CHAPTER I

INTRODUCTION AND EXAMPLES

Ergodic theory is the study of (semi–)group actions on spaces. The basic motivational example is that of a physical system evolving in time according to a fixed set of laws.

Examples

EXAMPLE 1.1. Let X denote the set of possible states of a physical system that evolves in time. The way in which a point $x \in X$ evolves in time is given by a mapping

$$(x, t) \mapsto T_t(x),$$

where $T_t(x)$ is the state that x moves to after t seconds. Notice that each T_t is a function $X \to X$. In particular, T_0 is the identity map on X. Assuming that each state x determines its future and its past completely, we can find the state that x reaches after $t + s$ seconds in two ways: firstly, it is by definition $T_{t+s}(x)$. On the other hand, the state x at time 0 has a unique future; after t seconds have elapsed it is at $T_t(x)$, and after a further s seconds it is at $T_s(T_t(x))$. It follows that

$$T_s \circ T_t = T_{s+t} \tag{1.1}$$

for all $s, t > 0$. Under our assumption (that the entire past and future of a state x is determined by x), each T_t must be a bijection on X. By (1.1), the assignment $t \mapsto T_t$ is a *homomorphism* $\mathbb{R} \to \mathrm{Bij}(X)$, *where* $\mathrm{Bij}(X)$ *is the group of bijections of* X. That is, T is an *action* of the group $\mathbb{R}$ by bijections of the set X.

If the physical system given by the action of T on X is sampled at times 0,1,2,3, ... (and for our mathematical convenience, at times –1,–2, ... as well), the transformation observed on X between each time is given by a single map, namely T_1. In this case we obtain an action of $\mathbb{Z}$ on X by bijections, via the homomorphism $n \mapsto (T_1)^n = T_n$.

Example (1.1) was completely general. A more genuine physical system is given by the following example.

EXAMPLE 1.2. Consider a system of n particles in $\mathbb{R}^3$, moving according to some mechanical law. The state of the system is completely determined (in the sense that we can – in principle – determine the complete future and past behaviour from this data) by the positions $\mathbf{q}_1, \dots, \mathbf{q}_n$ and momenta $\mathbf{p}_1, \dots, \mathbf{p}_n$ of the particles. Thus the allowed states form a subset $X \subset \mathbb{R}^{6n}$. Let the laws that govern the system be

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

given by a Hamiltonian $H : X \to \mathbb{R}$, so

$$\dot{\mathbf{q}}_i = \frac{\partial H}{\partial \mathbf{p}_i}$$
$$\dot{\mathbf{p}}_i = -\frac{\partial H}{\partial \mathbf{q}_i}. \tag{1.2}$$

The function H is constant along orbits of the $\mathbb{R}$ action given by time evolution (equivalently, is constant along the solutions of the differential equations (1.2)). Choose a suitable value for H, say e: we then obtain an $\mathbb{R}$–action T on the set $H^{-1}(e)$, which (under some reasonable assumptions) is a submanifold of $\mathbb{R}^{6n}$. A fundamental observation (due to Liouville) is that Lebesgue measure on $\mathbb{R}^{6n}$ restricts to give a measure on each $H^{-1}(e)$ that is *preserved by the* $\mathbb{R}$ *action*. The *ergodic hypothesis* was that the orbits of the action spread through the space so uniformly that for any continuous function $f : H^{-1}(e) \to \mathbb{R}$, the *space average* of f, $\int_{H^{-1}(e)} f$ would coincide with the *time average*, $\lim_{s\to\infty} \frac{1}{s}\int_0^s f(T_t(x))ds$ for almost any x.

We are therefore led to the following kind of models for studying dynamical systems: actions of a group G on a space X, preserving some natural structure that could range from almost nothing (cardinality of sets, group actions by bijections) to very specific structures (for example, group actions by automorphisms of compact groups, or actions by diffeomorphisms of a smooth manifold). The most important examples are the following.

[1] X is a probability space with σ–algebra $\mathfrak{B}$ and measure μ. The action T is by invertible measure–preserving transformations of $(X, \mathfrak{B}, \mu)$. This is classical *ergodic theory*.

