Throughout this chapter we assume \((X, \mathcal{F}, \mu)\) 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

\[
S \in \mathcal{F} \implies \varphi^{-1}(S) \in \mathcal{F} \quad \text{and} \quad \mu(\varphi^{-1}(S)) = \mu(S).
\]

The map \(\varphi\) defines a linear map \(T\) on functions:

\[
Tf(x) = f(\varphi(x)).
\]

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

\[
\int_X f(\varphi(x)) \, d\mu = \int_X f(x) \, d\mu.
\]

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:

\[
A_m f(x) = \frac{1}{m} \sum_{k=0}^{m-1} T^k f(x).
\]

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

\[
H = K \oplus \overline{R},
\]

where

\[
K = \text{Ker}(I - T^*) = \text{Ker}(I - T), \quad R = \text{Range}(I - T),
\]

and \( \overline{R} \) is the closure of \( R \).

Proof. First, note that

\[
R^\perp = \{ v \in H : (v, (I - T)w) = 0, \forall w \in H \}
\]

\[
= \{ v \in H : ((I - T^*)v, w) = 0, \forall w \in H \}
\]

\[
= \text{Ker}(I - T^*).
\]

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 \( \text{Ker}(I - T) \subseteq \text{Ker}(I - T^*) \). 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 \( f \in H \),

\[
A_m f = \frac{1}{m} \sum_{k=0}^{m-1} T^k f \to Pf,
\]

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

\[
\frac{1}{m} \sum_{k=0}^{m-1} T^k f = \frac{1}{m} (v - T^m v) \to 0, \quad \text{as } m \to \infty,
\]

and since the operator norm \( \|A_m\| \leq 1 \) 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 $\text{Ker}(I - T)$. Then, for $p \in [1, 2]$, $P$ extends to a continuous projection on $L^p(X, \mu)$, and

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

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

**Proof.** Note that the $L^p$-operator norm $\|A_m\|_{L^p(L^p)} \leq 1$ for each $m$, and since $\|g\|_{L^p} \leq \|g\|_{L^2}$ 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

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

Hence

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

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 $f$, $\|Qf\|_{L^p} \leq \|f\|_{L^p}$, 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 $A_n f(x)$, first by G. Birkhoff. This can be done via estimates of Yosida and Kakutani on the maximal function

$$A^# f(x) = \sup_{m \geq 1} A_m f(x) = \sup_{n \geq 1} A^#_n f(x),$$

where

$$A^#_n f(x) = \sup_{1 \leq m \leq n} A_m f(x).$$

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

\[(14.14)\quad E_n = \{x \in X : A_n^#f(x) \geq 0\}.\]

Then

\[(14.15)\quad \int_{E_n} f \, d\mu \geq 0.\]

Proof. For notational convenience, set

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

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

\[(14.16)\quad f \geq M_n f - T(M_n f)^+ .\]

Integrating (14.16) over $E_n$ yields

\[
\int_{E_n} f \, d\mu \geq \int_{E_n} (M_n f - T(M_n f)^+) \, d\mu \\
= \int_{E_n} (M_n f)^+ - T(M_n f)^+) \, d\mu \\
= \int_X (M_n f)^+ \, d\mu - \int_{E_n} T(M_n f)^+ \, d\mu \\
\geq \int_X (M_n f)^+ \, d\mu - \int_X T(M_n f)^+ \, d\mu = 0 ,
\]

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 $\lambda > 0$,

\begin{equation}
\mu(\{x \in X : A^\#_n f(x) \geq \lambda\}) \leq \frac{1}{\lambda} \|f\|_{L^1}.
\end{equation}

**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

\begin{equation}
\int_{E_{n\lambda}} (f - \lambda) \, d\mu \geq 0.
\end{equation}

Thus

\begin{equation}
\|f\|_{L^1} \geq \int_{E_{n\lambda}} f \, d\mu \geq \lambda \mu(E_{n\lambda}),
\end{equation}

as asserted in (14.18).

