Exercise 4.2.5. Let $T : X \to X$ be a continuous map on a compact metric space. Show that the measures in $\mathscr{E}^T(X)$ constrain all the ergodic averages in the following sense. For $f \in C(X)$, define

$$m(f) = \inf_{\mu \in \mathscr{E}^T(X)} \left\{ \int f \, \mathrm{d}\mu \right\}$$

and

$$M(f) = \sup_{\mu \in \mathscr{E}^T(X)} \left\{ \int f \, \mathrm{d}\mu \right\}.$$

Prove that

$$m(f) \leqslant \liminf_{N \to \infty} \mathsf{A}_N^f(x) \leqslant \limsup_{N \to \infty} \mathsf{A}_N^f(x) \leqslant M(f)$$

for any $x \in X$.

4.3 Unique Ergodicity

A natural distinguished class of transformations are those for which there is only one invariant Borel measure. This measure is automatically ergodic, and the uniqueness of this measure has several powerful consequences.

Definition 4.9. Let X be a compact metric space and let $T : X \to X$ be a continuous map. Then T is said to be *uniquely ergodic* if $\mathscr{M}^T(X)$ comprises a single measure.

Theorem 4.10. *For a continuous map $T : X \to X$ on a compact metric space, the following properties are equivalent.*

(1) *T is uniquely ergodic.*
(2) $|\mathscr{E}^T(X)| = 1$.
(3) *For every $f \in C(X)$,*

$$\mathsf{A}_N^f = \frac{1}{N} \sum_{n=0}^{N-1} f(T^n x) \longrightarrow C_f, \tag{4.4}$$

where C_f is a constant independent of x.

(4) *For every $f \in C(X)$, the convergence* (4.4) *is uniform across X.*
(5) *The convergence* (4.4) *holds for every f in a dense subset of $C(X)$.*

Under any of these assumptions, the constant C_f in (4.4) *is $\int_X f \, \mathrm{d}\mu$, where μ is the unique invariant measure.*

We will make use of Theorem 4.8 for the equivalence of (1) and (2); the equivalence between (1) and (3)–(5) is independent of it.

Proof of Theorem 4.10. (1) $\iff$ (2): If T is uniquely ergodic and μ is the only T-invariant probability measure on X, then μ must be ergodic by Theorem 4.4. If there is only one ergodic invariant probability measure on X, then by Theorem 4.8, it is the only invariant probability measure on X.

(1) $\implies$ (3): Let μ be the unique invariant measure for T, and apply Theorem 4.1 to the constant sequence (δ_x). Since there is only one possible limit point and $\mathscr{M}(X)$ is compact, we must have

$$\frac{1}{N}\sum_{n=0}^{N-1}\delta_{T^n x}\longrightarrow\mu$$

in the weak*-topology, so for any $f\in C(X)$

$$\frac{1}{N}\sum_{n=0}^{N-1}f(T^n x)\longrightarrow\int_X f\,\mathrm{d}\mu.$$

(3) $\implies$ (1): Let $\mu\in\mathscr{M}^T(X)$. Then by the dominated convergence theorem, (4.4) implies that

$$\int_X f\,\mathrm{d}\mu=\int_X\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}f(T^n x)\,\mathrm{d}\mu=C_f$$

for all $f\in C(X)$. It follows that C_f is the integral of f with respect to any measure in $\mathscr{M}^T(X)$, so $\mathscr{M}^T(X)$ can only contain a single measure.

Notice that this also shows $C_f=\int_X f\,\mathrm{d}\mu$ for the unique measure μ.