[2] X is a compact topological space, and T is an action by homeomorphisms of X. This is *topological dynamics*.

[3] X is a differentiable compact manifold, and T acts by diffeomorphisms of X. This is *smooth ergodic theory* or *differentiable dynamical systems*.

[4] Algebraic dynamical systems. For instance, X is a compact abelian group, which is a probability space with respect to Haar measure μ on the Borel σ–algebra $\mathfrak{B}$, and T is a G–action by measurable automorphisms of X. Notice that such an example sits in [1] and in [2]: a measurable automorphism is automatically continuous, so T is automatically an action by homeomorphisms of the compact topological space X.

More generally, the four families of examples are not unrelated: the following theorem is the first step towards using measure–theoretic ideas to understand homeomorphisms and diffeomorphisms.

THEOREM 1.3. *If X is a compact topological space, and $f : X \to X$ is a homeomorphism, then there is a probability m defined on the Borel sets of X which is f–invariant.*

This theorem may be saying very little (if X is a compact group, and f is an automorphism, then one possible candidate for m is the point mass at the identity), quite a lot (if X is the circle $\mathbb{S}^1$, and f is the rotation $f(z) = ze^{2\pi i\theta}$, $\theta \in \mathbb{Q}$, what are the possible measures m?) or a great deal (if X is the circle $\mathbb{S}^1$, and f is the rotation $f(z) = ze^{2\pi i\theta}$, $\theta \notin \mathbb{Q}$, what are the possible measures m?). Theorem 1.3 is proved in the exercises.

A *measurable space* is a set X with a collection of subsets $\mathfrak{B}$ of X such that

(1) $X \in \mathfrak{B}$,
(2) if $B \in \mathfrak{B}$ then $X \backslash B \in \mathfrak{B}$,
(3) $B_n \in \mathfrak{B}$ implies that $\bigcup_{n=1}^{\infty} B_n \in \mathfrak{B}$.

Such a collection $\mathfrak{B}$ is called a *σ–algebra* of subsets; the elements of $\mathfrak{B}$ are the *measurable sets*.

A *finite measure* on the measure space $(X, \mathfrak{B})$ is a map $m : \mathfrak{B} \to \mathbb{R}_{>0}$ with $m(\emptyset) = 0$ and $m(\bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} m(B_n)$ if $\{B_n\}$ is a pairwise disjoint collection of measurable sets. If, in addition, $m(x) = 1$, then $(X, \mathfrak{B}, m)$ is a probability space.

If X is a topological space, then the *Borel σ–algebra* is the smallest σ–algebra defined on X that contains all the open sets, and a measure m on X is a *Borel measure* if it is defined on the Borel σ–algebra.

EXAMPLE 1.4. The following are examples of probability spaces.

(1) Let $X = [0, 1]$, with m Lebesgue measure on the interval, and $\mathfrak{B}$ the σ–algebra of Lebesgue measurable sets.
(2) [COIN–TOSSING SPACE] Let $X_i = \{0, 1, \ldots, n-1\}$ for each $i \in \mathbb{Z}$; and let $m_i = (p_0, \ldots, p_{n-1})$ be a fixed probability vector (i.e. $p_j \geq 0$ and $\sum p_j = 1$). The discrete topology on X_i makes $X = \prod_{-\infty}^{\infty} X_i$ into a compact topological space. Subsets $A_j \subset X_j$ for $j = n, \ldots, m$ define a cylinder set in X:

$$C = \prod_{-\infty}^{n-1} X_i \times \prod_{j=n}^{m} A_j \times \prod_{m+1}^{\infty} X_i,$$

