Notes on ergodic theory

Michael Hochman[1]

July 5, 2017

Contents

Preface

These are notes from an introductory course on ergodic theory given at the Hebrew University of Jerusalem in the spring semester of 2017. The goal is to cover the foundations of ergodic theory – recurrence, ergodicity, ergodic theorems, ergodic decompositions, some topological dynamics, weak mixing vx. pure point spectrum, entropy as an isomorphism invariant, Shannon-McMillan-Breiman theorem, K-systems and the Pinsker factor. In addition we will hopefully cover some other subjects, among them possibly the Rohlin lemma, disjointness, and return times theorems.

Chapter 1

Introduction

At its most basic level, dynamical systems theory is about understanding the long-term behavior of a map $T : X \to X$ under iteration.

X is called the *phase space* and the points $x \in X$ may be imagined to represent the possible states of the "system".

The map T determines how the system evolves with time: time is discrete, and from state x it transitions to state Tx in one unit of time.

Thus if at time 0 the system is in state x, then the state at all future times $t = 1, 2, 3, \ldots$ are determined: at time $t = 1$ it will be in state Tx, at time $t = 2$ in state $T(Tx) = T^2x$, and so on; in general we define

$$T^n x = \underbrace{T \circ T \circ \ldots \circ T}_{n}(x)$$

so $T^n x$ is the state of the system at time n, assuming that at time zero it is in state x.

The "future" trajectory of an initial point x is called the **(forward) orbit**, denoted

$$O_T(x) = \{x, Tx, T^2x, \ldots\}$$

When T is invertible, $y = T^{-1}x$ satisfies $Ty = x$, so it represents the state of the world at time $t = -1$, and we write $T^{-n} = (T^{-1})^n = (T^n)^{-1}$. The one can also consider the **full-** or **two-sided orbit**

$$O_T^{\pm}(x) = \{T^n x \,:\, n \in \mathbb{Z}\}$$

There are many questions one can ask. Does a point $x \in X$ necessarily return close to itself at some future time, and how often this happens? If we fix another set A, how often does x visit A? If we cannot answer this for all points, we would like to know the answer at least for typical points. What is the behavior of pairs of points $x, y \in X$: do they come close to each other? given another pair x', y', is there some future time when x is close to x' and y is close to y'? If $f : X \to \mathbb{R}$, how well does the value of f at time 0 predict its value at future times? How does randomness arise from deterministic evolution of time? And so on.

In the set-theoretic framework above is too general to say anything except trivialities, but things become more interesting when more structure is given to X and T. The most common asumptions are that X is a topological space, and T continuous (this is called topological dynamics); X is a compact manifold and T a once- or many-times differentiable map (this is called smooth dynamics); or that there is a measure on X and T may preserve it (this is called ergodic theory). We will come give precise definitions shortly.

One might ask why these assumptions are natural ones to make. First, in many important examples, all these structures are present. In particular, a theorem of Liouville from celestial mechanics states that for Hamiltonian systems, e.g. systems governed by Newton's laws, all these assumptions are satisfied. Another example comes from the algebraic settings, e.g. automorphisms of compact abelian groups, or flows on homogeneous spaces. On theother hand, in some situations only some of these structures is available. An example is can be found in the applications of ergodic theory to combinatorics, where there is no smooth structure in sight. Thus the study of these assumptions individually is motivated by more than mathematical curiosity.

In these notes we focus primarily on ergodic theory, which is in a sense the most general of these theories. It is also the one with the most analytical flavor, and a surprisingly rich theory emerges from fairly modest axioms. The purpose of this course is to develop some of these fundamental results. We will also touch upon some applications and connections with dynamics on compact metric spaces.

Chapter 2

Measure preserving systems

2.1 Measure preserving systems

Definition 2.1.1. A **measure preserving system** (m.p.s.) is a quadruple $\mathcal{X} = (X, \mathcal{B}, \mu, T)$ where $(X, \mathcal{B}, \mu)$ is a probability space, and $T : X \to X$ is a measurable, measure-preserving map: that is

$$T^{-1}A \in \mathcal{B} \qquad \text{and} \qquad \mu(T^{-1}A) = \mu(A) \qquad \text{for all } A \in \mathcal{B}$$

If T is invertible and T^{-1} is measurable then it satisfies the same conditions, and the system is called invertible.

Example 2.1.2. Let X be a finite set with the σ-algebra of all subsets and normalized counting measure μ, and $T : X \to X$ a bijection. This is a measure preserving system, since measurability is not a question, and

$$\mu(T^{-1}A) = \frac{1}{|X|}|T^{-1}A| = \frac{1}{|X|}|A| = \mu(A)$$

This example is very trivial but many of the phenomena we will encounter can already be observed (and usually are easy to prove) for finite systems. It is worth keeping this example in mind.

Example 2.1.3. The identity map on any measure space is measure preserving.

Example 2.1.4 (Circle rotation)**.** Let $X = S^1 = \{x \in \mathbb{C} : |x| = 1\}$ be the unit circle with the Borel sets $\mathcal{B}$ and normalized length measure μ (the image of Lebesgue measure on $[0, 1]$ unter $t \mapsto (\cos 2\pi t, \sin 2\pi t)$). Let $\alpha \in \mathbb{R}$ and let $R_\alpha : S^1 \to S^1$ denote the rotation by angle α, that is, $z \mapsto e^{2\pi i\alpha}z$ (if $\alpha \notin 2\pi\mathbb{Z}$ then this map is not the identity). Then R_α preserves μ, since it transforms intervals to intervals of equal length. More precisely, consider the algebra $\mathcal{A}$ of half-open intervals with endpoints in $\mathbb{Q}[\alpha]$. Then T preserves $\mathcal{A}$ and preserves the restriction of the measure to it, hence it preserves the extension of the measure to the σ-algebra generated by $\mathcal{A}$, which is the measure μ on $\mathcal{B}$.

This example is sometimes described as $X = \mathbb{R}/\mathbb{Z}$, then the map is written additively, $x \mapsto x + \alpha$.

Here is a generalization: let G be a compact group with normalized Haar measure μ, fix $g \in G$, and consider $R_g : G \to G$ given by $x \to gx$. To see that $\mu(T^{-1}A) = \mu(A)$, let $\nu(A) = \mu(g^{-1}A)$, and note that ν is a Borel probability measure that is right invariant: for any $h \in H$, $\nu(Bh) = \mu(g^{-1}Bh) = \mu(g^{-1}B) = \nu(B)$. This $\nu = \mu$.

Example 2.1.5 (Doubling map)**.** Let $X = [0,1]$ with the Borel sets and Lebesgue measure, and let $Tx = 2x \bmod 1$. This map is onto is not 1-1, in fact every point has two pre-images which differ by $\frac{1}{2}$, except for 1, which is not in the image. To see that T_2 preserves μ, note that for any interval $I = [a, a+r) \subseteq [0,1)$,

$$T_2^{-1}[a, a+r) = [\frac{a}{2}, \frac{a+r}{2}) \cup [\frac{a}{2} + \frac{1}{2}, \frac{a+r}{2} + \frac{1}{2})$$

which is the union of two intervals of length half the length; the total length is unchanged.

Note: TI is generally of larger length than I. the property of measure preservation is defined by $\mu(T^{-1}A) = \mu(A)$.

If we identify $[0,1)$ with $\mathbb{R}/\mathbb{Z}$ then the example above coincides with the endomorphism $x \mapsto 2x$ of the compact group $\mathbb{R}/\mathbb{Z}$. Equivalently, if we identify $[0,1]$ it with $S^1 \subseteq \mathbb{C}$, then the map is $z \mapsto z^2$.

This example generalizes easily to $T_a x = ax \bmod 1$ for any $1 < a \in \mathbb{N}$. For non-integer $a > 1$ Lebesgue measure is not preserved. More generally one can consider a compact group G with Haar measure μ and an endomorphism $T : G \to G$. Then from uniqueness of Haar measure one again can show that T preserves μ.

Example 2.1.6. (Symbolic spaces and product measures) Let A be a finite set, $|A| \geq 2$, which we think of as a discrete topological space. Let $X^+ = A^{\mathbb{N}}$ and $X = A^{\mathbb{Z}}$ with the product σ-algebras. In both cases there is a map which shifts "to the right",

$$(\sigma x)_n = x_{n+1}$$

In the case of X this is an invertible map (the inverse is $(\sigma x)_n = x_{n-1}$). In the one-sided case X^+, the shift is not 1-1 since for every sequence $x = x_1 x_2 \ldots \in A^{\mathbb{N}}$ we have $\sigma^{-1}(x) = \{x_0 x_1 x_2 \ldots : x_0 \in A\}$.

Let p be a probability measure on A and $\mu = p^{\mathbb{Z}}$, $\mu^+ = p^{\mathbb{N}}$ the product measures on X, X^+, respectively. By considering the algebra of cylinder sets $[a] = \{x : x_i = a_i\}$, where a is a finite sequence of symbols, one may verify that σ preserves the measure.

Example 2.1.7. (Stationary processes) In probability theory, a sequence $\{\xi_n\}_{n=1}^{\infty}$ of random variables is called **stationary** if the distribution of a consecutive n-tuple $(\xi_k, \ldots, \xi_{k+n-1})$ does not depend on where it begins; i.e. $(\xi_1, \ldots, \xi_n) = (\xi_k, \ldots, \xi_{k+n-1})$ in distribution for every k and n. Intuitively, this means that if we observe a finite sample from the process, the values that we see give no information about when the sample was taken.

From a probabilistic point of view it rarely matters what the sample space is and one may as well choose it to be $(X, \mathcal{B}) = (Y^{\mathbb{N}}, \mathcal{C}^{\mathbb{N}})$, where $(Y, \mathcal{C})$ is the range of the variables. On this space there is again defined the shift map $\sigma : X \to X$ given by $\sigma((y_n)_{n=1}^{\infty}) = (y_{n+1})_{n=1}^{\infty}$. For any $A_1, \ldots, A_n \in \mathcal{C}$ and k let

$$A^k = \underbrace{Y \times \ldots \times Y}_{k} \times A_1 \times \ldots \times A_n \times Y \times Y \times Y \times \ldots$$

Note that $\mathcal{B}$ is generated by the family of such sets. If P is the underlying probability measure, then stationarity means that for any $A_1, \ldots, A_n$ and k,

$$P(A^0) = P(A^k)$$

Since $A^k = \sigma^{-k} A^0$ this shows that the family of sets B such that $P(\sigma^{-1}B) = P(B)$ contains all the sets of the form above. Since this family is a σ-algebra and the sets above generate $\mathcal{B}$, we see that σ preserves P.

There is a converse to this: suppose that P is a σ-invariant measure on $X = Y^{\mathbb{N}}$. Define $\xi_n(y) = y_n$. Then (ξ_n) is a stationary process.

Example 2.1.8. (Hamiltonian systems) The notion of a measure-preserving system emerged from the following class of examples. Let $\Omega = \mathbb{R}^{2n}$; we denote $\omega \in \Omega$ by $\omega = (p, q)$ where $p, q \in \mathbb{R}^n$. Classically, p describes the positions of particles and q their momenta. Let $H : \Omega \to \mathbb{R}$ be a smooth map and consider the differential equation

$$\begin{aligned} \frac{d}{dt} p_i &= -\frac{\partial H}{\partial q_i} \\ \frac{d}{dt} \dot{q}_i &= \frac{\partial H}{\partial p_i} \end{aligned}$$

Under suitable assumptions, for every initial state $\omega = (p_0, q_0) \in \Omega$ and $t \in \mathbb{R}$ there is determines a unique solution $\gamma_\omega(t) = (p(t), q(t))$, and $\omega_t = \gamma_\omega(t)$ is the state of the world after evolving for a period of t started from ω.

Thinking of t as fixed, we have defined a map $T_t : \Omega \to \Omega$ by $T_t \omega = \gamma_\omega(t)$. Clearly

$$T_0(\omega) = \gamma_\omega(0) = \omega$$

We claim that this is an action of $\mathbb{R}$. Indeed, notice that $\sigma(s) = \gamma_\omega(t + s)$ satisfies $\sigma(0) = \gamma_\omega(t) = \omega_t$ and $\dot{\sigma}(s) = \dot{\gamma_{\omega_t}}(t + s)$, and so $A(\sigma, \dot{\sigma}) = A(\gamma_\omega(t + s), \dot{\gamma}_\omega(t + s)) = 0$. Thus by uniqueness of the solution, $\gamma_{\omega_t}(s) = \gamma_\omega(t + s)$. This translates to

$$T_{t+s}(\omega) = \gamma_\omega(t + s) = \gamma_{\omega_t}(s) = T_s \omega_t = T_s(T_t \omega)$$

and of course also $T_{t+s} = T_{s+t} = T_t T_s \omega$. Thus $(T_t)_{t \in \mathbb{R}}$ is action of $\mathbb{R}$ on Ω.

It often happens that Ω contains compact subsets which are invariant under the action. For example there may be a notion of energy $E : \Omega \to \mathbb{R}$ that is preserved, i.e. $E(T_t\omega) = E(\omega)$, and then the level sets $M = E^{-1}(e_0)$ are invariant under the action. E is nice enough, M will be a smooth and often compact manifold. Furthermore, by a remarkable theorem of Liouville, if the equation governing the evolution is a Hamiltonian equation (as is the case in classical mechanics) then the flow preserves volume, i.e. $\mathrm{vol}(T_tU) = \mathrm{vol}(U)$ for every t and open (or Borel) set U. The same is true for the volume form on M.

2.2 Recurrence

A deep and basic properties of measure preserving systems is that they display "recurrence", meaning, roughly, that for typical x, anything that happens along its orbit happens infinitely often. This phenomenon was first discovered by Poincaré and bears his name.

Given a set A and $x \in A$ it will be convenient to say that x returns to A if $T^n x \in A$ for some $n > 0$; this is the same as $x \in A \cap T^{-n}A$. We say that x returns for A infinitely often if there are infinitely many such n.

The following proposition is, essentially, the pigeon-hole principle.

Proposition 2.2.1. *Let A be a measurable set, $\mu(A) > 0$. Then there is an n such that $\mu(A \cap T^{-n}A) > 0$.*

Proof. Consider the sets $A, T^{-1}A, T^{-2}A, \ldots, T^{-k}A$. Since T is measure preserving, all the sets $T^{-i}A$ have measure $\mu(A)$, so for $k > 1/\mu(A)$ they cannot be pairwise disjoint mod μ (if they were then $1 \geq \mu(X) \geq \sum_{i=1}^{k} \mu(T^{-i}A) > 1$, which is impossible). Therefore there are indices $0 \leq i < j \leq k$ such that $\mu(T^{-i}A \cap T^{-j}A) > 0$. Now,

$$T^{-i}A \cap T^{-j}A = T^{-i}(A \cap T^{-(j-i)}A)$$

so $\mu(A \cap T^{-(j-i)}A) > 0$, as desired. □

Theorem 2.2.2 (Poincare recurrence theorem)**.** *If $\mu(A) > 0$ then μ-a.e. $x \in A$ returns to A.*

Proof. Let

$$E = \{x \in A \, : \, T^n x \notin A \text{ for } n > 0\} = A \setminus \bigcup_{n=1}^{\infty} T^{-n}A$$

Thus $E \subseteq A$ and $T^{-n}E \cap E \subseteq T^{-n}E \cap A = \emptyset$ for $n \geq 1$ by definition. Therefore by the previous proposition, $\mu(E) = 0$. □

Corollary 2.2.3. *If $\mu(A) > 0$ then μ-a.e. $x \in A$ returns to A infinitely often.*

Proof. Let E be as in the previous proof. For any k-tuple $n_1 < n_2 < \ldots < n_k$, the set of points $x \in A$ which return to A only at times $n_1, \ldots, n_k$ satisfy $T^{n_k}x \in E$. Therefore,

$$\{x \in A \, : \, x \text{ returns to } A \text{ finitely often}\} = \bigcup_{k} \bigcup_{n_1 < \ldots < n_k} T^{-n_k}E$$

Hence the set on the left is the countable union of set of measure 0. □

In order to discuss of recurrence for individual points we suppose now that X is a metric space.

Definition 2.2.4. Let X be a metric space and $T : X \to X$. Then $x \in X$ is called *forward recurrent* if there is a sequence $n_k \to \infty$ such that $T^{n_k}x \to x$.

Proposition 2.2.5. *Let $(X, \mathcal{B}, \mu, T)$ by a measure-preserving system where X is a separable metric space and the open sets are measurable. Thenμ-a.e. x is forward recurrent.*

Proof. Let $A_i = B_{r_i}(x_i)$ be a countable sequence of balls that generate the topology. By Theorem 2.2.2, there are sets $A'_i \subseteq A_i$ of full measure such that every $x \in A'_i$ returns to A_i. Let $X_0 = X \setminus \bigcup(A_i \setminus A'_i)$, which is of full μ-measure. For $x \in X_0$ if $x \in A_i$ then x returns to A_i, so it returns to within $|\operatorname{diam} A_n|$ of itself. Since x belongs to A_n of arbitrarily small diameter, x is recurrent. □

When the phenomenon of recurrence was discovered it created quite a stir. Indeed, by Liouville's theorem it applies to Hamiltonian systems, such as planetary systems and the motion of molecules in a gas. In these settings, Poincaré recurrence seems to imply that the system is stable in the strong sense that it nearly returns to the same configuration infinitely often. This question arose original in the context of stability of the solar system in a weaker sense, i.e., will it persist indefinitely or will the planets eventually collide with the sun, or fly off into deep space. Stability in the strong sense above contradicts our experience. One thing to note, however, is the time frame for this recurrence is enormous, and in the celestial-mechanical or thermodynamics context it does not say anything about the short-term stability of the systems.

Recurrence also implies that there are no quantities that only increase as time moves forwards; this is on the face of it in contradiction of the second law of thermodynamics, which asserts that the thermodynamic entropy of a mechanical system increases monotonely over time. We say that a function $f : X \to \mathbb{R}$ is **increasing (respectively, constant) along orbits** if $f(Tx) \geq f(x)$ a.e. (respectively $f(Tx) = f(x)$ a.e.). This is the same as requiring that for a.e. x the sequence $f(x), f(Tx), f(T^2x), \ldots$ is non-decreasing (respectively constant). Although superficially stronger, the latter condition follows because for fixed n,

$$\begin{aligned} \mu(x : f(T^{n+1}(x)) < f(T^n x)) &= \mu(T^{-n}\{x : f(Tx) < f(x)\}) \\ &= \mu(x : f(Tx) < f(x)) \\ &= 0 \end{aligned}$$

and so the intersection of these events is still of measure zero. The same argument works for functions constant along orbits.

Corollary 2.2.6. *In a measure preserving system any measurable function that is increasing along orbits is a.s. constant along orbits.*

Proof. Let f be increasing along orbits. For every real number s let

$$E_s = \{x : f(x) < s < f(Tx)\}$$

Almost every x satisfies $f(Tx) \geq f(x)$. If x has this property, then $f(Tx) > f(x)$ if and only if $x \in E_s$ for some $s \in \mathbb{Q}$. Thus, it is enough to show that $\mu(\bigcup_{s\in\mathbb{Q}} E_s) = 0$, and for this it is enough to show $\mu(E_s) = 0$ for every s. Now note that outside a set of measure zero we have, by the non-increasing property, $f(T^n x) \geq f(Tx)$ for all $n \geq 1$; thus if $x \in E_s$ then $T^n x \notin E_s$ for all $n \geq 1$, and by Poincare recurrence, $\mu(E_s) = 0$. □

The last result highlights the importance of measurability. In the purely set-theoretic context, one can always choose a representative x from each orbit, and using it define $f(T^n x) = n$ for $n \geq 0$ (and also $n < 0$ if T is invertible). Then we have a function which is strictly increasing along orbits; but by the corollary, it cannot be measurable.

2.3 The Koopman operator

To every measure preserving system one can associate an isometry (or at least, nor-preserving self-map) of L^2, which captures a lot of information about the dynamics and is amenable to the tools of functional analysis. This is called the Koopman operator.

Let us first make some general oservationw. A function $T : X \to Y$ between sets induces a map $\widehat{T}$ sending a function with domain Y to a function with domain X. Namely,

$$\widehat{T}f(x) = f(Tx)$$

On the spaces of functions $f : Y \to \mathbb{R}$ or $f : Y \to \mathbb{C}$, the operator $\widehat{T}$ is linear, positive ($f \geq 0$ implies $\widehat{T}f \geq 0$), and multiplicative ($\widehat{T}(fg) = \widehat{T}f \cdot \widehat{T}g$). Also $|\widehat{T}f| = \widehat{T}|f|$ and $\widehat{T}(f^c) = (\widehat{T}f)^c$. When $(X, \mathcal{B})$ and $(Y, \mathcal{C})$ are measurable spaces and T is measurable, the induced map $\widehat{T}$ acts on the space of measurable functions on Y. Then they are topological spaces and T is continuous, it takes continuous functions to continuous functions.

For a map $T : X \to Y$ between measure spaces, we say that T maps the measure $\mu \in \mathcal{P}(X)$ to $\nu \in \mathcal{P}(Y)$ if $\mu(T^{-1}E) = \nu(E)$ for every measurable set $E \subseteq Y$.

Lemma 2.3.1. *Let $(X, \mathcal{B}, \mu)$, $(Y, \mathcal{C}, \nu)$ be a probability space, and $T : X \to Y$ measurable. Then T maps μ to ν if and only if $\int \widehat{T}f\, d\mu = \int f\, d\nu$ for every bounded measurable function $f : Y \to \mathbb{R}$ (or for every $f \in C(Y)$ if X is a compact).*

Proof. For $A \in \mathcal{C}$. Then

$$\int 1_A d\mu = \nu(A)$$

and, on the other hand, $\widehat{T}1_A(x) = 1_A(Tx) = 1_{T^{-1}A}(x)$, hence

$$\int \widehat{T}1_A \, d\mu = \mu(T^{-1}A)$$

Thus $\mu(T^{-1}A) = \nu(A)$ if and only if $\int 1_A d\nu = \int \widehat{T}1_A d\mu$. The latter condition clealry is equivalent to $\int f d\nu = \int \widehat{T} f d\mu$ for all simple functions (i.e. finite linear sums of indicator functions); and since since every bounded measurable function is a bounded limit of simple functions and $\widehat{T}$ preserves monotone limits, by the bounded convergence theorem, this is equivalent to the desired property.

In the case of a compact metric space, one begins the same way, and completes the proof using that $C(X) \subseteq L^1(\mu)$ densely. □

Proposition 2.3.2. *Let $T : X \to Y$ be a map between measurable spaces, mapping $\mu \in \mathcal{P}(X)$ to $\nu \in \mathcal{P}(Y)$. Then $\widehat{T}$ maps $L^p(\nu)$ isometrically into $L^p(\mu)$ for every $1 \leq p \leq \infty$. If the spaces are compact and f continuous and onto, then $\widehat{T}$ maps $C(Y)$ isometrically into $C(X)$ (with respect to the supremum norm).*

Proof. First note that if f is an a.e. defined function then $\widehat{T}f$ is also, because if E is the nullset where f is not defined then $T^{-1}E$ is the set where $\widehat{T}f$ is not defined, and $\mu(T^{-1}E) = \nu(E) = 0$. Thus $\widehat{T}$ acts on equivalence classes of measurable functions mod μ. Now, for $1 \leq p < \infty$ we have

$$\left\|\widehat{T}f\right\|_p^p = \int |\widehat{T}f|^p \, d\mu = \int \widehat{T}(|f|^p) \, d\mu = \int |f^p| \, d\nu = \|f\|_p^p$$

For $p = \infty$ the claim follows from the identity $\|f\|_\infty = \lim_{p\to\infty} \|f\|_p$.

In the topological setting, since T is assumed to be onto, it is clear that $\sup_{x\in X} |f(Tx)| = \sup_{y\in Y} |f(y)|$, and the claim follows. □

Corollary 2.3.3. *In a measure preserving system $(X, \mathcal{B}, \mu, T)$, the induced map $\widehat{T}$ on functions is a norm-preserving linear operator of L^p, and if T is invertible then $\widehat{T}$ is an isometry of L^p. If X is compact and T continuous and onto, the same holds with respect to $(C(X), \|\cdot\|_\infty)$.*

Definition 2.3.4. The induced operator $\widehat{T} : L^2(\mu) \to L^2(\mu)$ in a measure preserving system is called the **Koopman operator**.

When T is invertible the Koopman operator is a unitary operator and opens up the door for using spectral techniques to study the underlying system. We will return to this idea later. We usually write T instead of $\widehat{T}$. This introduces slight ambiguity but the meaning should usually be clear from the context.

As a simple demonstration we end this section by re-formulating Poincare recurrence in a functional form:

Corollary 2.3.5. *Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system and $f \in L^2(\mu)$. Then for a.e. x, if $f(x) > 0$ then $\sum_{i=1}^{n} T^i f(x) \to 0$ as $n \to \infty$.*

Proof. For $\varepsilon > 0$ let $E_\varepsilon = \{f > \varepsilon\}$. Then by Poincare recurrence, for a.e. $x \in E_\varepsilon$ we have $T^i f(x) = f(T^i x) \in E_\varepsilon$ for infinitely many i, so $f(T^i x) > \varepsilon$ for infinitely many i and the series diverges. Thus the series diverges for a.e. $x \in \bigcup_{n=1}^\infty E_{1/n}$, and the union is equal to $\{f > 0\}$ up to measure zero. $\square$

Chapter 3

Ergodicity

In this section and the following ones we will study how it may be decomposed into simpler systems.

3.1 Ergodicity

Definition 3.1.1. Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system. A measurable set $A \subseteq X$ is **invariant** if $T^{-1}A = A$. The system is **ergodic** if there are no non-trivial invariant sets; i.e. every invariant set has measure 0 or 1.

If A is invariant then so is $X \setminus A$. Indeed,

$$T^{-1}(X \setminus A) = T^{-1}X \setminus T^{-1}A = X \setminus A$$

Thus, ergodicity is an irreducibility condition: a non-ergodic system the dynamics splits into two (nontrivial) parts which do not "interact", in the sense that an orbit in one of them never enters the other.

Example 3.1.2. Let X be a finite set with normalized counting measure, and $T : X \to X$ a 1-1 map. If X consists of a single orbit then the system is ergodic, since any invariant set that is not empty contains the whole orbit. In general, X splits into the disjoint (finite) union of orbits, and each of these orbits is invariant and of positive measure. Thus the system is ergodic if and only if it consists of a single orbit.

Note that every (invertible) system splits into the disjoint union of orbits. However, these typically have measure zero, so do not in themselves prevent ergodicity.

Example 3.1.3. By taking disjoint unions of measure preserving systems with the normalized sum of the measures, one gets many examples of non-ergodic systems.

Definition 3.1.4. A function $f : X \to Y$ for some set Y is invariant if $f(Tx) = f(x)$ for all $x \in X$.

The primary example is 1_A when A is invariant.

Lemma 3.1.5. *The following are equivalent:*

1. $(X, \mathcal{B}, \mu, T)$ *is ergodic.*
2. *If* $T^{-1}A = A \bmod \mu$ *then* $\mu(A) = 0$ *or* 1.
3. *Every measurable invariant function is constant a.e.*
4. *If* $f \in L^1$ *and* $f \circ T = f$ *a.e. then* f *is constant a.e..*

Proof. (1) and (3) are equivalent since an invariant set A produces the invariant function 1_A, while if f is invariant and not a.e. constant then there is a measurable set U in the range of f such that $0 < \mu(f^{-1}U) < 1$. But this set is clearly invariant.

Exactly the same argument shows that (2) and (4) are equivalent.

We complete the proof by showing the equivalence of (3) and (4). Clearly (4) implies (3). Conversely, suppose that $f \in L^1$ and $Tf = f$ a.e. Let $g = \limsup f(T^n x)$. Clearly g is T-invariant (since $g(Tx)$ is the limsup of the shifted sequence $f(T^{n+1}x)$, and is the same as the limsup of $f(T^n x)$, which is $g(x)$). The proof will be done by showing that $g = f$ a.e. This is true at a point x if $f(T^n x) = f(x)$ for all $n \geq 0$, and for this it is enough that $f(T^{n+1}x) = f(T^n x)$ for all $n \geq 0$; equivalently, that $T^n x \in \{f \circ T = f\}$ for all n, i.e. that $x \in \bigcap T^{-n}\{f \circ T = f\}$. But this is an intersection of sets of measure 1 and hence holds for a.e. x, as desired. □

Example 3.1.6 (Irrational circle rotation)**.** Let $R_\alpha(x) = e^{2\pi i \alpha}x$ be an irrational circle rotation ($\alpha \notin \mathbb{Q}$) on S^1 with Lebesgue measure. We claim that this system is ergodic. Indeed, let $\chi_n(z) = z^n$ (the characters of the compact group S^1) and consider an invariant function $f \in L^\infty(\mu)$. Since $f \in L^2$, it can be represented in L^2 as a Fourier series $f = \sum a_n \chi_n$. Now,

$$T\chi_n(z) = (e^{2\pi i \alpha}z)^n = e^{2\pi i n \alpha}z^n = e^{2\pi i n \alpha}\chi_n$$

so from

$$f = Tf = \sum a_n T\chi_n = \sum e^{2\pi i n \alpha} a_n \chi_n$$

Comparing this to the original expansion we have $a_n = e^{2\pi i n \alpha}a_n$. Thus if $a_n \neq 0$ then $e^{2\pi i n \alpha} = 1$, which, since $\alpha \notin \mathbb{Q}$, can occur only if $n = 0$. Thus $f = a_0\chi_0$, which is constant.

Non-ergodicity means that one can split the system into two parts that don't "interact". The next proposition reformulates this in a positive way: ergodicity means that every pair of non-trivial sets do "interact".

Proposition 3.1.7. *The following are equivalent:*

1. $(X, \mathcal{B}, \mu, T)$ *is ergodic.*
2. *For any* $B \in \mathcal{B}$*, if* $\mu(B) > 0$ *then* $\bigcup_{n=N}^{\infty} T^{-n}B = X \bmod \mu$ *for every* N.

3. *If $A, B \in \mathcal{B}$ and $\mu(A), \mu(B) > 0$ then $\mu(A \cap T^{-n}B) > 0$ for infinitely many n.*

Proof. (1) implies (2): Given B let $B' = \bigcup_{n=N}^{\infty} T^{-n}B$ and note that

$$T^{-1}(B') = \bigcup_{n=N}^{\infty} T^{-1}T^{-n}B = \bigcup_{n=N+1}^{\infty} T^{-n}B \subseteq B'$$

Since $\mu(T^{-1}B') = \mu(B')$ we have $B' = T^{-1}B' \bmod \mu$, hence by ergodicity $B' = X \bmod \mu$.

(2) implies (3): Given A, B as in (3) we conclude from (2) that, for every N, $\mu(A \cap \bigcup_{n=N}^{\infty} T^{-n}B) = \mu(A)$, hence there some $n > N$ with $\mu(A \cap T^{-n}B) > 0$. This implies that there are infinitely many such n.

Finally if (3) holds and if A is invariant and $\mu(A) > 0$, then taking $B = X \setminus A$ clearly $A \cap \bigcup T^{-n}B = \emptyset$ for all n so $\mu(B) = 0$ by (3). Thus every invariant set is trivial. □

3.2 Mixing

Although a wide variety of ergodic systems can be constructed or shown abstractly to exist, it is surprisingly difficult to verify ergodicity of naturally arising systems. In most cases where ergodicity can be proved, it is because the system satisfies a stronger "mixing" property.

Definition 3.2.1. $(X, \mathcal{B}, \mu, T)$ is called *mixing* if for every pair A, B of measurable sets,

$$\mu(A \cap T^{-n}B) \to \mu(A)\mu(B) \qquad \text{as } n \to \infty$$

It is immediate from the definition that mixing systems are ergodic. The advantage of mixing over ergodicity is that it is enough to verify it for a "dense" family of sets A, B. It is better to formulate this in a functional way.

Lemma 3.2.2. *For fixed n, and $f, g \in L^2$, the map $(f, g) \mapsto \int f \cdot T^n g \, d\mu$ is multilinear and $\int f \cdot T^n g \, d\mu_2 \leq \|f\|_2 \|g\|_2$.*

Proof. Multilinearity is immediate using linearity of the induced map T on functions. Using Cauchy-Schwartz,

$$\int f \cdot T^n g \, d\mu \leq \|f\|_2 \|T^n g\|_2 = \|f\|_2 \|g\|_2$$ □

Proposition 3.2.3. *A measure preserving system $(X, \mathcal{B}, \mu, T)$ is mixing if and only if for every $f, g \in L^2$,*

$$\int f \cdot T^n g \, d\mu \to \int f \, d\mu \cdot \int g \, d\mu \qquad \text{as } n \to \infty \tag{3.1}$$

Furthermore this limit holds for ail $f, g \in L^2$ if and only if it holds for f, g in a dense subset of L^2, if and only if it holds for f, g in a set whose space is dense in $L^2(\mu)$.

Proof. We prove the second statement first. Note that both sides of (3.1) are multilinear in f and g, from which it follows that if we know it for $f, g \in \mathcal{F}$ then we know it for $f, g \in \operatorname{span} \mathcal{F}$. Thus we must show that if it holds for f, g in a dense subset $\mathcal{F} \subseteq L^2(\mu)$, then it holds for all $f, g \in L^2(\mu)$.

Also, in order to prove it for all $f, g \in L^2(\mu)$, it is enough to show it for functions with mean zero, because if we know it in this case, then for general $f, g \in L^2(\mu)$ if follows by replacing f by $f - \int f d\mu$ and g by $g - \int g d\mu$.

Thus, suppose that the desired limit holds for all pairs of functions in a dense subset $\mathcal{F} \subseteq L^2$. Let $f, g \in L^2$ with mean zero, and for $\varepsilon > 0$ choose $f', g' \in \mathcal{F}$ with $\|f - f'\| < \varepsilon$ and $\|g - g'\| < \varepsilon$. Note, in particular, that $\left|\int (f' - f)\, d\mu\right| \leq \|f - f'\|\, \|1\|$ by Cauchy-Schwartz, so since f has mean zero, $\left|\int f' d\mu\right| < \varepsilon$; and similarly $\left|\int g' d\mu\right| < \varepsilon$. Therefore,

$$
\begin{aligned}
\left|\int f \cdot T^n g \, d\mu\right| &\leq \left|\int (f - f' + f') \cdot T^n (g - g' + g')\, d\mu\right| \\
&\leq \left|\int (f - f') \cdot T^n g \, d\mu\right| + \left|\int f \cdot T^n (g - g')\, d\mu\right| + \\
&\quad + \left|\int (f - f') \cdot T^n (g - g')\, d\mu\right| + \left|\int f' \cdot T^n g' \, d\mu\right| \\
&\leq \varepsilon \|g\| + \|f\| \varepsilon + \left|\int f' \cdot T^n g' \, d\mu\right| \\
&\to \varepsilon(\|f\| + \|g\|) + \left|\int f' d\mu\right| \left|\int g' d\mu\right| \\
&< \varepsilon(\|f\| + \|g\|) + \varepsilon^2
\end{aligned}
$$

where in the second to last inequality we used the previous lemma twice, and in the limit we used $f', g' \in \mathcal{F}$. Since ε was arbitrary, this shows that $\left\|\int f \cdot T^n g \, d\mu\right\| \to 0 = \int f d\mu \cdot \int g d\mu$, as claimed.

For the first part, using the identities $\int 1_A \, d\mu = \mu(A)$, $T^n 1_A = 1_{T^{-n}A}$ and $1_A 1_B = 1_{A \cap B}$, we see that mixing is equivalent to (3.1) for indicator functions. If it holds for all L^2 functions it certainly does for indicators; convesely, the linear span of indicators is dense in L^2, so by the first part, if it holds for indicators, then it holds in all $L^2(\mu)$. ◻

Example 3.2.4. Let $X = A^{\mathbb{Z}}$ for a finite set A, take the product σ-algebra, and μ a product measure with marginal given by a probability vector $p = (p_a)_{a \in A}$. Let $\sigma : X \to X$ be the shift map $(\sigma x)_n = x_{n+1}$. We claim that this map is mixing and hence ergodic.

To prove this note that if $f(x) = \widetilde{f}(x_1, \ldots, x_k)$ depends on the first k coordinates of the input, then $\sigma^n f(x) = \widetilde{f}(x_{n+1}, \ldots, x_{n+k})$. If f, g are two such functions then for n large enough, $\sigma^n g$ and f depend on different coordinates, and hence, because μ is a product measure, they are independent in the sense of probability theory:

$$
\int f \cdot \sigma^n g \, d\mu = \int f \, d\mu \cdot \int \sigma^n g \, d\mu = \int f \, d\mu \cdot \int g \, d\mu
$$

so the same is true when taking $n \to \infty$. Mixing follows from the previous proposition.

We note that one cannot check ergodicity in a similar way. That is, it is not true that if a dense subset of functions $\mathcal{F} \subseteq L^2(X, \mathcal{B}, \mu)$ does not contain invariant functions (modulo μ), then the system is ergodic.

3.3 Kac's return time formula

We pause to give a nice application of ergodicity to estimation of the "recurrence rate" of points to a set.

Let $(X, \mathcal{B}, \mu, T)$ be ergodic and let $\mu(A) > 0$. Set $X_0 = \bigcup_{n=1}^{\infty} T^{-n}A$; we have seen that $\mu(X_0) = 1$ (Proposition 3.1.7). Thus for a.e. x there is a minimal $n \geq 1$ with $T^n x \in A$; we denote this number by $r_A(x)$ and note that r_A is measurable, since

$$\{r_A < k\} = \bigcup_{1 \leq i < k} T^{-i}A$$

Theorem 3.3.1 (Kac's formula). *Assume that T is invertible. Then $\int_A r_A \, d\mu = 1$; in particular, $\mathbb{E}(r_A|A) = 1/\mu(A)$, so the expected time to return to A starting from A is $1/\mu(A)$.*

Proof. Let $A_n = A \cap \{r_A = n\}$. Then

$$\int_A r_A d\mu = \sum_{n=1}^{\infty} n\mu(A_n) = \sum_{n=1}^{\infty} \sum_{k=1}^{n} \mu(T^k A_n)$$

The proof will be completed by showing that the sets $\{T^k A_n : n \in \mathbb{N}, 1 \leq k \leq n\}$ are pairwise disjoint and that their union has full measure. Indeed, for a.e. $x \in X$ there is a least $m \geq 1$ such that $y = T^{-m}x \in A$. Let $n = r_A(y)$. Clearly $m \leq n$, since if $n < m$ and $T^n y \in A$ then $T^n y = T^n T^{-m} x = T^{-(m-n)}x \in A$ and $m - n \geq 1$ is smaller than m. Thus $x \in T^m A_n$. This shows that the union of the given family is X up to a null set.

