22. Consequences of Birkhoff's Ergodic Theorem

§22.1 Statement of Birkhoff's Ergodic Theorem

Recall Birkhoff's ergodic theorem for an ergodic transformation:

Theorem 22.1 (Birkhoff's Ergodic Theorem)
Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T : X \to X$ be an ergodic measure preserving transformation. Let $f \in L^1(X, \mathcal{B}, \mu)$ be an integrable function. Then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j x) = \int f \, d\mu$$

for μ-a.e. $x \in X$.

§22.2 Consequences of Birkhoff's Ergodic Theorem

Here we give some simple corollaries of Birkhoff's Ergodic Theorem. The first result says that, for a typical orbit of an ergodic dynamical system, 'time averages' equal 'space averages'.

Corollary 22.2
If T is ergodic and if $B \in \mathcal{B}$ then for μ-a.e. $x \in X$, the frequency with which the orbit of x lies in B is given by $\mu(B)$, i.e.,

$$\lim_{n\to\infty} \frac{1}{n} \mathrm{card}\{j \in \{0, 1, \ldots, n-1\} \mid T^j x \in B\} = \mu(B) \quad \mu\text{-a.e.}$$

Proof. Apply the Birkhoff Ergodic Theorem with $f = \chi_B$. □

It is possible to characterise ergodicity in terms of the behaviour of sets, rather than points, under iteration. The next result deals with this.

Theorem 22.3
Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T : X \to X$ be a measure-preserving transformation. The following are equivalent:

(i) T is ergodic;

(ii) for all $A, B \in \mathcal{B}$,

$$\frac{1}{n} \sum_{j=0}^{n-1} \mu(T^{-j} A \cap B) \to \mu(A)\mu(B),$$

as $n \to \infty$.

Proof. (i) $\Rightarrow$ (ii): Suppose that T is ergodic. Since $\chi_A \in L^1(X, \mathcal{B}, \mu)$, Birkhoff's Ergodic Theorem tells us that

$$\frac{1}{n}\sum_{j=0}^{n-1} \chi_A \circ T^j \to \mu(A), \text{ as } n \to \infty,$$

μ-a.e. Multiplying both sides by χ_B gives

$$\frac{1}{n}\sum_{j=0}^{n-1} \chi_A \circ T^j \, \chi_B \to \mu(A)\chi_B, \text{ as } n \to \infty,$$

μ-a.e. Since the left-hand side is bounded (by 1), we can apply the Dominated Convergence Theorem to see that

$$\begin{aligned}\frac{1}{n}\sum_{j=0}^{n-1} \mu(T^{-j}A \cap B) &= \frac{1}{n}\sum_{j=0}^{n-1} \int \chi_A \circ T^j \, \chi_B \, d\mu \\ &= \int \frac{1}{n}\sum_{j=0}^{n-1} \chi_A \circ T^j \, \chi_B \, d\mu \to \mu(A)\mu(B),\end{aligned}$$

as $n \to \infty$.

(ii) $\Rightarrow$ (i): Now suppose that the convergence holds. Suppose that $T^{-1}A = A$ and take $B = A$. Then $\mu(T^{-j}A \cap B) = \mu(A)$ so

$$\frac{1}{n}\sum_{j=0}^{n-1} \mu(A) \to \mu(A)^2,$$

as $n \to \infty$. This gives $\mu(A) = \mu(A)^2$. Therefore $\mu(A) = 0$ or 1 and so T is ergodic. □

§22.3 Applications

§22.3.1 Normal numbers

A number $x \in [0, 1)$ is called *normal* (in base 2) if it has a unique binary expansion, the digit 0 occurs in its binary expansion with frequency 1/2, and the digit 1 occurs in its binary expansion with frequency 1/2. We will show that Lebesgue a.e. $x \in [0, 1)$ is normal.

To see this, observe that Lebesgue almost every $x \in [0, 1)$ has a unique binary expansion $x = \cdot x_1 x_2 \ldots$, $x_i \in \{0, 1\}$. Define $Tx = 2x \bmod 1$. Then $x_n = 0$ if and only if $T^{n-1}x \in [0, 1/2)$. Thus

$$\frac{1}{n}\text{card}\{1 \leq i \leq n \mid x_i = 0\} = \frac{1}{n}\sum_{i=0}^{n-1} \chi_{[0,1/2)}(T^i x).$$

Since T is ergodic (with respect to Lebesgue measure), for Lebesgue almost every point x the above expression converges to $\int \chi_{[0,1/2)}(x)\, dx = 1/2$. Similarly the frequency with which the digit 1 occurs is equal to 1/2. Hence Lebesgue almost every point in $[0,1)$ is normal.

