1.9. ELEMENTS OF ERGODIC THEORY

Let us consider a σ-finite and complete measure space $(\Omega, \mathcal{F}, \mu)$, and on it a *measurable transformation* T; that is, a mapping $T : \Omega \to \Omega$ with the property that $T^{-1}E \in \mathcal{F}$ holds for every $E \in \mathcal{F}$. We shall call such a transformation *measure-preserving* (or 'morphism'), if $\mu(T^{-1}E) = \mu(E)$ holds for every $E \in \mathcal{F}$.

For instance, suppose we take $\Omega = \mathbf{R}$ along with its Borel sets and Lebesgue measure. The transformation $T(\omega) = 2\omega$ is measurable but not measure-preserving; in fact, we have $\mu(T^{-1}E) = \mu(E)/2$ for every $E \in \mathcal{B}(\mathbf{R}))$.

On the other hand, let us we take $\Omega = [0, 1)$ along with its Borel sets and Lebesgue measure. The *dyadic transformation* T defined by $T(\omega) = 2\omega$ for $0 \le \omega < 1/2$ and $T(\omega) = 2\omega - 1$ for $1/2 \le \omega < 1$; we write equivalently: $T(\omega) = 2\omega$ modulo 1. This transformation is measurable *and* measure-preserving. It is not hard to check that $\mu(T^{-1}I) = \mu(I)$ holds for every closed/open interval I with dyadic rational endpoints; for instance, with $I = [\frac{2}{8}, \frac{5}{8})$ we have $T^{-1}(I) = [\frac{2}{16}, \frac{5}{16}) \cup [\frac{1}{2}(\frac{2}{8}+1), \frac{1}{2}(\frac{5}{8}+1))$ so that $\mu(T^{-1}(I)) = \frac{3}{16} + \frac{3}{16} = \frac{3}{8} = \mu(I)$. Then it is easy to extend this property to all the Borel subsets of Ω.

As another example, consider a set $\Omega = \{\omega_1, \cdots, \omega_n\}$ with a finite number of elements and take μ to be normalized counting measure on the collection $\mathcal{F}$ of all subsets: $\mu(\omega_i) = 1/N$ for all $i = 1, \cdots, N$. Then $T(\omega_i) = \omega_{i+1}$ for $i = 1, \cdots, N-1$ and $T(\omega_N) = \omega_1$ defines a measure-preserving transformation.

Phenomenological Origin: The motivation for these considerations comes from Statistical Mechanics. Consider a system of k particles inside a container in three-dimensional space, and assume that the masses of these particles (molecules of a certain gas) and the forces they exert are known with precision. Then the resulting conservative mechanical system has $n = 3k$ degrees of freedom, and its evolution can be described by $2n$ coördinates consisting of the positions $p = (p_1, \cdots, p_n)$ and momenta $q = (q_1, \cdots, q_n)$ (equivalently, velocities) of the particles. Then the state of this system becomes a point $\omega = (p, q)$ in the $2n$-dimensional "phase space" Ω.

As time goes on, the state of the system changes according to the differential equations

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \qquad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}, \qquad i = 1, \cdots n$$

prescribed by the appropriate physical laws of motion, where $H(p, q)$ is the Hamiltonian of the system and does not depend on time t. The entire temporal evolution of the system is then represented by a trajectory in phase space. Classical Mechanics postulates that this trajectory can be determined, at least in principle, once we know with precision its initial condition and can solve this systems of differential equations exactly. But in practice it is impossible to possess enough information for such a complete and accurate measurement. The basic idea, then, of Statistical Mechanics as envisioned by Gibbs, is to abandon this mechanistic viewpoint in favor of a statistical study of an ensemble of states (subset of phase space). Instead of asking "what will the system do at time t?", one asks "what is the probability that at time t the state of the system will belong to a given subset of the phase space?".

