24. ERGODICITY

The ergodic theorems tell us that "ergodic averages" of any integrable f converge to a function $\bar{f}$ that is invariant under the action of the m.p.t. However, we do not learn much about how $\bar{f}$ depends on its argument and, taken statistically, how wildly is $\bar{f}$ "distributed" over the underlying measure space. This is what we will address here.

24.1 Invariant sets and ergodicity.

We start by introducing notions that encode the invariance property of $\bar{f}$:

Definition 24.1 *Let $(\mathscr{X}, \mathcal{G}, \mu, \varphi)$ be a measure-preserving system. A set $A \in \mathcal{G}$ is*

(1) invariant *if $\varphi^{-1}(A) = A$, and*
(2) almost invariant *if $\varphi^{-1}(A) = A$ μ-a.e., which means $\mu(\varphi^{-1}(A) \triangle A) = 0$.*

We then have:

Lemma 24.2 *The families*

$$\mathscr{I} := \{A \in \mathcal{G} \colon A \text{ invariant}\} \tag{24.1}$$

and

$$\mathscr{I}' := \{A \in \mathcal{G} \colon A \text{ almost invariant}\} \tag{24.2}$$

are σ-algebras. Moreover, for all $A \in \mathcal{G}$ and all measurable $f \colon \mathscr{X} \to \mathbb{R}$,

(1) $\forall A \in \mathscr{I}' \, \exists B \in \mathscr{I} \colon \mu(A \triangle B) = 0$,
(2) $f \circ \varphi = f$ *μ-a.e. $\Rightarrow$ there exists g such that $f = g$ μ-a.e. and g is $\mathscr{I}$-measurable.*

Proof. That $\mathscr{I}$ and $\mathscr{I}'$ are σ-algebras is checked directly. Fir (1) let $A \in \mathcal{F}$ and set

$$B := \bigcap_{n \geqslant 0} \bigcup_{k \geqslant n} \varphi^{-k}(A) = \big\{x \in \mathscr{X} \colon \varphi^n(x) \in A \text{ i.o.}\big\} \tag{24.3}$$

Then $\varphi^{-1}(B) = B$ and so $B \in \mathscr{I}$. Since $\varphi^{-k}(A) = A$ μ-a.e., we have $B = A$ μ-a.e. We leave the proof of (2) to a homework exercise. $\square$

The previous lemma shows that we can pretty much neglect the difference between almost invariant and invariant events. Next we will give a probabilistic interpretation to the limit object in the Pointwise Ergodic Theorem:

Corollary 24.3 (Birkhoff-Khinchin theorem) *Let $(\mathscr{X}, \mathcal{G}, \mu)$ be a probability space and let $\varphi \colon \mathscr{X} \to \mathscr{X}$ be a measure-preserving transformation. Then for all $X \in L^1$,*

$$\frac{1}{n} \sum_{k=0}^{n-1} f \circ \varphi^k \underset{n \to \infty}{\longrightarrow} E(f|\mathscr{I}), \qquad \text{a.s.} \tag{24.4}$$

Proof. Since $\mu(\mathscr{X})$ the convergence of ergodic averages takes place in L^1. If $A \in \mathscr{I}$, then (writing expectations for integrals with respect to μ) for all $k \geqslant 0$,

$$E(f 1_A) = E(f \circ \varphi^k \, 1_A \circ \varphi^k) = E(f \circ \varphi^k 1_A) \tag{24.5}$$

by the fact that μ is preserved by φ^k and that $1_A \circ \varphi = 1_A$. Averaging over $k = 0, \dots, n-1$ then gives

$$E(1_A f) = E\left(1_A \frac{1}{n} \sum_{k=0}^{n-1} f \circ \varphi^k\right) \underset{n\to\infty}{\longrightarrow} E(1_A \bar{f}). \tag{24.6}$$

But $\bar{f}$ has a $\mathscr{I}$-measurable version and so $\bar{f} = E(f|\mathscr{I})$ μ-a.s. □

The interpretation of the limit as conditional expectation now suggests:

Definition 24.4 (Ergodicity) *An m.p.t. φ on $(\mathscr{X}, \mathcal{G}, \mu)$ is said to be* ergodic with respect to μ *(or, equivalently, the measure μ is* ergodic with respect to a m.p.t. φ*) if μ is trivial on φ-invariant sets in the sense that*

$$\forall A \in \mathscr{I}: \quad \mu(A) = 0 \vee \mu(A^c) = 0 \tag{24.7}$$

If both φ and μ are given, we also say that that the m.p.s. is ergodic *if* (24.7) *holds.*

