3. Invariant measures

§3.1 Introduction

In Lecture 1 we remarked that ergodic theory is the study of the qualitative distributional properties of typical orbits of a dynamical system and that these properties are expressed in terms of measure theory. Measure theory therefore lies at the heart of ergodic theory. However, we will not need to know the (many!) intricacies of measure theory. We will give a brief overview of the basics of measure theory, before studying invariant measures. In the next lecture we then go on to study ergodic measures.

§3.2 Measure spaces

We will need to recall some basic definitions and facts from measure theory.

Definition. Let X be a set. A collection $\mathcal{B}$ of subsets of X is called a *σ-algebra* if:

(i) $\emptyset \in \mathcal{B}$,

(ii) if $A \in \mathcal{B}$ then $X \setminus A \in \mathcal{B}$,

(iii) if $A_n \in \mathcal{B}$, $n = 1, 2, 3, \ldots$, is a countable sequence of sets in $\mathcal{B}$ then their union $\bigcup_{n=1}^{\infty} A_n \in \mathcal{B}$.

Remark Clearly, if $\mathcal{B}$ is a σ-algebra then $X \in \mathcal{B}$. It is easy to see that if $\mathcal{B}$ is a σ-algebra then $\mathcal{B}$ is also closed under taking countable intersections.

Examples.

1. The trivial σ-algebra is given by $\mathcal{B} = \{\emptyset, X\}$.

2. The full σ-algebra is given by $\mathcal{B} = \mathcal{P}(X)$, i.e. the collection of *all* subsets of X.

3. Let X be a compact metric space. The *Borel σ-algebra* is the smallest σ-algebra that contains every open subset of X. As σ-algebras are closed under taking complements, the Borel σ-algebra is also the smallest σ-algebra that contains every closed subset of X. An element of the Borel σ-algebra is called a *Borel set*.

Let X be a set and let $\mathcal{B}$ be a σ-algebra of subsets of X.

Definition. A function $\mu : \mathcal{B} \to \mathbb{R}^+ \cup \{\infty\}$ is called a *measure* if:

(i) $\mu(\emptyset) = 0$;

(ii) if A_n is a countable collection of pairwise disjoint sets in $\mathcal{B}$ (i.e. $A_n \cap A_m = \emptyset$ for $n \neq m$) then

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

We call $(X, \mathcal{B}, \mu)$ a *measure space*. If $\mu(X) < \infty$ then we call μ a *finite measure*. If $\mu(X) = 1$ then we call μ a *probability* or *probability measure* and refer to $(X, \mathcal{B}, \mu)$ as a *probability space*.

Definition. We say that a property holds *almost everywhere* if the set of points on which the property fails to hold has measure zero.

Definition. Suppose that X is a compact metric space and $\mathcal{B}$ is the Borel σ-algebra. A measure on $\mathcal{B}$ is called a *Borel measure*.

Consider the set

$$\mathcal{U} = \bigcup \{U \mid U \text{ is open, } \mu(U) = 0\};$$

that is, $\mathcal{U}$ is the largest open set with zero measure. The *support of* μ is the complement $\operatorname{supp} \mu = X \setminus \mathcal{U}$.

§3.3 The Kolmogorov extension theorem

In order to define a measure, it is necessary to define the measure of every set in the σ-algebra under consideration. This is usually impractical, and instead we seek a method that allows us to define a measure on a tractable subcollection of subsets and then extend it to the required σ-algebra.

Definition. A collection $\mathcal{A}$ of subsets of X is called an *algebra* if:

(i) $\emptyset \in \mathcal{A}$,

(ii) if $A \in \mathcal{A}$ then $A^c \in \mathcal{A}$,

(iii) if $A_1, A_2 \in \mathcal{A}$ then $A_1 \cup A_2 \in \mathcal{A}$.

Thus an algebra is similar to a σ-algebra, except that $\mathcal{A}$ is closed under finite, rather than countable, unions.

Example. Take $X = [0, 1]$. Then the collection $\mathcal{A} = \{$all finite unions of subintervals$\}$ is an algebra.

Let $\mathcal{B}(\mathcal{A})$ denote the σ-algebra generated by $\mathcal{A}$, i.e., the smallest σ-algebra containing $\mathcal{A}$. (In the above example $\mathcal{B}(\mathcal{A})$ is the Borel σ-algebra. This follows from the fact that any open set is a countable union of open intervals.)