Now if $\omega_t \in \Omega$ represents the state of the system at time t and $\omega \equiv \omega_0 \in \Omega$ is the initial configuration, let us denote by T_t the transformation that sends ω into ω_t, namely $\omega_t = T_t(\omega)$. Clearly we have the composition $T_{t+s} = T_t T_s$, so $\{T_t\}_{0 \le t < \infty}$ is a one-parameter semigroup of transformations, or "flow". A fundamental result of statistical mechanics is Liouville's theorem; its proof is quite simple, and can be based on the generalization of the divergence theorem to this $2n-$dimensional space. According to this theorem, if the coördinates used in the description of the system are chosen appropriately, then *the flow in phase space leaves all (2n-dimensional) volumes invariant.* In other words, $\{T_t\}_{0 \le t < \infty}$ is a flow of measure-preserving transformations. The problem at the heart of statistical mechanics is the study of the asymptotic properties of certain families of such flows. A towering result in this direction is Birkhoff's Pointwise Ergodic Theorem 9.2 below.

9.1 Definition: Recurrence. If T is a measurable transformation and $E \in \mathcal{F}$ a given non-empty, measurable set, we say that a point $\omega \in E$ is *recurrent* under T, if $T^n(\omega) \in E$ holds for some integer $n \in \mathbf{N}$. In other words, if "the orbit of successive actions of T started at this point, eventually visits the set E again".

A meta-theorem of this theory states that, whenever the underlying space Ω is "appropriately bounded", the orbits of the motion will necessarily exhibit some sort of recurrence, to wit, will return close to their initial position. A fundamental result in this direction follows; in it, the boundedness of Ω is expressed by the finiteness of the total measure.

9.1 Poincaré Recurrence Theorem: *If T is a measure-preserving transformation on a space $(\Omega, \mathcal{F}, \mu)$ of finite measure ($\mu(\Omega) < \infty$), and $E \in \mathcal{F}$ is a non-empty, measurable set, then a.e. point $\omega \in E$ is recurrent.*

Proof: Consider the set of points in E that never return to E under the orbit of successive actions of T, namely,

$$F := E \cap T^{-1}E^c \cap T^{-2}E^c \cap \cdots = \{\, \omega \in E \mid T^j(\omega) \notin E, \ \ \forall\, j \in \mathbf{N} \,\}.$$

If $\omega \in F$, then none on $T^j(\omega)$, $j \in \mathbf{N}$ belongs to F; equivalently, F is disjoint from all $T^{-j}F$, $j \in \mathbf{N}$. But this means that the sets F, $T^{-1}F$, $T^{-2}F$, $\cdots$ are all pairwise disjoint, since $T^{-n}F \cap T^{-(n+k)}F = T^{-n}\big(F \cap T^{-k}F\big) = \emptyset$ holds (mod. μ) for all integers

n and k. Therefore, we have

$$\sum_{n\in\mathbf{N}_0} \mu(T^{-n}F) \;=\; \mu\left(\bigcup_{n\in\mathbf{N}_0} T^{-n}F\right) \;<\; \infty\,,$$

since the measure is assumed to be finite. But the measure-preserving property of T implies $\mu(T^{-n}F) = \mu(F)$ for all $n \in \mathbf{N}$, and this forces $\mu(F) = 0$. ◇

It is even easier to prove the following weaker version of this result: *If T is a measure-preserving transformation on a space $(\Omega, \mathcal{F}, \mu)$ of finite measure, and if $B \in \mathcal{F}$ has $\mu(B) > 0$, then there exists some point $\omega_0 \in B$ and some integer $n \geq 1$ such that $T^n(\omega_0) \in B$.*

Discussion: Poincareé applied this result to the classical Hamiltonian systems discussed earlier. In that context, Liouville's theorem guarantees that the flow preserves preserves Euclidean volume, as well as surfaces of constant energy. Thus, surfaces of constant energy carry an invariant measure which is also finite, if the constant energy surface is bounded. Thus, by the Recurrence Theorem, almost every point $\omega \in \Omega$ has some $T^n(\omega)$, $n \in \mathbf{N}$ that eventually comes arbitrarily close to itself under the action of the flow. From this, one finds a sequence $n_k \to \infty$ such that $T^{n_k}(\omega) \to \omega$. This property is known as "stability in the sense of Poisson".