We adopt the convention (not always heeded in the literature) that ergodicity of a measure μ entails invariance under φ. (Otherwise, we would have to say "invariant and ergodic" which is longer.) Note that, for μ a probability measure, we can consolidate the defining property as

$$\mu \text{ is ergodic} \quad \Leftrightarrow \quad \exists A \in \mathscr{I}: \ \mu(A) \in \{0, 1\} \tag{24.8}$$

which gives ergodicity the structure of a zero-one law. The above definition is designed to include infinite measures as well.

24.2 Examples and non-examples.

Let us now give some examples in which ergodicity holds. We start with:

Lemma 24.5 *i.i.d. shifts are ergodic. More precisely, if $(S^{\mathbb{N}}, \Sigma^{\mathbb{N}}, \nu^{\otimes\mathbb{N}})$ is a product probability space and $\theta\colon S^{\mathbb{N}} \to S^{\mathbb{N}}$ is the left shift, then θ is ergodic.*

Proof. Let X_k be the coordinate projection on the k-th coordinate. Given $A \in \mathscr{I}$, we have $A = \theta^{-n}(A)$ for each $n \geqslant 1$. Since, equivalently, $A \in \sigma(X_0, X_1, \dots)$ this gives that $A = \theta^{-n}(A) \in \sigma(X_n, X_{n+1}, \dots)$ for all $n \geqslant 0$ and so

$$A \in \bigcap_{n \geqslant 1} \sigma(X_n, X_{n+1}, \dots) \tag{24.9}$$

i.e., A is a tail event. By Komogorov's zero-one law, A has probability either zero or one. Hence θ is ergodic. □

Hereby we get:

Corollary 24.6 *The "continued fractions" (Example 20.10) and "Baker's transform" (Example 20.12) m.p.s. are ergodic.*

This follows by the fact that both of these systems are "coded" (which technically means "bijectively bimesurable correspondence" with) i.i.d. sequences. We leave the details to the reader.

We note that encoding of a measure-preserving system in terms of a Bernoulli shift is a recurrent theme in ergodic theory. In fact, systems for which such an encoding is possible are referred to as *Bernoulli*. The question often asked is: What invariants (i.e., properties preserved by above measure-preserving bijections) are sufficient to check to see that a system is Bernoulli. One of these is entropy.

Another example we already analyzed comes in:

Lemma 24.7 *Irrational rotations of the circle are ergodic.*

Proof. Let $\varphi_\alpha\colon [0,1) \to [0,1)$ be defined by $\varphi_\alpha(x) := x + \alpha \bmod 1$ for some $\alpha \in \mathbb{R} \smallsetminus \mathbb{Q}$. In Theorem 21.5(1) we showed that $f \circ \varphi_\alpha = f$ implies that f is constant Lebesgue a.e.. Using this for $f := 1_A$, where $A \in \mathscr{I}$ we get that the Lebesgue measure of A is either zero or one, as desired. □

Insofar we have only mentioned examples in which μ was a probability measure. Here is an example in which μ is infinite:

Theorem 24.8 (Boole transform) *Consider the measure space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ where λ is the Lebesgue measure. Let $\varphi\colon \mathbb{R} \to \mathbb{R}$ be the map*

$$\varphi(x) := \begin{cases} x - 1/x, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0, \end{cases} \tag{24.10}$$

Then φ is a m.p.t. which, moreover, is ergodic.

The proof of the measure-preserving property is relegated to homework. The proof of ergodicity is more difficult and can be found in a paper by R.L. Adler and B. Weiss "The ergodic measure-preserving transformation of Boole" in Israel J. Math, 1973.

Having given some examples, let us move to non-examples:

Lemma 24.9 *Let (S, Σ) be a standard Borel space and let $\{X_k\}_{k\geqslant 1}$ be exchangeable S-valued random variables. Use μ to denote their law on $(S^{\mathbb{N}}, \Sigma^{\otimes\mathbb{N}})$ and let θ denote the left shift. Then θ preserves μ and*

$$\mu \text{ is ergodic} \quad\Leftrightarrow\quad \mu \text{ is i.i.d.} \tag{24.11}$$

In particular, all exchangeable measures on $(S^{\mathbb{N}}, \Sigma^{\otimes\mathbb{N}})$ that are not i.i.d. are not ergodic.