Theorem 3.1 (Kolmogorov Extension Theorem)
Let $\mathcal{A}$ be an algebra of subsets of X. Suppose that $\mu : \mathcal{A} \to \mathbb{R}^+ \cup \{\infty\}$ satisfies:

(i) $\mu(\emptyset) = 0$;

(ii) if $A_n \in \mathcal{A}$, $n \geq 1$, are pairwise disjoint and if $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$ then

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

(iii) there exists finitely or countably many sets $X_n \in \mathcal{A}$ such that $X = \bigcup_n X_n$ and $\mu(X_n) < \infty$;

Then there is a unique measure $\mu : \mathcal{B}(\mathcal{A}) \to \mathbb{R}^+ \cup \{\infty\}$ which is an extension of $\mu : \mathcal{A} \to \mathbb{R}^+ \cup \{\infty\}$.

Remarks.

(i) The important hypotheses are (i) and (ii) (hypothesis (iii) says that the space X is σ-finite, a common technical assumption). Thus the Kolmogorov Extension Theorem says that if we have a function μ that looks like a measure on an algebra $\mathcal{A}$, then it is indeed a measure when extended to $\mathcal{B}(\mathcal{A})$.

(ii) We will often use the Kolmogorov Extension Theorem as follows. Take $X = [0, 1]$ and take $\mathcal{A}$ to be the algebra consisting of all finite unions of sub-intervals of X. We then define the 'measure' μ of a subinterval in such a way as to be consistent with the hypotheses of the Kolmogorov Extension Theorem. It then follows that μ does indeed define a measure on the Borel σ-algebra.

(iii) Here is another way in which we shall use the Kolmogorov Extension Theorem. Suppose we have two measures, μ and ν, and we want to see if $\mu = \nu$. A priori we would have to check that $\mu(B) = \nu(B)$ for all $B \in \mathcal{B}$. The Kolmogorov Extension Theorem says that it is sufficient to check that $\mu(A) = \nu(A)$ for all A in an algebra $\mathcal{A}$ that generates $\mathcal{B}$. For example, to show that two measures on $[0, 1]$ are equal, it is sufficient to show that they give the same measure to each subinterval.

The following result can be deduced from the proof of the Kolmogorov Extension Theorem. It allows us to approximate a set in $\mathcal{B}(\mathcal{A})$ by sets in $\mathcal{A}$. Recall that the symmetric difference between two sets A, B is the set $A \triangle B = (A \setminus B) \cup (B \setminus A)$.

Corollary 3.2
Let $\mathcal{A}$ be an algebra and $\mu : \mathcal{A} \to \mathbb{R}^+$ be a function satisfying the hypotheses of the Kolmogorov Extension Theorem. Then for all $B \in \mathcal{B}(\mathcal{A})$ and all $\varepsilon > 0$ there exists $A \in \mathcal{A}$ such that $\mu(A \triangle B) < \varepsilon$.

§3.4 Examples of measures

§3.4.1 Lebesgue measure

Take $X = [0, 1]$ and let $\mathcal{A}$ denote the algebra of all finite unions of intervals. For an interval $[a, b]$ define $\mu([a, b]) = b - a$ and extend this to $\mathcal{A}$. This satisfies the hypotheses of the Kolmogorov Extension Theorem, and so defines a Borel probability measure. This is Lebesgue measure on $[0, 1]$.

In a similar way we can define Lebesgue measure on $\mathbb{R}/\mathbb{Z}$.

Take $X = \mathbb{R}^k/\mathbb{Z}^k$ to be the k-dimensional torus. A k-dimensional cube is a set of the form $[a_1, b_1] \times \cdots \times [a_k, b_k]$. Let $\mathcal{A}$ denote the algebra of all finite unions of k-dimensional cubes. For a k-dimensional cube $[a_1, b_1] \times \cdots \times [a_k, b_k]$ define

$$\mu([a_1, b_1] \times \cdots \times [a_k, b_k]) = \prod_{j=1}^{k} (b_j - a_j)$$

and extend this to $\mathcal{A}$. This satisfies the hypotheses of the Kolmogorov Extension Theorem and defines k-dimensional Lebesgue measure on the k-dimensional torus.