Note that

\begin{equation}
E_{n\lambda} \cap \{x \in X : A^\# f(x) \geq \lambda\} = E_{\lambda},
\end{equation}

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

\begin{equation}
A^\# f(x) = \sup_{m \geq 1} |A_m f(x)| \leq A^\# |f|(x).
\end{equation}

We have

\begin{equation}
\mu(\{x \in X : A^\# f(x) \geq \lambda\}) \leq \frac{1}{\lambda} \|f\|_{L^1}.
\end{equation}

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 $\varphi$ is a measure-preserving map, then, given $f \in L^1(X, \mu)$,

\begin{equation}
\lim_{m \to \infty} A_m f(x) = P f(x), \quad \mu\text{-a.e.}
\end{equation}

**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

\begin{equation}
g \in \text{Ker}(I - T), \quad h = (I - T)v, \quad \|f_1 - (g + h)\|_{L^2} \leq \frac{\varepsilon}{2}.
\end{equation}
Here $v \in L^2(X, \mu)$. It follows that
\[(14.26) \quad \|f - (g + h)\|_{L^1} \leq \varepsilon,\]
and we have
\[(14.27) \quad A_m f = A_m g + A_m h + A_m (f - g - h) = g + \frac{1}{m} (v - T^m v) + A_m (f - g - h).\]
Clearly $v(x)/m \to 0$, $\mu$-a.e., as $m \to \infty$. Also
\[(14.28) \quad \int \sum_{m \geq 1} \left| \frac{1}{m} T^m v(x) \right|^2 d\mu = \|v\|_{L^2}^2 \sum_{m \geq 1} \frac{1}{m^2} < \infty,\]
which implies $T^m v(x)/m \to 0$, $\mu$-a.e., as $m \to \infty$. We deduce that for each $\lambda > 0$,
\[(14.29) \quad \mu \left( \{ x \in X : \limsup A_m f(x) - \liminf A_m f(x) > \lambda \} \right)
\quad = \mu \left( \{ x \in X : \limsup A_m (f - g - h) - \liminf A_m (f - g - h) > \lambda \} \right)
\quad \leq \mu \left( \{ x \in X : A^\# (f - g - h) > \frac{\lambda}{2} \} \right)
\quad \leq \frac{2}{\lambda} \|f - g - h\|_{L^1}
\quad \leq \frac{2\varepsilon}{\lambda}.
\]
Since $\varepsilon$ can be taken arbitrarily small, this implies that $A_m f(x)$ converges as $m \to \infty$, $\mu$-a.e. We already know it converges to $P f(x)$ 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
\[(14.30) \quad \|A^\# f\|_{L^\infty} \leq \|f\|_{L^\infty}.\]
Now the Marcinkiewicz Interpolation Theorem (see Appendix D) applied to (14.23) and (14.30) yields
\[(14.31) \quad \|A^\# f\|_{L^p} \leq C_p \|f\|_{L^p}, \quad 1 < p < \infty.\]
Using this, we prove the following.
Proposition 14.7. In the setting of Proposition 14.3, we have, for all \( p \in [1, \infty) \),
\[
f \in L^p(X, \mu) \implies A_m f \to Pf, \quad \text{in } L^p\text{-norm},
\]
as \( m \to \infty \).

Proof. Take \( p \in (1, \infty) \). Given \( f \in L^p(X, \mu) \), we have
\[
|A_m f(x)| \leq A^# f(x), \quad A^# f \in L^p(X, \mu).
\]
Since the convergence (14.24) holds pointwise \( \mu\text{-a.e.} \), (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
\[
f \in L^p(X, \mu) \implies A^*_m f \to Pf, \quad \text{weak}^* \text{ in } L^p(X, \mu),
\]
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 \( \mu\text{-a.e.} \)),
for \( p \in [1, \infty) \). Of course if \( \varphi \) is invertible, then such a result is a simple
application of the results given above, with \( \varphi \) replaced by \( \varphi^{-1} \).