To show that the family is disjoint, suppose $x \in T^{m'} A_{n'}$ for some $(m', n') \neq (m, n)$. We cannot have $m' < m$ because then $T^{-m'}x \in A_{n'} \subseteq A$ would contradict minimality of m. We cannot have $m' > m$ because this would imply that $r_A(T^{-m'}x) \geq m' > m$, and at the same time $T^{-m}x = T^{m'-m}(T^{-m'}x) \in A_n \subseteq A$, implying $r_A(T^{-m'}x) \leq m' - m < m'$, a contradiction. Finally, $m = m'$ and $n \neq n'$ is impossible because then then $T^{-m}x \in A_n \cap A_{n'}$, despite $A_n \cap A_{n'} \neq \emptyset$. □

Even under the stated ergodicity assumption this result strengthens Poincare recurrence. It shows not only that a.e. $x \in A$ returns to A, but that it does so in finite expected time, and identifies the expectation. Simple examples show that the formula is incorrect in the non-ergodic case.

The invertability assumption is not necessary, it can be removed by "making the system invertible". We return to this later.

3.4 Ergodic measures as extreme points

For a measurable space $(X, \mathcal{B})$ and measurable map $T : X \to X$ let $\mathcal{P}_T(X)$ denote the set of T-invariant probability measures (if the σ-algebra is relevant we write $\mathcal{P}_T(X, \mathcal{B})$, but we usually omit it). It is clear that $\mathcal{P}_T(X)$ is a convex set. Recall that a point in a convex set is an extreme point if it cannot be written as a convex combination of other points in the set. In this section we shall see that the ergodic invariant measures as precisely the extreme points of $\mathcal{P}_T(X)$.

For a measure μ on X we write $T\mu$ for the measure $T\mu(A) = \mu(T^{-1}A)$. The map T sends $\mathcal{P}(X, \mathcal{B})$ into itself and clearly, μ is T-invariant if and only if $T\mu = \mu$.

Lemma 3.4.1. *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic measure preserving system and suppose that $\nu \in \mathcal{P}_T(X)$ and $\nu \ll \mu$. Then $\mu = \nu$.*

Proof. Let $f = d\nu/d\mu$. We claim that $f = 1$ μ-a.e. Given t let $E = \{f < t\}$; it suffices to show that this set is invariant μ-a.e. We first claim that the sets $E \setminus T^{-1}E$ and $T^{-1}E \setminus E$ are of the same μ-measure. Indeed,

$$\begin{aligned} \mu(E \setminus T^{-1}E) &= \mu(E) - \mu(E \cap T^{-1}E) \\ \mu(T^{-1}E \setminus E) &= \mu(T^{-1}E) - \mu(E \cap T^{-1}E) \end{aligned}$$

and since $\mu(E) = \mu(T^{-1}E)$, the right hand sides are equal, and hence also the left hand sides.

Now

$$\nu(E) = \int_E f\, d\mu = \int_{E \cap T^{-1}E} f\, d\mu + \int_{E \setminus T^{-1}E} f\, d\mu$$

On the other hand

$$\nu(E) = \nu(T^{-1}E) = \int_{T^{-1}E \cap E} f\, d\mu + \int_{(T^{-1}E) \setminus E} f\, d\mu$$

Subtracting we find that

$$\int_{E \setminus T^{-1}E} f\, d\mu = \int_{T^{-1}E \setminus E} f\, d\mu$$

On the left hand side the integral is over a subset of E, where $f < t$, so the integral is $< t\mu(E \setminus T^{-1}E)$; on the right it is over a subset of $X \setminus E$, where $f \geq t$, so the integral is $\geq t\mu(T^{-1}E \setminus E)$. Equality is possible only if the measure of these sets is 0, and since $\mu(E) = \mu(T^{-1}E)$, the set difference can be a μ-nullset if and only if $E = T^{-1}E \bmod \mu$, which is the desired invariance of E.

It follows that f is constant a.e., and since $\int f d\mu = \int 1 d\mu = 1$, we have $f \equiv 1$ μ-a.e. □

Remark 3.4.2. If T is invertible, there is an easier argument: for any measurable set A,

$$\begin{aligned} \nu(A) &= \int_A f d\mu \\ \nu(T^{-1}A) &= \int 1_{T^{-1}A} d\nu \\ &= \int T1_A \, d\nu \\ &= \int T1_A \cdot f d\mu \\ &= \int 1_A \cdot T^{-1} f \, d\mu \end{aligned}$$

where in the last line we used $\int g d\mu = \int T^{-1} g d\mu$ applied to $g = T1_A \cdot f$. We have found that $\int_A f d\mu = \int_A T^{-1} f d\mu$ for all measurable A, hence $f \circ T^{-1} = f$, and f is invariant.

Proposition 3.4.3. *The ergodic invariant measures are precisely the extreme points of* $\mathcal{P}_T(X)$.

Proof. If $\mu \in \mathcal{P}_T(X)$ is non-ergodic then there is an invariant set A with $0 < \mu(A) < 1$. Then $B = X \setminus A$ is also invariant. Let $\mu_A = \frac{1}{\mu(A)}\mu|_A$ and $\mu_B = \frac{1}{\mu(B)}\mu|_B$ denote the normalized restriction of μ to A, B. Clearly $\mu = \mu(A)\mu_A + \mu(B)\mu_B$, so μ is a convex combination of μ_A, μ_B, and these measures are invariant:

$$\begin{aligned} \mu_A(T^{-1}E) &= \frac{1}{\mu(A)}\mu(A \cap T^{-1}E) \\ &= \frac{1}{\mu(A)}\mu(T^{-1}A \cap T^{-1}E) \\ &= \frac{1}{\mu(A)}\mu(T^{-1}(A \cap E)) \\ &= \frac{1}{\mu(A)}\mu(A \cap E) \\ &= \mu_A(E) \end{aligned}$$

Thus μ is not an extreme point of $\mathcal{P}_T(X)$.

Conversely, suppose that $\mu = \alpha\nu + (1-\alpha)\theta$ for $\nu, \theta \in \mathcal{P}_T(X)$ and $\nu \neq \mu$. Clearly $\mu(E) = 0$ implies $\nu(E) = 0$, so $\nu \ll \mu$, and by the previous lemma $f = d\nu/d\mu \in L^1(\mu)$ is invariant. Since $1 = \nu(X) = \int f d\mu$, we know that $f \neq 0$, and since $\nu \neq \mu$ we know that f is not constant. Hence μ is not ergodic by Lemma 3.1.5. □

As an application we find that distinct ergodic measures are also separated at the spacial level:

Corollary 3.4.4. *Let μ, ν be ergodic measures for a measurable map T of a measurable space $(X, \mathcal{B})$. Then either $\mu = \nu$ or $\mu \perp \nu$.*

Proof. Suppose $\mu \neq \nu$ and let $\theta = \frac{1}{2}\mu + \frac{1}{2}\nu$. Since this is a nontrivial representation of θ as a convex combination, it is not ergodic, so there is a nontrivial invariant set A. By ergodicity, A must have μ-measure 0 or 1 and similarly for ν. They cannot be both 0 since this would imply $\theta(A) = 0$, and they cannot both have measure 1, since this would imply $\theta(A) = 1$. Therefore one is 0 and one is 1. This implies that A supports one of the measures and $X \setminus A$ the other, so $\mu \perp \nu$. □

This last proposition hints at the fact that we can perhaps partition the entire space into disjoint subsets, each supporting a single ergoic measure, This is often the case.

In the previous proposiiton we saw that ergodic measures are precisely the extreme points of $\mathcal{P}_T(X)$. In finite-dimensional vector spaces, a compact convex set is the convex hull of its extreme points; i.e., every point in it is a convex combination of extreme points. This gives some reason to believe that perhaps every invariant measure is a "convex combination" of ergodic ones. Of course this does not apply to $\mathcal{P}_T(X)$, which is not generally finite dimensional and does not a-priori carry a topology. But if one makes suitable assumptions, the Choquet theory of convex sets in topological vector spaces can be applied to get such a result. The topological issue can be addressed by assuming that X is a compact metric space; then the weak-* topology on measures makes $\mathcal{P}(X)$ compact, and assuming also that T is invariant, $\mathcal{P}_T(X)$ is closed. Then indeed every $\mu \in \mathcal{P}_T(X)$ is a "convex combination" of extreme points – i.e., of ergodic measures – in the sense that given μ there exists a measure ν on the (Borel) set of extreme points such that μ is the barycenter of ν. In other words, μ is an integral of ergodic measures. Then, for dynamical reasons, one can show that the representation is also unique.

We will prove all this and somewhat more later on, using more measure-theoretic tools.

Chapter 4

The ergodic theorem

4.1 A baby case

Let X be a finite set and $T : X \to X$ a bijection. For a function $f : X \to X$, consider the averages of f along orbits of T, specifically, set

$$S_N f(x) = \frac{1}{N} \sum_{n=0}^{N-1} f(T^n x)$$

We consider the limiting behavior of this average. First, suppose that X consists of a single T-orbit: $X = \{x_0, Tx_0, \ldots, T^{k-1}x_0\}$, so $|X| = k$. Then for any $x \in X$, we we have that $x, Tx, \ldots, T^{k-1}x$ ranges once over every point in X, and for each $y = T^i x$, by the same resoning the elements $T^i x, T^{i+1}x, \ldots, T^{i+k-1}x$ range once over all elements of X. Therefore, writing $N = nk+r$ with $n = [N/k]$ and $0 \leq r < k$, we have that eachpoint in X appears among $x, Tx, \ldots, T^{nk}x$ exactly n times; thus

$$\begin{aligned} \sum_{i=0}^{N-1} f(T^i x) &= n \sum_{y \in X} f(y) + \sum_{i=nk}^{nk+p-1} f(T^i x) \\ &= [\frac{N}{k}] \cdot k \cdot \frac{1}{k} \sum_{y \in X} f(y) + O(k \|f\|_\infty) \end{aligned}$$

Dividing by N and using $[N/k] = N/k + O(1)$, we find that

$$S_N f(x) = \frac{1}{k} \sum_{y \in X} f(y) + O_{k,f}(\frac{1}{N}) \to \frac{1}{k} \sum_{y \in X} f(y)$$

ad $N \to \infty$.

Now suppose that X is finite but not a single orbit. Then it decomposes into a finite union of orbits, $X = \bigcup X_j$. For $x \in X_j$ we have by the above that

$$S_N f(x) \to \frac{1}{|X_j|} \sum_{y \in X_j} f(y)$$

But the right hand side is nothing other than the conditional expectation of f on the σ-algebra $\mathcal{I}$ of T-invariant sets (namely, sets that are unions of complete orbits). Thus we have found that $S_N f(x) \to \mathbb{E}(f|\mathcal{I})(x)$.

4.2 Preliminaries

We have seen that in a measure preserving system, a.e. $x \in A$ returns to A infinitely often. Now we will see that more is true: these returns occur with a definite frequency which, in the ergodic case, is just $\mu(A)$; in the non-ergodic case the limit is $\mu_x(A)$, where μ_x is the ergodic component to which x belongs.

This phenomenon is better formulated at an analytic level in terms of averages of functions along an orbit. To this end let us introduce some notation. Let $T : V \to V$ be a linear operator of a normed space V, and suppose T is a contraction, i.e. $\|Tf\| \leq \|f\|$. This is the case when T is induced from a measure-preserving transformation (in fact we have equality). For $v \in V$ define

$$S_N v = \frac{1}{N} \sum_{n=0}^{N-1} T^n v$$

Note that in the dynamical setting, the frequency of visits x to A up to time N is $S_N 1_A(x) = \frac{1}{N} \sum_{n=0}^{N-1} 1_A(T^n x)$. Clearly S_N is linear, and since T is a contraction $\|T^n v\| \leq \|v\|$ for $n \geq 1$, so by the triangle inequality, $\|S_N v\| \leq \frac{1}{N} \sum_{n=0}^{N-1} \|T^n v\| \leq \|v\|$. Thus S_N are also contractions. This has the following useful consequence.

Lemma 4.2.1. *Let $T : V \to V$ as above and let $S : V \to V$ be another bounded linear operator. Suppose that $V_0 \subseteq V$ is a dense subset and that $S_N v \to Sv$ as $N \to \infty$ for all $v \in V_0$. Then the same is true for all $v \in V$.*

Proof. Let $v \in V$ and $w \in V_0$. Then

$$\limsup_{N\to\infty} \|S_N v - Sv\| \leq \limsup_{N\to\infty} \|S_N v - S_N w\| + \limsup_{N\to\infty} \|S_N w - Sv\|$$

Since $\|S_N v - S_N w\| = \|S_N(v - w)\| \leq \|v - w\|$ and $S_N w \to Sw$ (because $w \in V_0$), we have

$$\limsup_{N\to\infty} \|S_N v - Sv\| \leq \|v - w\| + \|Sw - Sv\| \leq (1 + \|S\|) \cdot \|v - w\|$$

Since $\|v - w\|$ can be made arbitrarily small, the lemma follows. □

4.3 The mean ergodic theorem

Historically, the first ergodic theorem is von-Neuman's "mean" ergodic theorem, which can be formulated in a purely Hilbert-space setting (and it is not hard to adapt it to L^P). Recall that if $T : V \to V$ is a bounded linear operator of a Hilbert space then $T^* : V \to V$ is the adjoint operator, characterized by $\langle v, Tw \rangle = \langle T^* v, w \rangle$ for $v, w \in V$, and satisfies $\|T^*\| = \|T\|$.

Lemma 4.3.1. *Let $T : V \to V$ be a contracting linear operator of a Hilbert space. Then $v \in V$ is T-invariant if and only if it is T^*-invariant.*

Remark 4.3.2. When T is unitary (which is one of the main cases of interest to us) this lemma is trivial. Note however that without the contraction assumption this is false even in $\mathbb{R}^d$.

Proof. Since $(T^*)^* = T$ it suffices to prove that $T^*v = v$ implies $Tv = v$.

$$\begin{aligned}\|v - Tv\|^2 &= \langle v - Tv, v - Tv\rangle \\ &= \|v\|^2 + \|Tv\|^2 - \langle Tv, v\rangle - \langle v, Tv\rangle \\ &= \|v\|^2 + \|Tv\|^2 - \langle v, T^*v\rangle - \langle T^*v, v\rangle \\ &= \|v\|^2 + \|Tv\|^2 - \langle v, v\rangle - \langle v, v\rangle \\ &= \|Tv\|^2 - \|v\|^2 \\ &\leq 0\end{aligned}$$

where the last inequality is because T is a contraction. □

Theorem 4.3.3 (Hilbert-space mean ergodic theorem)**.** *Let T be a linear contraction of a Hilbert space V, i.e. $\|Tv\| \leq \|v\|$. Let $V_0 \leq V$ denote the closed subspace of T-invariant vectors (i.e. $V_0 = \ker(T - I)$) and π the orthogonal projection to V_0. Then*

$$\frac{1}{N}\sum_{n=0}^{N-1} T^n v \to \pi v \qquad \textit{for all } v \in V$$

Proof. If $v \in V_0$ then $S_N v = v$ and so $S_N v \to v = \pi v$ trivially. Since $V = V_0 \oplus V_0^\perp$ and S_N is linear, it suffices for us to show that $S_N v \to 0 = \pi v$ for $v \in V_0^\perp$. The key insight is that $V_0^\perp$ can be identified as the space of *co-boundaries*,

$$V_0^\perp = \overline{\{v - Tv \,:\, v \in V\}}$$

or equivalently,

$$V_0 = \overline{\{v - Tv \,:\, v \in V\}}^\perp \tag{4.1}$$

Assuming this, by Lemma 4.2.1 we must only show that $S_N(v - Tv) \to 0$ for $v \in V$, and this follows from

$$\begin{aligned}S_N(v - Tv) &= \frac{1}{N}\sum_{n=0}^{N-1} T^n(v - Tv) \\ &= \frac{1}{N}(w - T^{N+1}w) \\ &\to 0\end{aligned}$$

where in the last step we used $\left\|w - T^{N+1}w\right\| \leq \|w\| + \left\|T^{N+1}w\right\| \leq 2\|w\|$.

To prove (4.1) it suffices to show that $w \perp \{v - Tv : v \in V\}$ if and only if $w \in V_0$. For any w and $v \in V$ we have the identity

$$\begin{aligned} \langle w, v - Tv \rangle &= \langle w, v \rangle - \langle w, Tv \rangle \\ &= \langle w, v \rangle - \langle T^* w, v \rangle \\ &= \langle w - T^* w, v \rangle \end{aligned}$$

We conclude that $w \perp v - Tv$ for all $v \in V$ if and only if $\langle w - T^* w, v \rangle = 0$ for all $v \in V$, if and only if $w - T^* w = 0$, which by the previous lemma happens if and only if $w \in V_0$. □

Now let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system and let T denote also the Koopman operator induced on L^2 by T. Then the space V_0 of T-invariant vectors is just $L^2(X, \mathcal{I}, \mu)$, where $\mathcal{I} \subseteq \mathcal{B}$ is the σ-algebra of invariant sets, and the orthogonal projection π to V_0 is just the conditional expectation operator, $\pi f = \mathbb{E}(f|\mathcal{I})$ (see the Appendix). We derive the following:

Corollary 4.3.4 (Dynamical mean ergodic theorem)**.** *Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system, let $\mathcal{I}$ denote the σ-algebra of invariant sets, and let π denote the orthogonal projection from $L(X, \mathcal{B}, \mu)$ to the closed subspace $L^2(X, \mathcal{I}, \mu)$. Then for every $f \in L^2$,*

$$\frac{1}{N} \sum_{n=0}^{N-1} T^n f \to \mathbb{E}(f|\mathcal{I}) \qquad \text{in } L^2$$

In particular, if the system is ergodic then the limit is constant:

$$\frac{1}{N} \sum_{n=0}^{N-1} T^n f \to \int f \, d\mu \qquad \text{in } L^2$$

Specializing to $f = 1_A$, and noting that L^2-convergence implies, for example, convergence in probability, the last result says that on an arbitrarily large part of the space, the frequency of visits of an orbit to A up to time N is arbitrarily close to $\mu(A)$, if N is large enough.

Corollary 4.3.5. *Let $(X, \mathcal{B}, \mu, T)$ be a dynamical system. Let $\mathcal{F} \subseteq L^2(\mu)$ be a dense set of functions. Then the system is ergodic if and only if $\lim \frac{1}{N} \sum_{n=0}^{N-1} f$ is a constant function for all $f \in \mathcal{F}$.*

Proof. The limit is the projection of f to $L^2(\mathcal{I})$. Since projection is continuous, and by assumption the image of $\mathcal{F}$ is contained in the constant functions, the image of the projection operator must be contained in the constant functions; thus $L^2(\mathcal{I}, \mu)$ consists only of constant functions, and the system is ergodic. □

4.4 The pointwise ergodic theorem

Very shortly after von Neumann's discovery of the mean ergodic theorem (and appearing in print before it), Birkhoff proved a stronger version in which convergence takes place a.e. and in L^1.

Theorem 4.4.1 (Pointwise ergodic theorem)**.** *Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system, let $\mathcal{I}$ denote the σ-algebra of invariant sets. Then for any $f \in L^1(\mu)$,*

$$\frac{1}{N}\sum_{n=0}^{N-1} T^n f \to \mathbb{E}(f|\mathcal{I}) \qquad \text{a.e. and in } L^1$$

In particular, if the system is ergodic then the limit is constant:

$$\frac{1}{N}\sum_{n=0}^{N-1} T^n f \to \int f\,d\mu \qquad \text{a.e. and in } L^1$$

We shall see several proofs of this result. The first and most "standard" proof follows the same scheme as the mean ergodic theorem: one first establishes the statement for a dense subspace $V \subseteq L^1$, and then uses some continuity property to extend to all of L^1. The first step is nearly identical to the proof of the mean ergodic theorem.

Proposition 4.4.2. *There is a dense subspace $V \subseteq L^1$ such that the conclusion of the theorem holds for every $f \in V$.*

Proof. We temporarily work in L^2. Let V_1 denote the set of invariant $f \in L^2$, for which the theorem holds trivially because $S_N f = f$ for all N. Let $V_2 \subseteq L^2$ denote the linear span of functions of the form $f = g - Tg$ for $g \in L^\infty$. The theorem also holds for these, since

$$\left\|g + T^{N+1}g\right\|_\infty \le \|g\|_\infty + \left\|T^{N+1}g\right\|_\infty = 2\|g\|_\infty$$

and therefore

$$\frac{1}{N}\sum_{n=0}^{N-1} T^n(g - Tg) = \frac{1}{N}(g - T^{N+1}g) \to 0 \qquad \text{a.e. and in } L^1$$

Set $V = V_1 + V_2$. By linearity of S_N, the theorem holds for $f \in V_1 + V_2$. Now, L^∞ is dense in L^2 and T is continuous on L^2, so $\overline{V}_2 = \overline{\{g - Tg : g \in L^2\}}$. In the proof of the mean ergodic theorem we saw that $L^2 = V_1 \oplus \overline{V}_2$, so $V = V_1 \oplus V_2$ is dense in L^2, and hence in L^1, as required. □

By Lemma 4.2.1, this proves the ergodic theorem in the sense of L^1-convergence for all $f \in L^1$. In order to similarly extend the pointwise version to all of L^1 we need a little bit of "continuity", which is provided by the following.

Theorem 4.4.3 (Maximal inequality)**.** *Let $f \in L^1$with $f \geq 0$ and $S_N f = \frac{1}{N}\sum_{n=0}^{N-1} T^n f$. Then for every t,*

$$\mu\left(x : \sup_N S_N f(x) > t\right) \leq \frac{1}{t}\int f\,d\mu$$

Before giving the proof let us show how this finishes the proof of the ergodic theorem. Write $S = \mathbb{E}(\cdot|\mathcal{I})$, which is a bounded linear operator on L^1, let $f \in L^1$ and $g \in V$. Then

$$\begin{aligned} |S_N f - Sf| &\leq |S_N f - S_N g| + |S_N g - Sg| \\ &\leq S_N|f-g| + |S_N g - Sf| \end{aligned}$$

Now, $S_N g \to Sg$ a.e., hence $|S_N g - Sf| \to |S(g-f)| \leq S|f-g|$ a.e. Thus,

$$\limsup_{N\to\infty} |S_N f - Sf| \leq \limsup_{N\to\infty} S_N|f-g| + S|g-f|$$

If the left hand side is $> \varepsilon$ then at least one of the terms on the right is $> \varepsilon/2$. Therefore,

$$\mu\left(\limsup_{N\to\infty}|S_N f - Sf| > \varepsilon\right) \leq \mu\left(\limsup_{N\to\infty} S_N|f-g| > \varepsilon/2\right) + \mu\left(S|g-f| > \varepsilon/2\right)$$

Now, by the maximal inequality, the first term on the right side is bounded by $\frac{1}{\varepsilon/2}\|f-g\|$. By the fact that S is conditional expectation we have $\int Sh\,d\mu = \int h\,d\mu$ for all $h \in L^1$, so by Markov's inequality, the second term is bounded by $\frac{1}{\varepsilon/2}\|g-f\|$ as well. Thus, for any $\varepsilon > 0$ and $g \in V$ we have found that

$$\mu\left(\limsup_{N\to\infty}|S_N f - Sf| > \varepsilon\right) \leq \frac{4}{\varepsilon}\|f-g\|$$

For each fixed $\varepsilon > 0$, the right hand side can be made arbitrarily close to 0, hence $\limsup|S_N f - Sf| = 0$ a.e. which is just $S_N f \to Sf = \mathbb{E}(f|\mathcal{I})$, as claimed.

We now return to the maximal inequality which will be proved by reducing it to a purely combinatorial statement about functions on the integers. Given a function $\widehat{f} : \mathbb{N} \to [0,\infty)$ and a set $\emptyset \neq I \subseteq \mathbb{N}$, the average of $\widehat{f}$ over I is denoted

$$S_I \widehat{f} = \frac{1}{|I|}\sum_{i\in I}\widehat{f}(i)$$

In the following discussion we write $[i,j]$ also for integer segments, i.e. $[i,j] \cap \mathbb{Z}$.

Proposition 4.4.4 (Discrete maximal inequality)**.** *Let $\widehat{f} : \mathbb{N} \to [0,\infty)$. Let $J \subseteq I \subseteq \mathbb{N}$ be finite intervals, and for each $j \in J$ let $I_j \subseteq I$ be a sub-interval of I whose left endpoint is j. Suppose that $S_{I_j}\widehat{f} > t$ for all $j \in J$. Then*

$$S_I \widehat{f} > t \cdot \frac{|J|}{|I|}$$

Proof. Suppose first that the intervals $\{I_j\}$ are disjoint. Then together with $U = I \setminus \bigcup I_j$ they form a partition of I, and by splitting the average $S_I\widehat{f}$ according to this partition, we have the identity

$$S_I\widehat{f} = \frac{|U|}{|I|}S_U\widehat{f} + \sum \frac{|I_j|}{|I|}S_{I_j}\widehat{f}$$

Since $\widehat{f} \geq 0$ also $S_U\widehat{f} \geq 0$, and so

$$S_I\widehat{f} \geq \sum \frac{|I_j|}{|I|}S_{I_j}\widehat{f} \geq \frac{1}{|I|}\sum t|I_j| \geq t\frac{|\bigcup I_j|}{|I|}$$

Now, $\{I_j\}_{j\in J}$ is not a disjoint family, but the above applies to every disjoint sub-collection of it. Therefor we will be done if we can extract from $\{I_j\}_{j\in J}$ a disjoint sub-collection whose union is of size at least $|J|$. This is the content of the next lemma. □

Lemma 4.4.5 (Covering lemma)**.** *Let $I, J, \{I_j\}_{j\in J}$ be intervals as above. Then there is a subset $J_0 \subseteq J$ such that (a) $J \subseteq \bigcup_{i\in J_0} I_j$ and (b) the collection of intervals $\{J_i\}_{i\in J_0}$is pairwise disjoint.*

Proof. Let $I_j = [j, j+N(j)-1]$. We define $J_0 = \{j_k\}$ by induction using a greedy procedure. Let $j_1 = \min J$ be the leftmost point. Assuming we have defined $j_1 < \ldots < j_k$ such that $I_{j_1}, \ldots, I_{j_k}$are pairwise disjoint and cover $J \cap [0, j_k+N(j_k)-1]$. As long as this is not all of J, define

$$j_{k+1} = \min\{I \setminus [0, j_k + N(j_k) - 1]\}$$

It is clear that the extended collection satisfies the same conditions, so we can continue until we have covered all of J. □

We return now to the dynamical setting. Each $x \in X$ defines a function $\widehat{f} = \widehat{f}_x : \mathbb{N} \to [0,\infty)$ by evaluating f along the orbit:

$$\widehat{f}(i) = f(T^i x)$$

Let

$$A = \{\sup_N S_N f > t\}$$

and note that if $T^j x \in A$ then there is an $N = N(j)$ such that $S_N f(T^j x) > t$. Writing

$$I_j = [j, j + N(j) - 1]$$

this is the same as

$$S_{I_j}\widehat{f} > t$$

Fixing a large M (we eventually take $M \to \infty$), consider the interval $I = [0, M-1]$ and the collection $\{I_j\}_{j\in J}$, where

$$J = J_x = \{0 \leq j \leq M-1 \,:\, T^j x \in A \text{ and } I_j \subseteq [0, M-1]\}$$

The proposition then gives

$$S_{[0,M-1]}\widehat{f} > t \cdot \frac{|J|}{M}$$

In order to estimate the size of J we will restrict to intervals of some bounded length $R > 0$ (which we eventually will send to infinity). Let

$$A_R = \{\sup_{0\leq N\leq R} S_N f > t\}$$

Then

$$J \supseteq \{0 \leq j \leq M - R - 1 \,:\, T^j x \in A_R\}$$

and if we write $h = 1_{A_R}$, then we have

$$\begin{aligned} |J| &\geq \sum_{j=0}^{M-R-1} \widehat{h}(j) \\ &= (M - R - 1)S_{[0,M-R-1]}\widehat{h} \end{aligned}$$

With this notation now in place,the above becomes

$$S_{[0,M-1]}\widehat{f_x} > t \cdot \frac{M-R-1}{M} \cdot S_{[0,M-R-1]}\widehat{h_x} \tag{4.2}$$

and notice that the average on the right-hand side is just frequency of visits to A_R up to time M.

We now apply a general principle called the *transference principle*, which relates the integral $\int g\,d\mu$ of a function $g : X \to \mathbb{R}$ its discrete averages $S_I\widehat{g}$ along orbits: using $\int g = \int T^n g$, we have

$$\begin{aligned} \int g\,d\mu &= \frac{1}{M}\sum_{m=0}^{M-1}\int T^m g\,d\mu \\ &= \int \left(\frac{1}{M}\sum_{m=0}^{M-1} T^m g\right) d\mu \\ &= \int S_{[0,M-1]}\widehat{g}_x\,d\mu(x) \end{aligned}$$

Applying this to f and using 4.2, we obtain

$$\begin{aligned} \int f\,d\mu &= S_{[0,M-1]}\widehat{f_x} \\ &> t \cdot \frac{M-R-1}{M} \cdot \int h\,d\mu \\ &= t \cdot (1 - \frac{R-1}{M}) \cdot \int 1_{A_R}\,d\mu \\ &= t \cdot (1 - \frac{R-1}{M}) \cdot \mu(A_R) \end{aligned}$$

Letting $M \to \infty$, this is

$$\int f \, d\mu > t \cdot \mu(A_R)$$

Finally, letting $R \to \infty$ and noting that $\mu(A_R) \to \mu(A)$, we conclude that $\int f \, d\mu > t \cdot \mu(A)$, which is what was claimed.

Example 4.4.6. Let $(\xi_n)_{n=1}^{\infty}$ be an independent identically distributed sequence of random variables represented by a product measure on $(X, \mathcal{B}, \mu) = (\Omega, \mathcal{F}, P)^{\mathbb{N}}$, with $\xi_n(\omega) = \xi(\omega_n)$ for some $\xi \in L^1(\Omega, \mathcal{F}, P)$. Let $\sigma : X \to X$ be the shift, which preserves μ and is ergodic, and $\xi_n = \xi_0(\sigma^n)$. Since the shift acts ergodically on product measures, the ergodic theorem implies

$$\frac{1}{N} \sum_{n=0}^{N-1} \xi_n = \frac{1}{N} \sum_{n=0}^{N-1} \sigma^n \xi_0 \to \mathbb{E}(\xi_0 | \mathcal{I}) = \mathbb{E}\xi_0 \qquad \text{a.e.}$$

Thus the ergodic theorem generalizes the law of large numbers. However it is a very broad generalization: it holds for any stationary process $(\xi_n)_{n=1}^{\infty}$ without any independence assumption, as long as the process is ergodic.

When T is invertible it is also natural to consider the two-sided averages $\overline{S}_N = \frac{1}{2N+1} \sum_{n=-N}^{N} T^n f$. Up to an extra term $\frac{1}{2N+1} f$, this is just $\frac{1}{2} S_N(T, f) + \frac{1}{2} S_N T^{-1}, f)$, where we write $S_N(T, f)$ to emphasize which map is being used. Since both of these converge in L^1 and a.e. to the same function $\mathbb{E}(f|\mathcal{I})$, the same is true for $\overline{S}_N f$.

4.5 (*) Sub-additive ergodic theorem

Theorem 4.5.1 (Subadditive ergodic theorem)**.** *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic measure-preserving system. Suppose that $f_n \in L^1(\mu)$ satisfy the subadditivity relation*

$$f_{m+n}(x) \le f_m(x) + f_n(T^m x)$$

and are uniformly bounded above, i.e. $f_n \le L$ for some L. Then $\lim_{n\to\infty} \frac{1}{n} f_n(x)$ *exists a.e. and is equal to the constant* $\lim_{n\to\infty} \frac{1}{n} \int f_n$.

Before giving the proof we point out two examples. First, if $f_n = \sum_{k=0}^{n-1} T^k g$ then f_n satisfies the hypothesis, so this is a generalization of the usual ergodic theorem (for ergodic T).

For a more interesting example, let $A_n = A(T^n x)$ be a stationary sequence of $d \times d$ matrices (for example, if the entries are i.i.d.). Let $f_n = \log \|A_1 \cdot \ldots \cdot A_n\|$ satisfies the hypothesis. Thus, the subadditive ergodic theorem implies that random matrix products have a Lyapunov exponent – their norm growth is asymptotically exponential.

Proof. Let us first make a simple observation. Suppose that $\{1, \ldots, N\}$ is partitioned into intervals $\{[a_i, b_i)\}_{i \in I}$. Then subadditivity implies

$$f_N(x) \le \sum_{i \in I} f_{b_i - a_i}(T^{a_i} x)$$

Let

$$a = \liminf \frac{1}{n} f_n$$

We claim that a is invariant. Indeed,

$$\frac{1}{n} f_n(Tx) \geq \frac{1}{n}(f_{n+1}(x) - f_1(x))$$

From this it follows that $a(Tx) \geq a(x)$ so by ergodicity a is constant.

Fix $\varepsilon > 0$. Since $\liminf \frac{1}{n} f_n = a$ there is an N such that the set

$$A = \{x \,:\, \frac{1}{n} f_n(x) < a + \varepsilon \text{ for some } 0 \leq n \leq N\}$$

satisfies $\mu(A) > 1 - \varepsilon$.

Now fix a typical point x. By the ergodic theorem, for every large enough M,

$$\frac{1}{M} \sum_{n=0}^{M-1} 1_A(T^n x) > 1 - \varepsilon$$

Fix such an M and let

$$I_0 = \{0 \leq n \leq M - N \,:\, T^n x \in A\}$$

For $i \in I_0$ there is a $0 \leq k_i \leq N$ such that $\frac{1}{k} f_{k_i}(T^i x) < a + \varepsilon$. Let $U_i = [i, n+k_n)$. Applying the covering lemma, Lemma 4.4.5, there is a subset $I_1 \subseteq I_0$ such that $\{U_i\}_{i \in I_1}$ are pairwise disjoint and $|\bigcup_{i \in I_1} U_i| \geq |I_0| > (1-\varepsilon)M$. By construction also $\bigcup_{i \in I_1} U_i \subseteq [0, M)$.

Choose an enumeration $\{U_i\}_{i \in I_2}$ of the complementary intervals in $[0, M) \setminus \bigcup_{i \in I_i} U_i$, so that $\{U_i\}_{i \in I_1 \cup I_2}$ is a partition of $[0, M)$. Writing $U_i = [a_i, b_i)$ and using the comment above, we find that (recall that L is a pointwise upper bound on f_n):

$$\begin{aligned}
\frac{1}{M} f_M(x) &\leq \frac{1}{M}\left(\sum_{i \in I_1} f_{b_i - a_i}(T^{a_i} x) + \sum_{i \in I_2} f_{b_i - a_i}(T^{a_i} x)\right) \\
&\leq \frac{\sum_{n \in I_1} |U_i|}{M}(a + \varepsilon) + \frac{\sum_{n \in I_2} |U_i|}{M} L \\
&\leq (a + \varepsilon) + \varepsilon L
\end{aligned}$$

Since this holds for all large enough M, for all ε, we conclude that $\limsup \frac{1}{M} f_M \leq a = \liminf \frac{1}{n} f_n$ so the limit exists and is equal to a.

It remains to identify $a = \lim \frac{1}{n} \int f_n$. First note that

$$\int f_{m+n} \leq \int f_m d\mu + \int f_n \circ T^m d\mu = \int f_m d\mu + \int f_n d\mu$$

so $a_n = \int f_n d\mu$ is subadditive, hence the limit $a' = \lim \frac{1}{n} a_n$ exists. By Fatou's lemma (since $f_n \leq L$ we can apply it to $-f_n$) we get

$$a = \int \limsup \frac{1}{n} f_n d\mu \geq \limsup \int \frac{1}{n} f_n d\mu = \lim \frac{1}{n} a_n = a'$$

Suppose the inequality were strict, $a' < a - \varepsilon$ for some $\varepsilon > 0$ and let n be such that $a_n < a - \varepsilon$. Note that for every $0 \le p \le n-1$ we have the identity

$$f_N(x) \le f_p(x) + \sum_{k=0}^{[N/n]-1} f_n(T^{kn+p}x) + f_{N-p-n([N/n]-1))}(T^{p+n([N/n]-1))}x)$$

Averaging this over $0 \le p < n$, we have

$$\frac{1}{N} f_N \le S_N(\frac{1}{n} f_n) + O(\frac{n}{N})$$

This by the ergodic theorem,

$$\lim_{N\to\infty} \frac{1}{N} f_N \le \lim_{N\to\infty} S_N(\frac{1}{n} f_n) = \int \frac{1}{n} f_n < a - \varepsilon$$

which is a contradiction to the definition of a. □

4.6 (*) Further generalizations

4.6.1 Group actions

Let G be a countable group. A measure preserving action of G on a measure space $(X, \mathcal{B}, \mu)$ is, first of all, an action, that is a map $G \times X \to X$, $(g, x) \mapsto gx$, such that $g(hx) = (gh)(x)$ for all $g, h \in G$ and $x \in X$. In addition, for each $g \in G$ the map $T_g : x \mapsto gx$ must be measurable and measure-preserving. It is convenient to denote the action by $\{T_g\}_{g\in G}$.

An invariant set for the action is a set $A \in \mathcal{B}$ such tat $T_g A = A$ for all $g \in G$. If every such set satisfies $\mu(A) = 0$ or $\mu(X \setminus A) = 0$, then the action is ergodic. There is an ergodic decomposition theorem for such actions, but for simplicity (and without loss of generality) we will assume that the action is ergodic.