§3.4.2 Stieltjes measures

Let $X = [0, 1]$ and let $\rho : [0, 1] \to \mathbb{R}^+$ be a non-decreasing function. Let $\mathcal{A}$ again denote the algebra of finite unions of intervals. For an interval $[a, b]$ define $\mu_\rho([a, b]) = \rho(b) - \rho(a)$ and extend this to $\mathcal{A}$. The Kolmogorov Extension Theorem then extends μ_ρ to a Borel measure.

Lebesgue measure can be viewed as a special, if somewhat trivial, example of this construction: take $\rho(x) = x$. A more interesting example that will prove useful when studying continued fractions is given by taking

$$\rho(x) = \frac{1}{\log 2} \int_0^x \frac{1}{1 + x} \, dx.$$

The resulting measure μ_ρ is called *Gauss' measure*.

A wide range of measures can be constructed using this method.

Definition. Suppose that μ_1, μ_2 are two measures on $(X, \mathcal{B})$. We say that μ_1 is *absolutely continuous* with respect to μ_2 (and write $\mu_1 \ll \mu_2$) if $\mu_2(B) = 0$ implies $\mu_1(B) = 0$. We say that μ_1 μ_2 are *equivalent* if $\mu_1 \ll \mu_2$ and $\mu_2 \ll \mu_1$. (Thus two measures are equivalent if they have the same sets of measure zero.)

We say that two probability measures μ_1, μ_2 on $(X, \mathcal{B})$ are *mutually singular* if there exist two disjoint sets $B_1, B_2 \in \mathcal{B}$ such that $B_1 \cup B_2 = X$ and $\mu_1(B_2) = 0, \mu_2(B_1) = 0$. (Thus the support of μ_1 is contained in B_1, and the support of μ_2 is contained in B_2.)

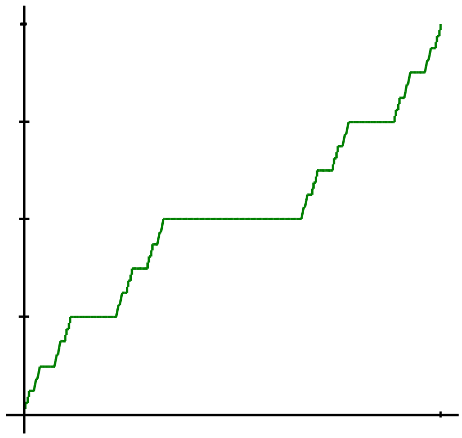

Figure 3.1: The graph of the Devil's staircase

If ρ is differentiable at Lebesgue-a.e. point then we call $\rho'(x)$ the *density* of μ_ρ. If $\rho'(x)$ is continuous then μ_ρ is absolutely continuous with respect to Lebesgue measure. Moreover, if in addition $\rho'(x) > 0$ (so that ρ is strictly increasing), then μ_ρ and Lebesgue measure are equivalent. Thus Gauss' measure and Lebesgue measure are equivalent.

However, it can happen that ρ is differentiable on a large set but μ_ρ and Lebesgue measure are mutually singular. Let E_1 denote the unit interval with the middle-third removed; thus $E_1 = [0, 1/3] \cup [1/3, 2/3]$, two intervals of length $1/3$. Construct E_n inductively by removing the middle third of each of the 2^{n-1} intervals of E_{n-1}, leaving 2^n intervals each of length $1/3^n$. Let $E = \cap_{n=1}^{\infty} E_n$ denote the standard middle-third Cantor set. It is well-known that E is an uncountable, perfect, nowhere dense subset of $[0, 1]$.

We define a function $\rho : [0, 1] \to \mathbb{R}$ as follows. Let $[a, b]$ be one of the intervals deleted in the construction of E_n from E_{n-1}. Then $a = k/3^n, b = (k+1)/3^n$ for some integer k. Write $k = \sum_{j=1}^{n} r_j 3^j$ in base 3, where $r_j \in \{0, 1, 2\}$. Define $s_j = 0$ if $r_j = 0$ and $s_j = 1$ if $r_j = 1, 2$. We then define ρ on the interval $[a, b]$ by taking $\rho(x) = \sum_{j=1}^{n} s_j 2^{-j}$. This defines a function that is uniformly continuous on the complement of E, and so extends uniquely to a continuous function ρ defined on $[0, 1]$. The function ρ is an example of a class of function known as *Devil's staircases* and has some remarkable properties: it is a continuous function, increasing from $\rho(0) = 0$ to $\rho(1) = 1$, is differentiable Lebesgue-a.e. but has zero derivative Lebesgue-a.e. See Figure 3.1.