9.1 Exercise: Strong Recurrence. (a) Prove the above claim.
(b) Under the assumptions of the Poincaré Recurrence Theorem, show that for every $E \in \mathcal{F}$ and a.e. $\omega \in E$ we have the following *strong recurrence property:* $T^j(\omega) \in E$ for infinitely many $j \in \mathbf{N}$, or equivalently

$$\sum_{j\in\mathbf{N}_0} f\left(T^j(\omega)\right) \;=\; \lim_{n\to\infty} \sum_{j=0}^{n-1} f\left(T^j(\omega)\right) \;=\; \infty \tag{9.1}$$

for the function $f = \chi_E$. Then show that (9.1) holds also for a.e. $\omega \in \Omega$ in the set $\{f > 0\}$, for an arbitrary measurable function $f : \Omega \to [0, \infty)$.

These strong recurrence results suggest now some fairly obvious questions:

(i) What is the rate at which $\sum_{j=0}^{n-1} f\left(T^j(\omega)\right)$ grows to infinity as $n \to \infty$, for a generic point $\omega \in \{f > 0\}$?

(ii) For $f = \chi_E$, what is the long-term proportion $\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} \chi_E\left(T^j(\omega)\right)$ of time spent in the set E, by an orbit of successive actions of T that starts at a generic point ω of this set?

Both questions are answered by the following fundamental result of this theory.

9.2 BIRKHOFF's POINTWISE ERGODIC THEOREM: *Suppose that* $T : \Omega \to \Omega$ *is a measure-preserving transformation on the* σ-finite *space* $(\Omega, \mathcal{F}, \mu)$, *and that the function* $f : \Omega \to \mathbf{R}$ *is integrable. Then the limit*

$$f^*(\omega) := \lim_{n\to\infty} \widehat{f}_n(\omega) \qquad \text{of the averages} \qquad \widehat{f}_n(\omega) := \frac{1}{n}\sum_{j=0}^{n-1} f\left(T^j(\omega)\right) \tag{9.2}$$

exists for a.e. $\omega \in \Omega$ *(we set* $f^*(\omega) := 0$ *on the exceptional set) and defines a measurable function* $f^* : \Omega \to \mathbf{R}$ *which is integrable and* $T-$invariant: $f^*(T(\omega)) = f^*(\omega)$ *for a.e.* $\omega \in \Omega$.

If in addition $\mu(\Omega) < \infty$, *then* $\int_\Omega f^* \, d\mu = \int_\Omega f \, d\mu$. *In this case we have also convergence in mean:* $\lim_{n\to\infty} \int_\Omega \left|\widehat{f}_n - f^*\right| d\mu = 0$.

The assumption of finite measure space cannot be removed in the last assertion of the theorem. For instance, if $\Omega = \mathbf{R}$ with Lebesgue measure, and if T is the translation $T(\omega) = \omega + 1$, then for $f = \chi_{[0,1)}$ we have $\int_\Omega f \, d\mu = 1$ whereas $f^*(\omega) = 0$ for all $\omega \in \Omega$.

9.2 Definition: Invariance and Ergodicity. A set $E \in \mathcal{F}$ is called *invariant* for the transformation T, if $T^{-1}E = E$ modulo μ; to wit, $\mu(E \triangle T^{-1}E) = 0$.

A measure-preserving transformation $T : \Omega \to \Omega$ is called **ergodic**, if for $T-$invariant set E satisfies either $\mu(E) = 0$ or $\mu(E^c) = 0$; equivalently, if every $T-$invariant function is a.e. constant. $\diamond$

Ergodicity is a precise formulation (though not the only one) of the requirement that a transformation "do a good job stirring things up" in the state-space it acts on; Exercises 9.5 and 9.6 provide examples in this regard.

The translation $T(\omega) = \omega + 1$ for $\omega \in \mathbf{Z}$ is ergodic, whereas the translation $T(\omega) = \omega + 2$ for $\omega \in \mathbf{Z}$ is not (the set of even integers is invariant). The translation $T(\omega) = \omega + 1$ for $\omega \in \mathbf{R}$ is not ergodic: the set $\cup_{n\in\mathbf{Z}}\{\omega \in \mathbf{R} \,|\, n < \omega < n + (1/2)\}$ is both invariant and non-trivial.