Exercise 22.1

(i) *Let $r \geq 2$. What would it mean to say that a number $x \in [0,1)$ is normal in base r?*

(ii) *Prove that for each r, Lebesgue a.e. $x \in [0,1)$ is normal in base r.*

(iii) *Conclude that Lebesgue a.e. $x \in [0,1)$ is simultaneously normal in every base $r = 2, 3, 4, \ldots$.*

Exercise 22.2

Prove that the arithmetic mean of the digits appearing in the base 10 expansion of Lebesgue-a.e. $x \in [0,1)$ is equal to 4.5, i.e. prove that if $x = \sum_{j=0}^{\infty} x_j/10^{j+1}$, $x_j \in \{0,1,\ldots,9\}$ then

$$\lim_{n\to\infty} \frac{1}{n}(x_0 + x_1 + \cdots + x_{n-1}) = 4.5 \text{ a.e.}$$

§22.3.2 Continued fractions

We will show that for Lebesgue a.e. $x \in (0,1)$ the frequency with which the natural number k occurs in the continued fraction expansion of x is given by

$$\frac{1}{\log 2} \log\left(\frac{(k+1)^2}{k(k+2)}\right).$$

Let λ denote Lebesgue measure and let μ denote Gauss' measure. Then λ-a.e. and μ-a.e. $x \in (0,1)$ is irrational and has an infinite continued fraction expansion

$$x = \cfrac{1}{x_1 + \cfrac{1}{x_2 + \cfrac{1}{x_3 + \cfrac{1}{x_4 + \cdots}}}}.$$

Let T denote the continued fraction map. Then $x_n = [1/T^{n-1}x]$.

Fix $k \in \mathbb{N}$. Then $x_n = k$ precisely when $[1/T^{n-1}x] = k$, i.e.

$$k \leq \frac{1}{T^{n-1}x} < k+1$$

which is equivalent to requiring

$$\frac{1}{k+1} < T^{n-1}x \leq \frac{1}{k}.$$

Hence

$$\begin{aligned}
\frac{1}{n}\text{card}\{1 \leq i \leq n \mid x_i = k\} &= \frac{1}{n}\sum_{i=0}^{n-1} \chi_{(1/(k+1),1/k]}(T^i x) \\
&\to \int \chi_{(1/(k+1),1/k]}\, d\mu \text{ for } \mu\text{-a.e. } x \\
&= \frac{1}{\log 2}\left[\log\left(1+\frac{1}{k}\right) - \log\left(1+\frac{1}{k+1}\right)\right] \\
&= \frac{1}{\log 2}\log\frac{(k+1)^2}{k(k+2)}.
\end{aligned}$$

As μ and λ are equivalent, this holds for Lebesgue almost every point.

Exercise 22.3
For $x \in (0,1) \setminus \mathbb{Q}$ write its infinite continued fraction expansion as

$$x = \cfrac{1}{x_1 + \cfrac{1}{x_2 + \cfrac{1}{x_3 + \cdots}}}.$$

Show that for Lebesgue almost every $x \in (0,1)$ we have

$$\frac{1}{n}(x_1 + x_2 + \cdots + x_n) \to \infty$$

as $n \to \infty$. (That is, for a typical point x, the average value of the coefficients in its continued fraction expansion is infinite.)

§22.4 A card trick

§22.4.1 The trick

Here is a way to impress your non-mathematical friends with Birkhoff's Ergodic Theorem!

Take two or three packs of cards and keep only the aces, twos, threes and fours (an ace counts as a one for this trick). Shuffle the cards.

Ask each member of the audience to pick a number—their 'initial number'—between one and four, and to keep that number secret. The rules of the game are as follows. Deal out the cards, one at a time. Each member of the audience counts out their initial number of cards; the value of this card becomes their new number. Now the audience member counts out their new number of cards, whereupon the value of this card becomes their new number. Repeat until the pack of cards is exhausted.

Thus, suppose that the audience member initially chooses 3. The dealer starts dealing out the cards, say with values 1,3,2. The value 2 of the third card becomes the audience member's new number. Now the dealer counts out another 2 cards: 3,1, say. The value 1 becomes the new number, etc.

When the end of the pack is reached, ask each member of the audience which number is their current number. Hopefully(!), everybody will have the same number.

§22.4.2 Why it works

Suppose the sequence of cards $4, 3, 2, 1, 1$ occurs somewhere in the pack. Then, no matter what number a person has, they will eventually enter this finite block; once this finite block is entered, the person will always end up at the final 1, after which all audience members will always have the same number. Call such a block a 'synchronising block'. There are lots of synchronising blocks (for example, $2, 1, 2, 1, 1$ is another synchronising block).

The shift map on the one-sided full 4-shift Σ_4 is ergodic. Hence almost every point in Σ_4 visits a synchronising block infinitely often. Hence it is highly likely (exactly how likely can be made precise) that a particular shuffle of cards (which corresponds to the initial few symbols in a typical point of Σ_4) contains at least one synchronising block.