For a function $f : X \to \mathbb{R}$ the function $T_g f = f \circ T_{g^{-1}} : X \to \mathbb{R}$ has the same regularity, and $\{T_g\}_{g\in G}$ gives an isometric action on L^p for all $1 \le p \le \infty$. Given a finite set $E \subseteq G$ let $S_E f$ be the functions defined by

$$S_E f(x) = \sum_{g\in E} f(T_g x)$$

As before, this is a contraction in L^p. We say that a sequence $E_n \subseteq G$ of finite sets *satisfies the ergodic theorem along* $\{E_n\}$ if $S_{E_n} f \to \int f$, in a suitable sense (e.g. in L^2 or a.e.) for every ergodic action and every suitable f.

Definition 4.6.1. A group G is *amenable* if there is a sequence of sets $E_n \subseteq G$ such that for every $g \in G$,

$$\frac{|E_n g \Delta E_n|}{|E_n|} \to 0$$

Such a sequence $\{E_n\}$ is called a *Følner sequence*.

For example, $\mathbb{Z}^d$ is a amenable because $E_n = [-n,n]^d \cap \mathbb{Z}^d$ satisfies

$$|(E_n + u) \cap E_n| = |E_{n-\|u\|_\infty}| = |E_n| + o(1)$$

The class of amenable groups is closed under taking subgroups and countable increasing unions, and if G and $N \lhd G$ are amenable so is G/N. Groups of sub-exponential growth are amenable; the free group is not amenable, but there are amenable groups of exponential growth.

Theorem 4.6.2. *If $\{E_n\}$ is a Følner sequence in an amenable group G then the ergodic theorem holds along $\{E_n\}$ in the L^2 sense (the mean ergodic theorem).*

Proof. Let

$$V_0 = \text{span}\{f - T_g f \,:\, f \in L^2\,,\, g \in G\}$$

One can show exactly as before that $V_0^\perp$ consists of the invariant functions (in this case, the constant functions, because we are assuming the action is ergodic). Then one must only show that $S_{E_n}(f - T_g f) \to 0$ for $f \in L^2$. But this is immediate from the Følner property, since

$$S_{E_n} f - S_{E_n} T_g f = S_{E_n \setminus E_n r^{-1}} f$$

and therefore

$$\left\| \frac{1}{|E_n|} S_{E_n}(f - T_g f) \right\|_2 \le \frac{1}{|E_n|} |E_n \setminus E_n g^{-1}| \cdot \|f\|_2 \le \frac{|E_n \Delta E_n g^{-1}|}{|E_n|} \|f\|_2 \to 0$$

This proves the mean ergodic theorem. □

The proof of the pointwise ergodic theorem for amenable groups is more delicate and does not hold for every Følner sequence. However, one can reduce it as before to a maximal inequality. What one then needs is an analog of the discrete maximal inequality, which now concerns functions $\widehat{f} : G \to [0,\infty)$, and requires an analog of the covering Lemma 4.4.5. Such a result is known under a stronger assumption on $\{E_n\}$, namely assuming that $|\bigcup_{k<n} E_k^{-1} E_n| \le C|E_n|$ for some constant C and all n. Every Følner sequence has a subsequence that satisfies this, and so every amenable group has a sequence along which the pointwise ergodic theorem holds a.e. and in L^1.

Outside of amenable groups one can also find ergodic theorems. The simplest to state is for the free group $\mathbb{F}_s$ on s generates $g_1^{\pm 1}, \ldots, g_s^{\pm 1}$. This is a non-amenable group which can be identified with the set of words in the generators that don't contain any occurrence of uu^{-1}. The group operation is concatenation follows by reduction, that is, repeatedly deleting any pair ss^1. For example the product of words $aba^{-1}c$ and $c^{-1}abb$ is

$$aba^{-1}cc^{-1}abb = aba^{-1}abb = abbb$$

the right hand side is reduced.

Let $E_n \subseteq \mathbb{F}_s$ denote the set of reduced words of length $\le n$.

Theorem 4.6.3 (Nevo-Stein, Bufetov). *If $\mathbb{F}_s$ acts ergodically by measure preserving transformations on $(X, \mathcal{B}, \mu)$ then for every $S_{E_n} f \to \int f$ for every $f \in L^1(\mu)$.*

There is a major difference between the proof of this result and in the amenable case. Because $|E_n \Delta E_n g^{-1}|/|E_n| \not\to 0$, the there is no trivial reason for the averages of co-boundaries to tend to 0. Consequently there is no natural dense set of functions in L^1 for which convergence holds. In any case, the maximal inequality is not valid either. The proof in non-amenable cases takes completely different approaches (but we will not discuss them here).

4.6.2 Hopf's ergodic theorem

Another generalization is to the case of a measure-preserving transformation T of a measure space $(X, \mathcal{B}, \mu)$ with $\mu(X) = \infty$ (but σ-finite). Ergodicity is defined as before – all invariant sets are of measure 0 or their complement is of measure 0. It is also still true that $T : L^2(\mu) \to L^2(\mu)$ is norm-preserving, and so the mean ergodic theorem holds: $S_N f \to \pi f$ for $f \in L^2$, where π is the projection to the subspace of invariant L^2 functions. Now, however, the only constant function that is integrable is 0, and we find that $S_N f \to 0$ in L^2. In fact this is true in L^1 and a.e. The meaning is, however, the same: if we take a set of finite measure A, this says that the fraction of time an orbit spends in A is the same as the relative size of A compared to Ω; in this case $\mu(A)/\mu(\Omega) = 0$.

Instead of asking about the absolute time spent in A, it is better to consider two sets A, B of positive finite measure. Then an orbit visits both with frequency 0, but one may expect that the frequency of visits to A is $\mu(A)/\mu(B)$-times the frequency of visits to B. This is actually he case:

Theorem 4.6.4 (Hopf). *If T is an ergodic measure-preserving transformation of $(X, \mathcal{B}, \mu)$ with $\mu(X) = \infty$, and if $f, g \in L^1(\mu)$ and $\int g d\mu \neq 0$, then*

$$\frac{\sum_{n=0}^{N-1} T^n f}{\sum_{n=0}^{N-1} T^n g} \xrightarrow[N \to \infty]{} \frac{\int f d\mu}{\int g d\mu} \qquad a.e.$$

Since the right hand side is usually not 0, one cannot expect this to hold in L^1.

Hopf's theorem can also be generalized to group actions, but the situation there is more subtle, and it is known that not all amenable groups have sequences E_n such that $\sum_{E_n} T^g f / \sum_{E_n} T^g h \to \int f / \int h$.

Chapter 5

The ergodic decomposition theorem

5.1 Ergodic decomposition: overview

Having described those systems that are "indecomposable", we now turn to study how a non-ergodic system may decompose into ergodic ones.

Example 5.1.1. If $T : X \to X$ is a bijection of a finite set X, and μ an invariant measure. We claim that μ is a convex combination of invariant measures.

We give two proos of this. The first approach is to decompose the space into "minimal" oinvariant sets, and condition μ on each one. To begin, note that μ is just a function $\mu : X \to [0, 1]$ such that $\sum_{x\in X} \mu(x) = 1$. By measure preservation, $\mu(Tx) = \mu(x)$, so μ is constant along orbits. Now, X is partitiones into disjoint T-orbits $O_1, \ldots, O_k$, and by the above, μ is constnat on each orbit: $\mu|_{O_i} \equiv p_i$ for some $p_i \geq 0$. If we write μ_i for the normalized counting measure on O_i, then

$$p_i \cdot |O_i| \cdot \mu_i = \mu|_{O_i}$$

Thus

$$\mu = \sum \mu|_{O_i} = \sum p_i|O_i| \cdot \mu_i$$

The weights $p_i|O_i|$ sum to one (e.g. evaluate the equation above on X), and each μ_i is ergodic, because a cyclic permutation has no invariant sets. Thus μ is a convex combination of ergodic measures.

Our second proof uses convex geometry. $\mathcal{P}_T(X)$ is a finite-dimensional compact convex set, being just the set of $\mu \in \mathbb{R}^X$ such that $\sum_{x\in X} \mu(x) = 0$ and $\mu(Tx) = \mu(x) \geq 0$ for all $x \in X$. We have seen already that the ergodic measures are its extreme points. It is a theorem that in a finite dimensional vector space, a compact convex set is the convex hull of its extreme points. Thus, every invariant measure is a convex combination of ergodic measures.

In general, it is too much to ask that every invariant measure be a convex combination of ergodic ones:

Example 5.1.2. Let $X = [0,1]$ with Borel sets and Lebesgue measure μ, and T the identity map. Then the only ergodic measures are point masses δ_x, and we cannot write μ as a finite convex combination of ergodic measures.

It is possible to generalize both approaches from the first example to the context of general measure preserving systems. The second approach requires one to develop the Choquet theory of representations of points in infinite-dimensional convex sets. While elegant, this would take us a bit out of the way. Instead we choose the first approach, which involves conditioning the measure on the σ-algebra of invariant sets. In the next few sections we develop the machinery for doing this, and then begin the proof itself.

5.2 Measure integration

A convex combination of probability measures is again a probability measure, Below we introduce temrinology for dealing with integrals of measures, rather than finite sums of them.

Given a measurable space $(X, \mathcal{B})$, a family $\{\nu_x\}_{x \in X}$ of probability measures on $(Y, \mathcal{C})$ is measurable if for every $E \in \mathcal{C}$ the map $x \mapsto \nu_x(E)$ is measurable (with respect to $\mathcal{B}$). Equivalently, for every bounded measurable function $f : Y \to \mathbb{R}$, the map $x \mapsto \int f(y)\, d\nu_x(y)$ is measurable.

Given a measure $\mu \in \mathcal{P}(X)$ we can define the probability measure $\nu = \int \nu_x d\mu(x)$ on Y by

$$\nu(E) = \int \nu_x(E)\, d\mu(x)$$

For bounded measurable $f : Y \to \mathbb{R}$ this gives

$$\int f\, d\nu = \int (\int f\, d\nu_x)\, d\mu(x)$$

and the same holds for $f \in L^1(\nu)$ by approximation (although f is defined only on a set E of full ν-measure, we have $\nu_x(E) = 1$ for μ-a.e. x, so the inner integral is well defined μ-a.e.).

Example 5.2.1. Let X be finite and $\mathcal{B} = 2^X$. Then

$$\int \nu_x\, d\mu(x) = \sum_{x \in X} \mu(x) \cdot \nu_x$$

Any convex combination of measures on Y can be represented this way, so the definition above generalizes convex combinations.

Example 5.2.2. Any measure μ on $(X, \mathcal{B})$ the family $\{\delta_x\}_{x \in X}$ is measurable since $\delta_x(E) = 1_E(x)$, and $\mu = \int \delta_x\, d\mu(x)$ because

$$\mu(X) = \int 1_E(x) d\mu(x) = \int \nu_x(E)\, d\mu(x)$$

In this case the parameter space was the same as the target space.

In particular, this representation shows that Lebesgue measure on $[0, 1]$ is an integral of ergodic measures for the identity map.

Example 5.2.3. $X = [0, 1]$ and $Y = [0, 1]^2$. For $x \in [0, 1]$ let ν_x be Lebesgue measure on the fiber $\{x\} \times [0, 1]$. Measurability is verified using the definition of the product σ-algebra, and by Fubini's theorem

$$\nu(E) = \int \nu_x(E) d\mu(x) = \int_0^1 \int_0^1 1_E(x, y) dy\, dx = \int \int_E 1 dx dy$$

so ν is just Lebesgue measure on $[0, 1]^2$.

One could also represent ν as $\int \nu_{x,y}\, d\nu(x, y)$ where $\nu_{x,y} = \nu_x$. Written this way each fiber measure appears many times.

5.3 Measure disintegration

We now turn to the problem of defining conditional measures. This is easy to do if we want to condition on sets of positive measure, or when we condition on a finite partition.

Example 5.3.1. Let $(X, \mathcal{B}, \mu)$ be a probability space and let $\mathcal{P} = \{P_1, \ldots, P_n\}$ be finite partition of it, i.e. P_i are measurable, $P_i \cap P_j = \emptyset$ for $i \neq j$, and $X = \bigcup P_i$. For simplicity assume also that $\mu(P_i) > 0$. Write $\mu_i = \frac{1}{\mu(P_i)}\mu|_{P_i}$, which is a probability meausre on P_i. Then

$$\mu = \sum_i \mu(P_i) \cdot \mu_i$$

and we have written μ as a convex combination of probability measures, one on each atom of $\mathcal{P}$. We can write this as an integral as follows: Let $i = i(x)$ denote the unique index such that $x \in P_i$ and let $\mu_x = \mu_{i(x)}$. Then $\mu\{x : i(x) = j\} = \mu(P_j)$, and since $x \mapsto \mu_x$ takes on the value μ_j on P_j we have

$$\mu = \int \mu_x\, d\mu(x)$$

Our goal is to give a similar decomposition of a measure with respect to an infinite (usually uncountable) partition $\mathcal{E}$ of X. Then the partition elements $E \in \mathcal{E}$ typically have measure 0, and the formula $\frac{1}{\mu(E)}\mu|_E$ no longer makes sense. As in probability theory one can define the conditional probability of an event E given that $x \in E$ as the conditional expectation $\mathbb{E}(1_E|\mathcal{P})$ evaluated at x (conditional expectation is reviewed in the Appendix). This would appear to give the desired decomposition: define $\mu_x(E) = \mathbb{E}(1_E|\mathcal{E})(x)$. For any countable algebra this does give a countably additive measure defined for μ-a.e. x. The problem is that for each set E, the measure $\mu_x(E)$ is defined only for a.e. x, but we want to define $\mu_x(E)$ for all measurable sets. Overcoming this problem is a technical but nontrivial chore which will occupy us for the rest of the section.

For a measurable space $(X, \mathcal{B})$ and a sub-σ-algebra $\mathcal{E} \subseteq \mathcal{B}$ generated by a countable sequence $\{E_n\}$. Write $x \sim_{\mathcal{E}} y$ if $1_E(x) = 1_E(y)$ for every $E \in \mathcal{E}$, or equivalently, $1_{E_n}(x) = 1_{E_n}(y)$ for all n. This is an equivalence relation. The atoms of $\mathcal{E}$ are by definition the equivalence classes of $\sim_{\mathcal{E}}$, which are measurable, being intersections of sequences F_n of the form $F_n \in \{E_n, X \setminus E_n\}$. We denote $\mathcal{E}(x)$ the atom containing x.

In the next theorem we assume that the space is compact, which makes the Riesz representation theorem available as a means for of defining measures. We shall discuss this restriction afterwards.

Theorem 5.3.2. *Let X be compact metric space, $\mathcal{B}$ the Borel algebra, and $\mathcal{E} \subseteq \mathcal{B}$ a countably generated sub-σ-algebra. Then there is an $\mathcal{E}-$measurable family $\{\mu_y\}_{y \in X} \subseteq \mathcal{P}(X)$ such that μ_y is supported on $\mathcal{E}(y)$ and*

$$\mu = \int \mu_y \, d\mu(y)$$

Furthermore if $\{\mu'_y\}_{y \in X}$ is another such system then $\mu_y = \mu'_y$ a.e.

Note that $\mathcal{E}$-measurability has the following consequence: For μ-a.e. y, for every $y' \in \mathcal{E}(y)$ we have $\mu_{y'} = \mu_y$ (and, since $\mu_y(\mathcal{E}(y)) = 1$, it follows that $\mu_{y'} = \mu_y$ for μ_y-a.e y').

Definition 5.3.3. The representation $\mu = \int \mu_y \, d\mu(y)$ in the proof is often called the **disintegration** of μ over $\mathcal{E}$.

We sketch the main steps of the proof. A detailed proof can be found in the appendix.

Step 1 Choose a countable dense $\mathbb{Q}$-linear subspace $V \leq C(X)$, and for $f \in V$ define

$$\overline{f} = \mathbb{E}(f|\mathcal{E})$$

Since V is countable, for μ-a.e. y the function $\widetilde{f}$ is defined at y for all $\widetilde{f} \in V$ and we can define $\Lambda_y : V \to \mathbb{R}$ by

$$\Lambda_y(f) = \overline{f}(y)$$

Furthermroe the linearity and positivity properties of conditional expectation hold for a.e. y, so for a.,e. y the function Λ_y is a bounded positive $\mathbb{Q}$-linear functional $(V, \|\cdot\|_\infty)$, and extends to a bounded linear functional $\Lambda_y : C(X) \to \mathbb{R}$. which, by the Riesz representation theorem, corresponds to a measure $\mu_y \in \mathcal{P}(X)$ with

$$\Lambda_y f = \int f(x) \, d\mu_y(x) \qquad \text{for } f \in C(X)$$

Step 2 One shows for $A \in \mathcal{B}$ that $\mu_y(A) = \mathbb{E}(1_A|\mathcal{E})(y)$ a.e., and in particular $y \to \mu_y$ is measurable. For this, one begins from the relation

$\int f d\mu_y = \mathbb{E}(f|\mathcal{E})(y)$ a.e., which we know for continuous f. For closed A one approximates 1_A as a pointwise limit of continuous functions to get the desired conclusion. Then one shows that the set of A with the desired property is a monotone class.

Step 3 Using monotone convergence and simple functions, one shows that $\int f d\mu_y = \mathbb{E}(f|\mathcal{E})(y)$ a.e., for all $f \in L^1(\mu)$.

Step 4 For $E \in \mathcal{E}$ we now know that $\mu_y(E) = \mathbb{E}(1_E|\mathcal{E})(y) = 1_E(y)$ a.s. Applying this to a sequence of sets $\{E_i\}$ generting $\mathcal{E}$, we deduce that a.s. μ_y is supported on an atom of $\mathcal{E}$, that is, every set in $\mathcal{E}$ has μ_y-mass 0 or 1 (this is where we use the countable generation assumption).

Step 5 We prove uniqueness: given another family $\{\mu'_y\}$ with the same properties, we define an operator $L^1(\mathcal{B},\mu) \to L^1(\mathcal{E},\mu)$ by $\widetilde{f}(y) = \int f d\mu_y$. This operator is clearly linear and one can show that it has theproperty that $\widetilde{gf} = g \cdot \widetilde{f}$ when $g \in L^\infty(\mathcal{E},\mu)$. This implies that $f \mapsto \widetilde{f}$ is just the conditional operator on $\mathcal{E}$, and this implies that $\int f d\mu'_y = \int f d\mu_y$ a.e. for every continuous function. From this we conclude $\mu'_y = \mu_y$ a.e.

It remains to address the compactness assumption on X. Examples show that one the disintegration theorem does require some assumption; it does not hold for arbitrary measure spaces and sub-σ-algebras. So we will not eliminate the compactness assumption, but it is not a large restriction.

Definition 5.3.4. A **Polish space** is a separable complete metric space.

Definition 5.3.5. A **standard Borel space** is a measurable space $(X,\mathcal{B})$ for which there exists a metric on X under which X is Polish and B is he σ-algebra of Borel sets.

In "real life", essentially all spaces are standard.

Theorem 5.3.6. *Any two uncountable standard Borel spaces are isomorphic, i.e. there is a measurable bijective map betwen them with measurable inverse, and which maps the σ-algebras to each other.*

In particular, if we are given a standard Borel space, then we may always assume that the underlying space is compact metric and the σ-algebra is the algebra of Borel sets. Thus, we can re-formulate the disintegration theorem as follows.

Theorem 5.3.7. *Let μ be a probability measure on a standard Borel space $(X,\mathcal{B},\mu)$ and $\mathcal{E} \subseteq \mathcal{B}$ a countably generated sub-σ-algebra. Then there is an $\mathcal{E}$-measurable family $\{\mu_y\}_{y\in Y} \subseteq \mathcal{P}(X,\mathcal{B})$ such that μ_y is supported on $\mathcal{E}(y)$ and*

$$\mu = \int \mu_y \, d\mu(y)$$

Furthermore if $\{\mu'_y\}_{y\in X}$ is another such system then $\mu_y = \mu'_y$.

5.4 The ergodic decomposition

Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system on a Borel space. Let $\mathcal{I} \subseteq \mathcal{B}$ denote the family of T-invariant measurable sets. It is easy to check that $\mathcal{I}$ is a σ-algebra.

The σ-algebra $\mathcal{I}$ in general is not countably generated. Consider the case of an ergodic system, and suppose that $\mathcal{I}$ were countably generated by sets $\{I_n\}_{n=1}^{\infty}$. By ergodicity, for each n either $\mu(I_n) = 1$ or $\mu(X \setminus I_n) = 1$; set $F_n = I_n$ or $F_n = X \setminus I_n$ according to these possibilities. Then $F = \bigcap F_n$ is an invariant set of measure 1 and is an atom of $\mathcal{I}$. Now note that if $x \in F$ then the orbit of x is invariant, and also measuable, since it is countable. Thus $O_T(x) \in \mathcal{I}$. But also $O_T(x) \subseteq F$, and F is an atom of $\mathcal{I}$, so $F = O_T(x)$. But now F is a countable set of full measure, contradicting non-atomicity of μ.

We shall work instead with a fixed countably generated μ-dense sub-σ-algebra $\mathcal{I}_0$ of $\mathcal{I}$. Let $L^1(X, \mathcal{I}, \mu)$ be a closed subspace of $L^1(X, \mathcal{B}, \mu)$. Since the latter is separable (by standardness of the space), so is the former. Choose a dense countable sequence $f_n \in L^1(X, \mathcal{I}, \mu)$, choosing representatives of the functions that are genuinely $\mathcal{I}$ measurable, not just modulo a $\mathcal{B}$-measurable nullset. Let $\mathcal{I}_0$ denote the smallest σ-algebra with respect the f_n are all measurable, so $\mathcal{I}_0$ is countably generated (e.g. it the algebra generated by the sets $A_{n,p,q} = \{p < f_n < q\}$ with $p, q \in \mathbb{Q}$. Clearly $\mathcal{I}_0 \subseteq \mathcal{I}$ and all of the f_n are $\mathcal{I}_0$-measurable, so $L^1(X, \mathcal{I}_0, \mu) = L^1(X, \mathcal{I}, \mu)$. In particular, $\mathcal{I}$ is contained in the μ-completion of $\mathcal{I}_0$.

The choice of $\mathcal{I}_0$ is highly non-cannonical. It is possible to make a more cannonical choice; see Section 6.

Theorem 5.4.1 (Ergodic decomposition theorem)**.** *Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system on a standard space, and let $\mathcal{I}, \mathcal{I}_0$ be as above. Then there is an $\mathcal{I}_0$-measurable (and in particular $\mathcal{I}$-measurable) disintegration $\mu = \int \mu_x \, d\mu(x)$ of μ such that a.e. μ_y is T-invariant, ergodic, and supported on $\mathcal{I}_0(y)$. Furthermore the representation is unique in the sense that if $\{\mu'_y\}$ is any other family with the same properties then $\mu_y = \mu'_y$ for μ-a.e. y.*

Let $\{\mu_y\}_{y \in X}$ be the disintegration of μ relative to $\mathcal{I}_0$, we need only show that for μ-a.e. y the measure μ_y is T-invariant and ergodic.

Claim 5.4.2. For μ-a.e. y, μ_y is T-invariant.

Proof. Define $\mu'_y = T\mu_y$. This is an $\mathcal{I}_0$ measurable family since for any $E \in \mathcal{B}$, $\mu'_y(E) = \mu_y(T^{-1}E)$ so measurability of $y \mapsto \mu'_y(E)$ follows from that of $y \mapsto \mu_y(T^{-1}E)$. We claim that $\{\mu'_y\}_{y \in X}$ is a disintegration of μ over $\mathcal{I}_0$. Indeed, for any $E \in \mathcal{B}$,

$$\begin{aligned} \int (\int \mu'_y(E)) \, d\mu(y) &= \int (\int \mu_y(T^{-1}E)) \, d\mu(y) \\ &= \mu(T^{-1}E) \\ &= \mu(E) \end{aligned}$$

Also $T^{-1}\mathcal{I}_0(y) = \mathcal{I}_0(y)$ (since $\mathcal{I}_0(y) \in \mathcal{I}$) so

$$\mu'_y(\mathcal{I}_0(y)) = \mu_y(T^{-1}\mathcal{I}_0(y)) = \mu_y(\mathcal{I}_0(y)) = 1$$

so μ'_y is supported on $\mathcal{I}_0(y)$. Thus, $\{\mu'_y\}_{y\in X}$ is an $\mathcal{I}_0$-measurable disintegration of μ, hence $\mu'_y = \mu_y$ a.e. This is exactly the same as a.e. T-invariance of μ_y. □

Claim 5.4.3. For μ-a.e. y, μ_y is ergodic.

Proof. Fix a countable dense family $\mathcal{F} \subseteq C(X)$. Then by the ergodic theorem applied to $(X, \mathcal{B}, \mu, T)$, for every $f \in \mathcal{F}$ we have

$$S_N f(y) \to \mathbb{E}_\mu(f|\mathcal{I})(y) = \mathbb{E}_\mu(f|\mathcal{I}_0)(y) \qquad \mu\text{-a.e.}$$

where the equality is because $\mathcal{I} = \mathcal{I}_0 \bmod \mu$. Since $\mathcal{F}$ is countable, for μ-a.e. y we have

$$S_N f(y) \to \mathbb{E}_\mu(f|\mathcal{I}_0)(y) \qquad \text{for all } f \in \mathcal{F}$$

Since $\mu = \int \mu_x d\mu(x)$, for μ-a.e. z, the limit above holds for μ_z-a.e. y; in addition for μ_z-a.e. y we know that μ_z is supported on $\mathcal{I}_0(y) = \mathcal{I}_0(z)$, and that $\mathbb{E}_\mu(f|\mathcal{I}_0)(z)$ is constant on $\mathcal{I}_0(z)$. It follows that for μ-a.e. z, for every $f \in \mathcal{F}$, the ergodic averages $S_N f$ converge μ_z-a.e. to a constant, and hence in L^2 (since f is bounded, and so are its averages). By Corollary 4.3.5, this implies that $(X, \mathcal{B}, \mu_z, T)$ is an ergodic system, as claimed. □

Our formulation of the ergodic decomposition theorem represents μ as an integral of ergodic measures parametrized by $y \in X$ (in an $\mathcal{I}$-measurable way). Sometimes the following formulation is given, in which $\mathcal{P}_T(X)$ is given the σ-algebra generated by the maps $\mu \mapsto \mu(E)$, $E \in \mathcal{B}$; this coincides with the Borel structure induced by the weak-* topology when X is given the structure of a compact metric space. One can show that the set of ergodic measures is measurable, for example because in the topological representation they are the extreme points of a weak-* compact convex set.

Theorem 5.4.4 (Ergodic decomposition, second version)**.** *Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system on a Borel space. Then there is a unique probability measure θ on $\mathcal{P}_T(X)$ supported on the ergodic measure and such that $\mu = \int \nu \, d\theta(\nu)$.*

Chapter 6

Topological dynamical systems

6.1 Topological dynamical systems

Ergodic theory has a close relative, in which instead of a measure being preserved, it is a topology.

Definition 6.1.1. A **topological dynamical system** (t.d.s) is a pair (X, T) where X is a compact metric space and $T : X \to X$ is continuous and onto.

Many of the examples we gave for measure preserving systems were also topological dynamical systems in the sense above. It is sometimes useful to allow compact or non-metrizable spaces but in this course we shall not encounter them.

Example 6.1.2.

1. $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ with a rotation $R_\alpha x = e^{2\pi i\alpha}z$.
2. $Tx = mx \bmod 1$ on $\mathbb{R}/\mathbb{Z}$ (for $m \in \mathbb{N}$).
3. $X = A^{\mathbb{N}}$ or $X = A^{\mathbb{Z}}$, where A is a compact metric space and X is given the product topology; and T is the shift map, $(Tx)_i = x_{i+1}$. Fo $A^{\mathbb{N}}$ this is a continuous map and for $A^{\mathbb{Z}}$ it is a homeeomorphism.

We shall later see that there is a close link between topological and measure-preserving systems: every topologicla system admits invariant measures, and every reasonable m.p.s. is isomorphic to an invariant measure on a t.d.s. This connection is useful because topological argument sometimes are useful in studying measure preserving systems. This usually is not strictly necessary, and the same results could be achieved using purely measure-theoretic means. However, introducing a topology – and with it, a natural separable class of test functions, namely, the continuous functions – often simplifies arguments, and we shall not hesitate to do this when it is is beneficial to us.

6.2 The weak-* topology on $\mathcal{P}(X)$

In this section we prove that every topological system admits an invariant probability measures. We first recall some basic tools for studying measures on compact spaces.

Let X be a compact metric space and let $\mathcal{M}(X)$ denote the linear space of signed (finite) Borel measures, and $\mathcal{P}(X) = \mathcal{P}(X, \mathcal{B}) \subseteq \mathcal{M}(X)$ the convex space of Borel probability measures. Two measures $\mu, \nu \in \mathcal{M}(X)$ are equal if and only if $\int f d\mu = \int f d\nu$ for all $f \in C(X)$, and the maps $\mu \mapsto \int f d\mu$, $f \in C(X)$, separate points.

Definition 6.2.1. Let X be a compact metric space. The weak-* topology on $\mathcal{M}(X)$ (or $\mathcal{P}(X)$) is the weakest topology that make the maps $\mu \mapsto \int f\, d\mu$ continuous for all $f \in C(X)$. In particular,

$$\mu_n \to \mu \text{ if and only if } \int f d\mu_n \to \int f d\mu \text{ for all } f \in C(X)$$

Lemma 6.2.2. *Let X be a compact metric space and $\mathcal{F} \subseteq C(X)$ a dense set of funtions. Suppose that $\mu_n \in \mathcal{P}(X)$ and $\lim \int f d\mu_n$ exists for all $f \in \mathcal{F}$. Then there exists $\mu \in \mathcal{P}(X)$ such that $\mu_n \to \mu$, that is, $\int f d\mu = \lim \int f d\mu_n$ for all $f \in C(X)$.*

Proof. Let $V = \operatorname{span} \mathcal{F}$. By assumption $\lim\int f d\mu_n$ exists for all $f \in \mathcal{F}$, and hence for all $f \in V$, since integrals and limits are finitely additive. For $f \in V$ denote the limit by $\Lambda(f)$. This is positive, linear, bounded function on V and so extends to such a function on $C(X) = \overline{V}$, which we denote also by Λ. By the Riesz representation theorem there exists $\mu \in \mathcal{P}(X)$ such that $\Lambda(f) = \int f d\mu$ for all $f \in C(X)$. We now claim $\int f d\mu_n \to \int f d\mu$ for all $f \in C(X)$. We already know this for , and V is dense. Fixing any $f \in C(X)$ let $\varepsilon > 0$ and $g \in (V)$ with $\|f - g\|_\infty < \varepsilon$. We get

$$\begin{aligned} |\int f d\mu_n - \int f d\mu| &< |\int f d\mu_n - \int g d\mu_n| + |\int g d\mu_n - \int g d\mu| + |\int g d\mu - \int f d\mu| \\ &< \varepsilon + |\int g d\mu_n - \int g d\mu| + \varepsilon \\ &\to 2\varepsilon \qquad \text{as } n \to \infty \end{aligned}$$

Since ε was arbitrary, we get $\int f d\mu_n \to \int f d\mu$. □

Proposition 6.2.3. *The weak-* topology is metrizable and compact.*

Proof sketch. Using the Stone-Weierstrass theorem shoose a dense sequence $\{f_i\}_{i=1}^\infty \subseteq C(X)$. Define a metric on $\mathcal{P}(X)$ by

$$d(\mu, \nu) = \sum_{i=1}^\infty 2^{-i} \left| \int f_i d\mu - \int f_i d\nu \right|$$

One shows that this metric is compatible with the topology. Next, if $\mu_n \in \mathcal{P}(X)$ is a sequence of measures, a diagonal argument can be used to show that there is a subsequence $\mu_{n(k)}$ such that for every i, the limit $\lim \int f_i d\mu_{n(k)}$ exists. The previous lemma now shows that $\mu_{n(k)} \to \mu$ for some measure μ. This proves sequential compactness, which, by metrizability, is compactness. $\square$

Let (X, T) be a topological dynamical system. Recall that the induced map $T = \widehat{T} : C(X) \to C(X)$ is an isometry of $(C(X), \|\cdot\|_\infty)$ (Corollary 2.3.3). We also get an induced map $T = \widehat{\widehat{T}} : \mathcal{P}(X) \to \mathcal{P}(X)$ by $\mu \mapsto \mu \circ T^{-1}$.

Lemma 6.2.4. *Let (X, T) be a topological dynamical system. Then the induced map $T : \mathcal{P}(X) \to \mathcal{P}(X)$ is continuous.*

Proof. If $\mu_n \to \mu$ then for $f \in C(X)$,

$$\int f \, dT\mu_n = \int f \circ T \, d\mu_n \to \int f \circ T \, d\mu = \int f \, dT\mu$$

This shows that $T\mu_n \to T\mu$, so T is continuous. $\square$

6.3 Existence of invariant measures

In a dynamical system (X, T), for $x \in X$ we write

$$\mu_{x,T} = \frac{1}{N} \sum_{n=0}^{N-1} \delta_{x,N}$$

This is a probability measure and we note that

$$S_N f(x) = \int f d\mu_{x,N}$$

for all $f \in C(X)$.

Proposition 6.3.1. *Every topological dynamical system (X, T) admits invariant measures.*

Proof. Let $x \in X$ be an arbitrary initial point and let $\mu_N = \mu_{x,N}$ be as above. Passing to a subsequence $N(k) \to \infty$ we can assume by compactness that $\mu_{N(k)} \to \mu \in \mathcal{P}(X)$. We must show that $\int f \, d\mu = \int f \circ T \, d\mu$ for all $f \in C(X)$. Now,

$$\begin{aligned}
\int f \, d\mu - \int f \circ T \, d\mu &= \lim_{k\to\infty} \int (f - f \circ T) \, d\mu_{N(l)} \\
&= \lim_{k\to\infty} \frac{1}{N(k)} \sum_{n=0}^{N(k)-1} \int (f - f \circ T)(T^n x) \\
&= \lim_{k\to\infty} \frac{1}{N(k)} \left(f(T^{N(k)-1}x) - f(x) \right) \\
&= 0
\end{aligned}$$

because f is bounded. □

There are a number of common variations of this proof. We could have defined $\mu_N = \frac{1}{N}\sum_{n=0}^{N-1}\delta_{T^n x_N}$ (with the initial point x_N varying with N), of begun with an arbitrary measure μ and $\mu_N = \frac{1}{N}\sum_{n=0}^{N-1} T^n\mu$. The proof would then show that any accumulation point of μ_N is T-invariant.

We denote the space of T-invariant measures by $\mathcal{P}_T(X)$.

Corollary 6.3.2. *If (X,T) is a topological dynamical system then $\mathcal{P}_T(X)$ is non-empty, compact and convex.*

Proof. The last proposition shows that it is non-empty, and convexity is trivial. For compactness, since it is a subset of the compact set $\mathcal{P}(X)$, we need only show it is closed. We know that

$$\mathcal{P}_T(X) = \bigcap_{f\in C(X)} \{\mu \in \mathcal{P}(X) \,:\, \int (f - Tf)d\mu = 0\}$$

Each of the sets in the intersection is the pre-image of 0 under the map $\mu \mapsto \int(f - Tf)d\mu$; since $f - Tf$ is continuous this map is continuous and so $\mathcal{P}_T(X)$ is the intersection of closed sets, hence closed. □

Corollary 6.3.3. *Every topological dynamical system (X,T) contains recurrent points.*

Proof. Choose any invariant measure $\mu \in \mathcal{P}_T(X)$ and apply Proposition 2.2.5 to the measure preserving system $(X,\mathcal{B},\mu,T)$. □

6.4 Generic points

The ergodic theorem is an a.e. statement relative to a given L^1 function, and such functions are defined a.e.; thus, it does not say anything about the distribution of single orbits. In a topological system, the continuous functions provide a natural class of test functions by which to examine an orbit.

Definition 6.4.1. Let (X,T) be a topological dynamical system. A point $x \in X$ is **generic** for a Borel measure $\mu \in \mathcal{P}(X)$ if it satisfies the conclusion of the ergodic theorem for every continuous function, i.e.

$$S_N f(x) = \frac{1}{N}\sum_{n=0}^{N-1} T^n f(x) \to \int f\,d\mu \qquad \text{for all } f \in C(X) \tag{6.1}$$

Thus, x is generic for μ if and only if $\mu_{x,N} \to \mu$ in the weak-* topology. We have already seen that if x is generic for μ then μ is T-invariant. Also, $\mu_{Tx,N}, \mu_{x,N}$ are clearly asymptotic, since $\|\mu_{Tx,N} - \mu_{x,N}\| \le 2/N$ in the total variation distance. Thus x is generic for μ if and only if Tx is.

Lemma 6.4.2. *Let $\mathcal{F} \subseteq C(X)$ be a countable dense set. If $S_N f(x)$ converges for every $f \in \mathcal{F}$ then x is generic for a measure $\mu \in \mathcal{P}_T(X)$ satisfying 6.1 for every $f \in C(X)$.*

Proof. Immediate from Lemma 6.2.2 applied to $(\mu_{x,N})_{N=1}^{\infty}$. □

In general, a generic point need not generate an ergodic measure, nor even be generic for any measure:

Example 6.4.3. Let $X = \{0,1\}^{\mathbb{N}}$ and let $\mu_0 = \delta_{000\ldots}$ and $\mu_1 = \delta_{111\ldots}$. These are ergodic measures for the shift σ. Now let $x \in X$ be the point such that $x_n = 0$ for $k^2 \leq n < (k+1)^2$ if k is even, and $x_n = 1$ for $k^2 \leq n < (k+1)^2$ if k is odd. Thus

$$x = 1110000011111110000000000111\ldots$$

We claim that x is generic for the non-ergodic measure $\mu = \frac{1}{2}\mu_0 + \frac{1}{2}\mu_1$. It suffices to prove that for any ℓ,

$$\begin{aligned} \frac{1}{N}\sum_{n=0}^{N-1} 1_{0^\ell}(T^n x) &\to \frac{1}{2} \\ \frac{1}{N}\sum_{n=0}^{N-1} 1_{1^\ell}(T^n x) &\to \frac{1}{2} \end{aligned}$$

where $0^\ell, 1^\ell$ are the sets of points beginning with ℓ consecutive 0s and ℓ consecutive 1s, respectively. The proofs are similar so we show this for 0^ℓ. Notice that $1_{0^\ell}(T^n x) = 1$ if $k^2 \leq n < (k+1)^2 - \ell$ and k is even, and $1_{0^\ell}(T^n x) = 0$ otherwise. Now, each N satisfies $k^2 \leq N < (k+2)^2$ for some even k. Then

$$\sum_{n=0}^{N-1} 1_{0^\ell}(T^n x) = \sum_{j=1}^{k/2}((2j+1)^2 - \ell) - (2j)^2) = \sum_{j=1}^{k/2}(4j + 1 - \ell) = (\frac{1}{2}k^2 + O(k))$$

Also $N - k^2 \leq (k+1)^2 - k^2 = O(k)$. Therefore $S_N 1_{0^\ell}(x) \to \frac{1}{2}$ as claimed.