9.2 Exercise: (i) Show that the collection $\mathcal{I}$ of $T-$invariant sets is a σ-algebra, and that a function $f : \Omega \to \mathbf{R}$ is $\mathcal{I}-$measurable if and only if it is $T-$invariant, that is, $f(T(\omega)) = f(\omega)$ holds for a.e. $\omega \in \Omega$.
(ii) Show that a measure-preserving transformation $T : \Omega \to \Omega$ is ergodic, if and only if every measurable, $T-$invariant function is a.e. constant.
(iii) If $\mu(\Omega) < \infty$, show that a measure-preserving transformation $T : \Omega \to \Omega$ is ergodic, if and only if every measurable, $T-$invariant function in $\mathbf{L}^1$ (or $\mathbf{L}^2$) is a.e. constant.

9.3 Exercise: Consider the circle group $\Omega = \{\omega \in \mathbf{C} : |\omega| = 1\}$, take a fixed $c \in \Omega$, and define $T(\omega) = c\,\omega$. This T is not ergodic if c is root of unity (that is, if $c^n = 1$ for some $n \in \mathbf{N}$), and is ergodic if c is not a root of unity.

9.4 Exercise: On $\Omega = [0,1]$ with its Borel sets and Lebesgue measure, define $T(\omega) = \omega + \xi$, modulo 1. Show that T is ergodic if and only if ξ is irrational.

Now let us go back to Theorem 9.2. The first thing to notice is that if the measure-preserving transformation T is ergodic, then the limiting function f^* of (9.2) is a constant.

- If $\mu(\Omega) = \infty$, then this constant is $f^* = 0$. The reason is that f^* is integrable, and the only integrable constant in this case is zero.
- If $\mu(\Omega) < \infty$, then this constant is

$$f^* = \frac{1}{\mu(\Omega)} \int_\Omega f \, d\mu \,. \tag{9.2$'$}$$

We have the following consequence, which is of great significance in the physical aspects of ergodic theory.

9.1 Corollary: *For any measure-preserving and ergodic transformation* $T : \Omega \to \Omega$ *on a space of finite measure, the time averages* $\frac{1}{n}\sum_{j=0}^{n-1} f\left(T^j(\omega)\right)$ *converge as* $n \to \infty$ *to the ensemble average* f^* *of* $(9.2)'$ *for a.e.* $\omega \in \Omega$.

To embark on the proof of the Pointwise Ergodic Theorem, we shall need some notation and an auxiliary result. Let us introduce an operator $U : \mathbf{L}^1 \to \mathbf{L}^1$ by

$$(Uf)(\omega) := f(T(\omega)), \qquad \omega \in \Omega$$

and note that the measure-preserving character $\mu(T^{-1}\,\cdot) = \mu(\cdot)$ of the transformation T make this U an *isometry* on $\mathbf{L}^1$; in fact,

$$\int_\Omega Uf\, d\mu = \int_\Omega f(T(\omega))\, d\mu(\omega) = \int_\Omega f(\omega)\, d\mu(T^{-1}(\omega)) = \int_\Omega f\, d\mu \tag{9.3}$$

by the change-of-variable formula of Exercise 2.7. Furthermore, this operator is linear and monotone:

$$U(\alpha f + \beta g) = \alpha\, Uf + \beta\, Ug\,, \qquad \text{and} \qquad Uf \le Uh$$

for any elements f, g and h of $\mathbf{L}^1$ with $f \le h$ and any real numbers α, β.

Some additional notation: We set $S_0 f := 0$ and $S_n f := f + Uf + \cdots + U^{n-1} f = n\widehat{f}_n$, $M_n f := \max_{0 \le k \le n}(S_k f)$, as well as $M_\infty f := \sup_{n \in \mathbf{N}_0}(S_n f) = \sup_{n \in \mathbf{N}_0}(M_n f)$. Note that $M_n f \in \mathbf{L}^1$ for all $n \in \mathbf{N}$.