and the collection of all such cylinders forms a basis of open sets for the topology on X. Define a Borel measure m on X by setting $m(C) = \prod_{j=n}^{m} m_j(A_j)$ and extending to all Borel sets. As an illustration, let $n = 2$, and choose the measure $m_i = (\frac{1}{2}, \frac{1}{2})$. If we identify heads with the symbol 0 and tails with the symbol 1, then (X, m) is a probability space representing a fair coin–toss repeated infinitely often. Cylinder sets correspond to specifying the outcome of finitely many of the independent coin–tosses, and the measure of the cylinder set is the probability of that event.
(3) [COMPACT GROUPS] Let X be an abelian group that is also a compact Hausdorff space, and assume that the map $(x, y) \to x - y$ is a continuous map from $X \times X$ in the product topology to X. Then X carries a unique translation invariant Borel probability μ (that is, a measure μ with

$$\mu(A + x) = \mu(A)$$

for any Borel set A and $x \in X$). The measure μ is called *Haar measure*.

Let $(X_1, \mathfrak{B}_1, m_1)$ and $(X_2, \mathfrak{B}_2, m_2)$ be two measure spaces. A transformation $T : X_1 \to X_2$ is *measurable* if $T^{-1}(\mathfrak{B}_2) \subset \mathfrak{B}_1$, *measure–preserving* if it is measurable and $m_1(T^{-1}(A)) = m_2(A)$ for any $A \in \mathfrak{B}_2$, and is an invertible measure–preserving transformation if it is measure–preserving, bijective, and T^{-1} is also measure–preserving.

Example 1.5. Let X be a compact abelian group with Haar measure μ and Borel σ–algebra $\mathfrak{B}$, and let T be an endomorphism of X. Then T is a measure–preserving transformation of $(X, \mathfrak{B}, \mu)$ (equivalently, T *preserves* the Haar measure). This may be seen as follows: define a new measure m on X by setting $m(A) = \mu(T^{-1}(A))$. Then $m(Tx + A) = \mu(T^{-1}(Tx + A)) = \mu(x + T^{-1}(A)) = \mu(T^{-1}(A)) = m(A)$, so that m is a probability invariant under translation by anything in the image of T; since T is onto we deduce from the uniqueness of Haar measure that $m = \mu$.

Three illustrations of Example 1.5. (1) Consider the additive circle $\mathbb{T}$ and the endomorphism $T(x) = 3x$ (mod 1). The pre–image under T of an interval comprises exactly three copies of the interval, each one third the size of the original. (2) Consider the coin–tossing space of Example 1.4(2), where we specialize to have $X_i = \{0, 1\}$ viewed as a group of two elements. Let m_i be the $(\frac{1}{2}, \frac{1}{2})$ measure (i.e. Haar measure on the finite group X_i. Define the shift transformation by $T(\mathbf{x})_k = x_{k+1}$. Then T is an automorphism of the compact group X, and preserves Haar measure. (3) Let X be the d–dimensional torus $(S^1)^k$, and let T be an automorphism of X. The map T is given by a matrix $[T] \in GL(d, \mathbb{Z})$ as usual (for instance, if $d = 2$ then an automorphism of $(\mathbb{S}^1)^2$ is a map of the form $(z, w) \mapsto (z^a w^b, z^c w^d)$ with $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in GL(2, \mathbb{Z})$).

Example 1.6. Let X be a compact abelian group with Haar measure μ and Borel σ–algebra $\mathfrak{B}$, and let T be a *rotation* of X, $T(x) = x + g$ for some fixed $g \in X$. Then T is a measure–preserving transformation of $(X, \mathfrak{B}, \mu)$ by definition of Haar measure.

Example 1.7. Let X be the coin–tossing space described in Example 1.4. Define a map $T : X \to X$ by the left shift: $T(\mathbf{x})_k = x_{k+1}$, where $\mathbf{x} = (x_k)_{k \in \mathbb{Z}} \in X$. Then T is an invertible measure–preserving transformation of X with the measure m given in Example 1.4. The transformation T is called a *Bernoulli shift*, or a *Bernoulli* $(p_0, \dots, p_{n-1})$*–shift*. Bernoulli shifts are abstract versions of independent identically distributed processes.

Isomorphism

The next step is to decide when two measure–preserving group actions are measurably indistinguishable. To motivate the definition, consider the following example of an action of $\mathbb{N}$.