Example 6.4.4. With (X, σ) as in the previous example, let $y_n = 0$ if $2^k \leq n < 2^{k+1}$ for k even and $y_n = 1$ otherwise. Then one can show that x is not generic for any measure, ergodic or not.

Theorem 6.4.5. *If μ is an ergodic measure in a topological dynamical system (X,T), then μ-a.e. x is generic for μ. More generally, if μ it T-invariant with ergodic decomposition $\mu = \int \mu_x d\mu(x)$, then μ-a.e. x is generic for measure μ_x.*

Proof. Suppose that μ is ergodic. Fix $\mathcal{F} \subseteq C(X)$ countable and dense. By the ergodic theorem, and using ergodicity, there exists a set $X_0 \subseteq X$ with $\mu(X_0) = 1$ and

$$S_N f(x) \to \int f d\mu \qquad \text{for } x \in X_0$$

Since $\mathcal{F}$ is countable, for every $x \in X_0$ this holds for all $f \in \mathcal{F}$ simultaneously. By Lemma 6.4.2 applied to the measures $\mu_{x,N}$, we conclude that $\mu_{x,N} \to \mu$ for $x \in X_0$, i.e., for μ-a.e. x.

In the non-ergodic case let $\mu = \int \mu_x d\mu(x)$ denote the ergodic decomposition. Then by the previous discussion, μ_x-a.e. y is generic for μ_x. Also, μ_x-a.e. y satisfies $\mu_y = \mu_x$; thus, μ_x-a.e. y is generic for μ_y. Letting E denote the set of y that are generic for μ_y, we have shown that $\mu_x(E) = 1$ for μ-a.e. x, hence $\mu(E) = \int \mu_x(E) d\mu(x) = 1$, as claimed. □

Definition 6.4.6. For a topological dynamical system (X, T), write G_T for the set of generic points, $G_T(\mu)$ for the set of points that are generic for a given $\mu \in \mathcal{P}_T(X)$, and μ_x for the measure for which a point $x \in G_T$ is generic (define μ_x for $x \notin G_T$ to be some fixed value, or leave it undefined).

Lemma 6.4.7. *The set G_T and the map $x \mapsto \mu_x$, defined on G_T, are Borel measurable.*

Proof. Fix $\mathcal{F} \subseteq C(X)$ countable. Then G_T is the set of points x such that $S_N f(x)$ converges for all $f \in \mathcal{F}$, which is measurable. Next, let $\pi : \mathcal{P}(X) \to \mathbb{R}^{\mathcal{F}}$ denote the map $\nu \mapsto (\int f d\nu)_{f \in \mathcal{F}} \in \mathbb{R}^{\mathcal{F}}$. This is a continuous embedding of $\mathcal{P}(X)$ into $\mathbb{R}^{\mathcal{F}}$ and hence has continuous, and therefore measurable, inverse from its image to $\mathcal{P}(X)$. For $x \in G_T$ the map $x \mapsto (\lim S_N f(x))_{f \in \mathcal{F}} \in \mathbb{R}^{\mathcal{F}}$ is measurable and sends x to $\pi(\mu_x)$. Thus, $x \mapsto \mu_x$ is measurable. □

Let $\mathcal{G} \subseteq \mathcal{I}$ denote the smallest σ-algebra with respect to which the set G_T and function $x \mapsto \mu_x$ are measurable; it consists of invarint sets becuase obviously $\mu_{Tx} = \mu_x$. Since $\mathcal{G}$ is the pull-back of the Borel σ-algebra of the compact metric space $\mathcal{P}_T(X)$, and the latter is countably generated, so is the former: $\mathcal{G} = \sigma(\mathcal{G}_0)$ where some countable family of sets $\mathcal{G}_0 \subseteq \mathcal{G}$.

Corollary 6.4.8. *Let $\mu \in \mathcal{P}_T(X)$. With the notation above, $\mathcal{G}$ coincides, mod μ, with the σ-algebra $\mathcal{I}$ of invariant sets.*

6.5 Unique ergodicity

When can the ergodic theorem be strengthened from a.e. point to every point? Once again the question does not make sense for L^1 functions, since these are only defined a.e., but it makes sense for continuous functions.

Definition 6.5.1. A topological system (X, T) is uniquely ergodic if there is only one invariant probability measure, which in this case is denoted μ_X.

Proposition 6.5.2. *Let (X, T) be a topological system and $\mu \in \mathcal{P}_T(X)$. The following are equivalent.*

1. *Every point is generic for μ.*
2. *$S_N f \to \int f \, d\mu$ uniformly, for every $f \in C(X)$.*

3. (X, T) is uniquely ergodic and μ is its invariant measure.

Proof. (1) implies (3): If $\nu \neq \mu$ were another invariant measure there would be points that are generic for it, contrary to (1).

(3) implies (2): Suppose (2) fails, so there is an $f \in C(X)$ such that $\left\|S_N f \not\to \int f d\mu\right\|_\infty \to 0$. Then there is some sequence $x_k \in X$ and integers $N_k \to \infty$ such that $S_{N_k} f(x_k) \to c \neq \int f d\mu$. Let ν be an accumulation point of $\frac{1}{N_k}\sum_{n=1}^{N_k} \delta_{T^n x_k}$. This is a T-invariant measure and $\int f d\nu = c$ so $\nu \neq \mu$, contradicting (3).

(2) implies (1) is immediate. □

Proposition 6.5.3. *Let $X = \mathbb{R}/\mathbb{Z}$ and $\alpha \notin \mathbb{Q}$. The map $T_\alpha x = x + \alpha$ on X is uniquely ergodic with invariant measure $\mu =$ Lebesgue.*

We give two proofs.

Proof number 1. We know that μ is ergodic for T_α so a.e. x is generic. Fix one such x. Let $y \in X$ be any other point. then there is a $\beta \in \mathbb{R}$ such that $y = T_\beta x$. For any function $f \in C(X)$,

$$\begin{aligned}
\frac{1}{N}\sum_{n=0}^{N-1} T_\alpha^n f(y) &= \frac{1}{N}\sum_{n=0}^{N-1} f(y + \alpha n) \\
&= \frac{1}{N}\sum_{n=0}^{N-1} f(x + \alpha_n + \beta) \\
&= \frac{1}{N}\sum_{n=0}^{N-1} (T_\beta f)(T_\alpha^n x) \\
&\to \int T_\beta f \, d\mu = \int f \, d\mu
\end{aligned}$$

Therefore every point is generic for μ and T_α is uniquely ergodic. □

Our second proof is based on a more direct calculation that does not rely on the ergodic theorem.

Definition 6.5.4. A sequence (x_k) in a compact metric space X equidistributes for a measure μ if $\frac{1}{N}\sum_{n=1}^{N} \delta_{x_n} \to \mu$ weak-*.

Lemma 6.5.5 (Weyl's equidistribution criterion)**.** *A sequence $(x_k) \subseteq \mathbb{R}/\mathbb{Z}$ equidistributes for Lebesgue measure μ if and only if for every m,*

$$\frac{1}{N}\sum_{n=0}^{N-1} e^{2\pi i m x_n} \to \begin{cases} 0 & m = 0 \\ 1 & m \neq 0 \end{cases}$$

Proof. Let $\chi_m(t) = e^{s\pi i m t}$. The linear span of $\{\chi_m\}_{m\in\mathbb{Z}}$ is dense in $C(\mathbb{R}/\mathbb{Z})$ by Fourier analysis so equidistribution of (x_k) is equivalent to $S_N\chi_m(x) \to \int \chi_m d\mu$ for every m. This is what the lemma says. □

Proof number 2. Fix $t \in \mathbb{R}/\mathbb{Z}$ and $x_k = t + \alpha k$. For $m = 0$ the limit in Weyl's criterion is automatic so we only need to check $m \neq 0$. Then

$$\frac{1}{N}\sum_{n=0}^{N-1} e^{2\pi i m x_n} = \frac{1}{N}e^{2\pi i m t} \cdot \sum_{n=0}^{N-1}(e^{2\pi i m \alpha})^n = \frac{1}{N}e^{2\pi i t} \cdot \frac{e^{2\pi i m \alpha N} - 1}{e^{2\pi i m \alpha} - 1} = 0$$

(note that $\alpha \notin \mathbb{Q}$ ensures that the denominator is not 0, otherwise the summation formula is invalid). □

Corollary 6.5.6. *For any open or closed set $A \subseteq \mathbb{R}/\mathbb{Z}$, for every $x \in \mathbb{R}/\mathbb{Z}$, $S_N 1_A(x) \to Leb(A)$.*

Proof. The boundary of an open or closed is countable and hence of Lebesgue measure 0. □

Example 6.5.7 (Benford's law)**.** Many samples of numbers collected in the real world exhibit the interesting feature that the most significant digit is not uniformly distributed. Rather, 1 is the most common digit, with frequency approximately 0.30; the frequency of 2 is about 0.18; the frequency of 3 is about 0.13; etc. More precisely, the frequency of the digit k is approximately $\log_{10}(1 + \frac{1}{d})$.

We will show that a similar distribution of most significant digits holds for powers of b whenever b is not a rational power of 10. The main observation is that the most significant base-10 digit of $x \in [1, \infty)$ is determined by $y = \log_{10} x \bmod 1$, and is equal to k if $y \in I_k = [\log_{10} k, \log_{10}(k + 1))$. Therefore, the asymptotic frequency of k being the most significant digits of b^n is

$$\begin{aligned}
\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} 1_{I_k}(\log_{10} b^n) &= \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} 1_{I_k}(n\frac{\ln b}{\ln 10}) \\
&= Leb(I_k) \\
&= Leb[\log_{10} k, \log_{10}(k+1)] \\
&= \log_{10}(1 + \frac{1}{k})
\end{aligned}$$

since this is just the frequency of visits of the orbit of 0 to $[\log_{10} k, \log_{10}(k+1)]$ under the map $t \mapsto t + \ln b/\ln 10 \bmod 1$, and $\ln b/\ln 10 \notin \mathbb{Q}$ by assumption (it would be rational if and only if b is a rational power of 10).

6.6 Topological models

Our last goal in this section is to show that most measure preserving systems are isomorphic to an invariant measure on a topological dynamical system. We first define isomorphism. As is usually the case in measure theory, we are interested in functions only up to changes on measure zero sets, and this is true also of measure preserving maps. Thus two systems will be considered isomorphic if, after discarding null sets, their points can beidentified so as for the measures and maps to agree. More precisely,

Definition 6.6.1. Two measure preserving systems $(X, \mathcal{B}, \mu, T)$ and $(Y, \mathcal{C}, \nu, S)$ are **isomorphic** if is a T-invariant measurable subsets $X_0 \subseteq X$ and an S-invariant measurable subset $Y_0 \subseteq Y$, and a function $\pi : X_0 \to Y_0$ such that

1. π is 1-1 and onto and both π, π^{-1} are measuable,
2. π is measure preserving, that is, $\nu = \pi\mu$, or equivalently, $\mu(\pi^{-1}C) = \nu(C)$ for all $C \in \mathcal{C}$;
3. π is **equivariant**, i.e.
$$S \circ \pi = \pi \circ T$$
or, equivalently, the following diagram commutes:
$$\begin{array}{ccc} X & \xrightarrow{T} & X \\ \downarrow \pi & & \downarrow \pi \\ Y & \xrightarrow{S} & Y \end{array}$$

Lemma 6.6.2. *If $(X, \mathcal{B}, \mu)$ is a measure space and $S, T : X \to X$ preserve μ and agree outside a nullset E, then $(X, \mathcal{B}, \mu, T)$ and $(X, \mathcal{B}, \mu, S)$ are isomorphic.*

Proof. Let G denote the (semi-)group generated by S, T and the identity, and let $G^{-1}E = \bigcup_{g \in G} g^{-1}E$. Then $G^{-1}E$ is a nullset which is S and T invariant and contains E, so S, T are isomorphic to their restrictions to $X \setminus G^{-1}E$, on which they agree. Hence they are isomorphic. □

Now, isomorphism between dynamical systems implies isomorphism, in the obvious sense, of the underlying measurable and probability spaces, and also of the associated L^p spaces (by an easy generalization of the Koopman operator). Not every probability space is isomorphic to a compact metric space with a Borel probability measure; for example, in the latter L^2 is separable, which in general need not hold. Thus, if we want a measure preserving system to be isomorphic to a topological one, we must avoid obstructions coming from the underlying measure spaces.

Definition 6.6.3. A measurable space $(X, \mathcal{B})$ is a **standard Botel space** if there exists a complete, separable metric on X for which $\mathcal{B}$ is the σ-algebra of Borel sets.

Fact 6.6.4. *$(X, \mathcal{B})$ is a strandard Borel space if and only if there is a compact metric on X whose Borel sets coincide with$\mathcal{B}$.*

Lemma 6.6.5. *Let $(X, \mathcal{B}, \mu)$, $(Y, \mathcal{C}, \nu)$ be probability spaces with $(X, \mathcal{B})$ and $(Y, \mathcal{C})$ standard Borel spaces. Let $\pi : X \to Y$ be a measuable map which preserves the measures in the sense that $\mu(\pi^{-1}C) = \nu(C)$ for all $C \in \mathcal{C}$, and suppose that π is injective. Then there exists a measurable $X_0 \subseteq X$ whose image $Y_0 = \pi X_0$ is measurable, and $\mu(X_0) = \nu(Y_0) = 1$.*

Proof. We may assume that X, Y are equipped with compact metrics. By Egorov's theorem, for every $\varepsilon > 0$ there exists $X_\varepsilon \in \mathcal{B}$ with $\mu(X_\varepsilon) > 1 - \varepsilon$ and $\pi|_{X_\varepsilon} : X_\varepsilon \to Y$ continuous. Let $Y_\varepsilon = \pi X_\varepsilon$; then Y_ε is compact and hence measurable. Let $X_0 = \bigcup_{n=1}^{\infty} X_{1/n}$, since $X_{1/n} \subseteq X_0$ we have $\mu(X_0) \geq 1 - 1/n$ for all n, hence $\mu(X_0) = 1$. Also, $\pi X_0 = \bigcup_{n=1}^{\infty} \pi X_n$ so $Y_0 = \pi X_0$ is measurable. Since $\nu(Y_0) = \mu(\pi^{-1} X_0) \geq \mu(X_0) = 1$ we have $\nu(Y_0) = 1$. □

Theorem 6.6.6. *Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system whose underlying measure space is standard Borel. Then it is isomoprhic to an invariant measure on a topological dynamical system.*

Proof. For simplicity we give the proof when T is invertible. We may by assumption assume that X is a compact metric space; but of course, T is not continuous.

Define $\pi : X \to X^{\mathbb{Z}}$ by $\pi(x) = (T^n x)_{n \in \mathbb{Z}}$. Endow $X^{\mathbb{Z}}$ with the product topology, which is compactly metrizable, and let $\mathcal{C}$ denote the Borel sets. Then π is Borel measurable, since its composition with each cordinate projection gives one of the measurable maps T^n. If we denote the shift map on $X^{\mathbb{Z}}$ by S, then S is continuous, and $\pi T = S \pi$. Define a Borel probability measure ν on $X^{\mathbb{Z}}$ by $\nu(E) = \mu(\pi^{-1} E)$. Then

$$\nu(S^{-1}E) = \mu(\pi^{-1}(S^{-1}E)) = \mu(T^{-1}(\pi^{-1}E)) = \mu(\pi^{-1}E) = \nu(E)$$

so $(X^{\mathbb{Z}}, \mathcal{C}, \nu, S)$ is a measure preserving system. Thus, π is almost an isomorphism;l it fails to be one only because it is not onto.

To correct this we apply the previous lemma. This gives sets $X_0 \subseteq X$ and $Y_0 \subseteq X^{\mathbb{Z}}$ of full measure, such that $Y_0 = \pi X_0$. The sets are not invariant but we can replace them with $X_1 = \bigcap_{n \in \mathbb{Z}} T^n X_0$ and $Y_1 = \bigcap_{n \in \mathbb{Z}} S^n Y_0$; clearly $Y_1 = \pi X_1$, both are measurable, both have full measure, and they are now invariant. This completes the proof. □

Chapter 7

Eigenvalues, group rotations and isometries

We have seen that 1 is an eigenvalue for the Koopman operator of a measure preserving system, , since the constant functions are eigenvactors for it. . We also saw that ergodicity is equivalent to simplicity of this eigenvalue. In this section we explore the significance of other eigenvalues, and their connections with algebraic and isometric factors of the system,

7.1 Eigenvalues of the Koopman operator

In this chapter and the next, all function spaces are complex, and $L^2(\mu)$ is equipped with the inner product $\langle f, g\rangle = \int f\overline{g}d\mu$.

Definition 7.1.1. Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system. λ is an **eigenvalue** of T if it is an eigenvalue for the associated Koopman operator, that is, if there is an $0 \neq f \in L^2(\mu, \mathcal{B})$ such that $Tf = \lambda f$. Such f is called an **eigenvector** for the eigenvalue λ, and the set of all eigenvectors for λ (together with 0) forms a closed linear subspace called the **eigenspace** of λ. We denote the set of eigenvalues by $\Sigma(T)$.

Eigenvalues and eigenvectors have some elementary properties:

1. Eigenvalues have modulus one. Indeed, the Koopman operator preserves norms and inner products, hence for an eigenvector v,

$$\|v\| = \|Tv\| = \|\lambda v\| = |\lambda|\,\|v\|$$

2. Similarly, if u, v are eigenvectors for α, β respectively, and $\alpha \neq \beta$, then $u \perp v$: indeed, since $\alpha\overline{\beta} = \alpha\beta^{-1} \neq 1$

$$\langle u, v\rangle = \langle Tu, Tv\rangle = \langle \alpha u, \beta v\rangle = \alpha\overline{\beta}\,\langle u, v\rangle$$

3. In particular, if L^2 is separable (e.g. if $(X, \mathcal{B})$ is stnadard Borel), then there are at most countably many eigenvalues.

4. If T is ergodic, and f is an eigenfunction for λ, then $|f|$ is a.s. constant. Indeed,
$$T|f| = |Tf| = |\lambda f| = |f|$$
hence $|f|$ is invariant and by ergodicity it is a.s. constant. By convention one usually takes eigenfunctions to have modulus one.

5. If T is ergodic then each eigenvalue has multiplicity one, that is, the space of eigenvectors is one-dimensional. Indeed, if f, g are eigenvectors for λ then $f/g : X \to S^1$ is invariant, because
$$T(f/g) = Tf/Tg = \lambda f/\lambda g = f/g$$
and hence f/g is a constant function of modulus one, so $f = \alpha g$ for some constant with $|\alpha| = 1$.

6. The set $\Sigma(T)$ of eigenvalues of the Koopman operator is a subgroup of the multiplicative group $S^1 = \{z \in \mathbb{C} : |z| = 1\}$. Indeed, if f, g are eigenfunctions for α, β then $f \cdot g$ is an eigenvalue for $\alpha\beta$, and $\overline{f}$ is an eigenvalue for $\overline{\alpha} = \alpha^{-1}$. Indeed, these functions are in L^2 because they are bounded, and
$$T(fg) = Tf \cdot Tg = \alpha f \cdot \beta g = \alpha\beta \cdot fg$$
and similarly $T\overline{f} = \overline{\alpha} \cdot \overline{f} = \alpha^{-1}\overline{f}$.

Example 7.1.2. Let $X = \{0, \ldots, n-1\}$ and $Tx = x + 1 \bmod n$. Then $f(x) = e^{2\pi i x/n}$ is an eigenfunction with eigenvalue $e^{2\pi i/n}$.

Example 7.1.3. Let $(S^1, \mathcal{B}, \mu, R_\alpha)$ be a circle rotation. Then $e^{2\pi i\alpha}$ is an eigenvalue, because for the identity map 2 $f(z) = z$,
$$f(Tx) = f(e^{2\pi\alpha}z) = e^{2\pi i\alpha}z = e^{2\pi i\alpha}f(x)$$
Furthermore, β is an eigenvalue if and only if $\beta = \alpha^n$ for some $n \in \mathbb{Z}$. Indeed, $\{z^n\}_{n\in\mathbb{Z}}$ are eigenvectors with eigenvalues $\{\alpha^n\}_{n\in\mathbb{Z}}$, and they are dense in L^2 (since they are dense in $C(S^1)$. Since an eigenvalue for $\beta \notin \{\alpha^n\}$ would be perpendicular to all z^n, it must be 0.

7.2 Group rotations

Eigenvectors and eigenvalues are algebraic features of the Koopman operator. But, as we shall see, their existence is explained by a structurel property of the dynamical system in question. More precisely, we shall see that the presence of nontrivial eigenvalues implies that the dynamical system has a compact group translation as a factor.

Definition 7.2.1. A **compact metric group** (c.m.g.) is a group X together with a compact metric on X, such that the group operations are continuous: if $x_n \to x$ and $y_n \to y$ in X then $x_n y_n \to xy$ and $x_n^{-1} \to x^{-1}$.

For instance, S^1 with the metric induced from $\mathbb{C}$ and the multiplication opraiton is a compact metric group. More generally, the group of orthogonal matrices with the operator norm is a compact group. Real or complex vector spaces and groups of matrices are metric groups but generally not compact.

Definition 7.2.2. A **character** of a compact metric group X is a continuous group homomorphism $\varphi : X \to S^1$, i.e. $\varphi(xy) = \varphi(x)\varphi(y)$. The characters form a group under pointwise multiplication. The group of characters is denoted $\widehat{X}$, and called the **dual group**.

Any non-trivial continuous group homomorphism $\varphi : X \to (\mathbb{C}, \times)$ must have image in S^1, because if $\varphi(x) \neq 0$ then $\varphi(x^n)=\varphi(x)^n$ and $\varphi(x^{-n}) = \varphi(x)^{-n}$, so unless $|\varphi(x)| = 1$ we conclude that φ is unbounded, contradicting continuity and compactness of X.

It is not obvious that there exist non-trivial characters, but in fact there are plenty:

Theorem 7.2.3 (Pontryagin/Peter-Weyl)**.** *Let X be an abelian compact metric group Then the characters separate points and their linear span of $\widehat{X}$ is dense in $C(X)$.*

We also need

Theorem 7.2.4 (Haar)**.** *A compact metric group X admits a unique Borel probability measure μ_X which is invariant under translations, i.e. for every $E \in \mathcal{B}$ and $g \in G$, $\mu(E) = \mu(gE) = \mu(Eg)$.*

We do not prove thes theorems here.

Definition 7.2.5. By a **group rotation** we will mean a dynamical system of the form $(X, \mathcal{B}, \mu, T)$ where X is an abelian compact metric group, $\mathcal{B}$ the Borel σ-algebra, μ the Haar measure, and T is a translation, i.e. $Tx = gx$ for a fixed element $g \in X$.

Note that translation maps preserve Haar measure, because $\mu \circ T^{-1}$ is again an invariant probability measure, and thus, by uniquenes of the Haar measure, equal to μ. It follows that a group translation is a m.p.s.

We now return to the Koopman operator:

Proposition 7.2.6. *Let $(X, \mathcal{B}, \mu, T)$ be group rotation with T translating by g. Then every character $\varphi \in \widehat{X}$ is an eigenfunction for T with eigenvalue $\varphi(g)$, and the eigenfunctions span $L^2(\mu, \mathcal{B})$.*

Proof. Let φ be a character. Then

$$T\varphi(x) = \varphi(Tx) = \varphi(gx) = \varphi(g)\varphi(x)$$

Thus φ is an eigenfunction for T. The second part follows from the Peter-Weyl theorem. □

7.3 Kronnecker factors and the Halmos-von Neumann theorem

Our goal is to show that nontrivial eigenvalues for the Koopman operator arise from group rotations "hiding" in the dynamics. We first introduce the notion of a factor.

Definition 7.3.1. Let $(X, \mathcal{B}, \mu, T)$ and $(Y, \mathcal{C}, \nu, S)$ be measure preserving systems. We say that Y is a factor of X if there is a measurable map $\pi : X \to Y$ defined μ-a.e., such that $\mu \circ \pi^{-1} = \nu$ and $\pi \circ T = S \circ \pi$, i.e. the following diagram commutes:

$$\begin{array}{ccc} X & \xrightarrow{T} & X \\ \downarrow \pi & & \downarrow \pi \\ Y & \xrightarrow{S} & Y \end{array}$$

π is called a **factor map**.

Example 7.3.2. If we start with two measure preserving systems $(X, \mathcal{B}, \mu, T)$ and $(Y, \mathcal{C}, \nu, S)$ and form the product system $X \times Y$ with the product measure $\mu \times \nu$ and the action $R(x, y) = (Tx, Sy)$, then the marginal maps $(x, y) \mapsto x$ and $(x, y) \mapsto y$ define factor maps from the product system to X, Y resp.

Example 7.3.3. Suppose that $(X, \mathcal{B}, \mu, T)$ is an invertible m.p.s. and there is a partition of X into n sets $X_0, X_1, \ldots, X_{n-1} \in \mathcal{B}$ such that T permutes them cyclically: $TX_i = X_{i+1 \bmod n}$. Then we can define $\pi : X \to \{0, \ldots, n-1\}$ with $f(x) = i$ iff $x \in X_i$, and taking $Si = i + 1 \bmod 1$, for $x \in X_i$ we have $\pi(Tx) = i + 1 \bmod 1$ by the defining property of the partition. Noticing that $\mu(X_i) = \mu(T^{-1}X_{i+!})$, so all X_i have the same mass, which must be $1/n$. Taking normalized counting measure ν on $\{0, \ldots, n-1\}$, we find that

$$\mu(\pi^{-1}(E)) = \mu(\bigcup_{i \in I} \pi^{-1}(i) = \mu(\bigcup_{i \in E} X_i) = \frac{1}{n}|E| = \nu(E)$$

Thus π is a factor map.

Conversely, suppose that $\pi : X \to \{0, \ldots, n-1\}$ is a factor map; then $X_i = \pi^{-1}(i)$ is a partition as above.

Proposition 7.3.4. *In an ergodic m.p.s.* $(X, \mathcal{B}, \mu, T)$ *admits a nontrivial group rotation as a factor then it admits nontrivial eigenfunctions. Furthermore, every eivenfunction of* Y *lifts to one of* X*, with the same eigenvalue.*

Proof. Let $(Y, \mathcal{C}, \nu, S)$ be a non-trivial group factor of X by the map $\pi : X \to Y$. Let φ be a non-trival eigenfunction character of Y (obtained from some character), with eigenvalue λ. Let $\psi(x) = \varphi(\pi x)$. Since φ is bounded, so if ψ, so it is in L^2, and

$$\psi(Tx) = \varphi(\pi(Tx)) = \varphi(S\pi(x)) = \lambda\varphi(\pi x) = \lambda\psi(x)$$

so ψ is an eigenfunction with eigenvalue λ. □

Proposition 7.3.5. *In an ergodic m.p.s. $(X, \mathcal{B}, \mu, T)$, if $f \in L^2(\mu, \mathcal{B})$ is an eigenfunction with eigenvalue $\lambda = e^{2\pi i\alpha}$, then X admits a group rotation factor. Specifically, letting $\lambda = e^{2\pi i\alpha}$,*

1. *If $\alpha \notin \mathbb{Q}$ then X factors to S^1 with the rotation R_α.*
2. *If $\alpha = k/m \in \mathbb{Q}$ in reduced form, then X factors to $\{0, \ldots, m-1\}$ with addition byt 1 mod m.*

Proof. The eigenfunction relation implies that

$$f(Tx) = \lambda f(x) = R_\alpha f(x)$$

Assume first $\alpha \notin \mathbb{Q}$, then this proves equivariance. Let ν normalized denote Lebesgue measure on S^1; we must show that $\nu = \mu \circ \varphi^{-1}$. To see this, define another measure ν' on S^1 by $\nu'(E) = \mu(f^{-1}E)$. This is an invariant measure for R_α because of the factor relation:

$$\nu'(R_\alpha^{-1}E) = \mu(f^{-1}R_\alpha^{-1}E) = \mu(((R_\alpha f)^{-1}E) = \mu((fT)^{-1}E) = \mu(T^{-1}(f^{-1}E)) = \mu(f^{-1}E) = \nu(E)$$

But (S^1, R_α) is uniquely ergodic, so $\nu' = \nu$.

Now suppose that $\alpha = k/m$. Then $f(Tx) = R_\alpha f(x)$ is still true but it will no longer be true that $\nu = \mu \circ \varphi^{-1}$. Instead we argue as follows. Let $\nu' = \mu \circ \varphi^{-1}$. Then by definition (S^1, ν', R_α) is a factor of X. This implies that X is ergodic: if $R_\alpha^{-1}E = E$ then

$$T^{-1}\varphi^{-1}E = \varphi^{-1}R_\alpha^{-1}E = \varphi^{-1}E$$

hence by ergodicity of$T$, $\mu(\varphi^{-1}E) = 0$ or 1. But then $\nu(E) = \mu(\varphi^{-1}E) = 0$ or 1, so (S^1, ν', R_α) is ergodic.

Now, each R_α-orbit is a coset of the grou $Y \subseteq S^1$ of m-th roots of unity in S^1, and the partition of S^1 into cosets of Y is countably generated (it is essentially the same as the partition of the arc $\{e^{2\pi it} : t \in [0, 1/m)\}$ into points). Thus ν' is supported on one coset $z_0 Y$. Let $\psi(x) = z_0^{-1}\varphi(x)$. Then ψ is an eigenfunction with eigenvalue λ (it is a scalar multiple of φ) and $\psi : X \to Y$ takes μ to counting $\nu' \circ R_{z_0}$ which is (by unique ergodicity of cyclic permutations) normalized counting measure on Y. This is the desired factor map (it is clearly the isomorphic to addiiton of 1 on $\{0, \ldots, m-1\}$). □

Combining the last two propositions we have:

Corollary 7.3.6. *A m.p.s. admits nontrivial eigenfunctions if and only if it admits a nontrivial group rotation factor.*

We will now take a closer look at systems that are isomorphic to group rotations. Let $\mathbb{T}^\infty = (S^1)^\mathbb{N}$, which is a compact metrizable group with the product topology. Let m be the infinite product of Lebesgue measure, which is invariant under translations in $\mathbb{T}^\infty$ and hence equal to Haar measure. Given $\alpha = (\alpha_1, \alpha_2, \ldots) \in \mathbb{T}^\infty$ let $L_\alpha : \mathbb{T}^\infty \to \mathbb{T}^\infty$ as usual be the translation map. Note that

$$L_\alpha^n x = \alpha^n x$$

Lemma 7.3.7. *The orbit closure G_α of $0 \in \mathbb{T}^\infty$ under L_α (that is, the closure of $\{\alpha^n : n \in \mathbb{N}\}$) is the closed subgroup generated by α.*

Proof. It is clear that it is contained in the group in question, is closed, contains α, and is a semigroup. The latter is because it is the closure of the semigroup $\{\alpha^n\}_{n\in\mathbb{N}}$; explicitly, if $x, y \in G_\alpha$ then $x = \lim \alpha^{n_i}$ and $y = \lim 0^{m_j}$, so

$$xy = (\lim \alpha^{m_j})(\lim \alpha^{n_i}) \lim \alpha^{m_i+n_i} \in G_\alpha$$

Also, if we knew that $g^{-1} \in G_\alpha$ we would have

$$x^{-1} = (\lim \alpha^{m_j})^{-1} \in \lim \alpha^{-m_j} = \lim(\alpha^{-1})^{m_j}$$

since G_α is closed under multiplication, $(\alpha^{-1})^{m_j} \in G_\alpha$, and since G_α is closed, the last line shows that $x^{-1} \in G_\alpha$.

It remains to show that $\alpha^{-1} \in G_\alpha$. For this choose $n_i \to \infty$ such that α^{n_j} converges, to some z. We can assume that $n_{i+1} > n_i - 1$ (since $n_j \to \infty$), and by passing to a subsequence, we can also assume that $\alpha^{n_{i+1}-n_i-1}$ converges, to some $\beta \in G_\alpha$. But:

$$\alpha\beta = \alpha \lim \alpha^{n_{i+1}-n_i-1} = \alpha \lim \alpha^{n_{i+1}} \lim \alpha^{-n_i} = zz^{-1} = 1$$

sp $\alpha^{-1} = \beta \in G_\alpha$, as claimed. □

Lemma 7.3.8. *Let $\alpha \in \mathbb{T}^\infty$ and G_α be as above. Let m_α be the Haar measure on G_α, equivalently, the unique L_α-invariant measure. Then in the m.p.s. (G_α, m, L_α), the spectrum $\Sigma(L_\alpha) = \langle \alpha_1, \alpha_2, \ldots \rangle \subseteq S^1$ is the discrete group generated by the coordinates α_i of α.*

Proof. Let $\pi_n : \mathbb{T}^\infty \to S^1$ denote the n-th coordinate projection. Clearly $\pi_n(L_\alpha x) = \alpha_n x_n = \alpha_n \pi_n(x)$, so the functions π_n are eigenfunctions of $(G_\alpha, m_\alpha, L_\alpha)$. Let $\mathcal{A}$ denote the $\mathbb{C}$-algebra generated by $\{\pi_n\}$. This is an algebra of continuous functions that separate points in $\mathbb{T}^\infty$ and certainly in G_α, so they are dense in $L^2(m_\alpha)$. Since this algebra consists precisely of the eigenfunctions with eigenvalues in $\langle \alpha_1, \alpha_2, \ldots \rangle$ we are done. □

Proposition 7.3.9. *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic measure preserving system on a standard Borel space. Let $\Sigma(T) = \{\alpha_1, \alpha_2, \ldots\}$ and $\alpha = (\alpha_1, \alpha_2, \ldots) \in \mathbb{T}^\infty$. Then*

1. *$(G_\alpha, m_\alpha, L_\alpha)$ is a factor of X.*
2. *Every eigenfunction of T arises as a lift of an eigenfunction of $(G_\alpha, m_\alpha, L_\alpha)$.*
3. *If $L^2(\mu)$ is spanned by eigenfunctions, then the factor map $X \to G_\alpha$ is an isomorphism.*
4. *If $\tau : X \to Y$ is a factor map to a group rotation (Y, S, ν), then π factors through G_α; that is, writing $\pi : X \to G_\alpha$ for the factor map, there exists a factor map $\tau' : G_\alpha \to Y$ such that $\tau = \tau'\pi$.*

Proof. Note that Σ is countable because eigenvectors of distinct eigenvalues are orthogonal, and $L^2(\mu, \mathcal{B})$ is separable by standardness of the space.

Let f_i be an eigenvector for α_i and define $F : X \to \mathbb{T}^\infty$ by

$$F(x) = (f_1(x), f_2(x), \ldots)$$

It is immediate that $F(Tx) = L_\alpha F(x)$.

Let $\nu = \mu \circ F^{-1}$. This is an L_α-invariant and ergodic measure on $\mathbb{T}^\infty$ (compare with the same property of ν' in Proposition 7.3.5).

Let $y_0 \in \operatorname{supp} \nu$ be a point with dense orbit in in $\operatorname{supp} \nu$ (which exists by ergodicity). Consider the map $L_{y^{-1}} x \mapsto y_0^{-1} x$, which commutes with L_α (because $\mathbb{T}^\infty$ is abelian), and note that $L_{y_0^{-1}}$ maps the L_α-orbit of y_0 to the L_α-orbit of 1, and so it maps the orbit closure of y_0 to G_α. Writing $\eta = \nu \circ L_{y_0^{-1}}^{-1}$ it follows that η is L_α-invariant and $\operatorname{supp} \eta = L_{y_0^{-1}} \operatorname{supp} \nu = G_\alpha$.

Since η is invariant under α^n for every n, it is invariant under every element of G_α, so by uniqueness of Haar measure, $\eta = m_\alpha$. Also it is easy to check that the map $\pi : x \mapsto L_{y^{-1}} F(x)$ is equivariant with L_α. Thus π is a factor map from X to $(G_\alpha, m_\alpha, R_\alpha)$.

Next, suppose that the eigenfunctions span $L^2(\mu)$. Since each eigenfunction is lifted by π from one of the eigenfunctions of G_α, we find that $L^2(\pi^{-1}\mathcal{B}_\alpha) = L^2(\mu)$, where $\mathcal{B}_\alpha$ is the Borel algebra of G_α. It follows that π is 1-1 a.e. and by standardness it is an isomorphism (this fact is beyond the scope of the course, but is not too hard).

Finally suppose $\tau : X \to Y$ with (Y, ν, S) a group rotation. The eigenvectors are dense in $L^2(\nu)$, and we have an isomorphism $\sigma : Y \to G_\beta$ where $\beta = (\beta_1, \beta_2, \ldots)$ enumerates $\Sigma(S)$. Now, each eigenvector f of (G_β, S) lifts to one of X and since the multiplicity is 1, this is *the* eigenvector for its eigenvalue, up to multiplication by a constant phase. Thus the coordinates of α include those of β. Let $\tau' : G_\alpha \to \mathbb{T}^\infty$ denote projection to the coordinates corresponding to eigenvalues of S; then the map $\tau' \circ F = \tau$. The claim follows. □

Definition 7.3.10. An ergodic measure preserving system has discrete spectrum if L^2 is spanned by eigenfunctions.

Corollary 7.3.11. *Discrete spectrum systems are isomorphic if and only if the induced unitary operators are unitarily equivalent.*

Proof. This follows from the theorem above and the fact that two diagonalizable unitary operators are unitarily equivalent if and only if they have the same eigenvalues (counted with multiplicities), and ergodicity implies that all eigenvalues are simple. □

Part (3) of the theorem above shows that every measure preserving system has a maximal isometric factor. This factor is called the **Kronecker factor**. The factor is canonical, although the factor map is not – one can always post-compose it with a translation of the group.