The measure μ_ρ is supported on the Cantor set E, which is easily seen to have Lebesgue measure zero. Hence μ_ρ and Lebesgue measure are mutually singular.

§3.4.3 Bernoulli and Markov measures

Let $\Sigma_k = \{(x_j)_{j=-\infty}^{\infty} \mid x_j \in \{1, \ldots, k\}\}$ denote the full two-sided k-shift. A cylinder set is a set of sequences where we fix which symbol can occur in a finite number of places. More specifically, fix $i_0, \ldots, i_n \in \{1, \ldots, k\}$. The cylinder $[i_0, \ldots, i_n]_m$ is the set of all sequences $(x_j)_{j=-\infty}^{\infty}$ with the restriction that symbol i_j must occur in the $(j + m)$th place of (x_j). That is,

$$[i_0, \ldots, i_n]_m = \{(x_j)_{j=-\infty}^{\infty} \in \Sigma_k \mid x_{j+m} = i_j,\ j = 0, 1, \ldots, n\}.$$

We can also define cylinders for one-sided shifts and for shifts of finite type.

Cylinders for shift spaces play the same role as intervals do for the unit interval. One can easily check that cylinders are open subsets of Σ_k (indeed, they are also closed subsets; this reflects the fact that Σ_k is totally disconnected). The collection $\mathcal{A}$ of finite unions of cylinders forms an algebra, and the σ-algebra generated by $\mathcal{A}$ is the Borel σ-algebra.

Therefore, by the Kolmogorov Extension Theorem, to define a Borel measure on Σ_k it is sufficient to define a measure on cylinders.

Let $P = (P_{i,j})$ be a $k \times k$ stochastic matrix. That is, $P_{i,j} \geq 0$ and each row of P sums to 1. Suppose we can find a left probability eigenvector $p = (p_1, \ldots, p_k)$ for P. That is, $p_j \geq 0$, $\sum p_j = 1$, and $pP = p$. Then we can define a probability measure μ_P on cylinders by setting

$$\mu_P[i_0, \ldots, i_n]_m = p_{i_0} P_{i_0,i_1} P_{i_1,i_2} \cdots P_{i_{n-1},i_n}. \tag{3.1}$$

This then extends to a Borel probability measure on Σ_k.

The motivation behind (3.1) is that we view p_i as the probability of starting with symbol i and $P_{i,j}$ as the probability of the symbol j occurring next, given that the preceding symbol was i. Hence the probability of a given symbol occurring depends only on a finite number of (indeed, one) preceding symbols; this is a characteristic of a Markov process. We call the measure μ_P a *Markov measure*.

If $P_{i,j} = p_i$ where $p = (p_1, \ldots, p_k)$ is a probability vector then P is a stochastic matrix. In this case, the above construction gives a Borel probability measure μ_p given on cylinders by

$$\mu_p[i_0, \ldots, i_n]_m = p_{i_0} p_{i_1} \cdots p_{i_n}.$$

This is called the *Bernoulli* $(p_1, \ldots, p_k)$*-measure*.

Let A be a $k \times k$ $0-1$ matrix with associated shift of finite type Σ_A (either one-sided or two-sided). Suppose that P is a stochastic matrix that is compatible with A, that is $P_{i,j} > 0$ if and only if $A_{i,j} = 1$. Then the above construction of a Markov measure gives a method for constructing Borel probability measures on Σ_A.

Of course, for a given stochastic matrix P, a left probability eigenvector may or may not exist. If A is irreducible, then the Perron-Frobenius theorem guarantees that such an eigenvector does exist. We shall revisit this in a more general form in a later lecture.

§3.4.4 Dirac measures

Let $(X, \mathcal{B})$ be any measure space. For $x \in X$ we define the measure

$$\delta_x(B) = \begin{cases} 1 & \text{if } x \in B \\ 0 & \text{otherwise.} \end{cases}$$

We call δ_x the *Dirac δ-measure supported at x*. Notice that the support of δ_x is the singleton $\{x\}$. For this reason, δ_x is often called a *point-mass at x*.