Let $X_1 = [0, 1]$, and let $T_1(x) = 2x$ (mod 1). Let $X_2 = \prod_{i=0}^{\infty}\{0, 1\}$, with the $(\frac{1}{2}, \frac{1}{2})$–measure – the one–sided coin tossing space. Let $T_2(\mathbf{x})_k = x_{k+1}$, the left shift map; this is a 2–to–1 measure–preserving transformation on X_2. Let $\theta : X_2 \to X_1$ be the map given by

$$\theta(\mathbf{x}) = \tfrac{x_0}{2} + \tfrac{x_1}{4} + \tfrac{x_2}{8} + \tfrac{x_3}{16} + \dots,$$

so that $\theta(T_2(\mathbf{x})) = T_1(\theta(\mathbf{x}))$. Notice that θ is an invertible measure–preserving transformation once we delete from X_2 all sequences of 0's and 1's that have finitely many 0's or finitely many 1's, and delete from X_1 the image under θ of all such sequences. We are therefore led to the following notion of measurable isomorphism: two (semi–)group actions are isomorphic if, *after deleting some null set in range and domain*, there is an invertible measure–preserving transformation that intertwines the actions.

DEFINITION 1.8. Let G be a countable group, and let $T^{(i)}$ on $(X_i, \mathfrak{B}_i, m_i)$ for $i = 1, 2$ be two actions of G by invertible measure–preserving transfromations. The actions $T^{(1)}$ and $T^{(2)}$ are isomorphic if there are null sets $N_1 \in \mathfrak{B}_1$, $N_2 \in \mathfrak{B}_2$, and an invertible measure–preserving transformation $\theta : X_1 \backslash N_1 \to X_2 \backslash N_2$ with $T_g^{(1)}\theta(x) = \theta(T_g^{(2)}(x))$ for all $x \in X_1 \backslash N_1$ and every $g \in G$.

The basic internal problem in ergodic theory is then the following.

PROBLEM 1.9. Given two G–actions, how can we decide whether or not they are isomorphic?

Problem 1.9 is intractable except in special cases (see Chapter 5). Ergodicity and mixing provide some crude invariants for isomorphism.

DEFINITION 1.10. A G–action T on $(X, \mathfrak{B}, m)$ is *ergodic* if any set $A \in \mathfrak{B}$ with $T_g^{-1}(A) = A$ for all $g \in G$ has $m(A) = 0$ or 1. Equivalently, T is ergodic if $f(T_g x) = f(x)$ almost everywhere for each $g \in G$, for a function $f \in L^2(m)$, implies that f is almost everywhere constant.

THEOREM 1.11. *An automorphism T of the n–dimensional torus $(\mathbb{S}^1)^n$ is ergodic (with respect to Haar measure m) if and only if the associated matrix $[T]$ has no unit root eigenvalues.*

PROOF. The family of functions $f_{^t(k_1,\dots,k_n)}(z_1, \dots, z_n) = z_1^{k_1} \cdots z_n^{k_n}$, $k_i \in \mathbb{Z}$, form an orthonormal basis for $L^2(m)$. The automorphism T sends one element of this set into another according to the rule $f_{\mathbf{k}}(T\mathbf{z}) = f_{^t[T]\mathbf{k}}(\mathbf{z})$.

If $[T]$ has a p^{th} root of unity eigenvalue (p minimal), then there is an integer vector $\mathbf{w} \in \mathbb{Z}^n \backslash \{0\}$ with

$$({}^t[T]^p - I_n)\mathbf{w} = 0.$$

Then the function

$$f = f_{\mathbf{w}} + f_{\mathbf{w}} \circ T + \cdots + f_{\mathbf{w}} \circ T^{p-1}$$

is T–invariant and non–constant (since it is a sum of distinct elements of the orthonormal basis). It follows that T is not ergodic.