We emphasize that in general it is false that unitary equivalence implies ergodic-theoretic isomorphism. The easiest example to state is that the product measures $(1/2, 1/2)^{\mathbb{Z}}$ and $(1/3, 1/3, 1/3)^{\mathbb{Z}}$ with the shift map have unitarily isomorphic induced actions on L^2, but they are not isomorphic.

7.4 Isometries and group rotations

Definition 7.4.1. A m.p.s. $(X, \mathcal{B}, \mu, T)$ is **isometric** if there exists a compact metric on the phase space for which $\mathcal{B}$ is the Borel σ-algebra and T acs as an isometry.

Group rotations are isometric: if d is a metric on a compact abelian group rotation (X, R_g) and μ is the Haar measure, we can define a new quivalent metric

$$\widetilde{d}(x, y) = \int d(gx, gy)d\mu(g)$$

(we leave the verification as an exercise). Writing $\delta(g) = d(gx, gy)$, we have

$$\begin{aligned} \widetilde{d}(hx, hy) &= \int d(ghx, ghy)d\mu(g) \\ &= \int d(hgx, hgy)d\mu(g) \\ &= \int \delta \circ R_h(g)d\mu(g) \\ &= \int \delta(g)d\mu(g) \\ &= \widetilde{d}(x, y) \end{aligned}$$

Thus $x \mapsto hx$ acts as an isometry.

The converse is, surprisingly, true as well, at least under the mild assumption that the isometry admits a dense orbit.

Proposition 7.4.2. *Let (Y, d) be a compact metric space and $S : Y \to Y$ an isometry with a dense orbit. Then there is a compact metric group G and $g \in G$ and a homeomorphism $\pi : Y \to G$ such that $L_g\pi = \pi S$. Furthermore if ν is an invariant measure on Y then it is ergodic and $\pi\nu$ is Haar measure on G.*

Proof. Consider the group Γ of isometries of Y with the sup metric,

$$d(\gamma, \gamma') = \sup_{y \in Y} d(\gamma(y), \gamma'(y))$$

Then (Γ, d) is a complete metric space, and note that it is right invariant: $d(\gamma \circ \delta, \gamma \circ \delta) = d(\gamma, \gamma')$.

Let $y_0 \in Y$ have dense orbit and set $Y_0 = \{S^n y_0\}_{n \in \mathbb{Z}}$. If the orbit is finite, $Y = Y_0$ is a finite set permuted cyclically by S, so the statement is trivial.

Otherwise $y \in Y_0$ uniquely determines n such that $S^n y_0 = y$ and we can define $\pi : Y_0 \to \Gamma$ by $y \mapsto S^n \in \Gamma$ for this n.

We claim that π is an isometry. Fix $y, y' \in Y_0$, so $y = S^n y_0$ and $y' = S^{n'} y_0$, so

$$d(\pi y, \pi y') = \sup_{z \in Y} d(S^n z, S^{n'} z)$$

Given $z \in Y$ there is a sequence $n_k \to \infty$ such that $S^{n_k} y_0 \to z$. But then

$$\begin{aligned}
d(S^n z, S^{n'} z) &= d(S^n(\lim S^{n_k} y_0), S^{n'}(\lim S^{n_k} y_0)) \\
&= \lim d(S^n S^{n_k} y_0, S^{n'} S^{n_k} y_0) \\
&= \lim d(S^{n_k}(S^n y_0), S^{n_k}(S^{n'} y_0)) \\
&= \lim d(S^n y_0, S^{n'} y_0) \\
&= d(S^n y_0, S^{n'} y_0) \\
&= d(y, y')
\end{aligned}$$

Thus $d(\pi y, \pi y') = d(y, y')$ and π is an isometry $Y_0 \hookrightarrow \Gamma$. Furthermore, for $y = S^n y_0 \in Y_0$,

$$\pi(Sy) == \pi(SS^n y) = S^{n+1} = L_S S^n = L_S \pi(y)$$

It follows that π extends uniquely to an isometry with $Y \hookrightarrow \Gamma$ also satisfying $\pi(Sy) = S(\pi y)$. The image $\pi(Y_0)$ is compact, being the continuous image of the compact set Y. Since $\pi(Y_0) = \{S^n\}_{n \in \mathbb{Z}}$ and this is a group its closure is also a group G.

Finally, suppose ν is an invariant measure on Y. Then $m = \pi\nu$ is L_S invariant on G. Since it is invariant under L_S it is invariant under $\{L_S^n\}_{n \in \mathbb{Z}}$, and this is a dense set of elements in G. Thus m it is invariant under every translation in G, and there is only one such measure up to normalization: Haar measure. The same argument applies to every ergodic components of m (w.r.t. L_S) and shows that the ergodic components are also Haar measure. Thus m is L_S-ergodic and since π is an isomorphism, (Y, ν, S) is ergodic. □

Corollary 7.4.3. *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic m.p.s. Then X is weak mixing if and only if it admits only trivial isometric systems as factors.*

Chapter 8

Weak mixing

Group rotations define a class of dynamical systems characterized by the presence of many sufficiently many eigenfunctions. We now turn to study the opposite property, defined in the next section. The remainder of the chapter is then devoted to exploring numerous other properties which are equivalent to weak mixing and shed new light on the phenomenon.

8.1 Weak mixing

Weak mixing, which we shall see is intermediate between these two, is a property wth many equivalent formulations, which make it resemble both ergodicity and mixing (and some altogether different). For the definition we choose the following.

Definition 8.1.1. A m.p.s. $(X, \mathcal{B}, \mu, T)$ is **weak mixing** if the only eigenfunctions are the constant ones.

We can right away give two equivalent definitions. Since the constant functions form a one-dimensional subspace of L^2 consisting of eigenfunctions of eigenvalue 1, we find that

> A system is weak mixing if and only if 1 is the only eigenvalue, and it is simple (the space of corresponding eigenfunctions is one-dimensional).

Also, since weak mixing implies ergodicity, and for ergodic m.p.s.'s the existence of non-trivial eigenfunctions is the same as the existence of non-trivial isometric factors, we find that

> A system is weak mixing iff it is ergodic and has no non-trivial isometric factors.

Proposition 8.1.2. *Mixing implies weak mixing, and weak mixing implies ergodicity.*

Proof. Let $(X, \mathcal{B}, \mu, T)$ denote the system in quesiton.

Weak mixing implies that there are no non-constant eigenfunctions, in partiular no non-constant invariant functions, hence it implies ergodicity.

Now suppose that T is mixing, so $\int f \cdot T^n g d\mu \to \int f d\mu \cdot \int g d\mu$ for all $f, g \in L^2$. Suppose that h is an eigenfunction with eigenvalue λ. Then taking $f = \overline{h}$ and $g = h$ the above becomes

$$\int f \cdot T^n g d = \int \overline{h} \cdot \lambda^n h d\mu = \lambda^n \|h\|^2 \to \int f \cdot g d\mu = \|h\|^2$$

We found that $\lambda^n \|h\|^2 \to \|h\|^2$. Since $\|h\| \neq 0$ this implies $\lambda^n \to 1$, which is possilble only if $\lambda = 1$. □

Neither implication in the proposition can be reversed. Here is an example showing that ergodicity does not imply weak mixing:

Example 8.1.3. Suppose that $T : S^1 \to S^1$ is the irrational rotation $z \mapsto \lambda z$. Then T is ergodic with respect to Lebesgue measure. But $F(x, y) = x\overline{y}$ is a non-trivial invariant function for $T \times T$: indeed,

$$((T \times T)F)(x, y) = F(Tx, Ty) = Tf \cdot \overline{Ty} = \lambda x \overline{\lambda y} = x\overline{y} = F(x, y)$$

since $\lambda\overline{\lambda} = 1$. Thus T is not weak mixing.

Examples exist of weakly mixing but not mixing m.p.s. but they are harder to construct. We may return to this later if time permits.

8.2 Double ergodicity

We have seen that weak mixing implies ergodicity, but is a stronger condition. It turns out to be equivalent to ergodicity of the cartesian product.

Theorem 8.2.1. *Let $(X, \mathcal{B}, \mu, T)$ be an invertible m.p.s. on a standard Borel space. Then it is weak mixing if and only if $(X \times X, \mathcal{B} \times \mathcal{B}, \mu \times \mu, T \times T)$ is ergodic.*

Recall that for an ergodic m.p.s., weak mixing is the same as admitting only trivial isometric factors. Thus one direction of the proof is provided by the following:

Proposition 8.2.2. *If a m.p.s. X admits a non-trivial isometric factor then $X \times X$ is not ergodic.*

Proof. First, observe that if Y is a nontrivial isometric system then it is not weak mixing, because, by definition, the metric is a nontrivial invariant function on $Y \times Y$. Now, if Y is an isometric factor of X then $X \times X$ factors onto $Y \times Y$, and so, since the latter is not ergodic, neither is the former (since $Y \times Y$ admits a non-trivial invariant set, and this lifts, via the factor map, to an invariant set in $X \times X$). □

For the converse direction we must show that if X is ergodic but $X \times X$ is not ergodic, then X admits a non-trivial isometric factors.

Recall that if $(X, \mathcal{B}, \mu)$ is a probability space then there is a pseudo-metric on $\mathcal{B}$ defined by

$$d(A, B) = \mu(A \Delta B) = \|1_A - 1_B\|_1$$

Identifying sets that differ in measure 0 dives a metric on equivalence classes, and the resulting space may be identified with the space of 0-1 valued functions in L^1, which is the same as the set of indicator functions. This is an isometry and since the latter space is closed in L^1, it is complete.

Now suppose that X is not weak mixing and let $A \subseteq X \times X$ be a non-trivial invariant set. Consider the map $X \to L^1$ given by $x \to 1_{A_x}$ where

$$A_x = \{y \in X \,:\, (x, y) \in A\}$$

The map is measurable, we use the fact that the Borel structure of the unit ball in L^1 in the norm and weak topologies coincide (this fact is left as an exercise). Then we only need to check that

$$x \mapsto \int 1_{A_x}(y) g(y)\, d\mu(y)$$

is measurable for every $g \in L^\infty$. For fixed g, this clearly holds when A is a product set or a union of product sets, and the general case follows from the monotone class theorem.

Now, notice that

$$\begin{aligned} TA_x &= \{Ty \,:\, (x, y) \in A\} \\ &= \{y \,:\, (x, T^{-1}y) \in A\} \\ &= \{y \,:\, (x, T^{-1}y) \in T^{-1}A\} \\ &= \{y \,:\, (Tx, y) \in A\} \\ &= A_{Tx} \end{aligned}$$

so $\pi : x \to A_x$ commutes with the action of T on X and L^1. Finally, the action of T on L^1 is an isometry. Therefore we have proved:

Claim 8.2.3. If T is not weak mixing then there is a complete metric space (Y, d), and isometry $T : Y \to Y$ and a Borel map $\pi : X \to Y$ such that $T\pi = \pi T$.

Let $\nu = \pi\mu$, the image measure; it is preserved. Thus (Y, ν, T) is almost the desired factor, except that the space Y is not compact (and there is another technicality we will mention later). To fix these problems we need a few general facts.

Definition 8.2.4. Let (Y, d) be a complete metric space. A subset $Z \subseteq Y$ is called ***totally bounded*** *if for every ε there is a finite set $Z_\varepsilon \subseteq Y$ such that $Z \subseteq \bigcup_{z \in Z_\varepsilon} B_\varepsilon(z)$.*

Lemma 8.2.5. *Let $Z \subseteq Y$ as above. Then $\overline{Z}$ is compact if and only if Z is totally bounded.*

Proof. This is left as an exercise. □

Lemma 8.2.6. *Let (Y, d) be a complete separable metric space, $S : Y \to Y$ an isometry and μ an invariant and ergodic Borel probability measure. Then* $\operatorname{supp}\mu$ *is compact.*

Proof. Let $C = \operatorname{supp}\mu$. This is a closed set and is clearly invariant so we only need to show that it is compact. For this it is enough to show that it is totally bounded.

Choose a μ-typical point y. By the ergodic theorem, its orbit is dense in C. Furthermore since S is an isometry, $B_r(S^n y) = S^n B_r(y)$. Now let $z \in C$. There is an n with $d(z, S^n y) < r$ so $B_r(T^n y) \subseteq B_{2r}(z)$, hence

$$\mu(B_r(y)) \leq \mu(B_{2r}(z))$$

This is true for every r.

Now, let $\{z_i\}$ be a maximal set of r-separated points in C. the set must be finite, because $B_{r/2}(z_i)$ are disjoint balls of mass uniformly bounded below. Therefore $B_{2r}(z_i)$ is a finite cover of C, and since r was arbitrary, C is totally bounded. □

Let ν be, as before, the image of μ under the map $\pi : X \to L^1$. Then $\operatorname{supp}\nu$ is compact and we can replace X by $\pi^{-1}(\operatorname{supp}\nu)$, which has full measure. We are done.

8.3 Mixing-in-density

Ergodicity and mixing are properties which measure how strongly the present and future of a dynamical system interact. Mixing means that the present and future are asymptotically independent: $\mu(A \cap T^{-n}B) \to \mu(A)\mu(B)$. To highlight the similarity between ergodicity and mixing, we begin by showing that ergodicity can is characterized as by same limit $\int fT^n g d\mu \to \int f d\mu \int g d\mu$, but with convergence understood in the Cesaro sense.

Lemma 8.3.1. *For a m.p.s. $(Y, \mathcal{C}, \nu, s)$ the following are equivalent:*

1. *Y is ergodic.*
2. *$\frac{1}{N}\sum_{n=0}^{N} \mu(A \cap T^{-n}B) \to \mu(A)\mu(B)$ for every $A, B \in \mathcal{C}$.*
3. *$\frac{1}{N}\sum_{n=0}^{N} \int f \cdot T^n g \to \int f d\mu \cdot \int g d\mu$ for every $f, g \in L^2(\nu)$.*

Proof. The equivalence of the last two conditions is standard via approximation of L^2 functions by step functions. We prove equivalence of the first two.

If (2) holds then for $A, B \in \mathcal{C}$ of positive measure we have $\mu(A \cap T^{-n}B) = \int 1_A \cdot T^n 1_B d\mu > 0$ infinitely often. This gives ergodicity.

Conversely, if the system is ergodic as in (1), then by the mean ergodic theorem, $S_N g = \frac{1}{N}\sum_{n=0}^{N-1} T^n g \to \int g$ in L^2 for any $g \in L^2$. So by continuity of the inner product,

$$\begin{aligned}
\lim_{N\to\infty} \sum_{n=0}^{N-1} \mu(A \cap T^{-n}B) &= \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N} \left(\int 1_A \cdot T^n 1_B d\mu \right) \\
&= \lim_{N\to\infty} \langle 1_A, S_N 1_B \rangle \\
&= \left\langle 1_A, \int 1_B d\mu \right\rangle \\
&= \langle 1_A, \mu(B) \rangle \\
& \quad \mu(A)\mu(B)
\end{aligned}$$

which is (2). □

Returning to weak mixing, we introduce a notion of convergence intermediate between Cesaro convergence and the standard convergence.

Definition 8.3.2. For a subset $I \subseteq \mathbb{N}$ we define the upper density to be

$$d(I) = \limsup_{N\to\infty} \frac{|I \cap \{1, \dots, N\}|}{N}$$

A sequence $a_n \in \mathbb{R}$ **converges in density** to $a \in \mathbb{R}$, denoted $a_n \xrightarrow{D} a$ or D-lim $a_n = a$, if

$$d(\{n : |a_n - a| > \varepsilon\}) = 0 \qquad \text{for all } \varepsilon > 0$$

Compare this to the usual notion of convergence, where we require the set above to be finite rather than 0-density. Since the union of finitely many sets of zero density has zero density, this notion of limit has the usual properties (with the exception that a subsequence may not have the same limit). One can also show the following:

Lemma 8.3.3. *For a bounded sequence a_n, the following are equivalent:*

1. $a_n \xrightarrow{D} a$.
2. $\frac{1}{N}\sum_{n=0}^{N} |a_n - a| = 0$.
3. $\frac{1}{N}\sum_{n=0}^{N} (a_n - a)^2 = 0$.
4. *There is a subset $I = \{n_1 < n_2 < \dots\} \subseteq \mathbb{N}$ with $d(I) = 1$ and $\lim_{k\to\infty} a_{n_k} = a$.*

We leave the proof of the lemma as an exercise.

Theorem 8.3.4. *For a m.p.s. $(X, \mathcal{B}, \mu, T)$ the following are equivalent:*

1. *X is weak mixing.*

2. $\frac{1}{N}\sum_{n=0}^{N-1}\left|\int f\cdot T^n g d\mu - \int f d\mu \int g d\mu\right| \to 0$ *for all* $f,g\in L^2(\mu)$.

3. $\frac{1}{N}\sum_{n=0}^{N-1}|\mu(A\cap T^{-n}B)-\mu(A)\mu(B)|\to 0$ *for every* $A,B\in\mathcal{B}$.

4. $\mu(A\cap T^{-n}B)\xrightarrow{D}\mu(A)\mu(B)$ *for every* $A,B\in\mathcal{B}$.

Proof. Since $|\mu(A\cap T^{-n}B)-\mu(A)\mu(B)|\leq 1$, the equivalence of (3) and (4) is Lemma 8.3.3. The equivalence of (2) and (3) is standard be approximating L^2 functions by simple functions and using Cauchy-Schwartz. So we have to prove that (1) $\iff$ (2).

We may suppose that X is ergodic, since otherwise (1) fails trivially and (2) fails already without absolute values by the lemma. Then for $f,g\in L^\infty(\mu)$ we know from the lemma that

$$\frac{1}{N}\sum_{n=0}^{N-1}\int f\cdot T^n g\to\int f\int g \tag{8.1}$$

Suppose that X is weak mixing. Let $f'=f(x)f(y)\in L^2(\mu\times\mu)$ and define $g'\in L^2(\mu\times\mu)$ similarly. By ergodicity of $X\times X$

$$\frac{1}{N}\sum_{n=0}^{N-1}\int f'\cdot(T\times T)^n g' d\mu\times\mu\to\int f'\int g'$$

but

$$\begin{aligned}\int f'(T\times T)^n g' d\mu\times\mu &= \int\int f(x)f(y)g(T^n x)g(T^n y)\,d\mu(x)d\mu(y)\\ &= (\int f\cdot T^n g d\mu)^2\end{aligned}$$

and

$$\int f' d\mu\times\mu\cdot\int g' d\mu\times\mu = (\int f d\mu)^2(\int g d\mu)^2$$

Thus we have proved:

$$\frac{1}{N}\sum_{n=0}^{N-1}(\int f\cdot T^n g d\mu)^2\to(\int f\int g)^2 \tag{8.2}$$

Combining this with $\frac{1}{N}\sum_{n=0}^{N-1}\int f\cdot T^n g\to\int f\int g$, we find that

$$\frac{1}{N}\sum_{n=0}^{N-1}(\int f\cdot T^n g d\mu-(\int f\int g))^2\to 0 \tag{8.3}$$

and since the terms are bounded, this implies (2).

In the opposite direction, assume (4), which is equivalent to (2). We must prove that $X \times X$ is ergodic, or equivalently, that for every $F, G \in L^2(\mu \times \nu)$,

$$\frac{1}{N}\sum_{n=0}^{N-1} \int F \cdot T^n G \, d\mu \times \nu \to \int F d\mu \times \nu \cdot \int G d\mu \times \nu$$

By approximation it is enough to prove this when $F(x_1, x_2) = f_1(x_1) f_2(x_2)$ and $G(x_1, x_2) = g_1(x_1) g_2(x_2)$. Furthermore we may assume that f_1, f_2, g_1, g_2 are simple, and even indicator functions $1_A, 1_{A'}, 1_B, 1_{B'}$. Thus we want to prove that for A, A', B, B',

$$\frac{1}{N}\sum_{n=0}^{N-1} \mu(A \cap T^{-n}B)\mu(A \cap T^{-n}B') \to \mu(A)\mu(B)\mu(A')\mu(B')$$

But $\mu(A \cap T^{-n}B) \to \mu(A)\mu(B)$ in density and $\mu(A' \cap T^{-n}B') \to \mu(A')\mu(B')$ in density, so the same is true for their product; and hence the averages converge as desired. □

8.4 A multiplier property

Weak mixing was defined by the property of $X \times X$, but it turns out that it can be characterized in terms of the behavior of products of X with arbitrary ergodic systems.

Proposition 8.4.1. *$(X, \mathcal{B}, \mu, T)$ is weak mixing if and only if $X \times Y$ is ergodic for every ergodic system $(Y, \mathcal{C}, \nu, S)$.*

Proof. One direction is trivial: if $X \times Y$ is ergodic whenever Y is ergodic then this is true in particular for the 1-point system. Then $X \times Y \cong X$ so X is ergodic. It then follows taking $Y = X$ that $X \times X$ is ergodic, so X is weak mixing.

In the opposite direction we must prove that for every $F, G \in L^2(\mu \times \nu)$,

$$\frac{1}{N}\sum_{n=0}^{N-1} \int F \cdot (T \times S)^n G \, d\mu \times \nu \to \int F d\mu \times \nu \int G d\mu \times \nu$$

As before it is enough to prove this when $F(x_1, x_2) = f_1(x_1) f_2(x_2)$ and $G(x_1, x_2) = g_1(x_1) g_2(x_2)$, and it reduces to

$$\frac{1}{N}\sum_{n=0}^{N-1} \int f_1(x) T^n f_2(x) d\mu(x) \cdot \int g_2(x) S^n g_2(x) d\nu(x) \to \int f_1 d\mu \int f_2 d\mu \int g_1 d\nu \int g_2 d\nu$$

Splitting $L^2(\mu)$ into constant functions and their orthogonal complement (functions of integral 0), it is enough to prove this for f_1 in each of these spaces. If

f_1 is constant then $\int f_1(x)T^n f_2(x)d\mu(x) = \int f_1 d\mu \int f_2 d\nu$ and the claim follows from ergodicity of S. On the other hand if $\int f_1 d\mu = 0$ we have

$$\left(\frac{1}{N}\sum_{n=0}^{N-1}\int f_1(x)T^n f_2(x)d\mu(x)\cdot\int g_2(x)S^n g_2(x)d\nu(x)\right)^2$$

$$\leq \frac{1}{N}\sum_{n=0}^{N-1}\left(\int f_1(x)T^n f_2(x)d\mu(x)\right)^2\cdot\frac{1}{N}\sum_{n=0}^{N-1}\left(\int g_2(x)S^n g_2(x)d\nu(x)\right)^2$$

but

$$\frac{1}{N}\sum_{n=0}^{N-1}\left(\int f_1(x)T^n f_2(x)d\mu(x)\right)^2 =\leq \frac{1}{N}\sum_{n=0}^{N-1}\left(\int f_1(x)T^n f_2(x)d\mu(x) - \int f_1\int f_2\right)^2 \to 0$$

by weak mixing of X and we are done. □

Corollary 8.4.2. *If X is weak mixing so is $X\times X$ and $X\times X\times\ldots\times X$.*

Proof. For any ergodic Y, $(X\times X)\times Y = X\times(X\times Y)$. Since $X\times Y$ is ergodic so is $X\times(X\times Y)$. The general claim follows in the same way. □

More generally,

Corollary 8.4.3. *If $X_1, X_2, \ldots$ are weak mixing so are $X_1\times X_2\times\ldots$.*

Also,

Corollary 8.4.4. *If $(X,\mathcal{B},\mu,T)$ is weak mixing then so is T^n for all $n\in\mathbb{N}$ (if T is invertible, also negative n).*

Proof. Since $T\times T$ is ergodic if and only if $T^{-1}\times T^{-1}$ is ergodic, weak mixing of T and T^{-1} are equivalent, so we only need to consider $n>0$.

First we show that T weak mixing implies that T^m is ergodic. Otherwise, let $f\in L^2$ be a T^m invariant and non-constant function. Consider the system $Y=\{0,\ldots,m-1\}$ and $S(y) = y+1 \bmod m$ with uniform measure. Let $F(x,i) = f(T^{m-i}x)$. Then

$$F(Tx, Si) = f(T^{m+1-(i+1)}x) = f(T^{m-i}x) = F(x,i)$$

so F is $T\times S$ invariant and non-constant. Hence $T\times S$ is not ergodic; but this contradicts the weak mixing property of T.

Applying the argument above to $T\times T$, which is itself weak mixing, we find that $(T\times T)^m$ is ergodic, equivalently $T^m\times T^m$ is ergodic, so T^m is weak mixing. □

8.5 (*) Spectral measures

Our characterization of weak mixing is, in the end, purely a Hilbert-space statement. Thus one should be able to prove the existence of eigenfunctions without use of the underlying dynamical system. This can be done with the help of the spectral theorem. Let us first give a brief review of the version we will use.

Let us begin with an example of a unitary operator. Let μ be a probability measure on the circle S^1 and let $M : L^2(\mu) \to L^2(\mu)$ be given by $(Mf)(z) = zf(z)$. Note that U preserves norm, since $|zf(z)| = |f(z)|$ for $z \in S^1$ and hence μ-a.e. z; it is invertible since the inverse is given by multiplication by $\overline{z}$.

The spectral theorem says that any unitary operator can be represented in this way on any invariant subspace for which it has an cyclic vector.

Theorem 8.5.1 (Spectral theorem for unitary operators). *Let $U : H \to H$ be a unitary operator and $v \in H$ a unit vector such that $\overline{\{U^n v\}_{n=-\infty}^{\infty}} = H$. Then there is a probability measure $\mu_v \in \mathcal{P}(S^1)$ and a unitary operator $V : L^2(\mu) \to H$ such that $U = VMV^{-1}$, where $M : L^2(\mu) \to L^2(\mu)$ is as above. Furthermore $V(1) = v$.*

We give the main idea of the proof. The measure μ_v is characterized by the statement because its Fourier transform $\widehat{\mu}_v : \mathbb{Z} \to \mathbb{R}$ is given by

$$\widehat{\mu}_v(n) = \int z^n \, d\mu_v = \langle M^n 1, 1\rangle_{L^2(\mu_v)} = \langle U^n v, v\rangle_H$$

Reversing this argument, in order to construct μ_v one starts with the sequence $a_n = \langle U^n 1, 1\rangle_H$. This sequence is positive definite in the sense that for any sequence $\lambda_i \in \mathbb{C}$ and any n, $\sum_{i,j=1}^{n} \lambda_i \overline{\lambda_j} a_{i-j} \geq 0$:

$$\begin{aligned}
\sum_{i,j=1}^{\infty} \lambda_i \overline{\lambda_j} a_{i-j} &= \sum_{i,j=-n}^{n} \lambda_i \overline{\lambda_j} \left\langle U^{i-j} v, v\right\rangle_H \\
&= \sum_{i,j=-n}^{n} \left\langle U^i \lambda_i v, U^j \lambda_j v\right\rangle_H \\
&= \left\langle \sum_{i=-n}^{n} U^i \lambda_i v, \sum_{j=-n}^{n} U^j \lambda_j v \right\rangle_H \\
&= \left\| \sum_{i=-n}^{n} U^i \lambda_i v \right\|_2^2 \\
&\geq 0
\end{aligned}$$

Therefore, by a theorem of Hergolz (also Bochner) a_n is the Fourier transform of a probability measure on S^1 (note that $a_0 == \|v\|^2 = 1$).

One first defines V on complex polynomials $p(z) = \sum_{n=0}^{d} b_n z^n$ by $Vp = \sum_{n=0}^{d} b_n U^n v$. One can check that this preserves inner products; it suffices to check for monomials, and indeed

$$\langle V(bz^m), V(cz^n)\rangle = b\overline{c}\,\langle U^m v, U^n v\rangle = b\overline{c}\cdot a_{m-n} = b\overline{c}\int z^{m-n}d\mu_v = \int (bz^m)(\overline{cz^n})d\mu_v = \langle bz^m, bz^n\rangle_{L^2(\mu_v)}$$

Since polynomials are dense in $L^2(\mu)$ it remains to extend V to measurable functions. The technical details of carrying this out can be found in many textbooks.

Lemma 8.5.2. *Let $U : H \to H$ be unitary and v a cyclic unit vector for U with spectral measure μ. Then $\alpha \in \Sigma(U)$ if and only if $\mu_v(\alpha) > 0$.*

Proof. If α is an atom of μ_v let $f = 1_{\{\alpha\}}$. This is a non-zero vector in $L^2(\mu_v)$, and $Mf(z) = zf(z) = \alpha f(z)$. Hence $\alpha \in \Sigma(M)$ and by the spectral theorem $\alpha \in \Sigma(U)$.

Conversely, suppose that $\mu_v(\{\alpha\}) = 0$. Consider the operator $U_\alpha(w) = \overline{\alpha}Uw$, which can easily be seen to be unitary. Clearly w is an eigenfunction with eigenvalue α if and only if $U_\alpha w = w$. Thus it suffices for us to prove that $\frac{1}{N}\sum_{n=0}^{N-1} U_\alpha^n w \to 0$ for all w, and, since v is cyclic and the averaging operator is linear and continuous, it is enough to check this for v. Transferring the problem to (S^1, μ_v, M), we must show that $\frac{1}{N}\sum_{n=0}^{N-1} \overline{\alpha}^n z^n \to 0$ in $L^2(\mu_v)$. We have

$$\frac{1}{N}\sum_{n=0}^{N-1} \overline{\alpha}^n z^n = \frac{1}{N}\frac{(\overline{\alpha}z)^N - 1}{\overline{\alpha}z - 1}$$

This converges to 0 at every point $z \neq \alpha$, hence μ_v-a.e., and it is bounded. Hence by bounded convergence, it tends to 0 in $L^2(\mu_v)$, as required. □

Proposition 8.5.3. *Let U, H, v, μ_v be as above. If μ_v is continuous (has no atoms), then*

$$\frac{1}{N}\sum_{n=0}^{N-1} |(w, U^n w')| \to 0$$

for every $w, w' \in H_v$.

Proof. Using the fact that w, w' can be approximated in L^2 by linear combinations of $\{U^n v\}$, it is enough to prove this for $w, w' \in \{U^n v\}$. Since the statement we are trying to prove is formally unchanged if we replace w by $U^{\pm 1}w$ or w' by $U^{\pm 1}w'$, we may assume that $w = w' = v$. Also, we may square the summand, as we have seen this does not affect the convergence to 0 of the averages. Passing to the spectral setting, we have

$$\begin{aligned}
\frac{1}{N}\sum_{n=0}^{N-1}|(v,U^n v)|^2 &= \frac{1}{N}\sum_{n=0}^{N-1}\left|\int z^n d\mu_v\right|^2 \\
&= \frac{1}{N}\sum_{n=0}^{N-1}\left(\int z^n d\mu_v \cdot \int \overline{z}^n d\mu_v\right) \\
&= \frac{1}{N}\sum_{n=0}^{N-1}\left(\int\int z^n \overline{y}^n d\mu_v(y) d\mu_v(z)\right) \\
&= \int\int \left(\frac{1}{N}\cdot\frac{(z\overline{y})^N - 1}{z\overline{y}-1}\right) d\mu_v \times \mu_v(z,y)
\end{aligned}$$

The integrand is bounded by 1 and tends pointwise to 0 off the diagonal $\{y = z\}$, which has $\mu_v \times \mu_v$-measure 0, since μ_v is non-atomic. Therefore by bounded convergence, the expression tends to 0. $\square$

Corollary 8.5.4. *If $(X, \mathcal{B}, \mu, T)$ is ergodic then it is weak mixing if and only if μ_f is continuous (has no atoms) for every $f \perp 1$ (equivalently, the maximal spectral type is non-atomic except for an atom at 1), if and only if $\Sigma(T) = \{1\}$.*

Proof. Suppose for $f \perp 1$ the spectral measure μ_f is continuous. By the last proposition, $\frac{1}{N}\sum_{n=0}^{N-1}|\int f \cdot T^n f\, d\mu| \to 0$. For general f, g we can write $f = f' + \int f$, $g = g' + \int g$, where $f', g' \perp 1$. Substituting $f = f' + \int f$ into

$$\frac{1}{N}\sum_{n=0}^{N-1}|\int f\cdot T^n f\, d\mu - (\int f\, d\mu)(\int g d\mu)|$$

we obtain the expression

$$\frac{1}{N}\sum_{n=0}^{N-1}|\int f'\cdot T^n g'\, d\mu|$$

which by assumption $\to 0$. This was one of our characterizations of weak mixing.

Conversely suppose T is weak mixing. Then it has no eigenfunctions except 1 (this was the trivial direction of the eigenfunction characterization), so if $f \perp 1$ also $\overline{\text{span}\{U^n f\}} \perp 1$ and so, since on this subspace T has no eigenfunctions, μ_f is continuous.

We already know that weak mixing implies $\Sigma(T) = \{1\}$. In the other direction, if T is not weak mixing, we just saw that there is some $f \perp 1$ with $\mu_f(\alpha) > 0$ for some α, and by the previous lemma, $\alpha \in \Sigma(T)$. $\square$

In a certain sense, we can now "split" the dynamics of a non-weak-mixing system into an isometric part, and a weak mixing part:

Corollary 8.5.5. *Let (X, B,μ, T) be ergodic. Then $L^2(\mu) = U \oplus V$, where $U = L^2(\mu, \mathcal{E})$ for $\mathcal{E} \subseteq \mathcal{B}$ the Kronecker factor, and V is an invariant subspace such that $T|_V$ is a weak-mixing in the sense that $\frac{1}{N}\sum_{n=0}^{N-1} |\int f \cdot T^n g \, d\mu| \to 0$ for $g \in V$.*

One should note that, in general, the subspace V in the corollary does not correspond to a factor in the dynamical sense.

An important consequence is the following:

Theorem 8.5.6. *Let $(X, \mathcal{B}, \mu, T)$ and $(Y, \mathcal{C}, \nu, S)$ be ergodic measure preserving systems. Then $X \times Y$ is ergodic if and only if $\Sigma(T) \cap \Sigma(S) = \{1\}$.*

Proof. Let $Z = X \times Y$, $R = T \times S$, $\theta = \mu \times \nu$. If $\alpha \neq 1$ is a common eigenvalue of T, S with eigenfunctions f, g, then $h(x, y) = \overline{g}(y) \cdot f(x)$ is a non-trivial invariant function, since

$$h(R(z, y)) = \overline{g}(Sy) \cdot f(Tx) = \overline{\alpha g}(y)\alpha f(x) = h(x, y)$$

and so Z is not ergodic.

Conversely, write $L^2(\mu) = V_{wm} \oplus V_{pp}$, where $T|_{V_{pp}}$ as in the previous corollary, where $T|_{V_{pp}}$ has no eigenvalues, and decompose $L^2(\nu) = W_{wm} \oplus W_{pp}$ similarly. We must show that

$$\frac{1}{N}\sum_{n=0}^{N-1} \int h \cdot R^n h d\theta \to (\int h)^2$$

for every $h \in L^2(\theta)$ and it suffices to check this for $h = fg$, $f \in L^2(\mu)$, $g \in L^2(\nu)$, since the span of these is dense in L^2. Then $\int hR^n hd\theta = \int fT^n f d\mu \cdot \int gS^n g d\nu$. Also, since we can write $f = f_{wm} + f_{pp}$ and $g = g_{wm} + g_{pp}$ for $f_{wm} \in V_{wm}$ etc. we can expand the expression above, and obtain a sum of terms of the form

$$\frac{1}{N}\sum_{n=0}^{N-1} \int h \cdot R^n h d\theta \to (\int h)^2 = \sum_{i,j,s,t \in \{wm,pp\}} \left(\frac{1}{N}\sum_{n=0}^{N-1} (\int f_i T^n f_j d\mu)(\int g_s S^n g_t d\nu) \right) \tag{8.4}$$

Consider the terms in parentheses; they are all bounded independently of n. So if $i, j = wm$ we can bound

$$\left| \frac{1}{N}\sum_{n=0}^{N-1} (\int f_i T^n f_j d\mu)(\int g_s S^n g_t d\nu) \right| \leq C \cdot \frac{1}{N}\sum_{n=0}^{N-1} |\int f_i T^n f_j d\mu| \to 0$$

and similarly if $s, t = wm$. Also if $i = wm$, $s = pp$, then $T^n f_j = \alpha^n f_j$ for some α, and $\overline{f}_j \perp f_i$. Hence

$$\int f_i T^n f_j d\mu = \alpha^n \int f_i f_j d\mu = \alpha \left\langle f_i, \overline{f}_j \right\rangle = 0$$

Thus in 8.4 we only need to consider the case $i, j, s, t = pp$. In this case we can expand each of the functions as a series in normalized, distinct eigenfunctions:

$f_{pp} = \sum \varphi_k$ and $g_{pp} = \sum \psi_k$ where $T\varphi_k = \alpha_k \varphi_k$ and $S\psi_k = \beta_k \psi_k$. We assume $\varphi_0 = const$ and $\psi_0 = const$. Expanding again using linearity, we must consider terms of the form

$$\frac{1}{N}\sum_{n=0}^{N-1}(\int \varphi_i T^n \varphi_j d\mu)(\int \psi_s S^n \psi_t d\nu) = \frac{1}{N}\sum_{n=0}^{N-1}(\alpha_j^n \int \varphi_i \varphi_j d\mu)(\beta_t^n \int \psi_s \psi_t d\nu)$$

Now, the first integral is 0 unless $\varphi_j = \overline{\varphi}_i$ and the second is 0 unless $\psi_t = \overline{\psi}_s$. If this is the case we have, writing $c_{i,s} = \|\varphi_j\|^2 \|\psi_j\|^2$

$$= \frac{1}{N}\sum_{n=0}^{N-1} \alpha_j^n \beta_t^n c_{i,s} = \begin{cases} c_{i,s} & \alpha_j = \overline{\beta}_t \\ c_{i,s}\frac{1}{N}\frac{(\alpha\beta)^N - 1}{\beta - 1} & \text{otherwise} \end{cases} \xrightarrow[N\to\infty]{} \begin{cases} c_{i,s} & \alpha_j = \overline{\beta}_t \\ 0 & \text{otherwise} \end{cases}$$

Since $\Sigma(T) \cap \Sigma(S) = \{1\}$ the limit is thus 0 except for $i = j = s = t = 0$. In the latter case, $c_{0,0} = \int \varphi_0^2 d\mu \int \psi_0^2 d\nu = (\int f)^2 (\int g)^2$, so this was the limit we wanted. □

Chapter 9

Shannon Entropy

Intuitively, a fair dice is "more random" than a fair coin (it has 6 equally likely outcomes, versus 2 for the coin), which in turn is "more random" than a biased coin (which has two outcomes but they are not equally likely). Entropy, which will occupy us in one form or another for the next few chapters, is a numerical quantity associated to a random variable, quantifying the amount of randomness inherrent in it. In this chapter we define entropy and establish some of its basic properties, as well as some applications.