§3.5 The Lebesgue integral

Let $(X, \mathcal{B}, \mu)$ be a measure space. We give a brief introduction to the definition of the Lebesgue integral on $(X, \mathcal{B}, \mu)$. In the special case where $X = [0, 1]$ and μ is Lebesgue measure, this extends the definition of the Riemann integral.

Definition. A function $f : X \to \mathbb{R}$ is *measurable* if $f^{-1}(D) \in \mathcal{B}$ for every Borel subset D of $\mathbb{R}$, or, equivalently, if $f^{-1}(c, \infty) \in \mathcal{B}$ for all $c \in \mathbb{R}$.

A function $f : X \to \mathbb{C}$ is *measurable* if both the real and imaginary parts, $\mathrm{Re} f$ and $\mathrm{Im} f$, are measurable.

We define integration via simple functions.

Definition. A function $f : X \to \mathbb{R}$ is *simple* if it can be written as a linear combination of characteristic functions of sets in $\mathcal{B}$, i.e.:

$$f = \sum_{i=1}^{r} a_i \chi_{A_i},$$

for some $a_i \in \mathbb{R}$, $A_i \in \mathcal{B}$, where the A_i are pairwise disjoint.

For a simple function $f : X \to \mathbb{R}$ we define

$$\int f \, d\mu = \sum_{i=1}^{r} a_i \mu(A_i)$$

(which can be shown to be independent of the representation of f as a simple function). Thus for simple functions, the integral can be thought of as being defined to be the area underneath the graph.

If $f : X \to \mathbb{R}$, $f \geq 0$, is measurable then one can show that there exists an increasing sequence of simple functions f_n such that $f_n \uparrow f$ pointwise[1] as $n \to \infty$ and we define

$$\int f\,d\mu = \lim_{n\to\infty} \int f_n\,d\mu.$$

This can be shown to be independent of the choice of sequence f_n.

For an arbitrary measurable function $f : X \to \mathbb{R}$, we write $f = f^+ - f^-$, where $f^+ = \max\{f, 0\} \geq 0$ and $f^- = \max\{-f, 0\} \geq 0$ and define

$$\int f\,d\mu = \int f^+\,d\mu - \int f^-\,d\mu.$$

Finally, for a measurable function $f : X \to \mathbb{C}$, we define

$$\int f\,d\mu = \int \mathrm{Re} f\,d\mu + i \int \mathrm{Im} f\,d\mu.$$

We say that f is *integrable* if

$$\int |f|\,d\mu < +\infty.$$

A real- or complex-valued function f defined on X is said to be *integrable* if $\int |f|\,d\mu < \infty$. The space of integrable functions is defined to be

$$L^1(X, \mathcal{B}, \mu) = \{f : X \to \mathbb{R} \mid f \text{ is measurable}, \int |f|\,d\mu < \infty\}.$$

More generally, for $p \geq 1$ we define the L^p-spaces to be

$$L^p(X, \mathcal{B}, \mu) = \{f : X \to \mathbb{R} \mid f \text{ is measurable}, \int |f|^p\,d\mu < \infty\}.$$

§3.6 Invariant measures

Let $(X, \mathcal{B}, \mu)$ be a probability space. A transformation $T : X \to X$ is said to be *measurable* if $T^{-1}B \in \mathcal{B}$ for all $B \in \mathcal{B}$.

Definition. We say that T is a *measure-preserving transformation* (m.p.t.) or, equivalently, μ is said to be a *T-invariant* measure, if $\mu(T^{-1}B) = \mu(B)$ for all $B \in \mathcal{B}$.

Lemma 3.3
The following are equivalent:

(i) T is a measure-preserving transformation;

(ii) for each $f \in L^1(X, \mathcal{B}, \mu)$, we have

$$\int f\,d\mu = \int f \circ T\,d\mu.$$

[1] $f_n \uparrow f$ pointwise means: for every x, $f_n(x)$ is an increasing sequence and $f_n(x) \to f(x)$ as $n \to \infty$.

Remark In Lemma 3.3(ii) we can replace the requirement that $f \in L^1(X, \mathcal{B}, \mu)$ by $f \in L^2(X, \mathcal{B}, \mu)$.

