (f, 1)_{L^2}(1, g)_{L^2}, \quad \text{as} \quad k \to \infty, \tag{14.55}$$

   for all $f, g \in L^2(X, \mu)$.

4. Show that a mixing transformation is ergodic.
   Hint. Show that

$$(A_k f, g) = \frac{1}{k}\sum_{j=0}^{k-1}(T^j f, g) \to (f, 1)(1, g). \tag{14.56}$$

Deduce that P in (14.7) is the orthogonal projection of $L^2(X, \mu)$ onto the space of constant functions. Alternatively, just apply (14.45) in case $Tf = f$.

5. Show that the map $\varphi : S^1 \to S^1$ in Exercise 2 is mixing.

6. Show that the two-sided and one-sided shifts Σ and Σ_0, given in (14.41) and (14.48), are mixing.
Hint. Verify (14.55) when f and g are elements of the orthonormal basis $\{v_\alpha\}$ described in (14.42). Alternatively, verify (14.55) when f and g are functions of x_j for $|j| \le M$.

7. Show that the maps R_α in (14.38) and in Exercise 1 are not mixing.

8. We assert that the ergodic transformation $\varphi : S^1 \to S^1$ in Exercise 2 is "equivalent" to the one-sided shift (14.48) for $X = \{0, 1\}$, with $\mu(\{0\}) = \mu(\{1\}) = 1/2$. Justify this.
Hint. Regard an element of Ω_0 as giving the binary expansion of a number $x \in [0, 1)$.

9. Let φ be an ergodic measure-preserving map on a probability space $(X, \mathfrak{F}, \mu)$, and take T as in (14.2). Show that

$$f \in \mathcal{M}^+(X), \ \int\limits_X f \, d\mu = +\infty \Longrightarrow \lim_{k\to\infty} \frac{1}{k} \sum_{j=1}^{k} T^j f(x) = +\infty, \ \mu\text{-a.e.}$$

Exercises 10–14 extend the Mean Ergodic Theorem to the following setting. Let $\mathcal{S}$ be a countably infinite semigroup, represented by a family of isometries on a Hilbert space H, so we have $\{T_\alpha : \alpha \in \mathcal{S}\}$, satisfying $T_\alpha : H \to H$, $T_\alpha^* T_\alpha = I$, $T_\alpha T_\beta = T_{\alpha\beta}$, for $\alpha, \beta \in \mathcal{S}$. Let $M_k \subset \mathcal{S}$ be a sequence of finite subsets of $\mathcal{S}$, of cardinality $\#M_k$. Assume that for each fixed $\gamma \in \mathcal{S}$,

$$\lim_{k\to\infty} \frac{\#(M_k \triangle M_k\gamma)}{\#M_k} = 0, \tag{14.57}$$

where $M_k\gamma = \{\alpha\gamma : \alpha \in M_k\}$ and $M_k \triangle M_k\gamma$ is the symmetric difference. Set

$$S_k f = \frac{1}{\#M_k} \sum_{\alpha \in M_k} T_\alpha f, \quad f \in H. \tag{14.58}$$

10. Show that there is an orthogonal direct sum decomposition

$$H = K \oplus \overline{R}, \tag{14.59}$$

where

$$K = \{f \in H : T_\alpha f = f, \forall\, \alpha \in \mathcal{S}\},$$
$$R = \bigoplus_{\alpha \in \mathcal{S}} \text{Range}\,(I - T_\alpha).$$

Hint. Show that $R^\perp = \bigcap_{\alpha \in \mathcal{S}} \text{Ker}\,(I - T_\alpha^*)$ and that $\text{Ker}\,(I - T_\alpha^*) = \text{Ker}\,(I - T_\alpha)$.

11. Show that $f \in K \Rightarrow S_k f \equiv f$.

12. Show that

$$\begin{aligned} f = (I - T_\gamma)v \Rightarrow S_k f &= \frac{1}{\# M_k} \sum_{\alpha \in M_k} (T_\alpha v - T_{\alpha\gamma} v) \\ &= \frac{1}{\# M_k} \sum_{\alpha \in M_k \triangle M_k \gamma} (\pm T_\alpha v). \end{aligned} \tag{14.60}$$

Use hypothesis (14.57) to deduce that $S_k f \to 0$ as $k \to \infty$.

13. Now establish the following mean ergodic theorem, namely, under the hypothesis (14.57),

$$f \in H \Longrightarrow S_k f \to Pf, \tag{14.61}$$

in H-norm, where P is the orthogonal projection of H onto K.

14. In case $\mathcal{S} = \mathcal{S}_\infty$ is the group arising in Proposition 14.12, with action on $H = L^2(\Omega, \omega)$ given by (14.51), if we set

$$M_k = \{\sigma \in \mathcal{S}_\infty : \sigma(\ell) = \ell \text{ for } |\ell| > k\}, \tag{14.62}$$

show that hypothesis (14.57) holds, and hence the conclusion (14.61) holds. In this case, $Pf = \int_\Omega f\, d\omega$, by Proposition 14.12.