Proof. Exchangeability ensures that

$$(X_1, \dots, X_n) \overset{\text{law}}{=} (X_2, \dots, X_{n+1}) = (X_1, \dots, X_n) \circ \theta \tag{24.12}$$

proving that $\mu \circ \theta = \mu$ on cylinder sets. Since the cylinder sets generate the σ-algebra $\Sigma^{\otimes\mathbb{N}}$, the equality holds on $\Sigma^{\otimes\mathbb{N}}$.

For the second part of the statement, by Lemma 24.5 it suffices to prove the implication $\Rightarrow$. Let $A \in \Sigma$, set $S_n := \sum_{k=1}^n 1_A(X_k)$ and note that $S_n \circ \theta = \sum_{k=2}^{n+1} 1_A(X_k)$. Letting $U := \limsup_{n\to\infty} \frac{S_n}{n}$, de Finetti's Theorem and a simple argument then tell us

$$\frac{S_n}{n} \underset{n\to\infty}{\longrightarrow} U \text{ a.s.} \;\wedge\; \frac{S_n}{n} \circ \theta \underset{n\to\infty}{\longrightarrow} U \text{ a.s.} \tag{24.13}$$

It follows that $U \circ \theta = U$ μ-a.s. Since μ is assumed ergodic, we have $U = \mu(A)$ a.s.

Now recall that the de Finetti-Hewitt-Savage Theorem (Theorem 18.8) gives

$$\mu(\cdot) = \int \nu^{\otimes\mathbb{N}}(\cdot)\rho(\mathrm{d}\nu) \tag{24.14}$$

Since

$$\nu^{\otimes\mathbb{N}}\Bigl(\lim_{n\to\infty}\frac{S_n}{n} = \nu(A)\Bigr) = 1 \tag{24.15}$$

we have

$$\forall t\in[0,1]\colon \quad \mu(U \leqslant t) = \rho\bigl(\nu\in\mathcal{M}_1(S)\colon \nu(A)\leqslant t\bigr) \tag{24.16}$$

But $t \mapsto \mu(U \leqslant t)$ jumps from 0 to 1 at $\mu(A)$ which is only possible if $\nu(A) = \mu(A)$ for ρ-a.e. ν. This means that ρ is a singleton and μ is thus i.i.d. □

Another non-example comes in:

Lemma 24.10 *Rational rotations (by α) of the circle are NOT ergodic w.r.t. the Lebesgue measure but are ergodic with respect to measures of the form*

$$\frac{1}{q}\sum_{k=0}^{q-1}\delta_{z+k/q \bmod 1} \tag{24.17}$$

where $z\in\mathbb{R}$ and $q\in\mathbb{N}$ is such that $\alpha = p/q$ with $0\leqslant p<q$ and $gcd(p,q) = 1$. (We thus have a continuum of ergodic measures in this case.)

Proof. Let $\alpha = p/q$ with p and q as above. Then for each $x\in[0,1)$ and each $f\colon[0,1)\to\mathbb{R}$,

$$\frac{1}{n}\sum_{k=0}^{n-1} f\circ\varphi_\alpha(x) \underset{n\to\infty}{\longrightarrow} \frac{1}{q}\sum_{k=0}^{q-1} f(x+k/q \bmod 1) \tag{24.18}$$

To see this observe that $kp/q\in\mathbb{N}$ and $\gcd(p,q)=1$ force $k\in q\mathbb{N}$. Stopping the sum on the left at the largest multiple of q less or equal than $n-1$ then gives the left-hand side plus errors of order q/n. Since the right-hand side is definitely NOT constant Lebesgue a.e., the Lebesgue measure is NOT ergodic. But the quantity is constant w.r.t. the measure (24.17) a.s. and so that measure is ergodic. □

A non-example with infinite mass is easy to construct: Take $(\mathbb{R},\mathcal{B}(\mathbb{R}),\lambda)$ with λ the Lebesgue measure and let $\varphi(x) := x+1$. Then any 1-periodic function is invariant and so φ is NOT ergodic.