9.1 Motivation: Optimal compression

Claude Shannon introduction entropy in his landmark 1948 paper on information theory, where it appears naturally in defining the complexity of a signal and the capacity of a transmission channel. To motivate the definition we follow a similar course, namely, we examine the problem of source coding. In this problem we are given a "signal", modeled by a random variable whose outcome is in some finite set (which we denote usually by Σ), and we want to represent the signal a finite binary sequence (or more generally, a sequence over some other alphabet), in such a way as to minimize the average length of the encoding word, and to be able to recover the original signal. We begin with some basic notation.

Definition 9.1.1. Let Σ be a finite set of symbols.

A **word** of length n over Σ is a finite sequence $a = a_1 \dots a_n \in \Sigma^n$, whose length we denote by $|a| = n$.

The set of all finite words (including the empty word) is denoted $\Sigma^* = \bigcup_{n=0}^{\infty} \Sigma^n$.

The **subword** of $\sigma \in \Sigma^n$ from index i to index $j > i$ is the word $\sigma_i^j = \sigma_i \sigma_{i+1} \dots \sigma_j$.

If a, b are words then ab is their concatenation.

Definition 9.1.2. A **code** on a finite set Σ is a map $c : \Sigma \to \{0,1\}^*$. We write $c^* : \Sigma^* \to \{0,1\}^*$ for the repeated applicaiton of c, given by $c^*(\sigma_1, \ldots, \sigma_n) = c(\sigma_1)\ldots c(\sigma_n)$ (we usually drop the superscript and just write $c(\sigma_1 \ldots \sigma_n)$).

One could just as well consider conding using other alphabets than $\{0,1\}$ but there is little reason to do so. The main consequence of the choice of a 2-symbol alphabet will be that we shall take our logarithms to be in base 2.

It is obviously desirable that a code be injective, i.e. for $c(\sigma)$ to determine σ. But if one is also interested in repeated coding of sequences of elements of Σ, one should ask slightly more:

Definition 9.1.3. A code c is **uniquely decodable** if c^* is inective.

Injectivity of c is not enough to ensure that it is uniquely decodable. For example if $\Sigma = \{\sigma_1, \sigma_2\}$ and if $c(\sigma_1) = 0$ and $c(\sigma_2) = 00$, then c is injective, but 00 could be the image of either σ_2 or of $\sigma_1\sigma_1$.

There are various ways to construct uniquely decodable codes. For example we can "mark" the end of each codeword with some sequence which appears nowhere else. But there is cleaner way to do it.

Definition 9.1.4. A code c is a **Prefix code** if $\sigma \neq \sigma'$ implies that $c(\sigma)$ is not a prefix of $c(\sigma')$ or vice versa.

Lemma 9.1.5. *Prefix codes are uniquely decodable.*

Proof. If $u = c(\sigma_1)\ldots c(\sigma_m) = c(\sigma'_1)\ldots c(\sigma'_{m'})$, suppose that for some i we have $c(\sigma_i) \neq c(\sigma'_i)$. Deleting the prefix $c^*(\sigma_1 \ldots \sigma_i) = c^*(\sigma'_1 \ldots \sigma'_i)$ from u, we get $c(\sigma_i)c(\sigma_{i+1})\ldots c(\sigma_m) = c(\sigma'_i)c(\sigma'_{i+1})\ldots c(\sigma'_{m'})$. If $|c(\sigma_i)| = |c(\sigma'_i)|$ then both $c(\sigma_i)$ and $c(\sigma'_i)$ would constitute the same prefix of u, and would be equal, contrary to assumption. But if $|c(\sigma_i)| > |c(\sigma'_i)|$ then σ'_i is a prefix of σ_i, and vice versa, so the code is not a prefix code. □

The following result gives a condition that ensures that a prefix code can be constructed using words of given lengths ℓ_i. It also shows that this is the same condition that ensures existence of a uniquely decodable code, so using prefix codes is optimal, at least in terms of the lengths of the code words.

Theorem 9.1.6. *Generalized Kraft inequality Let $\ell_1, \ldots, \ell_n \geq 1$. Then the following are equivalent:*

1. $\sum 2^{-\ell_i} \leq 1$.
2. There is a prefix code with lengths ℓ_i.
3. There is a uniquely decodable code with lengths ℓ_i.

Proof. (2) $\implies$ (3) was the previous lemma.

(1) $\implies$ (2): Let $L = \max \ell_i$ and order $\ell_1 \leq \ell_2 \leq \ldots \leq \ell_n = L$.

It is useful to identify $\bigcup_{i \leq L}\{0,1\}^i$ with the full binary tree of height L, so each vertex has two children, one connected to the vertex by an edge marked

0 and the other by an edge marked 1. Each vertex v in the tree is identified with the labels along the edges of the path connecting the root to v; the root corresponds to the empty word and the leaves (at distance L from the root) correspond to words of length L.

We define codewords $c(i) = a^i$ by induction. Assume we have defined $a^i, i < k$ with $|a^i| = \ell_i$. Let

$$A_i = \{a^i b \,:\, b \in \{0,1\}^{L-\ell_i}\}$$

Thus A_i is the set of leaves descended from a^i, or equivalently, the set of words of length L of which a^i is a prefix. We have $|A_i| = 2^{L-\ell_i}$. The total number of leaves descended form $a^1, \ldots, a^{k-1}$ is

$$|\bigcup_{i<k} A_i| \leq \sum_{i<k} |A_i| = \sum_{i<k} 2^{L-\ell_i} < 2^L$$

The strict inequality is because $\sum 2^{-\ell_i} \leq 1$, and the sum above includes at least one term less than the full sum.

Let $a \in \Sigma^L \setminus \bigcup A_i$ and $a^k = a_1 \ldots a_{\ell_k}$ the length-ℓ_k prefix of a. For $i < k$, a^i if a^i is a prefix of a^k then, since $\ell_i \leq \ell_k$, a^k is a child of a^i and so $a \in A_i$, a contradiction. If a^k is a prefix of a^i then since $\ell_i \leq \ell_k$ we have $a^i = a^k$ and the same arrive at the same contradiction. Therefore $a^1, \ldots, a^k$ is a prefix code.

(3) $\implies$ (1): Suppose c is uniquely decodable. Fix m. Then

$$(\sum 2^{-\ell_i})^m = \sum_{(i_1 \ldots i_m) \in \Sigma^m} 2^{-\sum_{j=1}^m \ell_{i_j}} = \sum_{(i_1, \ldots, i_m) \in \Sigma^m} 2^{-|c(i_1, \ldots, i_m)|}$$

divide the codewords according to length:

$$= \sum_{\ell=1}^{Lm} \sum_{\sigma \in \Sigma^{\leq m} \,:\, c(\sigma) = \ell} 2^{-\ell} \leq \sum_{\ell=1}^{Lm} 2^{-\ell} 2^{\ell} = Lm$$

taking m-th roots and $m \to \infty$, this gives (1) . □

Definition 9.1.7. Let ξ be a random variable with values in Σ (i.e., $\mathbb{P}(\xi \in \Sigma) = 1$), and distribution $p \in \mathcal{P}(\Sigma)$. The **mean coding length** of ξ w.r.t. a code c is $\mathbb{E}|c(\xi)| = \sum_{\sigma \in \Sigma} p(\sigma)|c(\sigma)|$.

Now suppose we are given a finie set Σ and random variable ξ on Σ with distribution $p \in \mathcal{P}(\Sigma)$, and we want to find a uniquely decodable code of optimal average coding length. By the previous theorem, defining such a code c amounts to determining the lenghs $\ell_\sigma = |c(\sigma)|$ so that they satisfy $\sum 2^{-\ell_\sigma} \leq 1$. Thus, our problem is:

> **Problem:** Find $\{\ell_\sigma\}_{\sigma \in \Sigma} \in \mathbb{N}^\Sigma$ which minimize $\sum p_\sigma \ell_\sigma$ subject to $\sum 2^{-\ell_\sigma} \leq 1$.

The first way to solve this problem is via lagrange multipliers. Replace the integer variable ℓ_σ by continuous real x_σ. By adding a "dummy variables" corresponding to signals which occur with probability zero, we can replace the condition $\sum 2^{-x_\sigma} \leq 1$ with $\sum 2^{-x_\sigma} = 1$. We then have the lagrangian

$$L(x, \lambda) = \sum p_\sigma x_\sigma - \lambda \left(\sum 2^{-x_\sigma} - 1 \right)$$

Differentiating by x_σ and by λ and setting to zero, we get

$$\begin{aligned} p_\sigma - \lambda 2^{-x_\sigma} \ln 2 &= 0 \\ \sum 2^{-x_\sigma} &= 1 \end{aligned}$$

Summing the first equation over σ, and using $\sum p_\sigma = 1$ and the second equation, we find

$$\lambda = 1/\ln 2$$

so

$$x_\sigma = -\log p_\sigma$$

and the "expected coding length" at the critical point is

$$-\sum p_\sigma \log p_\sigma$$

This motivates the following definition:

Definition 9.1.8. The (Shannon) **entropy** of a probability vector $p = (p_1, \dots, p_n)$ is

$$H(p_1, \dots, p_n) = -\sum p_i \log p_i$$

The entropy of a random variable ξ on a finite set Σ is the distribution of the probability vector $p_\sigma = \mathbb{P}(\xi = \sigma)$, i.e.

$$H(\xi) = -\sum_{\sigma \in \Sigma} p_\sigma \log p_\sigma$$

By convention the logarithm is taken in base 2 and $0 \log 0 = 0$. Infinite vectors (or variables with infinite range) can have infinite entropy but we will not discus these here.

Shannon's fundamental result on source coding is the following:

Theorem 9.1.9. *If c is a uniquely decodable code for a discrete random variable ξ, then the expected coding length is $\geq H(\xi)$, and equality is achieved if and only if $p_\sigma = 2^{-\ell_\sigma}$ where $\ell_\sigma = |c(\sigma)|$. Furthermore, there exists a prefix code code for ξ with expected coding length $H(\xi) + 1$.*

Proof. We have essentially proved the first statement in the calculus exercise above, which showed how the expression $-\sum p_i \log p_i$ naturally comes about. Now that we have identified it wecan give another, more conceptual proof. Suppose that $\sum 2^{-\ell_\sigma} = 1$ (as explained before it we have $\leq$ we can obtain equality by adding lengths associated to probability zero outcomes, which does not change the entropy). Set $q_\sigma = \log 1/\ell_\sigma$, so $\sum q_\sigma = 1$. The function $\log t$ is concave ($\log'' t = -1/t^2 < 0$), so

$$\begin{aligned} H(\xi) - \sum p_\sigma \ell_\sigma &= -\sum p_\sigma \log p_\sigma + p_\sigma \log q_\sigma \\ &= \sum p_\sigma \log \frac{q_\sigma}{p_\sigma} \\ &\leq \log \sum p_\sigma \frac{q_\sigma}{p_\sigma} \\ &= \log 1 \\ &= 0 \end{aligned}$$

which shows that the mean coding length is never less than the entropy.

Finally, to build an almost-optimal code, choose a prefix code with $|c(\sigma)| = \ell_\sigma = \lceil -\log p_\sigma \rceil$, which is possible since

$$\sum 2^{-\ell_\sigma} \leq \sum 2^{-\log p_\sigma} = \sum p_\sigma = 1$$

Since $\ell_\sigma \leq -\log p_\sigma + 1$, the expected coding length is

$$\sum p_\sigma \ell_\sigma \leq H(p) + 1$$

□

We have thus shown that, **up to one extra bit, the optimal coding rate $H(\xi)$ can be achieved.**

9.2 Shannon entropy

In the previous section we discussed finite-valued random variables. We shift our perspective now, temporarily, to study the entropy of partitions. Let us briefly explain the connection. A random variable ξ defined on a probability space $(\Omega, \mathcal{F}, P)$, and taking values in a finite set A, gives rise to a partition of Ω into level sets (which are measurable, since ξ is):

$$\mathcal{A}_\xi = \{\xi^{-1}(a) \,:\, a \in A\}.$$

The correpsondence between ξ and $\mathcal{A}_\xi$ is 1-1 if we regard $\mathcal{A}_\xi$ as a labeled collection of sets with $\xi^{-1}(a)$ bearing the label a. Otherwise, $\mathcal{A}_\xi$ does not determine the values of ξ, but only its level sets (for example, any strictly increasing function on $\mathbb{R}$ defines the partition into points, so many functions define the same partition). But it is easy to see that if $\mathcal{A}_\zeta = \mathcal{A}_\xi$ then ζ is a.s. a function of ξ and vice versa, so the "distribution" of a random variable, if we do not care about its values, is captured entirely by the masses the measure assigns to the atoms of the partition; in other words, $\mathcal{A}_\xi$ determines the probability vector $p_\xi = (\mathbb{P}(\xi = a))_{a\in A}$ up to permutation of coordinates.

Definition 9.2.1. The entropy of a probability measure μ with respect to a partition $\mathcal{A}$ is the non-negative number

$$H(\mu, A) = -\sum_{A \in \mathcal{A}} \mu(A) \log \mu(A)$$

This is just the entropy of the associated probability vector $(\mu(A))_{A \in \mathcal{A}}$, and if ξ is a random variable then $H(\xi) = H_\mu(\mathcal{A}_\xi)$.

Example 9.2.2. For $\underline{p} = (p, 1-p)$ the entropy $H(p) = -p \log p - (1-p)\log(1-p)$ depends on the single variable p. It is an exercise in calculus to verify that $h(\cdot)$ is strictly concave on $[0,1]$, increasing on $[0,1/2]$ and decreasing on $[1/2,1]$, with a unique maximum value $h(1/2) = 1$ and minimal values $h(0) = h(1) = 0$. Thus, the entropy is minimal when all the mass is on one atom of $\mathcal{A}$, and maximal when it is uniformly distributed.

Properties of entropy

(E1) $0 \le H(\mu, \mathcal{A}) \le \log|\mathcal{A}|$, and furthermore

- (a) $H(\mu, \mathcal{A}) = 0$ if and only if $\mu(A) = 1$ for some $A \in \mathcal{A}$.
- (b) $H(\mu, \mathcal{A}) = \log|\mathcal{A}|$ if and only if μ is uniform on $\mathcal{A}$, that is, $\mu(A) = 1/|\mathcal{A}|$ for $A \in \mathcal{A}$.

(E2) $H(\cdot, \mathcal{A})$ is concave: for probability measures μ, ν on and $0 < \alpha < 1$,

$$H(\alpha\mu + (1-\alpha)\nu, \mathcal{A}) \ge \alpha H(\mu, \mathcal{A}) + (1-\alpha) H(\nu, \mathcal{A})$$

with equality if and only if $\mu(A) = \nu(A)$ for all $A \in \mathcal{A}$.

Proof. We first prove (E2). The function $f(t) = -t \log t$ is strictly concave (because $f''(t) = 1/t^2$), so by Jensen's inequality,

$$\begin{aligned} H(\alpha\mu + (1-\alpha)\nu, \mathcal{A}) &= \sum_{A \in \mathcal{A}} f(\alpha\mu(A) + (1-\alpha)\nu(A)) \\ &\ge \sum_{A \in \mathcal{A}} (\alpha f(\mu(A)) + (1-\alpha) f(\nu(A))) \\ &= \alpha H(\mu, \mathcal{A}) + (1-\alpha) H(\nu, \mathcal{A}) \end{aligned}$$

with equality if and only if $\mu(A) = \nu(A)$ for all $A \in \mathcal{A}$.

The left inequality of (E1) is trivial. For the right one consider the function $H(\underline{p}) = -\sum_{A \in \mathcal{A}} p_A \log p_A$ on the simplex Δ of probability vectors $\underline{p} = (p_A)_{A \in \mathcal{A}}$. By (E2) this function is strictly concave, so its minima are attainged at its extreme points; these are the vectors $\underline{p}$ which are concentrated on one coordinate, and there the entropy is zero. It remains to show that the (unique) maximum is attained at $\underline{p}^* = (1/|\mathcal{A}|, \ldots, 1/|\mathcal{A}|)$, since $H(\underline{p}^*) = \log|\mathcal{A}|$. Existence and uniqueness of the maximal point $\underline{p}^*$ follows because Δ is compact and convex and $H(\cdot)$ is strictly concave. Since $H(\cdot)$ is invariant under permutation of variables, the maximizing point $\underline{p}^*$ must also be invariant under coordinate permutations, and hence all its coordinates are equal. Since it is a probability vector they are are equal to $1/|\mathcal{A}|$. $\square$

For a set B of positive measure, let μ_B denote the conditional probability measure $\mu_B(C) = \mu(B \cap C)/\mu(B)$. Note that for a partition $\mathcal{B}$ we have the identity

$$\mu = \sum_{B \in \mathcal{B}} \mu(B) \cdot \mu_B \tag{9.1}$$

Definition 9.2.3. The **conditional entropy** of μ and $\mathcal{A}$ given another partition $\mathcal{B} = \{B_i\}$ is defined by

$$H(\mu, \mathcal{A}|\mathcal{B}) = \sum_{B \in \mathcal{B}} \mu(B) H(\mu_B, \mathcal{A})$$

This is just the average over $B \in \mathcal{B}$ of the entropy of $\mathcal{A}$ with respect to the conditional measure on B.

Definition 9.2.4. Let $\mathcal{A}, \mathcal{B}$ be partitions of the same space.

1. The *join* of $\mathcal{A}, \mathcal{B}$ is the partition

$$\mathcal{A} \vee \mathcal{B} = \{A \cap B \,:\, A \in \mathcal{A}\,,\, B \in \mathcal{B}\}$$

2. $\mathcal{A}$ *refines* $\mathcal{B}$ (up to measure 0) if every $A \in \mathcal{A}$ is contained in some $B \in \mathcal{B}$ (up to measure 0).
3. $\mathcal{A}, \mathcal{B}$ are *independent* if $\mu(A \cap B) = \mu(A)\mu(B)$ for $A \in \mathcal{A}$, $B \in \mathcal{B}$.

Properties of entropy (continued)

(E2') $H(\cdot, \mathcal{A}|\mathcal{B})$ is concave:

(E3) $H(\mu, \mathcal{A} \vee \mathcal{B}) = H(\mu, \mathcal{A}) + H(\mu, \mathcal{B}|\mathcal{A})$

(E4) $H(\mu, \mathcal{A} \vee \mathcal{B}) \geq H(\mu, \mathcal{A})$ with equality if and only if $\mathcal{A}$ refines $\mathcal{B}$ up to μ-measure 0.

(E5) $H(\mu, \mathcal{A} \vee \mathcal{B}) \leq H(\mu, \mathcal{A}) + H(\mu, \mathcal{B})$ with equality if and only if $\mathcal{A}, \mathcal{B}$ are independent. Equivalently, $H_\mu(\mathcal{B}|\mathcal{A}) \leq H(\mathcal{B})$ with equality if and only if $\mathcal{A}, \mathcal{B}$ are independent.

Proof. For (E3), by algebraic manipulation,

$$\begin{aligned}
H(\mu, \mathcal{A} \vee \mathcal{B}) &= \\
&= -\sum_{A \in \mathcal{A}, B \in \mathcal{B}} \mu(A \cap B) \log \mu(A \cap B) \\
&= \sum_{A \in \mathcal{A}} \mu(A) \sum_{B \in \mathcal{B}} \frac{\mu(A \cap B)}{\mu(A)} \left(-\log \frac{\mu(A \cap B)}{\mu(A)} - \log \mu(A) \right) \\
&= -\sum_{A \in \mathcal{A}} \mu(A) \log \mu(A) \sum_{B \in \mathcal{B}} \mu_A(B) - \sum_{A \in \mathcal{A}} \mu(A) \sum_{B \in \mathcal{B}} \mu_A(B) \log \mu_A(B) \\
&= H(\mu, \mathcal{A}) + H(\mu, \mathcal{B}|\mathcal{A})
\end{aligned}$$

The inequality in (E4) follows from (E3) since $H(\mu,\mathcal{B}|\mathcal{A})\geq 0$; there is equality if and only if $H(\mu_A,\mathcal{B})=0$ for all $A\in\mathcal{A}$ with $\mu(A)>0$. By (E1), this occurs precisely when, on each $A\in\mathcal{A}$ with $\mu(A)\neq 0$, the measure μ_A is supported on a single atom of $\mathcal{B}$, which means that $\mathcal{A}$ refines $\mathcal{B}$ up to measure 0.

For (E2'), let $\mu=\alpha\eta+(1-\alpha)\theta$. For $B\in\mathcal{B}$ let $\beta_B=\frac{\alpha\eta(B)}{\mu(B)}$. Then $(1-\beta_B)=\frac{(1-\alpha)\theta(B)}{\mu(B)}$ and

$$\mu_B=\beta_B\eta_B+(1-\beta_B)\theta_B$$

hence

$$\begin{aligned}
H(\mu,\mathcal{A}|\mathcal{B}) &= \\
&= \sum_{B\in\mathcal{B}}\mu(B)H(\mu_B,\mathcal{B}) && \text{by definition}\\
&\geq \sum_{B\in\mathcal{B}}\mu(B)\left(\beta_B H(\eta_B,\mathcal{A})+(1-\beta_B)H(\theta_B,\mathcal{A})\right) && \text{by concavity (E2)}\\
&= \sum_{B\in\mathcal{B}}\left(\alpha\eta(B)\cdot H(\eta_B,\mathcal{A})+(1-\alpha)\theta(B)\cdot H(\theta_B,\mathcal{A})\right)\\
&= \alpha H(\eta,\mathcal{A}|\mathcal{B})+(1-\alpha)H(\theta,\mathcal{A}|\mathcal{B})
\end{aligned}$$

Finally, (E5) follows from (E1) an (E2). First,

$$H(\mu,\mathcal{B}|\mathcal{A})=\sum_{B\in\mathcal{B}}\mu(B)H(\mu_B,\mathcal{A})\leq H(\sum_{B\in\mathcal{B}}\mu(B)\mu_B,\mathcal{A})=H(\mu,\mathcal{A})$$

It is clear that if $\mathcal{A},\mathcal{B}$ are independent there is equality. To see this is the only way it occurs, one again uses strict convexity of $H(\underline{p})$, which shows that the independent case is the unique maximizer. □

There are a few generalizations of these properties which are useful:

Properties of entropy (continued):

1. ([E3']) $H(\mathcal{A},\mathcal{B}|\mathcal{C})=H(\mathcal{B}|\mathcal{C})+H(\mathcal{A}|\mathcal{B}\vee\mathcal{C})$.

(E5') If $\mathcal{C}$ refines $\mathcal{B}$ then $H(\mathcal{A}|\mathcal{C})\leq H(\mathcal{A}|\mathcal{B})$, with equality if and only if $\mathcal{B}=\mathcal{C}$.

Proof. For (E3') expand both sides using (E3). For (E5') use (E3'), noting that $\mathcal{C}=\mathcal{C}\vee\mathcal{B}$ since $\mathcal{C}$ refines $\mathcal{B}$. □

We have already remarked that discrete random variables correspond to partitions of the underlying probability space, and the entropy of the random variable is that of the corresponding partition. Let us now say a few words about how relations and operations on partitions translate to random variables.

If $\xi:X\to I$ and $\zeta:X\to J$ are random variables corresponding partitions $\mathcal{A}_\xi,\mathcal{A}_\zeta$ of X, respectively, then the pair (ξ,ζ) is an $I\times J$-valued random variable corresponding to the partition $\mathcal{A}_\xi \vee \mathcal{A}_\zeta$. The random variables are independent if and only the corresponding partitions are. If $\mathcal{A}_\xi$ refines $\mathcal{A}_\zeta$ (modulo nullsets) , the the atom $\mathcal{A}_\zeta(x)$ is determined by $\mathcal{A}_\xi(x)$, hence ζ is a function of ξ (a.s.), that is, there is a function $I \to J$ such that $\zeta = f(\xi)$ (a.s.). The converse is trivially true, and by symmetry, ξ and ζ are functions of each other if and only if the partitions agree.

Using random variable notation, we define

$$H(\xi|\zeta) = H(\mu, \mathcal{A}_\xi|\mathcal{A}_\zeta)$$

(where μ is the underlying probability measure), and similarly deifne

$$H(\xi, \xi'|\zeta, \zeta') = H(\mu, \mathcal{A}_\xi \vee \mathcal{A}_{\xi'}|\mathcal{A}_\zeta \vee \mathcal{A}_{\zeta'})$$

etc.

One interprets $H(\xi)$ as a measure of the randomness of ξ: If it takes on 1 value e.s. then $H(\xi) = 0$, if it takes on n values then $H(\xi) \leq \log n$ with equality if and only if $\xi(a) = \frac{1}{n}$ for each of these values; etc. The inequality $H(\xi|\zeta) \leq H(\xi)$ (and equality if and only if ξ, ζ are independent) means that the amount of uncertanty about ξ can only be decreased if we learn the output of another ranodom variable ζ, and in fact it must decrease unless they are independent.

Remark. The definition of entropy may seem somewhat arbitrary. However, up to normalization, it is essentially the only possible definition if we wish (E1)–(E5) to hold. A proof of this can be found in Shannon's original paper on information theory and entropy, [**?**].

Chapter 10

Entropy of a stationary process

10.1 Stationary processes and measure preserving systems

Recal that a stochastic process (ξ_n) is a sequence of random variables (defined on a common probability space. A process is stationary if $\mathbb{P}(\xi_n\xi_{n+1}\ldots\xi_{n+k} \in A_0 \times A_1 \times \ldots \times A_k)$ is independent of n. In partcular, every ξ_n has the same distribution. In this section we are interested only in finite valued processes, i.e. those for which $\mathbb{P}(\xi_n \in \Sigma) =$ for some finite set Σ.

For a finite partition $\mathcal{A} = \{A_i\}$ of a set we write $\mathcal{A}(x) = i$ if $x \in A_i$. A measure preserving system $(X, \mathcal{B}, \mu, T)$ together with a finite partition $\mathcal{A} = \{A_i\}$ define a stationary process (ξ_n) by

$$\xi_n(x) = \mathcal{A}(T^n x)$$

that is, $\xi_n(x) = i$ if and only if $T^n x \in A_i$. This is defined for $n \geq 0$ and, it T is invertible, also for $n < 0$. Measure preservation implies that (ξ_n) is stationary. Note that the partition of X determined by ξ_n is just $T^{-n}\mathcal{A}$, and more generally, $\xi_m, \ldots, \xi_n$ determines hte partition

$$\mathcal{A}_m^n = \bigvee_{i=m}^{n} T^{-i}\mathcal{A}$$

Conversely, a finite valued stochastic process (ξ_n) with values in Σ induces a measure on $\Sigma^{\mathbb{Z}}$ (the distribution of the sequence (ξ_n), which by stationarity is shift invariant, and if we take $\mathcal{A} = \{A_\sigma\}_{\sigma\in\Sigma}$ to be the partition of $\Sigma^{\mathbb{Z}}$ according tot he time-0 coordinate, i.e. $A_\sigma = \{x : x_0 = \sigma\}$, then the process (ζ_n) arising from $\mathcal{A}$ as above has the same distribution as the original process (ξ_n). Thus, up to identifying processes with the same distribution, stationary sstochstic processes and measure preserving systems with a distinguished partition are equivalent models.

10.2 Entropy of a stationary process

Definition 10.2.1. Let $\xi = (\xi_n)$ be a finite-valued stationary stochasticprocess. The entropy $h(\xi)$ of the process is the limit

$$h(\xi) \quad = \quad \lim_{n\to\infty} \frac{1}{n} H(\xi_0, \ldots, \xi_{n-1})$$

Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system. The entropy of a finite partition $\mathcal{A}$ of the system is

$$\begin{aligned} h_\mu(T, \mathcal{A}) &= \lim_{n\to\infty} \frac{1}{n} H_\mu(\mathcal{A}_0^{n-1}) \\ &= \lim_{n\to\infty} \frac{1}{n} H_\mu(\bigvee_{k=0}^{n-1} T^{-k}\mathcal{A}) \end{aligned}$$

Lemma 10.2.2. *The limits in the definition exists.*

Proof. Set

$$a_n = H(\xi_0, \ldots, \xi_{n-1})$$

Then the existence of the limit will follow if we show that a_n is sub-additive, i.e. $a_{m+n} \leq a_m + a_n$. Indeed,

$$\begin{aligned} a_{m+n} &= H(\xi_0, \ldots, \xi_{m-1}, \xi_m, \ldots, \xi_{m+n-1}) \\ &= H(\xi_0, \ldots, \xi_{m-1}) + H(\xi_m, \ldots, \xi_{m+n-1} | \xi_0 \ldots \xi_{m-1}) \\ &\leq H(\xi_0, \ldots, \xi_{m-1}) + H(\xi_m, \ldots, \xi_{m+n-1}) \\ &= H(\xi_0, \ldots, \xi_{m-1}) + H(\xi_0, \ldots, \xi_{n-1}) \\ &= a_m + a_n \end{aligned}$$

where in the second to last line we used the fact that, by stationarity, $\xi_0, \ldots, \xi_{n-1}$ and $\xi_m, \ldots, \xi_{m+n-1}$ have the same distribution, and hence the same entropy. □

Example 10.2.3. Let $\xi = (\xi_n)$ be i.i.d. with marginal $p \in \mathcal{P}(\Sigma)$, that is, $\mathbb{P}(\xi_n = \sigma) = p_\sigma$ and the ξ_n are independent. Then ξ_i is independent of $\xi_{i+1}, \ldots, \xi_{i+k}$ and hence

$$H(\xi_i, \ldots, \xi_{i+k}) = H(\xi_i) + H(\xi_{i+1}, \ldots, \xi_{i+k}) = H(p) + H(\xi_{i+1}, \ldots, \xi_{i+k})$$

and so by induction

$$H(\xi_0, \ldots, \xi_{n-1}) = nH(p)$$

hence

$$h(\xi) = H(p)$$

Example 10.2.4. Let $X = S^1$, μ length measure and R_θ a rotation. Let $\mathcal{A}$ denote the partition of X into northern and southern hemispheres (with some convention for the endpoints). Then $R_\rho^{-n}\mathcal{A}$ is also a partition into two intervals. The partition $\mathcal{A}^n$ is then also a partition into intervals, and these are determined by the endpoints of the intervals $T^{-k}\mathcal{A}$, $k = 0, \ldots, n-1$. There are at most $2n$ such endpoints (exactly $2n$ if θ is irrational) and so $\mathcal{A}^n$ consists of at most $2n$ intervals. Hence $H_\mu(\mathcal{A}^n) \leq \log 2n$ by (E1) and

$$0 \leq h_\mu(T, \mathcal{A}) \leq \lim_{n\to\infty} \frac{\log 2n}{n} = 0$$

so $h_\mu(\mathcal{A}) = 0$.

The entropy of a process $\xi = (\xi_n)$ can be interpreted as the average number of bits per symbol needed to code long blocks. Indeed, if we want to encode $\xi_0 \ldots \xi_{n-1}$, we need at least $H(\xi_0 \ldots \xi_{n-1}) = nh(\xi) + o(n)$ bits on average, so the number of of bits *per symbol* is obtaind by dividing by n, giving $h(\xi) + o(1)$. In fact we have seen that prefix codes exist which code $\xi_0 \ldots \xi_{n-1}$ in $H(\xi_0 \ldots \xi_{n-1}) + 1$ but on average, which is also $h(\xi) + o(1)$ bits per symbol! Thus, for a stationary process, we can get arbitrarily close to the optimal coding rate if we code long blocks.

Entropy also has another intepretation, as the amoung of information of ξ_0 given the past, or future:

Proposition 10.2.5. *Let $\xi = (\xi_n)$ be a stationary process. Then*

$$h(\xi) = H(\xi_0|\xi_1, \xi_2, \ldots)$$

and if ξ is two-sided then also

$$h(\xi) = H(\xi_0|\xi_{-1}\xi_{-2} \ldots)$$

Proof. Using (E3) we have

$$\begin{aligned} H(\xi_0 \ldots \xi_{n-1}) &= H(\xi_{n-1}) + H(\xi_0 \ldots \xi_{n-2}|\xi_{n-1}) \\ &= H(\xi_{n-1}) + H(\xi_{n-2}|\xi_{n-1}) + H(\xi_0 \ldots \xi_{n-3}) \\ &\vdots \quad \vdots \\ &= \sum_{k=0}^{n-1} H(\xi_k|\xi_{k+1}, \xi_{k+2}, \ldots, \xi_{n-1}) \end{aligned}$$

Now, $H(\xi_0|\xi_1, \ldots, \xi_{n-1}) \to H(\xi_0|\xi_1, \xi_2 \ldots)$, so this is true also of the Cesaro averages, and we get

$$h(\xi) = \lim_{n\to\infty} \frac{1}{n} H(\xi_0, \ldots, \xi_{n-1}) = \lim_{n\to\infty} \sum_{k=0}^{n-1} H(\xi_k|\xi_{k+1}, \ldots, \xi_{n-1}) = H(\xi_0|\xi_1, \xi_2 \ldots)$$

The formula $h(\xi) = H(\xi_0|\xi_{-1}, \xi_{-2}, \ldots)$ is proved similarly, going backwards. □

Remark 10.2.6. From the proof and the fact that the sequence $H(\xi_0|\xi_1 \dots \xi_{n-1})$ is non-increasing in n we find that in fact $\frac{1}{n}H(\xi_0, \dots, \xi_{n-1})$ is non-increasing as well.

Definition 10.2.7. A process $\xi = (\xi_n)_{n=-\infty}^{\infty}$ is called **deterministic** if the past deermines the future, that is, if ξ_0 is a.s. determined by the past, $\xi_{-1}, \xi_{-2}, \dots$ (and hence $\xi_0, \xi_1, \dots$ are a.s. all determined by $\xi_{-1}, \xi_{-2} \dots$).

Lemma 10.2.8. *ξ is deterministic if and only if $h(\xi) = 0$.*

Proof. ξ_0 is deterministic if and only if it is measurable with respect to $\sigma(\xi_{-1}, \xi_{-2}, \dots)$, if and only if its conditional entropy on this σ-algebra is zero. By the proposition, this is just saying that the entropy is zero. □

The claim has a surprising consequence: A process is deterministic if and only if it is also deterministic with time reversed, i.e. the past determines the future, if and only if the future determines the past! This is because both imply zero entropy. Interestingly, there is no known proof of this fact which does not involve entropy.

10.3 An example: Decay of long words for Bernoulli measures

Given a stationary process $\xi = (\xi_n)$ with values in Σ, every word $a \in \Sigma^*$ is assigned a probability, namely $p(a_1 \dots a_n) = \mathbb{P}(\xi_1 \dots \xi_n = a_1 \dots a_n)$. If we observe the random sample $\xi_0 \dots \xi_{n-1}$, it makes sense to ask what its probability os occurring is, i.e., what is $p(\xi_0, \dots, \xi_{n-1})$. In general this probability depends on the sample, and besides being obviously decreasing as we increase the sample size, it is not clear what regularity it possesses. We shall see that in fact the rate of decay of this sequence is govorned by the entropy of the process. We begin with an example.

Example 10.3.1. Let $(\xi_n)_{n=0}^{\infty}$ be a $\{0,1\}$-valued n i.i.d. process with $\mathbb{P}(\xi_n = 0) = p$ and $\mathbb{P}(\xi_n = 1) = 1 - p$ for some $0 < p \leq \frac{1}{2}$. If $p = \frac{1}{2}$ then for every sequence $a \in \{0,1\}^n$, $\mathbb{P}(\xi_1 \dots \xi_n = a_1 \dots a_n) = 2^{-n}$, independent of choice of a. But if $0 < p < \frac{1}{2}$ then different sequences may yield different probabilities, the minimal one being $a = 00 \dots 0$ with probability p^n and the largest being $a = 11 \dots 1$ with probability $(1-p)^n$. In general, writing $p_0 = p$ and $p_1 = 1 - p$, we have

$$\begin{aligned} \mathbb{P}(\xi_1 \dots \xi_n = a_1 \dots a_2) &= \prod_{i=1}^{n} p_{a_i} \\ &= p^{\#\{1 \leq i \leq n \,:\, a_i = 0\}} \cdot p^{\#\{1 \leq i \leq n \,:\, a_i = 1\}} \end{aligned}$$

Now, for an infinite realization $a \in \{0,1\}^{\mathbb{N}}$ of the process, by the ergodic theorem (or law of large numbers),

$$\#\{1 \leq i \leq n \,:\, a_i = 0\} = n(p + o(1))$$

and

$$\#\{1 \le i \le n \,:\, a_i = 1\} = n(1-p+o(1))$$

Therefore with probability one over the choice of a,

$$\begin{aligned}\mathbb{P}(\xi_1 \ldots \xi_n = a_1 \ldots a_2) &= p^{n(p+o(1))}(1-p)^{n(1-p+o(1))} \\ &= 2^{n(-p\log p-(1-p)\log(1-p)+o(1))} \\ &= 2^{nH(p)+o(n)}\end{aligned}$$

In other words, with probability one over the choice of a,

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\xi_1 \ldots \xi_n = a_1 \ldots a_2) = H(p)$$

So, although different realizations have initial segments with different probabilities, asymptotically the probabilities are a.s. the same (when measured in this way).

In particular, for any $\varepsilon > 0$, the set of sequences

$$\Sigma_n = \{a \in \{0,1\}^n \,:\, 2^{-(H(p)+\varepsilon)n} \le \mathbb{P}(\xi_1 \ldots \xi_n = a_1 \ldots a_n) \le 2^{-(H(p)-\varepsilon)n}\}$$

satisfies that a.s., $x_1 \ldots x_n \in \Sigma_n$ for all large enough n; hence $P(\xi_1 \ldots \xi_n \in \Sigma_n) \to 1$ as $n \to \infty$. This tells us that *most* realizations of the first n variables occur with "comparable" probabilities.