Proof. For $B \in \mathcal{B}$, it is clear that $\chi_B \in L^1(X, \mathcal{B}, \mu)$. Note that $\chi_B \circ T = \chi_{T^{-1}B}$. Hence

$$\begin{aligned} \mu(B) &= \int \chi_B \, d\mu = \int \chi_B \circ T \, d\mu \\ &= \int \chi_{T^{-1}B} \, d\mu = \mu(T^{-1}B). \end{aligned}$$

This proves one implication.

Conversely, suppose that T is a measure-preserving transformation. For any characteristic function χ_B, $B \in \mathcal{B}$,

$$\int \chi_B \, d\mu = \mu(B) = \mu(T^{-1}B) = \int \chi_{T^{-1}B} \, d\mu = \int \chi_B \circ T \, d\mu$$

and so the equality holds for any simple function (a finite linear combination of characteristic functions). Given any $f \in L^1(X, \mathcal{B}, \mu)$ with $f \geq 0$, we can find an increasing sequence of simple functions f_n with $f_n \to f$ pointwise, as $n \to \infty$. For each n we have

$$\int f_n \, d\mu = \int f_n \circ T \, d\mu$$

and, applying the Monotone Convergence Theorem to both sides, we obtain

$$\int f \, d\mu = \int f \circ T \, d\mu.$$

To extend the result to general real-valued f, consider the positive and negative parts. This completes the proof. □

§3.7 Examples

We shall discuss two different methods for determining whether a dynamical system $T : X \to X$ preserves a given measure μ.

One method uses the Kolmogorov Extension Theorem. It is easy to see that if T is measurable then we can define a new measure $T_*\mu$ by $(T_*\mu)(B) = \mu(T^{-1}B)$. This is a probability measure on $(X, \mathcal{B})$. Thus to show that μ is T-invariant, we have to show that $\mu = T_*\mu$, i.e. $\mu(B) = \mu(T^{-1}B)$ for all $B \in \mathcal{B}$. By the Kolmogorov Extension Theorem, it is sufficient to check that $\mu(A) = \mu(T^{-1}A)$ for all $A \in \mathcal{A}$, where $\mathcal{A}$ is an algebra that generates $\mathcal{B}$.

We shall also discuss algebraic examples of dynamical systems and discuss when Haar measure is an invariant measure.

If the dynamical system has a periodic orbit then it is easy to find an invariant measure.

§3.7.1 Measures supported on a periodic orbits

Let $T : X \to X$ be a measurable dynamical system defined on a measure space $(X, \mathcal{B})$. Suppose that $x = T^n x$ is a periodic point with period n. Then the probability measure

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

is T-invariant. This is clear from Lemma 3.3, noting that for $f \in L^1(X, \mathcal{B}, \mu)$

$$\begin{aligned} \int f \circ T \, d\mu &= \frac{1}{n}(f(Tx) + \cdots + f(T^{n-1}x) + f(T^n x)) \\ &= \frac{1}{n}(f(x) + f(Tx) + \cdots + f(T^{n-1}x)) \\ &= \int f \, d\mu, \end{aligned}$$

using the fact that $T^n x = x$.

§3.7.2 Using the Kolmogorov Extension Theorem

We give three examples of using the Kolmogorov Extension Theorem to prove that a given measure is invariant for a given dynamical system

Proposition 3.4
Let $T : \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z}$ be the doubling map $T(x) = 2x \bmod 1$. Then Lebesgue measure μ is T-invariant.

Proof. Let $\mathcal{A}$ denote the algebra of finite unions of intervals. For an interval $[a, b]$ we have that

$$T^{-1}[a, b] = \{x \mid T(x) \in [a, b]\} = \left[\frac{a}{2}, \frac{b}{2}\right] \cup \left[\frac{a+1}{2}, \frac{b+1}{2}\right].$$

See Figure 3.2.