Conversely, if T is not ergodic, choose a non–constant function $f \in L^2(m)$ which is T–invariant. Enumerate the orthonormal basis $\{f_{\mathbf{k}}\}_{\mathbf{k} \in \mathbb{Z}^n}$ as $\chi_0, \chi_1 \dots$ where χ_0 is the constant function 1. Let $\sum_{i=0}^{\infty} a_i \chi_i$ be the Fourier series of f; by Plancherel,

$$\sum_{i=0}^{\infty} |a_i|^2 < \infty. \tag{1.3}$$

Since f is not constant, $a_s \neq 0$ for some $s \neq 0$. The T–invariance of f implies that the coefficients of $\chi_s \circ T$, $\chi_s \circ T^2, \dots$ are all a_s. By (1.3), we must therefore have $\chi_s \circ T^p = \chi_s \circ T^q$ for some $p > q \geq 0$. Since T is injective, we have therefore that $\chi_s \circ T^{p-q} = \chi_s$, so there exists $\mathbf{k}$ such that $\mathbf{k} =^t [T]^{p-q}\mathbf{k}$, so $[T]$ has a unit root eigenvalue. □

Since ergodicity is clearly preserved by measurable isomorphism, this gives examples of non–isomorphic transformations. The automorphisms of the 2–torus defined by the matrices $\begin{bmatrix} 0 & 1 \\ -1 & -1 \end{bmatrix}$ and $\begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$ cannot be isomorphic.

DEFINITION 1.12. A G–action T on $(X, \mathfrak{B}, m)$ is *k–fold mixing* if for any sets $B_0, \dots, B_k \in \mathfrak{B}$,

$$\lim_{g_i - g_j \to \infty} m(T_{-g_0}B_0 \cap T_{-g_1}B_1 \cap \dots \cap T_{-g_k}B_k) = m(B_0) \cdots m(B_k), \tag{1.4}$$

where $g_i - g_j \to \infty$ means that for each pair $i \neq j$ the difference $g_i - g_j$ leaves finite sets in G.

In formulating k–fold mixing we may assume that $g_0 = 0$.

The condition with 2 sets is also known as *strong mixing* or simply *mixing*; we should also point out that many people prefer to call the above property $(k+1)$–mixing. If a G action is k–fold mixing for every k, then it is *mixing of all orders*.

EXAMPLE 1.13. Let T be an ergodic automorphism of a compact group X. Then T is mixing of all orders.

We shall not prove this here; it follows at once from a much stronger property that an ergodic automorphism of a compact group must satisfy: it must behave measurably like the shift on a coin–tossing space (see Example 1.4) for some choice of n and $(p_0, \dots, p_{n-1})$.

The behaviour of several commuting compact group automorphisms is quite different, as shown by the following fundamental example due to Ledrappier [18].

EXAMPLE 1.14. [MIXING $\not\Rightarrow$ HIGHER MIXING] Let $X = \{\mathbf{x} \in \{0,1\}^{\mathbb{Z}^2} \mid x_{(n,m)} + x_{(n+1,m)} + x_{(n,m+1)} = 0 \bmod 2 \ \forall n, m \in \mathbb{Z}\}$; this is a compact totally disconnected group, with Haar measure μ say. Define a $\mathbb{Z}^2$ action on X by automorphisms as follows: $T_{(1,0)}$ is the horizontal shift $T_{(1,0)}(\mathbf{x})_{(n,m)} = x_{(n+1,m)}$ and $T_{(0,1)}$ is the vertical shift $T_{(0,1)}(\mathbf{x})_{(n,m)} = x_{(n,m+1)}$.

We claim that T is mixing: it is sufficient to check that for any finite sets F_1 and F_2 in $\mathbb{Z}^2$, and any allowed maps $f_1 : F_1 \to \{0,1\}$ and $f_1 : F_1 \to \{0,1\}$ (allowed meaning that each f_i is a restriction of an element of X, so $f_i(n,m) + f_i(n+1,m) + f_i(n,m+1) = 0 \bmod 2$ for all (n,m) with $(n,m), (n+1,m), (n,m+1) \in F_i$), there is an M with the property that $|(n,m)| \geq M$ implies that there is an $\mathbf{x} \in X$ with the property that $\mathbf{x}$ restricted to F_1 is f_1 and $T_{(n,m)}(\mathbf{x})$ restricted to F_2 is f_2. What this means is that the cylinder sets defined by specifying what we see on F_1 and F_2 become independent if they are moved sufficiently far apart. This is clear: from each F_i construct $\tilde{F}_i$, a triangle containing F_i of the shape $(a,b) + \{(c,d) \mid c \geq 0, d \geq 0, c + d \leq K\}$. Then if the shapes are moved far enough apart to ensure that the triangles do not touch, we can consistently fill in the two shapes.