10.4 Maker's theorem

Theorem 10.4.1. *Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system. Let $f_n \in L^1$ and $f_n \to f$ a.e. Suppose that $\sup_n |f_n| \in L^1$. Then*

$$\frac{1}{N} \sum_{n=0}^{N-1} T^n f_{N-n} \to \mathbb{E}(f|\mathcal{I})$$

a.e. and in L^1, where $\mathcal{I} \subseteq \mathcal{F}$ is the σ-algebra of T-invariant sets. Also,

$$\frac{1}{N} \sum_{n=0}^{N-1} T^n f_n \to \mathbb{E}(f|\mathcal{I})$$

Proof. We prove the first statement, and begin under the assumption that T is ergodic, so $\mathcal{I}$ is trivial.

We first claim that we may assume that $f \equiv 0$. By the ergodic theorem $\frac{1}{N}\sum_{n=0}^{N-1} T^n f \to \mathbb{E}(f|\mathcal{I})$ a.e. and in L^1, so in order to prove the theorem it is enough to show that $\frac{1}{N}\sum_{n=0}^{N-1} T^n(f_{N-n} - f) \to 0$ a.e. and in L^1. Since $\sup_n |f_n - f| \in L^1$, we have reduced to the case $f \equiv 0$.

Assume now $f \equiv 0$. Let $\varepsilon > 0$ and let

$$g = \sup_n |f_n|$$

By assumption $g \in L^1$, so we can fix $\delta > 0$ such that for any set E with $\mu(E) < \delta$ we have $\int_E g d\mu < \varepsilon$.

Since $f_n \to 0$ a.e., there is an n_0 and a set A with $\mu(A) > 1 - \delta$ such that $|f_n(x)| < \varepsilon$ for $x \in X$ and all for $n > n_0$.

Now consider $f'_n = 1_A f_n$ and $f''_n = 1_{X \setminus A} f_n$, so $f_n = f'_n + f''_n$. Since $|f'_n| < \varepsilon$ for $n > n_0$ and $|f'_n| \leq g$, we have

$$\begin{aligned} \frac{1}{N}\sum_{n=0}^{N-1} |T^n f'_{N-n}| &< \frac{1}{N}\sum_{n=0}^{N-n_0-1} \varepsilon + \frac{1}{N}\sum_{n=N-n_0-1}^{N-1} T^n g \\ &< \varepsilon + \frac{1}{N}\left(\sum_{n=0}^{N-1} T^n g - \sum_{n=0}^{N-n_0-1} T^n g\right) \qquad (10.1) \end{aligned}$$

The last term on the right tends to 0 a.e. and in L^1 as $N \to \infty$. On the other hand

$$\begin{aligned} \frac{1}{N}\sum_{n=0}^{N-1} T^n |f''_{N-n}| &\leq \frac{1}{N}\sum_{n=0}^{N-1} T^n |1_{X \setminus A} g| \\ &\to \int_{X \setminus A} g \, d\mu \\ &< \varepsilon \qquad (10.2) \end{aligned}$$

a.e. and in L^1, because $\mu(X \setminus A) < \delta$. Combining the two inequalities we conclude that

$$\begin{aligned} \limsup_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} T^n |f_{N-n}| &\leq \limsup_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} T^n |f'_{N-n}| + \limsup_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} T^n |f''_{N-n}| \\ &\leq 2\varepsilon \end{aligned}$$

so $\frac{1}{N}\sum_{n=0}^{N-1} T^n |f_{N-n}| \to 0$ a.e., and similarly, taking the L^1-norm, in L^1.

In the case that $\mathcal{I}$ is non-trivial we proceed in the same manner, but in (10.2) the conclusion becomes

$$\frac{1}{N}\sum_{n=0}^{N-1} T^n |f''_{N-n}| \to \mathbb{E}(1_{X \setminus A} g | \mathcal{I})$$

Now, $\int \mathbb{E}(1_{X \setminus A} g|\mathcal{I}) d\mu = \int 1_{X \setminus A} g d\mu < \varepsilon$, and since $1_{X \setminus A} g \geq 0$ and conditional expectation is a positive operator, $\mathbb{E}(1_{X \setminus A} g|\mathcal{I}) \geq 0$ a.s. Thus by Markov's inequality

$$\mu(\mathbb{E}(1_{X \setminus A} g|\mathcal{I}) \geq \sqrt{\varepsilon}) \leq \frac{\int \mathbb{E}(1_{X \setminus A} g|\mathcal{I}) d\mu}{\sqrt{\varepsilon}} < \sqrt{\varepsilon}$$

We find that

$$\mu\left(x : \limsup_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} T^n |f''_{N-n}|(x) > \sqrt{\varepsilon}\right) < \sqrt{\varepsilon}$$

and so as before, combining the above with (10.1),

$$\mu\left(x : \limsup_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}T^n|f_{N-n}|(x) > \varepsilon+\sqrt{\varepsilon}\right) < \sqrt{\varepsilon}$$

Taking $\varepsilon_k = 2^{-k}$, we have $\sum\sqrt{\varepsilon_k} < \infty$. Applying Borel-Cantelli, we find that a.e. x is in finitely many of the events above and hence a.s. $\frac{1}{N}\sum_{n=0}^{N-1}T^n|f_{N-n}|(x) \to 0$ as desired. □

10.5 The Shannon-McMillan-Breiman theorem

Recal that for a partition $\mathcal{B}$ we write $\mathcal{B}(x)$ for the element of $\mathcal{B}$ containing x. Recall that $\mathcal{A}_m^n = \bigvee_{i=m}^n T^{-i}\mathcal{A}$ and note that

$$\mathcal{A}_0^{k-1}(x) = \bigcap_{i=0}^{k-1}(T^{-i}\mathcal{A})(x)$$

Theorem 10.5.1. *Let $\xi = (\xi_n)_{n=0}^{\infty}$ be an ergodic stationary process with values in a finite set I. Let $p(a_1\dots a_n) = \mathbb{P}(\xi_1\dots\xi_n = a_1\dots a_n)$. Then*

$$\lim_{n\to\infty}\frac{1}{n}\log p(\xi_1\dots\xi_n) = h(\xi)$$

almost-surely and in L^1.

Equivalently, let $(X,\mathcal{F},\mu,T)$ be an ergodic measure preserving system and $\mathcal{A}$ a finite partition. Then for a.e. x,

$$\lim_{n\to\infty}\frac{1}{n}\log\mu(\mathcal{A}_0^{n-1}(x)) = h_\mu(T,\mathcal{A})$$

Proof. The two versions are related in the usual way, noting that if $\xi_k = \xi_0\circ T^k$ and $\mathcal{A} = \mathcal{A}_{\xi_0}$ then

$$p(\xi_0(x),\dots,\xi_{k-1}(x)) = \mu(\bigcap_{i=0}^{k-1}(T^{-i}\mathcal{A})(x)) = \mu(\mathcal{A}_0^{k-1}(x))$$

We shall work with the partition formulation.

Fix x. Defining $\mathcal{A}_m^n = \{X\}$ to be the trivial partition when $m > n$, the law of total probability tells us that

$$\begin{aligned}\mu(\mathcal{A}_0^{n-1}(x)) &= \mu(\bigcap_{k=0}^{n-1}(T^{-k}\mathcal{A})(x)) \\ &\quad \prod_{k=0}^{n-1}\mu((T^{-k}\mathcal{A})(x)|\mathcal{A}_{k+1}^{n-1}(x)) \\ &= \prod_{k=0}^{n-1}\mu(\mathcal{A}(T^kx)|\mathcal{A}_1^{n-(k+1)}(T^kx))\end{aligned}$$

Hence

$$\log \mu(\mathcal{A}_0^{n-1}(x)) = \sum_{k=1}^{n} \mu(\mathcal{A}(T^k x)|\mathcal{A}_1^{n-(k+1)}(T^k x))$$

Therefore, if we define

$$f_k = -\log \mu(\mathcal{A}(x)|\mathcal{A}_1^{k-1}(x)),$$

then we have

$$-\frac{1}{n}\log \mu(\mathcal{A}_0^{n-1}(x)) = \frac{1}{n}\sum_{k=0}^{n-1} f_{n-k}(T^k x)$$

In order to complete the proof, we show that the f_k satisfy the hypothesis of maker's theorem, and identify the limit. Let

$$f(x) = -\log \mu_x(\mathcal{A}(x)|\bigvee_{i=1}^{\infty} T^{-i}\mathcal{A})$$

(that is, if $x \in A \in \mathcal{A}$ then $f(x) = \mu_x(A|\bigvee_{i=1}^{\infty} T^{-i}\mathcal{A})$).

Claim 10.5.2. $f_k \to f$ a.e.

Proof. Immediate from the martingale theorem, since fixing $A \in \mathcal{A}$, for μ-a.e. $x \in A$, we have

$$\begin{aligned}
\mu_x(\mathcal{A}(x)|\mathcal{A}_1^{k-1}) &= \mu_x(A|\bigvee_{i=1}^{k-1} T^{-i}\mathcal{A}) \\
&= \mathbb{E}(1_A|\bigvee_{i=1}^{k-1} T^{-i}\mathcal{A})(x) \\
&\xrightarrow[n\to\infty]{} \mathbb{E}(1_A|\bigvee_{i=1}^{\infty} T^{-i}\mathcal{A})(x) \\
&= f(x) \quad \square
\end{aligned}$$

Claim 10.5.3. $\sup_k |f_k| \in L^1$

Proof. Let

$$E_t = \{x \,:\, \sup_k f_k(x) > t\}$$

It suffices for us to show that $\mu(E_t) < C \cdot 2^{-t}$ where C is independent of t, since then

$$0 \leq \sup_k |f_k| \leq \sum_{n=0}^{\infty} 1_{E_n}$$

and the right hand side is integrable.

For each $A \in \mathcal{A}$ consider the family $\mathcal{U}_A$ of sequences $(A_i)_{i=1}^k$ of any length for which $A_i \in T^{-i}\mathcal{A}$, and such that

$$-\log \frac{\mu(A \cap \bigcap_{i=1}^k A_i)}{\mu(\bigcap_{i=1}^k A_i)} > t$$

but $(A_i)_{i=1}^\ell$ does not satisfy this for any $1 \leq \ell < k$. Evidently the sets $\bigcap_{i=1}^k A_i$ are pairwise disjoint as (A_i) ranges over $\mathcal{U}_A$, and every $x \in A \cap E_t$ belongs to such an intersection. Therefore it suffices for us to show that

$$\mu(A \cap \bigcup_{(A_i) \in \mathcal{U}_A} \bigcap A_i) < 2^{-t}$$

since then, summing over $A \in \mathcal{A}$, we have $\mu(E_t) < |\mathcal{A}| \cdot 2^{-t}$.

To show the last inequality above, observe that for each $(A_i) \in \mathcal{U}_A$ we have

$$\mu(\bigcap_{i=1}^k A_i \cap A) = \mu(\bigcap_{i=1}^k A_i) \cdot \frac{\mu(A \cap \bigcap_{i=1}^k A_i)}{\mu(\bigcap_{i=1}^k A_i)} < 2^{-t} \cdot \mu(\bigcap_{i=1}^k A_i)$$

Therefore, using the fact that the sets $\bigcap_{i=1}^k A_i$ are pairwise disjoint for $(A_i) \in \mathcal{U}_A$,

$$\begin{aligned}
\mu(A \cap \bigcup_{(A_i) \in \mathcal{U}_A} \bigcap A_i) &= \sum_{(A_i) \in \mathcal{U}_A} \mu(A \cap \bigcap A_i) \\
&< 2^{-t} \sum_{(A_i) \in \mathcal{U}_A} \mu(\bigcap A_i) \\
&\leq 2^{-t} \mu(\bigcup_{(A_i) \in \mathcal{U}_A} \bigcap A_i) \\
&\leq 2^{-t}
\end{aligned}$$

as desired. □

□

We can now apply Makers theorem and deduce that $-\frac{1}{n} \log \mu(\mathcal{A}^n(x)) \to \mathbb{E}(f|\mathcal{I})$ a.s. as $n \to \infty$, where $\mathcal{I}$ is the σ-algebra of T-invariant sets. Since our system is ergodic this is simply $\int f d\mu$, and we have already seen that this is the entropy of the system.

Remark 10.5.4. The proof shows that convergence holds also in the non-ergodic case, and the limit is $\mathbb{E}(f|\mathcal{I})$. If $\mu = \int \nu_x d\mu(x)$ is the ergodic decomposition of μ, then $\mathbb{E}(f|\mathcal{I})(x) = \int f \, d\nu_x$. It is also not too hard to show that $\int f d\nu_x = h_{\nu_x}(T, \mathcal{A})$ a.s. Therefore $\frac{1}{n} \log \mu(\mathcal{A}^n(x)) \to h_{\nu_x}(T)$ a.s.

10.6 Entropy-typical sequences

Let $\xi = (\xi_n)$ be an ergodic process with values in the finite set Σ and entropy $h = h(\xi)$. For $a \in \Sigma^n$ write

$$p(a) = \mathbb{P}(\xi_1 \dots \xi_n = a)$$

For $\varepsilon > 0$ let

$$\mathcal{T}_{n,\varepsilon} = \{a \in \Sigma^n \,:\, 2^{-n(h+\varepsilon)} < p(a) < 2^{-n(h-\varepsilon)}\}$$

We say that $a \in \mathcal{T}_{n,\varepsilon}$ is ε-entropy typical. It is clear that

$$|\mathcal{T}_{n,\varepsilon}| < 2^{n(h+\varepsilon)}$$

because

$$1 \geq \sum_{a \in \mathcal{T}_{n,\varepsilon}} p(a) > |\mathcal{T}_{n,\varepsilon}| \cdot 2^{-n(h+\varepsilon)}$$

Theorem 10.6.1 (asymptotic equipartition property)**.** *With the above notation, for every $\varepsilon > 0$, with probability one we have*

$$\xi_1 \dots \xi_n \in \mathcal{T}_{n,\varepsilon} \textit{ for all large enough } n$$

(In particular $\mathbb{P}(\xi_1 \dots \xi_n \in \mathcal{T}_{n,\varepsilon}) \to 1$). Also, for all large neough n,

$$\frac{1}{2} . 2^{n(h-\varepsilon)} < |\mathcal{T}_{n,\varepsilon}| < 2^{n(h+\varepsilon)}$$

Proof. The first statement is a reformulation of Shannon-McMillan Breiman, and since $f_n = 1_{(\xi_! \dots \xi_n) \in \mathcal{T}_{n,\varepsilon}}$ tends pointwise to 1, it also tends to one in the mean, hence $\mathbb{P}(\xi_1 \dots \xi_n \in \mathcal{T}_{n,\varepsilon}) \to 1$. Finally, we already saw that $|\mathcal{T}_{n,\varepsilon}| \leq 2^{n(h+\varepsilon)}$. For the other bound, note that by the first part of the theorem,

$$\mathbb{P}(p(\xi_1 \dots \xi_n) < 2^{-n(h-\varepsilon)}) > \frac{1}{2}$$

for all large n. Thus

$$\frac{1}{2} < \sum_{a \in \mathcal{T}_{n,\varepsilon}} p(a) < |\mathcal{T}_{n,\varepsilon}| \cdot 2^{-n(h-\varepsilon)}$$

from which the claim follows. □

The theorem says that for large n **typical samples of the process have essentially the same probability**. Indeed, with high probability any two independent words have probabilities which are within a subexponential multiple of each other, because

$$a, b \in \mathcal{T}_{n,\varepsilon} \qquad \Longrightarrow \qquad 2^{-\varepsilon n} < \frac{p(a)}{p(b)} < 2^{\varepsilon n} \qquad \text{or:} \quad \frac{1}{n}|\log p(a) - \log p(b)| < \varepsilon$$

We can derive the following conclusion:

> As $n \to \infty$ with probability $1 - o(1)$, the word $\xi_1 \dots \xi_n$ is drawn from a set of size $2^{n(h+o(1))}$ and has probability $2^{-n(h+o(1))}$.

This gives us a new way to understand the possibility of approximating Shannon's lower bound when coding long sample of the process. We can use the following procedure:

- Fix $\varepsilon > 0$ and n large large enough that $\mathbb{P}(\xi_1 \dots \xi_n \in \mathcal{T}_{n,\varepsilon}) > 1 - \varepsilon$ and $|\mathcal{T}_{n,\varepsilon}| < 2^{n(h+\varepsilon)}$.
- Let $m_0 = \lceil n(h+\varepsilon) \rceil$, and choose an injective map $c_0 : \mathcal{T}_{n,\varepsilon} \to \{0,1\}^{m_0}$, which we can do because of our bound on $|\mathcal{T}_{n,\varepsilon}|$.
- Let $m_1 = \lceil \log |\Sigma|^n \rceil \le n \lceil \log |\Sigma| \rceil$, and choose an injective map $c_1 : \Sigma^n \to \{0,1\}^{m_1}$.
- Define $c : \Sigma^n \to \{0,1\}^*$ by

$$c(a) = \begin{cases} 0c_0(a) & a \in \mathcal{T}_{n,\varepsilon} \\ 1c_1(a) & \text{otherwise} \end{cases}$$

This is clearly a prefix code: if $c(a)$ is a prefix of $c(b)$, then theh begin with the same letter, so either $c(a) = 0c_0(a)$ and $b = 0c_0(b)$, in which case $a = b$ because c_0 is injective, or $c(a) = 1c_1(a)$ and $c(b) = 1c_2(b)$ with the same conclusion. The mean coding length of c is easy to compute since

$$c(\xi_1 \dots \xi_n) = (1 + m_0) 1_{\{\xi_1 \dots \xi_n \in \mathcal{T}_{n,\varepsilon}\}} + (1 + m_1) 1_{\{\xi_1 \dots \xi_n \notin \mathcal{T}_{n,\varepsilon}\}}$$

hence

$$\begin{aligned} \mathbb{E}(|c(\xi_1 \dots \xi_n)|) &= (1 + m_0)\mathbb{P}(\xi_1 \dots \xi_n \in \mathcal{T}_{n,\varepsilon}) + (1 + m_1)\mathbb{P}(\xi_1 \dots \xi_n \notin \mathcal{T}_{n,\varepsilon}) \\ &\le (2 + n(h+\varepsilon)) + (2 + n \log |\Sigma|)\varepsilon \\ &= n(h + \varepsilon + o(1)) \end{aligned}$$

Chapter 11

Kolmogorov-Sinai entropy

11.1 Entropy of a measure preserving system

We have defined the entropy of a partition in a m.p.s. However, different partitions can give different entropies (each system gives rise to many processes). For example, in any system the trivial partition into one set has entropy zero. To obtain a number associated to the system alone we have the following.

Definition 11.1.1. The Kolmogorov-Sinai entropy (or just entropy) of a measure preserving system $(X, \mathcal{F}, \mu, T)$ is

$$h_\mu(T) = \sup\{h_\mu(T, \mathcal{A}) \,:\, \mathcal{A} \text{ a finite patition of } X\}$$

It is possible to have $h_\mu(T) = \infty$. Indeed the entropy $h_\mu(T, \mathcal{A})$ is finite when $\mathcal{A}$ is finite but the upper bound $\log |\mathcal{A}|$ tends to infinity when the size of the partition does, and it is possible for the dynamical entropy to approach infinite as well. For exmepe, let λ be Lebesgue measure on $[0, 1]$ and $\mu = \lambda^{\mathbb{Z}}$, which is shift-invariant. The partition $\mathcal{A}_n$ of $[0, 1]$ into n equal sub-intervals induces a partition of $[0, 1]^{\mathbb{Z}}$ by partitioning according to the time-0 coordinate. Then the sequence $T^{-k}(\mathcal{A}_n)$ are independent, so

$$h_\mu(T, \mathcal{A}_n) = H_\mu(T, \mathcal{A}_n) = \log n \to \infty$$

hence

$$h_\mu(T) = \infty$$

Proposition 11.1.2. *Entropy is an isomophism invariant, i.e. isomorphic systems have the same entropy.*

Proof. Suppose $(X_i, \mathcal{F}_i, \mu_i, T_i)$, $i = 1, 2$, are m.p.s.'s and $f : X_1 \to X_2$ an isomorphism between them. For any sets $B_0, \ldots, B_k \in \mathcal{F}_2$ and $B = \bigcap_{i=0}^{k} T^{-1} B_k$, we have

$$f^{-1}(B) = \bigcap_{i=0}^{k} T^{-i}(f^{-1} B_i)$$

It follows that for any partition $\mathcal{B}$ of $\mathcal{F}_2$ and $\mathcal{A} = f^{-1}\mathcal{B}$, there is a measure-preserving identification between the atoms of $\mathcal{B}_1^k$ and $\mathcal{A}_1^k$, given by f^{-1}, and therefore

$$H_\mu(\mathcal{A}_1^k) = H_\mu(\mathcal{B}_1^k)$$

which implies

$$h_\mu(T_1, \mathcal{A}) = h_\mu(T_2, \mathcal{B})$$

Thus

$$\begin{aligned} h_\mu(T!) &= \sup\{h_\mu(T_1, \mathcal{A}) \,:\, \mathcal{A} \text{ a finite patition of } X_1\} \\ &\geq \sup\{h_\mu(T_1, f^{-1}(\mathcal{B})) \,:\, \mathcal{B} \text{ a finite patition of } X_2\} \\ &= \sup\{h_\mu(T_2, \mathcal{B}) \,:\, \mathcal{B} \text{ a finite patition of } X_2\} \\ &= h_\mu(T_2) \end{aligned}$$

The reverse inequality follows by symmetry, proving the claim. □

Calculating entropy is potentially difficult, since one must take into account all partitions. In practice, it is enough to consider a dense family of partitions, and sometimes even a single one. The following proposition allows us to compare the entropy determeined by two partitions.

Proposition 11.1.3. *Let $\mathcal{A}, \mathcal{B}$ be partitions in an invertible measure preserving system $(X, \mathcal{F}, \mu, T)$. Then*

$$h_\mu(T, \mathcal{A} \vee \mathcal{B}) = h_\mu(\mathcal{B}) + H_\mu(\mathcal{A}|\mathcal{A}_1^\infty \vee \mathcal{B}_{-\infty}^\infty)$$

and in any system (even not invertible),

$$h_\mu(T, \mathcal{A} \vee \mathcal{B}) \leq h_\mu(\mathcal{B}) + H_\mu(\mathcal{A}|\mathcal{A}_1^\infty \vee \mathcal{B}_0^\infty)$$

Proof. For each n, using (E3) once and then again inductively as in Proposition **??**,

$$\begin{aligned} H_\mu((\mathcal{A} \vee \mathcal{B})_0^{n-1}) &= H_\mu(\mathcal{A}_0^{n-1} \vee \mathcal{B}_0^{n-1}) \\ &= H_\mu(\mathcal{B}_0^{n-1}) + H_\mu(\mathcal{A}_0^{n-1}|\mathcal{B}_0^{n-1}) \\ &= H_\mu(\mathcal{B}_0^{n-1}) + \sum_{m=0}^{n-1} H_\mu(T^{-m}\mathcal{A}|\mathcal{A}_{m+1}^{n-1} \vee \mathcal{B}_0^{n-1}) \\ &= H_\mu(\mathcal{B}_0^{n-1}) + \sum_{m=0}^{n-1} H_\mu(\mathcal{A}|\mathcal{A}_1^{n-m-1} \vee \mathcal{B}_{-m}^{n-m-1}) \end{aligned}$$

Dividing by n and taking $n \to \infty$ the left hand side and the first term on the right tend to $h_\mu(T, \mathcal{A} \vee \mathcal{B})$ and $h_\mu(T, \mathcal{B})$ respectively. To evaluate the limit

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=0}^{n-1} H_\mu(\mathcal{A}|\mathcal{A}_1^{n-m-1} \vee \mathcal{B}_{-m}^{n-m-1})$$

note that for every m we have

$$\sum_{m=0}^{n-1} H_\mu(\mathcal{A}|\mathcal{A}_1^{n-m-1} \vee \mathcal{B}_{-m}^{n-m-1}) \geq H_\mu(\mathcal{A}|\mathcal{A}_1^\infty \vee \mathcal{B}_{-\infty}^\infty)$$

hence the right hand side is a lower bound for the limit. On the other hand, for $m > \sqrt{n}$ we have

$$H_\mu(\mathcal{A}|\mathcal{A}_1^{n-m-1} \vee \mathcal{B}_{-m}^{n-m-1}) \leq H_\mu(\mathcal{A}|\mathcal{A}_1^{\sqrt{n}} \vee \mathcal{B}_{-\sqrt{n}}^{\sqrt{n}})$$

and for every m the terms are bounded by $\log|\mathcal{A}|$, hence

$$\sum_{m=0}^{n-1} H_\mu(\mathcal{A}|\mathcal{A}_1^{n-m-1} \vee \mathcal{B}_{-m}^{n-m-1}) \leq \frac{\sqrt{n}}{n} \cdot \log|\mathcal{A}| + \frac{n-\sqrt{n}}{n} \cdot H_\mu(\mathcal{A}|\mathcal{A}_1^{\sqrt{n}} \vee \mathcal{B}_{-\sqrt{n}}^{\sqrt{n}})$$

The right hand side tends to $H_\mu(\mathcal{A}|\mathcal{A}_1^\infty \vee \mathcal{B}_{-\infty}^\infty)$, completing the proof.

In the non-invertible case we start with the same identity and note that conditioning only on $T^{-k}\mathcal{B}$ for non-negative k only can only increase the entropy. The rest is the same. □

Definition 11.1.4. A partition $\mathcal{A}$ in an invertible measure preserving system $(X, \mathcal{F}, \mu, T)$ is a generating partition if $\bigvee_{n=-\infty}^{\infty} T^{-n}\mathcal{A} = \mathcal{F}$ up to μ-measure 0 (that is $\mathcal{F} = \sigma(\mathcal{A}_n : n \in \mathbb{Z})$). If $\bigvee_{n=0}^{\infty} T^{-n}\mathcal{A} = \mathcal{F}$ we say that $\mathcal{A}$ is a one-sided generator (this definition makes sense also when T is not invertible).

Theorem 11.1.5. *Let $\mathcal{B}$ be a generating (or one-sided generating) partition in a measure preserving system $(X, \mathcal{F}, \mu, T)$. Then $h_\mu(T) = h_\mu(T, \mathcal{B})$.*

Proof. We prove the case of an invertible system, the other is similar. We must show that $h_\mu(T, \mathcal{A}) \leq h_\mu(T, \mathcal{B})$ for any finite partition $\mathcal{A}$. Indeed, fixing $\mathcal{A}$,

$$\begin{aligned} h_\mu(T, \mathcal{A}) &\leq h_\mu(T, \mathcal{A} \vee \mathcal{B}) \\ &= h_\mu(T, \mathcal{B}) + h_\mu(T, \mathcal{A}| \bigvee_{k=1}^{\infty} T^{-k}\mathcal{A} \vee \bigvee_{k=-\infty}^{\infty} T^{-k}\mathcal{B}) \\ &= h_\mu(T, \mathcal{B}) \end{aligned}$$

because $\mathcal{A} \in \mathcal{B}_{-\infty}^\infty$. □

Corollary 11.1.6. *Let μ_0 be a measure on a finite set A. Then the entropy of the product system $\mu_0^{\mathbb{Z}}$ with the shift is $H(\mu_0)$. In particular the product measures $\{\frac{1}{2}, \frac{1}{2}\}^{\mathbb{Z}}$ and $\{\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\}^{\mathbb{Z}}$ with the shift maps are not isomorphic.*

Proof. For a finite set A, the partition according to the 0-coordinate generates in the system $(A^{\mathbb{Z}}, \sigma)$, so the entropy is the entropies of this partition, which is just $H(\mu_0)$. □

In the absence of a generating partition, entropy can also be computed as follows.

Theorem 11.1.7. *Let $\mathcal{E}$ be an algebra of sets which generate the σ-algebra $\mathcal{F}$ of a measure preserving system $(X, \mathcal{F}, \mu, T)$. Then $h_\mu(T) = \sup h_\mu(T, \mathcal{B})$, where the supremum is over $\mathcal{E}$-measurable partitions $\mathcal{B}$.*

Proof. Write $\alpha = \sup H_\mu(T, \mathcal{B})$ for $\mathcal{B}$ as above. Evidently $H_\mu(T) \geq \alpha$ so we only need to prove the opposite inequality. For any finite partition $\mathcal{A}$, it is possible to find a refining sequence $\mathcal{C}_n$, $n = 1, 2, \ldots$, of $\mathcal{E}$-measurable partitions such that $\mathcal{A} \in \bigvee_{n=1}^{\infty} \mathcal{C}_n$. Then the argument in the proof of the previous theorem shows that $H_\mu(T, \mathcal{A}) \leq \lim H_\mu(T, \mathcal{C}_n) \leq \alpha$. □

11.2 Formal properties of entropy

Lemma 11.2.1 (Elemntary properties). *1.* $0 \leq h_\mu(T, \mathcal{A}) \leq \log |\mathcal{A}|$

2. $h_\mu(T, \mathcal{A}) \leq h_\mu(T, \mathcal{A} \vee \mathcal{B}) \leq h_\mu(T, \mathcal{A}) + h_\mu(T, \mathcal{B})$

3. $h_\mu(T, \mathcal{A}) = h_\mu(T, \mathcal{A}^k)$ *for all* $k \geq 1$.

4. $h_\mu(T^k, \mathcal{A}_0^{k-1}) = h_\mu(T, \mathcal{A})$.

5. If T is invertible, then $h_\mu(T, \mathcal{A}) = h_\mu(T^{-1}, \mathcal{A})$.

Proof. These are all easy consequences of the properties of Shannon entropy. For example,. to prove (3) note that

$$\begin{aligned}(\mathcal{A}_0^{m-1})_0^{n-1} &= \bigvee_{k=0}^{n-1} T^{-k} \mathcal{A}_0^{m-1} \\ &= \bigvee_{k=0}^{n-1} T^{-k} (\bigvee_{j=0}^{m-1} T^{-j} \mathcal{A}) \\ &= \bigvee_{k=0}^{n+m-2} T^{-k} \mathcal{A} \\ &= \mathcal{A}_0^{n+m-2}\end{aligned}$$

so

$$\lim_{n \to \infty} \frac{1}{n} H_\mu((\mathcal{A}_0^{m-1})_0^{n-1}) = \lim_{n \to \infty} \frac{n+m-2}{n} \cdot \frac{1}{n+m-2} H_\mu(\mathcal{A}_0^{n+m-2}) = h_\mu(T, \mathcal{A})$$

□

For (4), let $k \in \mathbb{N}_+$. We saw that $h_\mu(T, \mathcal{A}) = h_\mu(T, \mathcal{A}_0^{k-1})$. Now,

$$\bigvee_{i=0}^{n-1} T^{-ki} (\bigvee_{j=0}^{k-1} T^{-j} \mathcal{A}) = \mathcal{A}_0^{kn-1}$$

so

$$h_\mu(T^k, \mathcal{A}_0^{k-1}) = \lim_{n \to \infty} \frac{1}{n} H_\mu(\bigvee_{i=0}^{n-1} T^{-ki} (\bigvee_{j=0}^{k-1} T^{-j} \mathcal{A})) = k \cdot \lim_{n \to \infty} \frac{1}{kn} H_\mu(\mathcal{A}_0^{kn-1}) = k h_\mu(T, \mathcal{A})$$

We leave the west to the reader.

Lemma 11.2.2. .

1. $h_\mu(T^k) = |k| h_\mu(T)$.
2. $h_\mu(T \times S) = h_\mu(T) + h_\mu(S)$

Proof. For any finite partition $\mathcal{A}$ we saw that $h_\mu(T, \mathcal{A}) = k h_\mu(T^k, \mathcal{A}_0^{k-1})$. This implies that

$$\begin{aligned} h_\mu(T^k) &= \sup_{\mathcal{B}} \{h_\mu(T^k, \mathcal{B})\} \\ &\geq h_\mu(T^k, \mathcal{A}_0^{k-1}) \\ &= \frac{1}{k} h_\mu(T, \mathcal{A}) \end{aligned}$$

On the other hand since $\mathcal{A}_0^{k-1}$ refines $\mathcal{A}$, so

$$h_\mu(T^k, \mathcal{A}) \leq h_\mu(T^k, \mathcal{A}_0^{k-1}) = k h_\mu(T, \mathcal{A})$$

and this holds for all finite partitions $\mathcal{A}$, which gives the reverse inequality.

For the second statement, let $\mathcal{A}$ be a partition of T and $\mathcal{B}$ a partition of S, identifid as partitions of the product system. These are independent partitions so

$$\begin{aligned} h_\mu(T \times S, \mathcal{A} \times \mathcal{B}) &= \lim \frac{1}{n} H_\mu(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{A} \vee \bigvee_{i=0}^{n-1} T^{-i}\mathcal{B}) \\ &= \lim \frac{1}{n} \left(H_\mu(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{A}) + H_\mu(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{B}) \right) \\ &= h_\mu(T, \mathcal{A}) + h_\mu(T, \mathcal{B}) \end{aligned}$$

This shows that $h_\mu(T \times S) \geq h_\mu(T) + h_\mu(S)$. On the other hand, the algebra generated by product sets is dense in the product σ-algebra, so $h_\mu(T \times X)$ is the supremum of partitions from this algebra, and every such partition is refined by a partition of the form $\mathcal{A} \vee \mathcal{B}$ as above; so $h_\mu(T \times S) \leq h_\mu(T) + h_\mu(S)$. □

11.3 Factors and relative entropy

Definition 11.3.1. Factor

Remark 11.3.2. Identification of factors with sub-σ-algebras

Example 11.3.3. Trivial factors, factor generated by a partition/family of sets, product systems and marginal projections.

Definition 11.3.4. The entropy of a partition $\mathcal{A}$ in $(X, \mathcal{B}, \mu, T)$ relative to a factor $\mathcal{E} \subseteq \mathcal{B}$ is

$$h_\mu(T, \mathcal{A}|\mathcal{E}) = \lim_{n\to\infty} \frac{1}{n} H_\mu(\mathcal{A}_0^{n-1}|\mathcal{E})$$

and the entropy of the system relative to $\mathcal{E}$ it

$$h_\mu(T|\mathcal{E}) = \sup\{h_\mu(T, \mathcal{A}) \,:\, \mathcal{A} \text{ a finite partition of } X\}$$

The limit in the definition exists by subadditivity using exactly the same calculation as without the factor. the same proof also gives:

Proposition 11.3.5. $h_\mu(T, \mathcal{A}|\mathcal{B}) = H(\mathcal{A}| \bigvee_{i=1}^\infty T^i\mathcal{A} \vee \mathcal{B})$.

Remark 11.3.6. The usual definition is relative to the trivial factor.

Recall that

$$h_\mu(T, \mathcal{A} \vee \mathcal{B}) = h_\mu(T, \mathcal{B}) + H(\mathcal{A}|\mathcal{A}_1^\infty \vee \mathcal{B}_{-\infty}^\infty)$$

Proposition 11.3.7. *If $\mathcal{E}$ is a factor of $(X, \mathcal{F}, \mu, T)$ then $h_\mu(T) = h_\mu(T|_\mathcal{E}) + h_\mu(T|\mathcal{E})$, assuming that $h_\mu(T|\mathcal{E}) < \infty$.*

Proof. For any partitions $\mathcal{A} \subseteq \mathcal{F}$ and $\mathcal{B} \subseteq \mathcal{E}$ we have

$$\begin{aligned} h_\mu(T) &\geq h_\mu(\mathcal{A} \vee \mathcal{B}) \\ &= h_\mu(T, \mathcal{B}) + H(\mathcal{A}|\mathcal{A}_1^\infty \vee \mathcal{B}_{-\infty}^\infty) \\ &\geq h_\mu(T, \mathcal{B}) + H(\mathcal{A}|\mathcal{A}_1^\infty \vee \mathcal{E}) \\ &= h_\mu(T, \mathcal{B}) + h_\mu(T, \mathcal{A}|\mathcal{E}) \end{aligned}$$

This shows that $h_\mu(T) \geq h_\mu(T|\mathcal{E}) + h_\mu(T|_\mathcal{E})$. On the other hand, by choosing $\mathcal{A}$ fine enough we can ensure that $|h_\mu(T) - h_\mu(\mathcal{A} \vee \mathcal{B})|$ is arbitrarily small and likewise $|h_\mu(T, \mathcal{A}|\mathcal{E}) - h_\mu(T|\mathcal{E})|$. Also choosing $\mathcal{B}$ fine enough we can ensure that $|h_\mu(T, \mathcal{B}) - h_\mu(T|_\mathcal{E})|$ is arbitrarily small. Finally, we can choose $\mathcal{B}$ so that $|H(\mathcal{A}|\mathcal{A}_1^\infty \vee \mathcal{E}) - h_\mu(T, \mathcal{B}) + h_\mu(T, \mathcal{A}|\mathcal{E})|$ is arbitrarily small. This controls all the inequalities above and allows us to reverse them with an arbitrarily small error. This proves the claim. □

Corollary 11.3.8. *If $\mathcal{A}$ generates the system and $\mathcal{B}$ generates a factor then the relative entropy is $H(\mathcal{A}| \bigvee_{i=1}^\infty T^i\mathcal{A} \vee \bigvee_{i=-\infty}^\infty T^i\mathcal{B})$.*

11.4 The Pinsker algebra

Given a set $A \subseteq X$ let $\mathcal{P}_A = \{A, X \setminus A\}$.

Definition 11.4.1. The **Pinsker algebra** (really a σ-algebra) of a m.p.s. $(X, \mathcal{B}, \mu, T)$ is

$$\Pi = \{A \in \mathcal{F} \,:\, h_\mu(T, \mathcal{P}_A) = 0\}$$

Clearly Π is T-invariant.