24.3 Necessary and sufficient conditions.

We now proceed to give necessary and sufficient conditions for ergodicity. We start with:

Lemma 24.11 *Let $(\mathscr{X},\mathcal{G},\mu,\varphi)$ be an m.p.s. Then the following are equivalent:*

(1) *φ is ergodic*
(2) *for all measurable $h\colon\mathscr{X}\to\mathbb{R}$,*

$$h = h\circ\varphi \text{ a.e.} \;\Rightarrow\; \exists c\in\mathbb{R}\colon h = c \text{ a.e.} \tag{24.19}$$

It actually suffices to check (2) for $h\in L^\infty$ and, in fact, h of the form $h = 1_A$.

Proof. If (2) holds then for each $A \in \mathscr{I}$, for which $1_A = 1_A \circ \varphi$, we have $1_A = c$ a.e. But this is only possible if $c \in \{0,1\}$. Hence, either $\mu(A) = 0$ or $\mu(A^c) = 0$ implying (1). Since only indicators are involved, it is enough to check (2) for indicators.

Conversely, suppose (1) and let h obey $h = h \circ \varphi$ a.e. Define

$$c := \sup\{a \in \mathbb{R} \colon \mu(h < a) = 0\} \tag{24.20}$$

Since $\{h < c\} = \bigcup_{n \geqslant 1}\{h < c - 2^{-n}\}$, we get $\mu(h < c) = 0$ by the Monotone Convergence Theorem. This rules out that $c = +\infty$ because h is assumed to be real valued. Next, $\mu(h \leqslant c + 2^{-n}) > 0$ and, since μ is ergodic, we thus have $\mu(h > c + 2^{-n}) = 0$ for all $n \geqslant 1$. Using that $\{h > c\} = \bigcup_{n \geqslant 1}\{h > c + 2^{-n}\}$, the Monotone Convergence Theorem again gives $\mu(h > c) = 0$. It follows that $\mu(h \neq c) = 0$, proving (2). □

For finite measure spaces we get the following criterion:

Lemma 24.12 *Let $(\mathscr{X}, \mathcal{G}, \mu, \varphi)$ be a m.p.s. with $\mu(\mathscr{X}) < \infty$. Then*

$$\varphi \text{ ergodic} \quad\Leftrightarrow\quad \forall f \in L^1\colon\ \frac{1}{n}\sum_{k=0}^{n-1} f \circ \varphi^k \underset{n\to\infty}{\longrightarrow} \frac{1}{\mu(\mathscr{X})}\int f \mathrm{d}\mu \ \text{ a.e.} \tag{24.21}$$

Proof. By the Pointwise Ergodic Theorem, the "ergodic averages" converge to $\bar f$ a.e. Since $\mu(\mathscr{X}) < \infty$, the convergence is in L^1 and so $\int \bar f \mathrm{d}\mu = \int f \mathrm{d}\mu$. Ergodicity then implies that $\bar f$ is constant a.e. and so $\mu(\mathscr{X})\bar f = \int f \mathrm{d}\mu$. On the other hand, if the claim on the right of (24.21) holds for $f = 1_A$ with $A \in \mathscr{I}$, then $1_A = \mu(A)/\mu(\mathscr{X})$, i.e., either $\mu(A) = 0$ or $\mu(A^c) = 0$ and so φ is ergodic. □

Note that the need for division by $\mu(\mathscr{X})$ is not why we had to restrict to $\mu(\mathscr{X}) < \infty$. Indeed, the implication $\Rightarrow$ holds in this case as well:

Lemma 24.13 *Let $(\mathscr{X}, \mathcal{G}, \mu, \varphi)$ be a m.p.s. with $\mu(\mathscr{X}) = \infty$. Then*

$$\varphi \text{ ergodic} \quad\Rightarrow\quad \forall f \in L^1\colon\ \frac{1}{n}\sum_{k=0}^{n-1} f \circ \varphi^k \underset{n\to\infty}{\longrightarrow} 0 \ \text{ a.e.} \tag{24.22}$$

Proof. The "ergodic averages" converge to $\bar f \in L^1$ a.e. which by Lemma 24.11 is constant a.e. and so must vanish a.e. because non-zero constants are not integrable. □

We note that the converse to (24.22) fails in general when $\mu(\mathscr{X}) = \infty$. This boils down to the familiar example $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda, \varphi)$, where λ is the Lebesgue measure and $\varphi(x) := x + 1$. Indeed, all φ-invariant functions are 1-periodic, and so is thus the limit of the "ergodic averages," but none of these is in L^1 except that which vanishes a.e. Hence, the limit on the right of (24.22) is zero a.e. for all $f \in L^1$ yet φ is not ergodic.