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 \( \text{Ker}(I - T) \), i.e., with the set of
invariant functions, where we say \( f \in L^p(X, \mu) \) is invariant if and only if
\[
f(x) = f(\varphi(x)), \quad \mu\text{-a.e.}
\]
We note that the following conditions are equivalent:
\[
\begin{align*}
\text{(a)} & \quad f \in L^1(X, \mu) \text{ invariant } \implies f \text{ constant (\( \mu\text{-a.e.} \)),} \\
\text{(b)} & \quad f \in L^2(X, \mu) \text{ invariant } \implies f \text{ constant (\( \mu\text{-a.e.} \)),} \\
\text{(c)} & \quad S \in \mathcal{F} \text{ invariant } \implies \mu(S) = 0 \text{ or } \mu(S) = 1.
\end{align*}
\]
Here we say \( S \in \mathcal{F} \) is invariant if and only if
\[
\mu(\varphi^{-1}(S) \triangle S) = 0,
\]
where \( A \triangle B = (A \setminus B) \cup (B \setminus A) \). Note that if \( S \in \mathcal{F} \) satisfies (14.35), then
\[
\tilde{S} = \bigcap_{j \geq 0} \bigcup_{k \geq j} \varphi^{-k}(S) \implies \varphi^{-1}(\tilde{S}) = \tilde{S} \quad \text{and} \quad \mu(\tilde{S} \triangle S) = 0.
\]
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)\( \implies \) (a). Meanwhile
clearly (a)\( \implies \) (b)\( \implies \) (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
\[
A_m f \longrightarrow \int_X f \, d\mu, \quad \text{in } L^p\text{-norm and } \mu\text{-a.e.}
\]

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

**Proposition 14.9.** The map \( R_\alpha \) 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
\[
\hat{T}f(k) = (Tf, e_k) = (f, T^{-1}e_k) = e^{i\alpha} \hat{f}(k).
\]
Thus
\[
Tf = f, \quad \text{if } \hat{f}(k) \neq 0 \implies e^{i\alpha} = 1.
\]
But \( e^{i\alpha} = 1 \) for some nonzero \( k \in \mathbb{Z} \) if and only if \( \alpha/2\pi \) is rational.

In the next example, let \( (X, \mathcal{B}, \mu) \) be a probability space, and form the two-sided infinite product
\[
\Omega = \prod_{k=-\infty}^{\infty} X,
\]
which comes equipped with a \( \sigma \)-algebra \( \mathcal{O} \) and a product measure \( \omega \), via the construction given at the end of Chapter 6. There is a map on \( \Omega \) called the two-sided shift:
\[
\Sigma : \Omega \to \Omega, \quad \Sigma(x)_k = x_{k+1}, \quad x = (\ldots, x_{-1}, x_0, x_1, \ldots).
\]

**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 \( \alpha = (\ldots, \alpha_{-1}, \alpha_0, \alpha_1, \ldots) \) such that \( \alpha_k \neq 0 \) for only finitely many \( k \). Set
\[
v_\alpha(x) = \prod_{k=-\infty}^{\infty} u_{\alpha_k}(x_k), \quad \alpha \in \mathcal{A},
\]
and note that for each \( \alpha \in \mathcal{A} \) only finitely many factors in this product are not \( \equiv 1 \). We have the following:

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

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

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

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

\[ (14.45) \quad \hat{f}(\alpha) = (f, v_\alpha) = (Tf, Tv_\alpha) = \hat{f}(\sigma(\alpha)), \]

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

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

and \( \{ \sigma^\ell(\alpha) : \ell \in \mathbb{Z}^+ \} \) is an infinite set except for \( \alpha = 0 = (\ldots, 0, 0, 0, \ldots) \), 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

\[ (14.47) \quad \Omega_0 = \prod_{k=0}^{\infty} X, \]

with \( \sigma \)-algebra \( \mathcal{O}_0 \) and product measure \( \omega_0 \) constructed in the same fashion. As in (14.41), one sets

\[ (14.48) \quad \Sigma_0 : \Omega_0 \to \Omega_0, \quad \Sigma_0(x)_k = x_{k+1}, \quad x = (x_0, x_1, x_2, \ldots). \]

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 \( f(x_1, x_2, x_3, \ldots) = f(x_{k+1}, x_{k+2}, x_{k+3}, \ldots) \). Multiplying both sides by \( g(x_1, \ldots, x_k) \) 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 \( g = g(x_1, \ldots, x_k) \), 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 $S$ of maps on $X$ satisfying (14.1) for each $\varphi \in S$ and such that

$$\varphi, \psi \in S \implies \varphi \circ \psi \in S.$$  

(14.49)