To see that T is not 2–fold mixing, it is sufficient to exhibit three sets that fail to mix. Let $B = B_0 = B_1 = B_2 = \{\mathbf{x} \mid x_{(0,0)} = 0\}$. Then $\mu(B) = \frac{1}{2}$. Now notice that for any n, $x_{(0,0)} + x_{(2^n,0)} + x_{(0,2^n)} = 0 \bmod 2$, so

$$B \cap T_{(2^n,0)}B \cap T_{(0,2^n)}B = B \cap T_{(2^n,0)}B,$$

and therefore $\mu(B \cap T_{(2^n,0)}B \cap T_{(0,2^n)}B) = \frac{1}{4}$ for all n, showing that T is not 2–fold mixing.

Definition 1.15. A G–action T on $(X, \mathfrak{B}, \mu)$ is a *factor* of the G–action S on $(Y, \mathfrak{C}, \nu)$ if there is a measure–preserving map $\theta : X \to Y$ with $\theta(T_g(x)) = S_g(\theta(x))$ almost everywhere, for each $g \in G$.

Lemma 1.16. *If the G–action T is k–fold mixing, and S is a factor of T, then S is also k–fold mixing.*

Two G actions are said to be *weakly isomorphic* if each is a factor of the other: a deep result is that there are non–isomorphic $\mathbb{Z}$–actions that are weakly isomorphic.

Notes. Most of the material in this chapter can be found in Walter's book [35]. Example 1.14 is originally due to Ledrappier, [18]. Further examples of this sort, and some of their deeper properties, are described by Kitchens and Schmidt in [13] and [14].

Exercises

1.1 Prove that if T is a measure–preserving map, then so is T^k for any $k \geq 1$. If T^k is measure–preserving for some $k \geq 2$, does it follow that T is measure–preserving? If T^k is measure–preserving for all $k \geq 2$, does it follow that T is measure–preserving?

1.2 Prove that an invertible measure–preserving transformation T is ergodic if and only if T^{-1} is ergodic.

1.3 Let $X = \{1, 2, \dots, n\}$, with measure μ defined by $\mu(\{k\}) = \frac{1}{n}$. For each element τ of the group S_n of permutations of X, show that the corresponding map $\tau : X \to X$ is measure–preserving. Is the action of S_n on X ergodic? Identify all the elements of S_n that are ergodic transformations of X. If τ is ergodic, then for which k is is true that τ^k is also ergodic?

1.4 Let $A \in GL(2, \mathbb{Z})$ be a matrix with no unit root eigenvalues, and let $T = T_A$ be the associated toral automorphism of $\mathbb{T}^2$. Given functions $f, g \in L^2(\mathbb{T}^2)$, prove directly that

$$\int_{\mathbb{T}^2} f(x)g(T^n x)d\mu(x) \to \int f d\mu \int g d\mu$$

as $n \to \infty$. [Hint: the trigonometric polynomials are dense in L^2]. Use this to show that T is mixing.

1.5 Let $T_\alpha(z) = e^{i\alpha}z$ be the rotation through α radians on $\mathbb{S}^1$. Prove that T_α is ergodic if and only if $\alpha/2\pi$ is irrational.