Another version of the criterion in Lemma 24.12 comes in:

Lemma 24.14 *Let $(\mathscr{X}, \mathcal{G}, \mu, \varphi)$ be a m.p.s. with $\mu(\mathscr{X}) = 1$. Then*

$$\varphi \text{ ergodic} \ \Leftrightarrow\ \forall A, B \in \mathcal{G}\colon\ \frac{1}{n}\sum_{k=0}^{n-1} \mu\big(\varphi^{-k}(A) \cap B\big) \underset{n\to\infty}{\longrightarrow} \mu(A)\mu(B) \tag{24.23}$$

Proof. Writing $\langle\cdot,\cdot\rangle$ for the inner product in L^2, we have

$$\frac{1}{n}\sum_{k=0}^{n-1}\mu\bigl(\varphi^{-k}(A)\cap B\bigr) = \Bigl\langle \frac{1}{n}\sum_{k=0}^{n-1} 1_A\circ\varphi^k, 1_B\Bigr\rangle \tag{24.24}$$

The first term converges to $P1_A$ in L^2 by the Mean Ergodic Theorem, where P is the projection on the subspace of invariant functions. The condition on the right is thus equivalent to

$$\forall A,B\in\mathcal{G}\colon\ \langle P1_A, 1_B\rangle = \mu(A)\mu(B) \tag{24.25}$$

Testing this against $B := \{Pf < \mu(A)\}$ and $B := \{Pf > \mu(A)\}$ implies and then is checked to be equivalent to

$$\forall A\in\mathcal{G}\colon\ P1_A = \mu(A) \tag{24.26}$$

By Lemma 24.12 this is equivalent to ergodicity. □

The last condition relies on spectral calculus:

Lemma 24.15 *Let $(\mathscr{X},\mathcal{G},\mu,\varphi)$ be a m.p.s. with $\mu(\mathscr{X})<\infty$. For $f\in L^2$ set $Tf := f\circ\varphi$. Then T is an isometry in L^2 such that the following are equivalent:*

(1) *φ is ergodic*
(2) *1 is a simple eigenvalue of T*
(3) *every eigenvalue of T is simple*

Moreover, the eigenvalues of T form a subgroup of $\{\mathrm{e}^{2\pi\mathrm{i}t}\colon t\in[0,1)\}$.

Proof. Recall that $\lambda\in\mathbb{C}$ is an eigenvalue of an operator T on L^2 if there exists $f\in L^2$ such that $Tf=\lambda f$. (Such an f is then called an eigenvector.) Alternatively, λ is an eigenvalue if $\operatorname{Ker}(\lambda - T)\neq\{0\}$. The eigenvalue λ is simple if $\dim\operatorname{Ker}(\lambda-T)=1$. Note also that the fact that T is an isometry means that all eigenvalues of T are complex modulus 1. This follows by noting that $Tf=\lambda f$ implies

$$\langle f,f\rangle = \langle Tf,Tf\rangle = |\lambda|^2\langle f,f\rangle \tag{24.27}$$

where the first line is by T being an isometry and the second line by the sequilinear character of the inner product (which, even in spaces over $\mathbb{R}$, has to be taken as over $\mathbb{C}$).

Note that 1 is always an eigenvalue because $1\in L^2$ and $T1=1$. If f is such that $Tf=f$ then f is invariant and so, since ergodicity is equivalent to f being a constant, it is equivalent to $\operatorname{Ker}(1-T)$ being one-dimensional. This proves (1) ⇔ (2). Since (3) ⇒ (2) is trivial, it remains to prove (1) ⇒ (3). Here we need to argue a bit more. First observe that, if $Tf=\lambda f$ then

$$|Tf| = |f\circ\varphi| = |f|\circ\varphi = T|f| \tag{24.28}$$

and so $|f|\in\operatorname{Ker}(1-T)$. But then $|f|$ is constant and, being non-zero, $f\neq 0$ a.s. Next set $Tg=\lambda' g$ and note that

$$T(g/f) = \frac{g\circ\varphi}{f\circ\varphi} = \frac{\lambda'}{\lambda}(g/f) \tag{24.29}$$

and so $g/f\in\operatorname{Ker}(\lambda'/\lambda - T)$. If $\lambda'=\lambda$ then $g/f\in\operatorname{Ker}(1-T)$ and g/f equals a constant, proving that $\operatorname{Ker}(\lambda-T)$ is one-dimensional as well, showing (1) ⇒ (3).