9.2 The Maximal Ergodic Theorem: *For every* $f \in \mathbf{L}^1$ *we have* $\int_{\{M_\infty f>0\}} f\,d\mu \geq 0$.

Proof: The sequence of sets $B_n := \{M_n f > 0\}$, $n \in \mathbf{N}$ increases to $\{M_\infty f > 0\}$, so by dominated convergence it is enough to show

$$\int_{B_n} f\,d\mu \,\geq\, 0\,, \qquad \forall \;\; n \in \mathbf{N}\,.$$

Now on the set B_n we have $M_n f = \max_{1\leq k\leq n}(S_k f)$, and the monotonicity of U gives $S_k f = f + U(S_{k-1} f) \leq f + U(M_n f)$ for $1 \leq k \leq n$, thus also $M_n f \leq f + U(M_n f)$. Therefore,

$$\int_{B_n} M_n f\,d\mu \leq \int_{B_n} \Big(f + U(M_n f)\Big)\,d\mu \,=\, \int_{B_n} f\,d\,\mu \,+\, \int_{B_n} M_n f\,d\mu$$

thanks to the isometry of U, and the result follows since, as we have already remarked, $M_n f \in \mathbf{L}^1$. $\diamond$

PROOF OF THEOREM 9.2: Let us start by considering an *invariant* set $E = T^{-1}E$ mod. μ, and note that for it we have $S_n(f\chi_E) = S_n(f)\chi_E$ and $M_\infty(f\chi_E) = M_\infty(f)\chi_E$ mod. μ, so by the maximal ergodic theorem

$$\int_{E\cap\{M_\infty f>0\}} f\,d\mu \,\geq\, 0\,. \tag{9.4}$$

If we take a real number λ and replace f by $f-\lambda$, then $\{M_\infty(f-\lambda) > 0\}$ is the set on which we have: $S_n(f-\lambda) > 0$, or equivalently $\widehat{f}_n = (S_n f)/n > \lambda$, for some $n \in \mathbf{N}$. Setting

$$F_\lambda \,:=\, \left\{\omega \in \Omega \,\middle|\, \sup_{n\in\mathbf{N}} \widehat{f}_n(\omega) > \lambda \right\} \,=\, \bigcup_{n\in\mathbf{N}} \left\{\omega \in \Omega \,\middle|\, \widehat{f}_n(\omega) > \lambda \right\} \,=\, \{M_\infty(f-\lambda) > 0\}$$

in the notation of (9.2), we observe from these considerations that (9.4) gives

$$\int_{E\cap F_\lambda} f\,d\mu \,\geq\, \lambda\cdot\mu(E\cap F_\lambda) \quad \text{for any invariant set } E\,. \tag{9.5}$$

- Now for any real numbers $\alpha < \beta$ consider the set

$$E_{\alpha\beta} \,:=\, \left\{ \underline{\lim}_{n\to\infty}\widehat{f}_n < \alpha < \beta < \overline{\lim}_{n\to\infty}\widehat{f}_n \right\}$$

$$=\, \left\{ \underline{\lim}_{n\to\infty}(S_n f/n) < \alpha < \beta < \overline{\lim}_{n\to\infty}(S_n f/n) \right\}\,.$$

We claim that *this set is invariant:* $T^{-1} E_{\alpha\beta} = E_{\alpha\beta}$ modulo $\mathbf{P}$.
To see this, note that $S_n(Uf) = S_{n+1}f - f$ and $\lim_{n\to\infty}(f/n) = 0$ hold a.e., so

$$\begin{aligned} T^{-1} E_{\alpha\beta} &= \left\{ \underline{\lim}_{\,n\to\infty}(S_n(Uf)/n) < \alpha < \beta < \overline{\lim}_{\,n\to\infty}(S_n(Uf)/n) \right\} \\ &= \left\{ \underline{\lim}_{\,n\to\infty}(S_{n+1}f/n) < \alpha < \beta < \overline{\lim}_{\,n\to\infty}(S_{n+1}f/n) \right\} = E_{\alpha\beta}\,, \qquad \text{mod. } \mathbf{P}\,. \end{aligned}$$

Now observe that $E_{\alpha\beta} \cap F_\beta = E_{\alpha\beta}$, so that (9.5) gives

$$\int_{E_{\alpha\beta}} f\,d\mu \geq \beta \cdot \mu(E_{\alpha\beta})\,, \qquad \text{as well as} \qquad \int_{E_{\alpha\beta}} f\,d\mu \leq \alpha \cdot \mu(E_{\alpha\beta})$$