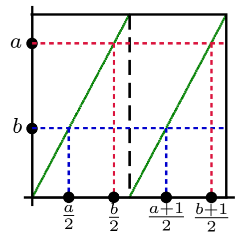

Figure 3.2: The pre-image of an interval under the doubling map

Hence

$$\begin{aligned} T_*\mu[a,b] &= \mu T^{-1}[a,b] \\ &= \mu\left[\frac{a}{2},\frac{b}{2}\right] \cup \left[\frac{a+1}{2},\frac{b+1}{2}\right] \\ &= \frac{b}{2} - \frac{a}{2} + \frac{(b+1)}{2} - \frac{(a+1)}{2} \\ &= b-a = \mu[a,b]. \end{aligned}$$

Hence $T_*\mu = \mu$ on the algebra $\mathcal{A}$. As $\mathcal{A}$ generates the Borel σ-algebra, by uniqueness in the Kolmogorov Extension Theorem we see that $T_*\mu = \mu$; i.e. Lebesgue measure is T-invariant. □

Proposition 3.5
The continued fraction map $T : [0,1] \to [0,1]$ *given by* $T(x) = 1/x \bmod 1$ *preserves Gauss' measure* μ *where*

$$\mu(B) = \frac{1}{\log 2}\int_B \frac{1}{1+x}\,dx.$$

Proof. As in the proof of Proposition 3.4, it is sufficient to check that $\mu([a,b]) = \mu(T^{-1}[a,b])$ for any interval $[a,b]$. First note that

$$T^{-1}[a,b] = \bigcup_{n=1}^{\infty}\left[\frac{1}{b+n},\frac{1}{a+n}\right].$$

Thus

$$\begin{aligned} &\mu(T^{-1}[a,b]) \\ &= \frac{1}{\log 2}\sum_{n=1}^{\infty}\int_{\frac{1}{b+n}}^{\frac{1}{a+n}} \frac{1}{1+x}\,dx \\ &= \frac{1}{\log 2}\sum_{n=1}^{\infty}\left[\log\left(1+\frac{1}{a+n}\right) - \log\left(1+\frac{1}{b+n}\right)\right] \\ &= \frac{1}{\log 2}\sum_{n=1}^{\infty}[\log(a+n+1) - \log(a+n) - \log(b+n+1) + \log(b+n)] \\ &= \lim_{N\to\infty}\frac{1}{\log 2}\sum_{n=1}^{N}[\log(a+n+1) - \log(a+n) - \log(b+n+1) + \log(b+n)] \\ &= \frac{1}{\log 2}\lim_{N\to\infty}[\log(a+N+1) - \log(a+1) - \log(b+N+1) + \log(b+1)] \\ &= \frac{1}{\log 2}\left(\log(b+1) - \log(a+1) + \lim_{N\to\infty}\log\left(\frac{a+N+1}{b+N+1}\right)\right) \\ &= \frac{1}{\log 2}\left(\log(b+1) - \log(a+1)\right) \end{aligned}$$

$$= \frac{1}{\log 2}\int_a^b \frac{1}{1+x}\,dx = \mu[a,b],$$

as required. □

Proposition 3.6
Let P be a stochastic matrix, let $p = (p_1, \ldots, p_k)$ be a probability vector such that $pP = p$, and let μ_P be the corresponding Markov measure on the full k-shift Σ_k. Let $\sigma : \Sigma_k \to \Sigma_k$ denote the shift map $(\sigma x)_k = x_{k+1}$. Then σ preserves μ_P.

Proof. By the Kolmogorov Extension Theorem, it is sufficient to prove that $\mu_P(C) = \mu_P(\sigma^{-1}C)$ for all cylinders C. Let $C = [i_0, \ldots, i_n]_m = \{x \in \Sigma_k \mid x_{j+m} = i_j \text{ for } j = 0, 1, \ldots, n\}$. Then $\sigma^{-1}C = [i_0, \ldots, i_n]_{m-1}$. Clearly,

$$\mu_P(C) = p_{i_0} P_{i_0,i_1} \cdots P_{i_{n-1},i_n} = \mu_P(\sigma^{-1}C).$$

□

§3.7.3 Haar measure

Let G be a compact topological group equipped with the Borel σ-algebra $\mathcal{B}$. For example, G could be the k-dimensional torus $\mathbb{R}^k/\mathbb{Z}^k$, or a matrix group such as $SO(n)$. When G is compact, it is well-known that there exists a unique probability measure μ that is invariant under left and right group multiplication, i.e. $\mu(gA) = \mu(Ag) = \mu(A)$ where $gA = \{gx \mid g \in G\}$, $Ag = \{xg \mid g \in G\}$. This measure is called *Haar measure*.