1.6 Prove the following uniform distribution theorem of Bohl, Sierpinski and Weyl. Let T_α be an ergodic rotation of the circle $\mathbb{S}^1$, and let $f : \mathbb{S}^1 \to \mathbb{C}$ be Riemann integrable. Then

$$\frac{1}{N} \sum_{n=0}^{N-1} f(T_\alpha^n z) \to \int f$$

for every $z \in \mathbb{S}^1$. First show this for functions of the form $f(z) = z^q$, $q \in \mathbb{Z}$, then for trigonometric polynomials, and finally for Riemann integrable functions. [Hint: if f is Riemann integrable, then for any $\varepsilon > 0$ there exist trigonometric polynomials p and P with $p(z) < f(z) < P(z)$ for all z, and $\int (p(z) - P(z))\, d\mu < \varepsilon$.]

1.7 Use 1.6 to find the distribution of the first digits of 2^n. Show that 2^n has first digit k if and only if $k \leq \frac{2^n}{10^r} < k+1$, for some integer r. Let $\alpha = \log_{10} 2$, and use () to denote the fractional part of a real number. Then 2^n has first digit k if and only if $(n\alpha) \in [\log_{10} k, \log_{10}(k+1))$. Deduce from Exercise 1.6 that if $\tau(n,k)$ is the number of times that k appears as the first digit in $1, 2, 4, \dots, 2^{n-1}$, then

$$\tfrac{1}{n}\tau(n,k) \to \log_{10}(1 + \tfrac{1}{k}).$$

1.8 Let T be a measure–preserving transformation of the Lebesgue space $(X, \mathcal{B}, \mu)$, and let $A_n(f)(x) = \frac{1}{n}\sum_{j=0}^{n-1} f(T^j x)$. Prove *von Neumann's mean ergodic theorem:* if $f \in L^2(\mu)$ then $A_n(f)$ converges in L^2 to a function $\bar{f}$ with $\bar{f}(Tx) = \bar{f}(x)$ for almost every $x \in X$. Do this by writing the Hilbert space $L^2(\mu)$ as a direct sum $\overline{range(U_T - I)} \oplus \overline{range(U_T - I)^{\perp}}$ and considering the two pieces of f separately. (Here $U_T(f)(x) = f(Tx)$ is the unitary operator on $L^2(\mu)$ associated to T).

1.9 Deduce from 1.8 that if T is an ergodic measure–preserving transformation, then $A_n(f)$ converges in $L^2(\mu)$ to the constant $\int f d\mu$ (that is, $\|A_n(f) - \int f d\mu\| \to 0$ as $n \to \infty$).

1.10 In fact much more is true: the *Birkhoff ergodic theorem* asserts that the convergence in Exercise 1.9 happens *pointwise almost everywhere.* Prove that this convergence implies that T is ergodic.

1.11 Prove that a measure–preserving transformation T of $(X, \mathfrak{B}, \mu)$ is ergodic if and only if

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

for every $A, B \in \mathfrak{B}$.

1.12 For each $k \geq 1$, construct an example of a $\mathbb{Z}^2$ action by automorphisms of a compact abelian group that is k–fold mixing but not $(k+1)$–fold mixing.

1.13 Prove directly (using trigonometric polynomials) that an ergodic automorphism of a torus is mixing. Prove that an ergodic automorphism of a torus is mixing of all orders.

1.14 Define the map $T : (0,1) \to (0,1)$ by $T(x) = \{\frac{1}{x}\}$ where $\{s\}$ denotes that fractional part of s. Check that T is well–defined if the rationals are deleted from $(0,1)$. Show that T preserves the probability measure μ defined by $\mu(A) = \frac{1}{\log 2}\int_A \frac{dx}{1+x}$, and prove that T is an ergodic transformation with respect to μ. Check that the rationals in $(0,1)$ form a null set with respect to μ.

The transformation T relates to continued fractions as follows. If x in $(0,1)\backslash\mathbb{Q}$ has continued fraction expansion

$$x = [a_1, a_2, \dots] = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \dots}}}}$$

then $a_n = (T^{n-1}(x))^{-1} - T^n(x)$. Using 1.10 above, prove that for almost every $x \in (0,1)$ the following hold:

a. $\lim_{n\to\infty} \frac{1}{n}(a_1 + a_2 + \cdots + a_n) = \infty$;

b. $\lim_{n\to\infty} \sqrt[n]{a_1 a_2 \dots a_n} = \prod_{k=1}^{\infty}(1 + \frac{1}{k^2+2k})^{\log k/\log 2} = 2.67\dots$;

c. $\lim_{n\to\infty} \frac{\log q_n}{n} = \frac{\pi^2}{12\log 2}$.