(upon replacing f, α, β by $-f, -\beta, -\alpha$). Together, these two inequalities imply $\mu(E_{\alpha\beta}) = 0$, which then leads to

$$\mu\left(\underline{\lim}_{\,n\to\infty}\widehat{f}_n < \overline{\lim}_{\,n\to\infty}\widehat{f}_n\right) = \mu(E) = 0 \qquad \text{for the set} \quad E := \bigcup_{\substack{\alpha<\beta \\ \alpha,\,\beta\in\mathbf{Q}}} E_{\alpha\beta} \in \mathcal{I}\,.$$

It follows that $f^*(\omega) = \lim_{n\to\infty} \widehat{f}_n(\omega)$ exists for every $\omega \in \Omega \setminus E$. We define $f^*(\omega) = 0$ for $\omega \in E$. The resulting function f^* is $T-$invariant ($\mathcal{I}$-measurable); this follows by arguments similar to those in the previous paragraph. We need to show that f^* is also integrable.

This is a rather easy consequence of the isometry property of the operator U; indeed, (9.3) gives

$$\int_\Omega \widehat{f}_n\,d\mu = \frac{1}{n}\int_\Omega S_n f\,d\mu = \frac{1}{n}\int_\Omega \left(f + Uf + \cdots U^{n-1}f\right) d\mu = \int_\Omega f\,d\mu \qquad (9.3)'$$

as well as

$$\int_\Omega |\widehat{f}_n|\,d\mu \leq \frac{1}{n}\int_\Omega \left(|f| + |Uf| + \cdots |U^{n-1}f|\right) d\mu = \int_\Omega |f|\,d\mu$$

for every $n \in \mathbf{N}$. From Fatou's lemma we obtain now that $\int_\Omega |f^*|\,d\mu \leq \int_\Omega |f|\,d\mu < \infty$.

• It remains to show that the two integrals $\int_\Omega f^*\,d\mu$ and $\int_\Omega f\,d\mu$ are actually equal and that we have $\lim_{n\to\infty} ||\widehat{f}_n - f^*||_1 = 0$, if the measure is finite. For this it suffices, by virtue of $(9.3)'$ and in the light of Theorem 5.2, to show that the sequence $\{\widehat{f}_n\}_{n\in\mathbf{N}}$ is uniformly integrable.

To this end, let us observe $\lambda \cdot \mu(F_\lambda) \leq \int_{F_\lambda} |f|\,d\mu$ from (9.5), and apply the same inequality to the function $-f$ in order to obtain

$$\lambda \cdot \mu(G_\lambda) \leq \int_{G_\lambda} 2|f|\,d\mu\,, \qquad \text{where} \qquad G_\lambda := \{\sup_{n\in\mathbf{N}} |\widehat{f}_n| > \lambda\}\,.$$

Therefore, for real constants $\lambda > \alpha > 1$ we get

$$\int_{\{|\widehat{f}_n|>\lambda\}} |\widehat{f}_n| \, d\mu \le \frac{1}{n} \sum_{j=0}^{n-1} \int_{G_\lambda} |f \circ T^j| \, d\mu \le \frac{1}{n} \sum_{j=0}^{n-1} \left(\int_{\{|f\circ T^j|>\alpha\}} |f \circ T^j| \, d\mu + \alpha \cdot \mu(G_\lambda) \right)$$

$$= \int_{\{|f|>\alpha\}} |f| \, d\mu + \alpha \cdot \mu(G_\lambda) \le \int_{\{|f|>\alpha\}} |f| \, d\mu + \frac{2\alpha}{\lambda} \int_\Omega |f| \, d\mu \, .$$

Now take $\alpha = \sqrt{\lambda}$ and let $\lambda \to \infty$; the required uniform integrability of $\{\widehat{f}_n\}_{n\in\mathbf{N}}$ follows from the integrability of f. ◇

9.5 Exercise: Mixing. A measure-preserving transformation T on a probability space $(\Omega, \mathcal{F}, \mu)$ is called *mixing*, if $\lim_{n\to\infty} \mu(A \cap T^{-n}B) = \mu(A)\mu(B)$ holds for all $A \in \mathcal{F}$, $B \in \mathcal{F}$. Show that such a transformation is ergodic.