For the k-dimensional torus, Haar measure is k-dimensional Lebesgue measure. To see this, let μ denote k-dimensional Lebesgue measure. For each $a \in \mathbb{R}^k/\mathbb{Z}^k$, let $T_a(x) = x + a$. Let $f \in L^1(\mathbb{R}^k/\mathbb{Z}^k, \mathcal{B}, \mu)$. Then, by the change of variable formula for Lebesgue integration,

$$\int f \circ T_a \, d\mu = \int f(x+a)\, d\mu = \int f \, d\mu.$$

Hence μ is T_a-invariant for each $a \in \mathbb{R}^k/\mathbb{Z}^k$.

The following result is immediate from the definition of Haar measure.

Proposition 3.7
Let G be a compact group with Haar measure μ and let $a \in G$. Define $T : G \to G$ by $T(x) = gx$. Then μ is a T-invariant measure.

Remark Thus Lebesgue measure is an invariant measure for a rotation on a circle or k-dimensional torus.

The uniqueness of Haar measure implies the following result.

Proposition 3.8
Let G be a compact group with Haar measure μ and let α be an autoomorphism of G. Define $T : G \to G$ by $T(x) = \alpha(x)$. Then μ is a T-invariant measure.

Proof. Define $T_*\mu(B) = \mu(T^{-1}B)$. It is sufficient to prove that $T_*\mu$ and μ define the same measure. Let L_g denote left-multiplication in G by g. Note that $\alpha^{-1}(L_g(x)) = L_{\alpha^{-1}g}\alpha^{-1}(x)$. As Haar measure is characterised by being the unique measure that is invariant under all group rotations, it is sufficient to check that $T_*\mu$ is invariant under group rotations. This follows as

$$T_*\mu(g(B)) = \mu(\alpha^{-1}L_g(B)) = \mu(L_{\alpha^{-1}g}\alpha^{-1}(B)) = \mu(\alpha^{-1}B) = T_*\mu(B).$$

Hence $T_*\mu$ is invariant under any group rotation, and so must be equal to Haar measure. □

Remark It follows that Lebesgue measure is an invariant measure for linear toral automorphisms, such as the Cat map.

§3.8 References

Most of the material in this lecture is standard in ergodic theory and can be found in many texts, including

P. Walters, *An introduction to ergodic theory*, Springer, Berlin, 1982.

W. Parry, *Topics in Ergodic Theory*, C.U.P., Cambridge, 1981.

Good texts on abstract measure theory include

P. Halmos, *Measure Theory*, Graduate Texts in Mathematics vol. 18, Springer-Verlag, New York, Berlin, 1984.

H.L. Royden, *Real Analysis*, Macmillan, New York, 1988.

The Perron-Frobenius theorem can be found in

Gantmacher, *The Theory of Matrices*, Vol. 2, Chelsea, New York, 1974.

§3.9 Exercises

Exercise 3.1
Let G be a compact abelian group with Haar measure μ. Let α be an automorphism of G and fix $a \in G$. Define the affine map $T : G \to G$ by $T(x) = \alpha(x) + a$. Prove that μ is T-invariant.

Exercise 3.2

Let $\beta > 1$ and consider the transformation $T : [0,1] \to [0,1]$ given by $T(x) = \beta x \bmod 1$. We have already seen that if β is an integer than T preserves Lebesgue measure. Suppose that $\beta > 1$ is not an integer.

(i) Show that the measure

$$\mu(B) = \int_B h(x)\,dx$$

is T-invariant if and only if

$$h(x) = \frac{1}{\beta} \sum_{y:Ty=x} h(y). \tag{3.2}$$

(ii) Show that the function

$$h(x) = \sum_{n=0, x<T^n 1}^{\infty} \frac{1}{\beta^n}$$

satisfies (3.2). (Here the sum is interpreted as follows: the term $1/\beta^n$ is included in the sum precisely when $x < T^n 1$.)

(iii) Show that if there exists n, m, $n \neq m$, such that $T^n 1 = T^m 1$ then h is a step function with finitely many jumps. Show that this is the case for the case when β is the golden mean ($\beta^2 = \beta + 1$) and find an explicit expression for a T-invariant measure that is equivalent to Lebesgue measure.

(Such transformations play an important role in the expansion of real numbers using non-integer bases.)