Chapter 20

Ergodicity

Ergodic theory can be described as the discipline that studies the *long run average* behavior of *dynamical systems*. There is a set S of possible *states* of the system, and the evolution of the system is usually modeled as a function $T\colon S \to S$. If the system is in state s at time t, then Ts is the state of the system at time $t{+}1$. The sequence $\{s, Ts, T^2s, \ldots\}$ is called the *orbit* of the state s.

There are several approaches to the mapping $T\colon S \to S$, depending on the structure of the state space S and the topological properties of the mapping T. In this chapter, we discuss briefly two approaches. In the first approach S is a probability measure space and T is a measure-preserving transformation. In the second case S is a Banach space and T is a continuous linear operator.

A real function $f\colon S \to \mathbb{R}$ (subject to some measurability or continuity requirement) is usually interpreted as some sort of measurement of the system. If a phenomenon follows the evolutionary orbit $\{s, Ts, T^2s, \ldots\}$, then the sequence of real numbers $\{f(s), f(Ts), f(T^2s), \ldots\}$ represents the values of the measurements of some quantity during the evolution of the phenomenon. The average of these measurements during the first n periods is given by

$$A_n(f) = \frac{1}{n}\sum_{i=0}^{n-1} f(T^i s).$$

As mentioned above, the concern of ergodic theory is the long run behavior of the sequence of time averages $\{A_n(f)\}$, especially the convergence of these averages. Results on the convergence of the sequence of averages are known as *ergodic theorems*. A limit of the sequence $\{A_n(f)\}$ can be interpreted as an "equilibrium" value of the measurement f.

There are several ergodic theorems in the literature, and our goal here is to describe a few of them that you may find useful. We have no intention of entering into the delicate details of ergodic theory at this time. There are many detailed and extensive treatments of the theory from various points of view and of varying degrees of obscurity. W. Parry [270] and K. E. Petersen [274] offer quite readable treatments of basic ergodic theorems. U. Krengel [217] and P. Walters [339] are highly operator theoretic in nature. R. Mañé [239] studies differentiable structures and ergodic theory. A. Lasota and M. C. Mackey [223] apply ergodic theory

to “chaotic” systems. Y. Kifer [202] studies ergodic theory in terms of random sequences of functions. The monograph by D. S. Ornstein [267] addresses the question of isomorphism of dynamical systems using the concepts of coding and entropy. C. A. Futia [132] discusses the use of ergodic theory in economic theory.

20.1 Measure-preserving transformations and ergodicity

In this section (Ω, Σ, μ), or simply Ω, denotes a probability space. We start with the definition of a measure-preserving transformation.

20.1 Definition *A transformation $\xi\colon \Omega \to \Omega$ is **μ-measure-preserving** (or simply **measure-preserving**) if it is measurable and*

$$\mu(A) = \mu(\xi^{-1}(A))$$

for each $A \in \Sigma$. In other words, ξ is measure-preserving if the measure $\mu\xi^{-1}$ induced by ξ on Σ coincides with μ.

Continuous μ-measure-preserving transformations are precisely those whose composition operators leave μ invariant.

20.2 Theorem *Let X be a compact metrizable space, and let μ be a Borel probability measure on X. For a continuous function $\xi\colon X \to X$, the following statements are equivalent.*

1. *The transformation ξ is μ-measure-preserving.*
2. *The measure μ is T_ξ-invariant, where $T_\xi\colon C(X) \to C(X)$ is the composition operator defined as usual by $T_\xi(f) = f \circ \xi$.*

Proof: By the Change of Variables Theorem 13.46, for each $f \in C(X)$ we have

$$\langle f, T'_\xi\mu\rangle = \langle T_\xi f, \mu\rangle = \langle f \circ \xi, \mu\rangle = \langle f, \mu\xi^{-1}\rangle.$$

Since every probability measure on X is regular, we infer that $T'_\xi\mu = \mu\xi^{-1}$. Consequently, $T'_\xi\mu = \mu$ if and only if $\mu = \mu\xi^{-1}$. ∎

For a function $f\colon X \to X$ we employ the following standard notation and terminology.

- The **iterates** of f are defined inductively by $f^0(x) = x$ and $f^{n+1}(x) = f(f^n(x))$ for each $x \in X$.
- If A is a subset of X, then we let $f^0(A) = A$ and

$$f^{-(n+1)}(A) = f^{-1}(f^{-n}(A)) = \{x \in X : f^{n+1}(x) \in A\}.$$

In other words, a point belongs to $f^{-n}(A)$ if and only if its n^{th} iterate lies in A.