9.6 Exercise: Show that a measure-preserving transformation T on the probability space $(\Omega, \mathcal{F}, \mu)$ is ergodic, if and only if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu\big(A \cap T^{-k}B\big) = \mu(A)\mu(B) \tag{9.6}$$

holds for all $A \in \mathcal{F}$, $B \in \mathcal{F}$; or equivalently, if and only if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \int_\Omega f(T^n(\omega))g(\omega) \, d\mu(\omega) = \int_\Omega f(\omega) \, d\mu(\omega) \cdot \int_\Omega g(\omega) \, d\mu(\omega) \tag{9.7}$$

holds for any functions f and g in $\mathbf{L}^2(\mu)$.

9.7 Exercise: In the case $\mu(\Omega) < \infty$, show that the result $\int_\Omega f^* \, d\mu = \int_\Omega f \, d\mu$ of Theorem 9.2 can be strengthened to

$$\int_\Lambda f^* \, d\mu = \int_\Lambda f \, d\mu \, , \qquad \forall \quad \Lambda \in \mathcal{I} \, .$$

• **FURTHER DEVELOPMENTS.** In 1977, Furstenberg proved the following *Multiple Recurrence result:*

If $T:\Omega\to\Omega$ *is a measure-preserving transformation on a finite measure space* $(\Omega,\mathcal{F},\mathbf{P})$*, if* $k\geq 1$ *is an integer, and if the set* $E\in\mathcal{F}$ *has positive measure, then*

$$\liminf_{N\to\infty}\frac{1}{N}\sum_{j=0}^{N-1}\mu\big(E\cap T^{-j}E\cap\cdots\cap T^{-(k-1)j}E\big)\;>\;0\,;$$

in particular, there exists then an $n\in\mathbf{N}$ *such that* $\mu\big(E\cap T^{-n}E\cap\cdots\cap T^{-(k-1)n}E\big)\;>\;0$.

He then made an unexpected connection with Combinatorics, showing that regularity properties of sets A of integers with positive upper density

$$d^*(A)\;:=\;\limsup_{N\to\infty}\frac{1}{N}\cdot\#\big(A\cap\{1,\cdots,N\}\big)$$

correspond to multiple recurrence results:

• *Let* $A\subset N$ *have positive upper density. There exists a probability space* $(\Omega,\mathcal{F},\mathbf{P})$*, a measure-preserving transformation* $T:\Omega\to\Omega$*, and a set* $E\in\mathcal{F}$ *with* $\mu(E)=d^*(A)$*, such that for any integer* $k\geq 1$ *we have*

$$d^*\big(A\cap(A+m_1)\cap\cdots\cap(A+m_k)\big)\;\geq\;\mu\big(E\cap T^{-m_1}E\cap\cdots\cap T^{-m_k}E\big)\,,\quad\forall\;(m_1,\cdots,m_k)\in\mathbf{N}^k\,.$$

As a corollary one then obtains the following result of Szemerédi, who had proved in 1975 by purely combinatorial arguments a long-standing conjecture of Erdös & Turán: *If a set of integers has positive upper density, then it contains arithmetic progressions of any given finite length.*

The set of primes has zero upper-density, so Szemerédi's theorem does not apply to it. In a spectacular recent development, Green & Tao (2004) used ergodic-theoretic ideas, along with deep results from harmonic analysis, combinatorics and number theory, to show that **the set of primes contains finite arithmetic progressions of any given length.** In fact, they showed the following stronger result: *If* A *is a set of primes with*

$$\limsup_{N\to\infty}\frac{1}{\pi(N)}\cdot\#\big(A\cap\{1,\cdots,N\}\big)\;>\;0\,,$$

where $\pi(N)$ *is the number of primes in* $\{1,\cdots,N\}$*, then for any given integer* $k\geq 1$ *the set* A *contains an arithmetic progression of length* k.