The last argument shows that if λ and λ' are eigenvalues, then so is λ'/λ. Since $Tf = \lambda f$ implies that the complex conjugate g of f obeys $Tg = \lambda^{-1}g$, for λ^{-1} being the complex conjugate of λ, we have shown that the set of eigenvalues is closed under multiplication and division and contains 1, and so is a subgroup of the circle group. $\square$

Two remarks are in order. First, the spectral representation of T gives us yet another way to prove the Mean Ergodic Theorem for invertible maps by methods similar to the proof of the Weil Equidistribution Theorem. (Indeed, if φ is invertible then T is unitary and, under the spectral measure, reduces to a multiplication operator by a λ with $|\lambda| = 1$ whose powers average to zero unless $\lambda = 1$.) Second, while the characterization using eigenvalues may appear useful, in practice this is obfuscated by the rest of the spectrum which is often the whole unit circle.

24.4 Ergodicity as irreducibility property.

As our final topic, we will note that ergodicity is basically an irreducibility property (similar as being i.i.d. was an irreducibility property among exchangeable measures). We start with:

Lemma 24.16 (Mutual singularity) *Let $(\mathscr{X}, \mathcal{G})$ be a measurable space, $\varphi\colon \mathscr{X} \to \mathscr{X}$ a measurable map and μ_1, μ_2 two φ-invariant ergodic probability measures. Then*

$$\mu_1 \perp \mu_2 \text{ on } \mathscr{I} \;\vee\; \mu_1 = \mu_2 \tag{24.30}$$

Proof. Let μ_1 and μ_2 be both φ-ergodic. If $\mu_1 \neq \mu_2$, then there exists $A \in \mathcal{G}$ such that $\mu_1(A) \neq \mu_2(A)$. For $i = 1, 2$ set

$$E_i := \left\{ \limsup_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} 1_A \circ \varphi^k = \mu_i(A) \right\} \tag{24.31}$$

Lemma 24.12 then gives

$$\mu_1(E_1) = 1 \;\wedge\; \mu_2(E_2) = 1 \tag{24.32}$$

But $E_1 \cap E_2 = \emptyset$ and $E_1, E_2 \in \mathscr{I}$ so $\mu_1 \perp \mu_2$ on $\mathscr{I}$. $\square$

As a consequence we get:

Lemma 24.17 (Extremality) *Let $(\mathscr{X}, \mathcal{G})$ be a measurable space, $\varphi\colon \mathscr{X} \to \mathscr{X}$ a measurable map and μ_1, μ_2, μ probability measures that are φ-invariant and such that*

$$\exists \alpha \in [0,1]\colon \quad \mu = \alpha\mu_1 + (1-\alpha)\mu_2 \tag{24.33}$$

Then

$$\mu \text{ ergodic } \Rightarrow\; \alpha \in \{0,1\} \;\vee\; \mu_1 = \mu_2 \tag{24.34}$$

In short, ergodic measures are extremal in the convex set of invariant measures.

Proof. Suppose μ is ergodic and $\alpha \in (0,1)$. Then $\mu(A) \in \{0,1\}$ implies $\mu_1(A) = \mu_2(A) \in \{0,1\}$ and so $\mu_1 = \mu_2$ on $\mathscr{I}$. Lemma 24.16 then gives $\mu_1 = \mu_2$ as desired. $\square$

One can push this further to show that every φ-invariant measure μ can be written as a unique convex combination of ergodic measures as

$$\mu(\cdot) = \int \nu(\cdot)\rho(\mathrm{d}\nu) \tag{24.35}$$

where

$$\rho\big(\{\nu \in \mathcal{M}_1(\mathscr{X})\colon \ \text{ergodic}\}\big) = 1 \tag{24.36}$$

Such a statement is referred to as *ergodic decomposition*. The precise formulation requires topological conditions on the underlying space and also arguments that would not be too illuminating at this point and so we omit it.