In (c) above, q_n is the nth partial quotient in the continued fraction of x: $[a_1, a_2, \dots] = \frac{p_n}{q_n}$ in lowest terms.

1.15 Prove Lemma 1.16.

1.16 Prove Theorem 1.3 by the following steps (consult any book on Functional Analysis or Linear Operators for the details). Let X be a compact Hausdorff space, and let $f : X \to X$ be a continuous map.

a. Define $T_f : C(X) \to C(X)$ by $T_f(g)(x) = g(f(x))$, where $C(X)$ is the algebra of continuous real–valued functions on X made into a complete normed linear space with norm $\|g\| = \max_{x\in X} |g(x)|$.

b. Show that T_f is *positive* (that is, $g \geq 0$ implies that $T_f(g) \geq 0$). Show that $\|T_f(g)\| \leq \|g\|$ and the constant function $g(x) = 1$ is fixed by T_f. Show that if f is onto, then $\|T_f(g)\| = \|g\|$.

c. Define a pairing $\langle\ ,\ \rangle$ between $C(X)$ and the space $M(X)$ of all finite signed Borel measures on X by setting $\langle g, \mu\rangle = \int_X g(x)d\mu(x)$.

d. Let T_f^* denote the operator *adjoint* to T_f, defined by $\langle g, T_f^*\mu\rangle = \langle T_f g, \mu\rangle$.

e. Introduce the total variation norm $\|\mu\| = |\mu|(X)$ on $M(X)$, and show that T_f^* is continuous on $M(X)$.

f. Translate (2) to deduce that $\mu \geq 0 \implies T_f^*\mu \geq 0$; $\|T_f^*\mu\| \leq \|\mu\|$; $(T_f^*\mu)(X) = \mu(X)$. Show that if f is injective, then $\|T_f^*\mu\| = \|\mu\|$.

g. Let $K = \{\mu \in M(X) \mid \mu \geq 0, \mu(X) = \|\mu\| = 1\}$. Show that K is non–empty, convex, weak*–compact, and preserved by T_f^*.

h. Deduce by the Kakutani–Markov fixed point theorem that T_f^* has a fixed point in $\mu \in K$.

i. Show that any fixed point for T_f^* is an invariant Borel probability measure for f.

1.17 We began this chapter by considering *flows*: actions of the reals. Very few of the transformations we have considered can come from flows. Prove the following.

a. Let X be a compact topological space. A flow $T : \mathbb{R} \to Homeo(X)$ is *continuous* if the map $(x,t) \mapsto T_t(x)$ is continuous. Show that if a homeomorphism S is the time one map of some continuous flow T, then $\pi_1(S)$ is trivial. Deduce that an ergodic toral automorphism cannot be the time–one map of a continuous flow.

b. Let X be a torus, and let S be an ergodic automorphism of X. A flow T on X is a *measurable algebraic flow* if the map $(x,t) \mapsto T_t(x)$ is measurable, and each T_t is an automorphism of X. Show that S cannot be the time–one map of any algebraic measurable flow.

1.18 Show that an automorphism of the 2–torus is ergodic if and only if the corresponding matrix in $GL_2(\mathbb{Z})$ has determinant -1 and trace exceeding 2 in modulus or determinant 1 and non–zero trace.

1.19 Consider an automorphism of the 3–torus. The corresponding integer matrix has characteristic equation $\chi(\lambda) = \lambda^3 - a\lambda^2 + b\lambda \pm 1$. Describe completely the condition for ergodicity in terms of a, b, and the determinant.

1.20 Justify the statement made in Definition 1.10.

1.21 Let S be the endomorphism of the 2–torus corresponding to the matrix $\begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$, and let T be the endomorphism of the 1–torus corresponding to the matrix $[\,2\,]$. Prove directly that S and T are measurably isomorphic.