In such a case, one says a function $f \in L^p(X, \mu)$ is invariant provided (14.33) holds for each $\varphi \in S$, one says $S \in \mathfrak{F}$ is invariant provided (14.35) holds for all $\varphi \in S$, and one says the action of $S$ 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 $S = \{\varphi^k : k \in \mathbb{Z}^+\}$. 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 $S_{\infty}$ 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 $\omega$. The group $S_{\infty}$ acts on $\Omega$ by

$$\varphi_{\sigma} : \varphi_{\sigma}(x) = x_{\sigma(k)},$$

(14.50)

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

**Proposition 14.12.** The action of $S_{\infty}$ on $\Omega$ defined by (14.50) is ergodic.

**Proof.** Let $\{v_\alpha : \alpha \in A\}$ 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

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

(14.51)

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

$$\hat{T}_\sigma f = f(\varphi_{\sigma}(x)) = f(\varphi_{\sigma}(x)) = \hat{f}(\sigma^\# \alpha), \quad \forall \alpha \in A, \quad \sigma \in S_{\infty}.$$  

(14.52)

Since $\|f\|_{L^2}^2 = \sum_{\alpha} \|\hat{f}(\alpha)\|^2 < \infty$ and $\{\sigma^\# \alpha : \sigma \in S_{\infty}\}$ is an infinite set, for each nonzero $\alpha \in A$, it follows that $\hat{f}(\alpha) = 0$ for nonzero $\alpha \in A$, 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, $A$ is the set defined in the beginning of the proof of Proposition 14.10.
Proposition 14.13. Let $G$ be a group of bijective maps on $\mathbb{Z}$ with the property that

$$\{\sigma^\# \alpha : \sigma \in G\}$$

is an infinite set, for each nonzero $\alpha \in A$, where $\sigma^\# \alpha$ is given by (14.51). Then the action of $G$ on $\Omega$, 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 $(e^{-\alpha_1}, \ldots, e^{i\alpha_n}) \in \mathbb{T}^n$, define

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

Give necessary and sufficient conditions that $R_\alpha$ 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 $\varphi$ is ergodic. 

*Hint.* Examine the Fourier series of an invariant function.

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

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

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

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

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

4. Show that a mixing transformation is ergodic.

*Hint.* Show that

$$\langle A_k f, g \rangle = \frac{1}{k} \sum_{j=0}^{k-1} \langle T^j f, g \rangle \to \langle f, 1 \rangle(1, g).$$
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 $\Sigma$ and $\Sigma_0$, given in (14.41) and (14.48), are mixing.  
   *Hint.* Verify (14.55) when $f$ and $g$ are elements of the orthonormal basis $\{\upsilon_\alpha\}$ described in (14.42). Alternatively, verify (14.55) when $f$ and $g$ are functions of $x_j$ for $|j| \leq M$.

7. Show that the maps $R_\alpha$ 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 $\Omega_0$ as giving the binary expansion of a number $x \in [0, 1)$.

9. Let $\varphi$ be an ergodic measure-preserving map on a probability space $(X, \mathcal{F}, \mu)$, and take $T$ as in (14.2). Show that
   \[
   f \in \mathcal{M}^+(X), \quad \int_X f \, d\mu = +\infty \implies \lim_{k \to \infty} \frac{1}{k} \sum_{j=1}^{k} T^j f(x) = +\infty, \ \mu\text{-a.e.}
   \]

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 $\gamma \in \mathcal{S}$,

\begin{equation}
\lim_{k \to \infty} \frac{\# (M_k \triangle M_k \gamma)}{\# M_k} = 0, 
\end{equation}

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

\begin{equation}
S_k f = \frac{1}{\# M_k} \sum_{\alpha \in M_k} T_\alpha f, \quad f \in H.
\end{equation}
10. Show that there is an orthogonal direct sum decomposition

\[ H = K \oplus \overline{R}, \]

where

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

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

\[ f = (I - T_\gamma)v \Rightarrow S_k f = \frac{1}{\#M_k} \sum_{\alpha \in M_k} (T_\alpha v - T_{\alpha, \gamma} v) \]

\[ = \frac{1}{\#M_k} \sum_{\alpha \in M_k, \Delta M_k, \gamma} (\pm T_\alpha v). \]

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),

\[ f \in H \implies S_k f \to Pf, \]

in \( H \)-norm, where \( Pf \) is the orthogonal projection of \( H \) onto \( K. \)

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

\[ M_k = \{ \sigma \in S_{\infty} : \sigma(\ell) = \ell \text{ for } |\ell| > k \}, \]

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.