(1) $\implies$ (4): Let $\mu\in\mathscr{M}^T(X)$, and notice that we must have $C_f=\int f\,\mathrm{d}\mu$ as above. If the convergence is not uniform, then there is a function g in $C(X)$ and an $\varepsilon>0$ such that for every N_0 there is an $N>N_0$ and a point $x_j\in X$ for which

$$\left|\frac{1}{N}\sum_{n=0}^{N-1}g(T^n x_j)-C_g\right|\geqslant\varepsilon.$$

Let $\mu_N=\frac{1}{N}\sum_{n=0}^{N-1}\delta_{T^n x_j}$, so that

$$\left|\int_X g\,\mathrm{d}\mu_N-C_g\right|\geqslant\varepsilon. \tag{4.5}$$

By weak*-compactness the sequence (μ_N) has a subsequence $\bigl(\mu_{N(k)}\bigr)$ with

$$\mu_{N(k)}\to\nu$$

as $k\to\infty$. Then $\nu\in\mathscr{M}^T(X)$ by Theorem 4.1, and

$$\left|\int_X g\,\mathrm{d}\nu-C_g\right|\geqslant\varepsilon$$

by equation (4.5). However, this shows that $\mu \neq \nu$, which contradicts (1).

(4) $\Longrightarrow$ (5): This is clear.

(5) $\Longrightarrow$ (1): If $\mu, \nu \in \mathscr{E}^T(X)$ then, just as in the proof that (3) $\Longrightarrow$ (1),

$$\int_X f \, \mathrm{d}\nu = C_f = \int_X f \, \mathrm{d}\mu$$

for any function f in a dense subset of $C(X)$, so $\nu = \mu$. □

The equivalence of (1) and (3) in Theorem 4.10 appeared first in the paper of Kryloff and Bogoliouboff [214] in the context of uniquely ergodic flows.

Example 4.11. The circle rotation $R_\alpha : \mathbb{T} \to \mathbb{T}$ is uniquely ergodic if and only if α is irrational. The unique invariant measure in this case is the Lebesgue measure $m_{\mathbb{T}}$. This may be proved using property (5) of Theorem 4.10 (or using property (1); see Theorem 4.14). Assume first that α is irrational, so $\mathrm{e}^{2\pi \mathrm{i} k\alpha} = 1$ only if $k = 0$. If $f(t) = \mathrm{e}^{2\pi \mathrm{i} kt}$ for some $k \in \mathbb{Z}$, then

$$\frac{1}{N}\sum_{n=0}^{N-1} f(R_\alpha^n t) = \frac{1}{N}\sum_{n=0}^{N-1} \mathrm{e}^{2\pi \mathrm{i} k(t+n\alpha)} = \begin{cases} 1 & \text{if } k = 0; \\ \dfrac{1}{N}\mathrm{e}^{2\pi \mathrm{i} kt}\dfrac{\mathrm{e}^{2\pi \mathrm{i} Nk\alpha} - 1}{\mathrm{e}^{2\pi \mathrm{i} k\alpha} - 1} & \text{if } k \neq 0. \end{cases} \tag{4.6}$$

Equation (4.6) shows that

$$\frac{1}{N}\sum_{n=0}^{N-1} f(R_\alpha^n t) \longrightarrow \int f \, \mathrm{d}m_{\mathbb{T}} = \begin{cases} 1 \text{ if } k = 0; \\ 0 \text{ if } k \neq 0. \end{cases}$$

By linearity, the same convergence will hold for any trigonometric polynomial, and therefore property (5) of Theorem 4.10 holds. For a curious application of this result, see Example 1.3.

If α is rational, then Lebesgue measure is invariant but not ergodic, so there must be other invariant measures.

Example 4.11 may be used to illustrate the ergodic decomposition of a particularly simple dynamical system.

Example 4.12. Let $X = \{z \in \mathbb{C} \mid |z| = 1 \text{ or } 2\}$, let α be an irrational number, and define a continuous map $T : X \to X$ by $T(z) = \mathrm{e}^{2\pi \mathrm{i}\alpha} z$. By unique ergodicity on each circle, any invariant measure μ takes the form

$$\mu = sm_1 + (1 - s)m_2,$$

where m_1 and m_2 denote Lebesgue measures on the two circles comprising X. Thus $\mathscr{M}^T(X) = \{sm_1 + (1-s)m_2 \mid s \in [0, 1]\}$, with the two ergodic measures given by the extreme points $s = 0$ and $s = 1$. The decomposition of μ is described by the measure $\nu = s\delta_{m_1} + (1 - s)\delta_{m_2}$. A convenient notation for this is $\mu = \int_{\mathscr{M}^T(X)} m \, \mathrm{d}\nu(m)$.