Proposition 11.4.2. *Π is a σ-algebra.*

Proof. Let $A \in \sigma(\Pi)$. Thus there are $A_n \in \Pi$ with $A \in \sigma(A_1, A_2, \ldots)$ up to measure 0. Letting $\mathcal{B}_n = \bigvee_{i=1}^{n} \mathcal{P}_{A_i}$, we have

$$h_\mu(T, \mathcal{B}_n) \leq \sum_{i=1}^{n} h_\mu(T, \mathcal{P}_{A_i}) = 0$$

so $h_\mu(T, \mathcal{B}_n) = 0$. Also

$$h_\mu(T, \mathcal{A}|\mathcal{B}_n) \to h_\mu(T, \mathcal{A}| \bigvee_{i=1}^{\infty} \mathcal{P}_{A_i}) = h_\mu(T, \mathcal{A}|\Pi) = 0$$

Hence

$$0 \leq h_\mu(T, \mathcal{P}_A) \leq h_\mu(T, \mathcal{P}_A \vee \mathcal{B}_n) = h_\mu(T, \mathcal{B}_n) + h_\mu(T, \mathcal{A}|\mathcal{B}_n) \to 0$$

so $A \in \Pi$. □

11.5 The tail algebra and Pinsker's theorem

Definition 11.5.1. For a stationary process $\xi = (\xi_n)_{n=-\infty}^{\infty}$, the **tail algebras** (some called the remote past and future algebras)

$$\begin{aligned} \mathcal{T}^-(\xi) &= \bigcap_{n=1}^{\infty} \sigma(\ldots \xi_{-n-1}\xi_{-n}) \\ \mathcal{T}^+(\xi) &= \bigcap_{n=1}^{\infty} \sigma(\xi_n, \xi_{n+1}, \ldots) \end{aligned}$$

For a partition $\mathcal{A}$ we similarly write

$$\begin{aligned} \mathcal{T}^-(\mathcal{A}) &= \bigcap_{n\in\mathbb{N}} \bigvee_{i=-\infty}^{-n} T^{-i}\mathcal{A} \\ \mathcal{T}^+(\mathcal{A}) &= \bigcap_{n\in\mathbb{N}} \bigvee_{i=n}^{\infty} T^{-i}\mathcal{A} \end{aligned}$$

Theorem 11.5.2. *If $\mathcal{A}$ generates then $\Pi = \mathcal{T}^{\pm}(\mathcal{A})$.*

Proof. Let $\mathcal{B} \in \mathcal{T}^-(\mathcal{A})$. Since $\mathcal{A}$ generates we have

$$h_\mu(T) = h_\mu(T, \mathcal{A}) \leq h_\mu(T, \mathcal{A} \vee \mathcal{B}) \leq h_\mu(T)$$

and so

$$\begin{aligned} h_\mu(T) &= h_\mu(T, \mathcal{A} \vee \mathcal{B}) \\ &= h_\mu(T, \mathcal{B}) + h_\mu(T, \mathcal{A}|\mathcal{B}) \\ &= h_\mu(T, \mathcal{B}) + H_\mu(\mathcal{A}| \bigvee_{i=-\infty}^{-1} T^{-i}\mathcal{A} \vee \bigvee_{i=-\infty}^{\infty} T^{-i}\mathcal{B}) \\ &= h_\mu(T, \mathcal{B}) + H_\mu(\mathcal{A}| \bigvee_{i=-\infty}^{-1} T^{-i}\mathcal{A}) \\ &= h_\mu(T, \mathcal{B}) + h_\mu(\mathcal{A}) \end{aligned}$$

where in the last transition we used that $\mathcal{T}^-(\mathcal{A}) \subseteq \bigvee_{i=-\infty}^{-1} T^{-i}\mathcal{A}$ and $T^j\mathcal{B} \in \mathcal{T}^-(\mathcal{A})$ for all j, hence $\bigvee_{i=-\infty}^{\infty} T^{-i}\mathcal{B} \subseteq \mathcal{T}^-(\mathcal{A})$. Subtracting $h_\mu(T, \mathcal{A})$ from both sides gives $h_\mu(T, \mathcal{B}) = 0$.

Now suppose that $\mathcal{B} \in \Pi$. Then we again have, for every k,

$$\begin{aligned} H_\mu(\mathcal{A}| \bigvee_{i=-\infty}^{-1} T^{-i}\mathcal{A}) &= h_\mu(T) \\ &= h_\mu(T, \mathcal{B}) + H_\mu(\mathcal{A}| \bigvee_{i=-\infty}^{-1} T^{-i}\mathcal{A} \vee \bigvee_{i=-\infty}^{\infty} T^{-i}\mathcal{B}) \\ &= H_\mu(\mathcal{A}| \bigvee_{i=-\infty}^{-1} T^{-i}\mathcal{A} \vee \bigvee_{i=-\infty}^{\infty} T^{-i}\mathcal{B}) \\ &\leq H_\mu(\mathcal{A}| \bigvee_{i=-\infty}^{-1} T^{-i}\mathcal{A} \vee T^{-k}\mathcal{B}) \\ &\leq H_\mu(\mathcal{A}| \bigvee_{i=-\infty}^{-1} T^{-i}\mathcal{A}) \end{aligned}$$

so we have for all k,

$$H_\mu(\mathcal{A}| \bigvee_{i=-\infty}^{-1} T^{-i}\mathcal{A}) = H_\mu(\mathcal{A}| \bigvee_{i=-\infty}^{-1} T^{-i}\mathcal{A} \vee T^{-k}\mathcal{B})$$

An elementary calculation using the conditional entropy formula shows that this implies for all k that

$$H_\mu(T^{-k}\mathcal{B}| \bigvee_{i=-\infty}^{-1} T^{-i}\mathcal{A}) = H_\mu(T^{-k}\mathcal{B}| \bigvee_{i=-\infty}^{0} T^{-i}\mathcal{A})$$

or equivalently, for all k,

$$H_\mu(\mathcal{B}| \bigvee_{i=-\infty}^{k} T^{-i}\mathcal{A}) = H_\mu(\mathcal{B}| \bigvee_{i=-\infty}^{k+1} T^{-i}\mathcal{A})$$

Now, since $\mathcal{A}$ generates, we know that

$$\lim_{n\to\infty} H(\mathcal{B}|\bigvee_{i=-\infty}^{n} T^{-i}\mathcal{A}) = H(\mathcal{B}|\bigvee_{i=-\infty}^{\infty} T^{-i}\mathcal{A}) = 0$$

but since

$$H(\mathcal{B}|\bigvee_{i=-\infty}^{n} T^{-i}\mathcal{A}) = H(\mathcal{B}|\bigvee_{i=-\infty}^{n-1} T^{-i}\mathcal{A}) = \ldots = H(\mathcal{B}|\bigvee_{i=-\infty}^{-1} T^{-i}\mathcal{A})$$

we find that

$$\lim_{n\to\infty} H(\mathcal{B}|\bigvee_{i=-\infty}^{-1} T^{-i}\mathcal{A}) = 0$$

so $\mathcal{B} \in \bigvee_{i=-\infty}^{-1} T^{-i}\mathcal{A}$. The same argument shows that $\mathcal{B} \in \bigvee_{i=-\infty}^{-k} T^{-i}\mathcal{A}$ for all k, so $\mathcal{B} \in \mathcal{T}^{-}(\mathcal{A})$. □

Corollary 11.5.3. $\mathcal{T}^{+} = \mathcal{T}^{-}$.

11.6 Systems with completely positive entropy

Definition 11.6.1. CPE (K) systems

Definition 11.6.2. A system has uniform mixing if for every partition $\mathcal{P}$, $h_\mu(T^n, \mathcal{P}) \to H_\mu(\mathcal{P})$ as $n \to \infty$. In other words,

$$\sup_N \left(\frac{1}{N} H_\mu(\bigvee_{i=1}^{N} T^{-nN}\mathcal{P}) - H_\mu(\mathcal{P}) \right) = o(1) \qquad \text{as } n \to \infty$$

Theorem 11.6.3. *A system is CPE if and only if it has uniform mixing.*

Proof. If $h_\mu(T, Q) = 0$ then $h_\mu(T^n, Q) = 0$ for all n so there is no uniform mixing.

In the other direction if the system is CPE, then $\mathcal{T}^{-}(\mathcal{P}) \subseteq \Pi$ is trivial, so from the martingale theorem,

$$H_\mu(\mathcal{P}|\bigvee_{i=-\infty}^{-n} T^{-i}\mathcal{P}) \to H_\mu(P) \qquad \text{as } n \to \infty$$

since $\bigvee_{i=-\infty}^{-1} T^{-ni}\mathcal{P} \subseteq \bigvee_{i=-\infty}^{-n} T^{-i}\mathcal{P}$ we have

$$H_\mu(\mathcal{P}|\bigvee_{i=-\infty}^{-n} T^{-i}\mathcal{P}) \leq H_\mu(\mathcal{P}|\bigvee_{i=-\infty}^{-1} T^{-ni}\mathcal{P}) \leq H_\mu(P)$$

hence

$$h_\mu(T^n, \mathcal{P}) = H(\mathcal{P}|\bigvee_{i=-\infty}^{-1} T^{-ni}\mathcal{P}) \to H_\mu(P) \qquad \text{as } n \to \infty$$ □

Proposition 11.6.4. *If T is uniformly mixing (equivalently CPE) then for any partition $\mathcal{P}$ and any k,*

$$H(\bigvee_{i=0}^{k-1} T^{-in}\mathcal{P}) \to kH(\mathcal{P})$$

In particular, for any functions $f_0, \ldots, f_{k-1} \in L^\infty(\mu)$,

$$\int f_0(x) \cdot f_1(T^n x) \cdot f_3(T^{-2n}x) \cdot \ldots \cdot f_{k-1}(T^{(k-1)n}x) d\mu(x) \to \prod \int f_i d\mu$$

and in particular T is strongly mixing.

Proof sketch. Let us do it for $k = 2$. First one shows that it is enough to prove this for simple functions, hence for indicator functions. Let $f_i = 1_{A_i}$ and let $\mathcal{P}$ the partition determines by A_1, A_2. Now,

$$\begin{aligned} \int f_0(x) \cdot f_1(T^n x) d\mu(x) &= \int \mathbb{E}_\mu(f_0 \cdot T^n f_1 | T^n f_1) d\mu(x) \\ &= \int \mathbb{E}_\mu(f_0 | T^n f_1) \cdot T^n f_1 d\mu(x) \end{aligned}$$

becayse $\mathbb{E}(ab|\mathcal{E}) = a\mathbb{E}(b|\mathcal{E})$ if a is $\mathcal{E}$-measurable. Because $H_\mu(\mathcal{P}|T^{-n}\mathcal{P}) \to H_\mu(\mathcal{P})$, the partition $T^{-n}\mathcal{P}$ becomes asymptotically independent of $\mathcal{P}$, in the sense that if $\mu(T^{-n}B \cap C) \to \mu(C)\mu(B)$ for $B, C \in \mathcal{P}$ (this is an exercise in the definition of entropy). Since f_0 is $\mathcal{P}$ measurable and $T^n f_1$ is $T^{-n}\mathcal{P}$-measurable it follows that $\mathbb{E}_\mu(f_0|T^n f_1) \to \int f_0 d\mu$. Then by bounded convergence and invariance we get that the equation above tends to $\int f_0 d\mu \int f_1 d\mu$ as $n \to \infty$, as claimed. □

Chapter 12

Appendix

12.1 The weak-* topology

Recall that for a compact metric space X, the weak-* topology on $\mathcal{P}(X)$ (the space of Borel probability meaures on X) is the weakest topology which makes $\mu \mapsto \int f d\mu$ continuous for every $f \in C(X)$.

Proposition 12.1.1. *Let X be a compact metric space. Then $\mathcal{P}(X)$ is metrizable and compact in the weak-* topology.*

Proof. Let $\{f_i\}_{i=1}^{\infty}$ be a countable dense subset of the unit ball in $C(X)$. Define a metric on $\mathcal{P}(X)$ by

$$d(\mu,\nu) = \sum_{i=1}^{\infty} 2^{-i} |\int f_i d\mu - \int f_i d\nu|$$

It is easy to check that this is a metric. We must show that the topology induced by this metric is the weak-* topology.

If $\mu_n \to \mu$ weak-* then $\int f_i d\mu_n - \int f_i d\mu \to 0$ as $n \to \infty$, hence $d(\mu_n, \mu) \to 0$.

Conversely, if $d(\mu_n, \mu) \to 0$, then $\int f_i d\mu_n \to \int f_i d\mu$ for every i and therefore for every linear combination of the f_is. Given $f \in C(X)$ and $\varepsilon > 0$ there is a linear combination g of the f_i such that $\|f - g\|_\infty < \varepsilon$. Then

$$\begin{aligned} |\int f d\mu_n - \int f d\mu| &< |\int f d\mu_n - \int g d\mu_n| + |\int g d\mu_n - \int g d\mu| + |\int g d\mu - \int f d\mu| \\ &< \varepsilon + |\int g d\mu_n - \int g d\mu| + \varepsilon \end{aligned}$$

and the right hand side is $< 3\varepsilon$ when n is large enough. Hence $\mu_n \to \mu$ weak-*.

Since the space is metrizable, to prove compactness it is enough to prove sequential compactness, i.e. that every sequence $\mu_n \in \mathcal{P}(X)$ has a convergent subsequence. Let $V = \mathrm{span}_{\mathbb{Q}}\{f_i\}$, which is a countable dense $\mathbb{Q}$-linear subspace of $C(X)$. The range of each $g \in V$ is a compact subset of $\mathbb{R}$ (since X is compact and g continuous) so for each $g \in V$ we can choose a convergent subsequence of $\int g d\mu_n$. Using a diagonal argument we may select a single subsequence $\mu_{n(j)}$ such that $\int g\mu_{n(j)} \to \Lambda(g)$ as $j \to \infty$ for every $g \in V$. Now, Λ is a $\mathbb{Q}$-linear functional because

$$\begin{aligned}\Lambda(af_i + bf_j) &= k \lim \int (af_i + bf_j)\, d\mu_{n(k)} \\ &= \lim_{k\to\infty} a \int f_i d\mu_{n(k)} + b \int f_j d\mu_{n(k)} \\ &= a\Lambda(f_i) + b\Lambda(f_j)\end{aligned}$$

Λ is also uniformly continuous because, if $\|f_i - f_j\|_\infty < \varepsilon$ then

$$\begin{aligned}|\Lambda(f_i - f_j)| &= \left| \lim_{k\to\infty} \int (f_i - f_j)\, d\mu_{n(k)} \right| \\ &\leq \lim_{k\to\infty} \int |f_i - f_j| d\mu_{n(k)} \\ &\leq \varepsilon\end{aligned}$$

Thus Λ extends to a continuous linear functional on $C(X)$. Since Λ is positive (i.e. non-negative on non-negative functions), sos is its extension, so by the Riesz representation theorem there exists $\mu \in \mathcal{P}(X)$ with $\Lambda(f) = \int f d\mu$. By definition $\int g d\mu - \int g d\mu_{n(k)} \to 0$ as $k \to \infty$ for $g \in V$, hence this is true for the f_i, so $d(\mu_{n(k)}, \mu) \to 0$ Hence $\mu_{n(k)} \to \mu$ weak-* . □

Sometimes when $\mu_n \to \mu$ one would like to say that $\mu_n(E) \to \mu(E)$ for some set E. This is not always true. For example, $\delta_{1/n} \to \delta_0$ in $\mathbb{R}$ but $\delta_{1/n}(\{0\}) = 0 \not\to 1 = \delta_0(\{0\})$. But there are some general things that can be said, and if the set interacts nicely with the limit measure, the limit behaves "correctly":

Lemma 12.1.2. *Let X be a compact metric space. If $\mu_n \to \mu$ weak-* in $\mathcal{P}(X)$, and if $U \subseteq X$ is open and $C \subseteq X$ is closed, then*

$$\begin{aligned}\liminf \mu_n(U) &\geq \mu(U) \\ \limsup \mu_n(C) &\leq \mu(C)\end{aligned}$$

Proof. Let $f_k \in C(X)$ with $f_k \nearrow 1_U$ (e.g. $f_n(y) = 1 - e^{-kd(y,U^c)}$). Then $1_U \geq f_n$ and so

$$\liminf \mu_n(U) \geq \lim \int f_k d\mu_n = \int f_k d\mu \to \mu(U)$$

The other inequality is proves similarly using $g_n \searrow 1_C$. □

Proposition 12.1.3. *Let X be a compact metric space. If $\mu_n \to \mu$ weak-* in $\mathcal{P}(X)$ and if $A \subseteq X$ satisfies $\mu(\partial A) = 0$ then $\mu_n(A) \to \mu(A)$.*

Proof. Let $U = interior(A)$ and $C = \overline{A}$, so $1_U \leq 1_A \leq 1_C$. By the lemma,

$$\liminf \mu_n(A) \geq \liminf \mu_n(U) \geq \mu(U)$$

and

$$\limsup \mu_n(A) \leq \limsup \mu_n(C) \leq \mu(C)$$

But by our assumption, $\mu(U) = \mu(C) = \mu(A)$, and we find that

$$\mu(A) \leq \liminf \mu_n(A) \leq \limsup \mu_n(A) \leq \mu(A)$$

So all are equalities, and $\mu_n(A) \to \mu(A)$. □

12.2 Regularity

I'm not sure we use this anywherem, but for the record:

Lemma 12.2.1. *A Borel probability measure on a complete (seperable) metric space is regular.*

Proof. It is easy to see that the family of sets A with the propertty that

$$\begin{aligned} \mu(A) &= \inf\{\mu(U) \,:\, U \supseteq A \text{ is open}\} \\ &= \sup\{\mu(C) \,:\, C \subseteq A \text{ is closed}\} \end{aligned}$$

contains all open and closed sets, and is a σ-algebra. Therefore every Boral set A has this property. We need to verify that in the second condition we can repace closed by compact. Clearly it is enough to show that for every closed set C and every $\varepsilon > 0$ there is a compact $K \subseteq C$ with $\mu(K >> \mu(C) - \varepsilon$.

Fix C and $\varepsilon > 0$. For every n we can find a finite family $B_{n,1}, \ldots, B_{n,k(n)}$ of δ-balls whose union $B_n = \bigcup B_{n,i}$ intersects A in a set of measure $> \mu(A) - \varepsilon/2^n$. Let $K_0 = C \cap \bigcap B_n$, so that $\mu(K_0) > \mu(C) - \varepsilon$. By construction K_0 is pre-compact, and $K = \overline{K_0} \subseteq C$, so K has the desired property. □

12.3 Conditional expectation

Whe (X, B, μ) is a probability space, $f \in L^1$, and A a set of positive measure, then the conditional expectation of f on A is usually defined as $\frac{1}{\mu(A)} \int_A f \, d\mu$. When A has measure 0 this formula is meaningless, and it is not clear how to give an alternative definition. But if $\mathcal{A} = \{A_i\}_{i \in I}$ is a partition of X into measurable sets (possibly of measure 0), one can sometimes give a meaningful definition of the conditional expectation of f on $\mathcal{A}(x)$ for a.e. x, where $\mathcal{A}(x)$ is the element A_i containing x. Thus the conditional expectation of$f$ on $\mathcal{A}$ is a function that assigns to a.e. x the conditional expectation of f on the set $\mathcal{A}(x)$. Rather than partitions, we will work with σ-algebras; the connection is made by observign that if $\mathcal{E}$ is a countably-generated σ-algebra then the partition of X into the atoms of $\mathcal{E}$ is a measurable partition.

Theorem 12.3.1. *Let $(X, \mathcal{B}, \mu)$ be a probability space and $\mathcal{E} \subseteq \mathcal{B}$ a sub-σ algebra. Then there is a linear operator $L^1(X, \mathcal{B}, \mu) \to L^1(X, \mathcal{E}, \mu)$ satisfying*

1. *Chain rule: $\int \mathbb{E}(f|\mathcal{E})\, d\mu = \int f\, d\mu$.*
2. *Product rule: $\mathbb{E}(gf|\mathcal{E}) = g \cdot \mathbb{E}(f|\mathcal{E})$ for all $g \in L^\infty(X, \mathcal{E}, \mu)$.*

Proof. We begin with existence. Let $f \in L^1(X, \mathcal{B}, \mu)$ and let μ_f be the finite signed measure $d\mu_f = f d\mu$. Then $\mu_f \ll \mu$ in the measure space $(X, \mathcal{B}, \mu)$ and this remains true in $(X, \mathcal{E}, \mu)$. Let $\mathbb{E}(f|\mathcal{E}) = d\mu_f/d\mu \in L^1(X, \mathcal{E}, \mu)$, the Radon-Nykodim derivative of μ_f with respect to μ in $(X, \mathcal{E}, \mu)$.

The domain of this map is $L^1(X, \mathcal{B}, \mu)$ and its range is in $L^1(X, \mathcal{E}, \mu)$ by the properties of $d\mu_f/d\mu$.

Linearity follows from uniqueness of the Radon-Nykodim derivative and the definitions. The chain rule is also immediate:

$$\int \mathbb{E}(f|\mathcal{E})\, d\mu = \int \frac{d\mu_f}{d\mu}\, d\mu = \int f\, d\mu$$

For the product rule, let $g \in L^\infty(X, \mathcal{E}, \mu)$. We must show that $g \cdot \frac{d\mu_f}{d\mu} = \frac{d\mu_{gf}}{d\mu}$ in $(X, \mathcal{E}, \mu)$. Equivalently we must show that

$$\int_E g \frac{d\mu_f}{d\mu}\, d\mu = \int_E \frac{d\mu_{gf}}{d\mu}\, d\mu \qquad \text{for all } E \in \mathcal{E}$$

Now, for $A \in \mathcal{E}$ and $g = 1_A$ we have

$$\begin{aligned}
\int_E 1_A \frac{d\mu_f}{d\mu}\, d\mu &= \int_{A\cap E} \frac{d\mu_f}{d\mu}\, d\mu \\
&= \mu_f(A \cap E) \\
&= \int_{A\cap E} f\, d\mu \\
&= \int_E 1_A f\, d\mu \\
&= \int_E \frac{d\mu_{1_A f}}{d\mu}\, d\mu
\end{aligned}$$

so the identity holds. By linearity of these integrals in the g argument it holds linear combinations of indicator functions. For arbitrary $g \in L^\infty$ we can take a uniformly bounded sequence of such functions converging pointwise to g, and pass to the limit using dominated convergence. This proves the product rule.

To prove uniqueness, let $T : L^1(X, \mathcal{B}, \mu) \to L^1(X, \mathcal{E}, \mu)$ be an operator with these properties. Then for $f \in L^1(X, \mathcal{B}, \mu)$ and $E \in \mathcal{E}$,

$$\begin{aligned} \int_E Tf \, d\mu &= \int 1_E Tf \, d\mu \\ &= \int T(1_E f) \, d\mu \\ &= \int 1_E f \, d\mu \\ &= \int_E f \, d\mu \end{aligned}$$

where the second equality uses the product rule and the third uses the chain rule. Since this holds for all $E \in \mathcal{E}$ we must have $Tf = d\mu_f / d\mu$. □

Proposition 12.3.2. *The conditional expectation operator satisfies the following properties:*

1. *Positivity:* $f \geq 0$ *a.e. implies* $\mathbb{E}(f|\mathcal{E}) \geq 0$ *a.e.*
2. *Triangle inequality:* $|\mathbb{E}(f|\mathcal{I})| \leq \mathbb{E}(|f|\,|\mathcal{I})$.
3. *Contraction:* $\|\mathbb{E}(f|\mathcal{E})\|_1 \leq \|f\|_1$*; in particular,* $\mathbb{E}(\cdot|\mathcal{E})$ *is* L^1*-continuous.*
4. *Sup/inf property:* $\mathbb{E}(\sup f_i|\mathcal{E}) \geq \sup \mathbb{E}(f_i|\mathcal{E})$ *and* $\mathbb{E}(\inf f_i|\mathcal{E}) \leq \inf \mathbb{E}(f_i|\mathcal{E})$ *for any countable family* $\{f_i\}$.
5. *Jensen's inequality: if* g *is convex then* $g(\mathbb{E}(f|\mathcal{E})) \leq \mathbb{E}(g \circ f|\mathbb{E})$.
6. *Fatou's lemma:* $\mathbb{E}(\liminf f_n|\mathcal{E}) \leq \liminf \mathbb{E}(f_n|\mathcal{E})$.

Remark 12.3.3. Properties (2)–(6) are consequences of positivity only.

Proof. (1) Suppose $f \geq 0$ and $\mathbb{E}(f|\mathcal{E}) \not\geq 0$, so $\mathbb{E}(f|\mathcal{E}) < 0$ on a set $A \in \mathcal{E}$ of positive measure. Applying the product rule with $g = 1_A$, we have

$$\mathbb{E}(1_A f|\mathcal{E}) = 1_A \mathbb{E}(f|\mathcal{E})$$

hence, replacing f by 1_A, we can assume that $f \geq 0$ and $\mathbb{E}(f|\mathcal{E}) < 0$. But this contradicts the chain rule since $\int f \, d\mu \geq 0$ and $\int \mathbb{E}(f|\mathcal{E}) \, d\mu < 0$.

(2) Decompose f ionto positive and negative parts, $f = f^+ - f^-$, so that $|f| = f^+ + f^-$. By positivity,

$$\begin{aligned} |\mathbb{E}(f|\mathcal{E})| &= |\mathbb{E}(f^+|\mathcal{E}) - \mathbb{E}(f^-|\mathcal{E})| \\ &\leq |\mathbb{E}(f^+|\mathcal{E})| + |\mathbb{E}(f^-|\mathcal{E})| \\ &= \mathbb{E}(f^+|\mathcal{E}) + \mathbb{E}(f^-|\mathcal{E}) \\ &= \mathbb{E}(f^+ + f^-|\mathcal{E}) \\ &= \mathbb{E}(|f|\,|\mathcal{E}) \end{aligned}$$

(3) We compute:

$$\begin{aligned}\|\mathbb{E}(f|\mathcal{E})\|_1 &= \int |\mathbb{E}(f|\mathcal{E})|\, d\mu \\ &\leq \int \mathbb{E}(|f|\,|\mathcal{E})|\, d\mu \\ &= \int |f|\, d\mu \\ &= \|f\|_1\end{aligned}$$

where we have used the triangle inequality and the chain rule.

(4) We prove the sup version. By monotonicity and continuity it suffices to prove this for finite familis and hence for two functions. The claim now follows from the identity $\max\{f_1, f_2\} = \frac{1}{2}(f_1 + f_2 + |f_1 - f_2|)$, linearity, and the triangle inequality.

(5) For an affine function $g(t) = at + b$,

$$\mathbb{E}(g \circ f|\mathcal{E}) = \mathbb{E}(af + b|\mathcal{E}) = a\mathbb{E}(f|\mathcal{E}) + b = g \circ \mathbb{E}(f|\mathcal{E})$$

If g is convex then $g = \sup g_i$ where $\{g_i\}_{i \in I}$ is a countable family of affine functions. Thus

$$\begin{aligned}\mathbb{E}(g \circ f|\mathcal{E}) &= \mathbb{E}(\sup_i g_i \circ f|\mathcal{E}) \\ &\geq \sup_i \mathbb{E}(g_i \circ f|\mathcal{E}) \\ &= \sup_i g_i \circ \mathbb{E}(f|\mathcal{E}) \\ &= g \circ \mathbb{E}(f|\mathcal{E})\end{aligned}$$

(6) Since $\inf_{k>n} f_k \nearrow \liminf f_k$ as $n \to \infty$ the convergence is also in L^1, so by continuity and positivity the same holds after taking the conditional expectation. Thus, using the inf property,

$$\begin{aligned}\liminf_{n\to\infty} \mathbb{E}(f_n|\mathcal{E}) &= \lim_{n\to\infty} \inf_{k>n} \mathbb{E}(f_k|\mathcal{E}) \\ &\geq \lim_{n\to\infty} \mathbb{E}(\inf_{k>n} f_k|\mathcal{E}) \\ &= \mathbb{E}(\liminf_{n\to\infty} f_n|\mathcal{E}) \quad \square\end{aligned}$$

Corollary 12.3.4. *The restriction of the conditional expectation operator to $L^2(X, \mathcal{B}, \mu)$ coincides with the orthogonal projection $\pi : L^2(X, \mathcal{B}, \mu) \to L^2(X, \mathcal{E}, \mu)$.*

Proof. Write $\pi = \mathbb{E}(\cdot|\mathcal{E})$. If $f \in L^2$ then by by convexity of $t \to t^2$ and Jensen's inequality (which is immediate for simple functions and hence holds for $f \in L^1$

by approximation),

$$\begin{aligned}
\|\pi f\|_2 &= \int |\mathbb{E}(f|\mathcal{E})|^2 \, d\mu \\
&\leq \int \mathbb{E}(|f|^2|\mathcal{E}) \, d\mu \\
&= \int |f|^2 \, d\mu && \text{by the chain rule} \\
&= \|f\|_2
\end{aligned}$$

Thus π maps L^2 into the subspacfe of $\mathcal{E}$-measurable L^2 functions, hence $\pi : L^2(X,\mathcal{B},m) \to L^2(X,\mathcal{E},\mu)$. We will now show that π is the identity on $L^2(X,\mathcal{E},\mu)$ and is π. Indeed, if $g \in L^2(X,\mathbb{E},\mu)$ then for every $A \in \mathcal{E}$

$$\begin{aligned}
\pi g &= \mathbb{E}(g \cdot 1|\mathcal{E}) \\
&= g \cdot \mathbb{E}(1|\mathcal{E})
\end{aligned}$$

Since $\int \mathbb{E}(1|\mathcal{E}) = \int 1 = 1$, this shows that π is the identity on $L^2(X,\mathcal{E},)$. Next if $f,g \in L^2$ then $fg \in L^1$, and

$$\begin{aligned}
\langle f, \pi g\rangle &= \int f \cdot \mathbb{E}(g|\mathcal{E}) \, d\mu \\
&= \int \mathbb{E}\left(f \cdot \mathbb{E}(g|\mathcal{E})\right) \, d\mu && \text{by the chain rule} \\
&= \int \mathbb{E}\left(f|\mathcal{E}\right) \mathbb{E}(g|\mathcal{E}) \, d\mu && \text{by the product rule} \\
&= \int \mathbb{E}\left(\mathbb{E}(f|\mathcal{E}) \cdot g\right) \, d\mu && \text{by the product rule} \\
&= \int \mathbb{E}(f|\mathcal{E}) \cdot g \, d\mu && \text{by the chain rule} \\
&= \langle \pi f, g\rangle
\end{aligned}$$

so π is self-adjoint. □

12.4 Measure disintegration

We give a detailed proof of the following Theorem, which appeared in Section 5.3:

Theorem 12.4.1. *Let X be compact metric space, $\mathcal{B}$ the Borel algebra, and $\mathcal{E} \subseteq \mathcal{B}$ a countably generated sub-σ-algebra. Then there is an $\mathcal{E}-$measurable family $\{\mu_y\}_{y\in X} \subseteq \mathcal{P}(X)$ such that μ_y is supported on $\mathcal{E}(y)$ and*

$$\mu = \int \mu_y \, d\mu(y)$$

Furthermore if $\{\mu'_y\}_{y\in X}$ is another such system then $\mu_y = \mu'_y$ a.e.

We begin the proof.

We adopt the convention that y denotes the variable of $\mathcal{E}$-measurable functions.

Let $V \subseteq C(X)$ be a countable dense $\mathbb{Q}$-linear subspace with $1 \in V$. For $f \in V$ let

$$\overline{f} = \mathbb{E}(f|\mathcal{E})$$

(see the Appendix for a discussion of conditional expectation). Since V is countable there is a subset $X_0 \subseteq X$ of full measure such that $\overline{f}$ is defined everywhere on X_0 for $f \in V$ and $f \mapsto \overline{f}$ is $\mathbb{Q}$-linear and positive on X_0, and $\overline{1} = 1$ on X_0. Thus, for $y \in X_0$ the functions $\Lambda_y : V \to \mathbb{R}$ given by

$$\Lambda_y(f) = \overline{f}(y)$$

are positive $\mathbb{Q}$-linear functionals on the normed space $(V, \|\cdot\|_\infty)$, and they are continuous, since by positivity of conditional expectation $\|\overline{f}\|_\infty \leq \|f\|_\infty$. Thus Λ_y extends to a positive $\mathbb{R}$-linear functional$\Lambda_y : C(X) \to \mathbb{R}$. Note that $\Lambda_y 1 = \overline{1}(y) = 1$. Hence, by the Riesz representation theorem, there exists $\mu_y \in \mathcal{P}(X)$ such that

$$\Lambda_y f = \int f(x)\, d\mu_y(x)$$

For $y \in X \setminus X_0$ define μ_y to be some fixed measure to ensure measurability.

Proposition 12.4.2. *$y \to \mu_y$ is $\mathcal{E}$-measurable and $\mathbb{E}(1_A|\mathcal{E})(y) = \mu_y(A)$ μ-a.e., for every $A \in \mathcal{B}$.*

Proof. Let $\mathcal{A} \subseteq \mathcal{B}$ denote the family of sets $A \in \mathcal{B}$ such that $y \mapsto \mu_y(A)$ measurable from $(X, \mathcal{E})$ to $(X, \mathcal{B})$ and $\mathbb{E}(1_A|\mathcal{E})(y) = \mu_y(A)$ μ-a.e. We want to show that $\mathcal{A} = \mathcal{B}$.

Let $\mathcal{A}_0 \subseteq \mathcal{B}$ denote the family of sets $A \subseteq X$ such that 1_A is a pointwise limit of a uniformly bounded sequence of continuous functions. First, $\mathcal{A}_0$ is an algebra: clearly $X, \emptyset \in \mathcal{A}$, if $f_n \to 1_A$ then $1 - f_n \to 1_{X \setminus A}$, and if also $g_n \to 1_B$ then $f_n g_n \to 1_A 1_B = 1_{A \cap B}$.

We claim that $\mathcal{A}_0 \subseteq \mathcal{A}$. Indeed, if $f_n \to 1_A$ and $\|f_n\|_\infty \leq C$ then

$$\int f_n\, d\mu_y \to \int 1_A\, d\mu_y = \mu_y(A)$$

by dominated convergence, so $y \mapsto \mu_y(A)$ is the pointwise limit of the functions $y \mapsto \int f_n\, d\mu_y$, which are the same a.e. as the measurable functions $\overline{f}_n = \mathbb{E}(f_n|\mathcal{E}) : (X, \mathcal{E}) \to (X, \mathcal{B})$. This establishes measurability of the limit function $y \mapsto \mu_y(A)$ and also proves that this function is $\mathbb{E}(1_A|\mathcal{E})$ a.e., since $\mathbb{E}(\cdot|\mathcal{E})$ is continuous in L^1 and $f_n \to 1_A$ boundedly. This proves $\mathcal{A}_0 \subseteq \mathcal{A}$.

Now, $\mathcal{A}_0$ contains the closed sets, since if $A \subseteq X$ then $1_A = \lim f_n$ for $f_n(x) = \exp(-n \cdot d(x, A))$. Thus $\mathcal{A}_0$ generates the Borel σ-algebra $\mathcal{B}$.

Finally, we claim that $\mathcal{A}$ is a monotone class. Indeed, if $A_1 \subseteq A_2 \subseteq \ldots$ belong to $\mathcal{B}'$ and $A = \bigcup A_n$, then $\mu_y(A) = \lim \mu_y(A_n)$, and so $y \mapsto \mu_y(A)$ is the pointwise limit of the measurable functions $y \mapsto \mu_y(A_n)$. The latter functions are just $\mathbb{E}(1_{A_n}|\mathcal{E})$ and, since $1_{A_n} \to 1_A$ in L^1, by continuity of conditional expectation, $\mathbb{E}(1_{A_n}|\mathcal{E}) \to \mathbb{E}(1_A|\mathcal{E})$ in L^1. Hence $\mu_y(A) = \mathbb{E}(1_A|\mathcal{E})$ a.e. as desired.

Since $\mathcal{A}$ is a monotone class containing the sub-algebra of $\mathcal{A}_0$ and $\mathcal{A}_0$ generates $\mathcal{B}$, by the monotone class theorem we have $\mathcal{B} \subseteq \mathcal{A}$. Thus $\mathcal{A} = \mathcal{B}$, as desired. □

Proposition 12.4.3. *$\mathbb{E}(f|\mathcal{E})(y) = \int f\,d\mu_y$ μ-a.e. for every $f \in L^1(\mu)$.*

Proof. We know that this holds for $f = 1_A$ by the previous proposition. Both sides of the claimed equality are linear and continuous under monotone increasing sequences. Approximating by simple functions this gives the claim for positive $f \in L^1$ and, taking differences, for all $f \in L^1$. □

Proposition 12.4.4. *μ_y is μ-a.s. supported $\mathcal{E}(y)$, that is, $\mu_y(\mathcal{E}(y)) = 1$ ν-a.e.*

Proof. For $E \in \mathcal{E}$ we have

$$1_E(y) = \mathbb{E}(1_E|\mathcal{E})(y) = \int 1_E\,d\mu_y = \mu_y(E)$$

and it follows that $\mu_y(E) = 1_E(y)$ a.e. Let $\{E_n\}_{n=1}^{\infty}$ generate $\mathcal{E}$, and choose a set of full measure on which the above holds for all $E = E_n$. For y in this set let $F_n \in \{E_n, X \setminus E_n\}$ be such that $\mathcal{E}(y) = \bigcap F_n$. By the above $\mu_y(F_n) = 1$, and so $\mu_y(\mathcal{E}(y)) = 1$, as claimed. □

Proposition 12.4.5. *If $\{\mu'_y\}_{y\in Y}$ is another family with the same properties then $\mu'_y = \mu_y$ for μ-a.e. y.*

Proof. For $f \in L^1(\mu)$ define $f'(y) = \int f\,d\mu'_y$. This is clearly a linear operator defined on $L^1(X, \mathcal{B}, \mu)$, and its range is $L^1(X, \mathcal{E}, \mu)$ because

$$\int |f'|\,d\mu \le \int (\int |f|\,d\mu_y)\,d\mu(y) = \int |f|\,d\mu = \|f\|_1$$

The same calculation shows that $\int f'\,d\mu = \int f\,d\mu$. Finally, for $E \in \mathcal{E}$ we know that μ_y is supported on E for μ-a.e. $y \in E$ and on $X \setminus E$ for μ-a.e. $y \in X \setminus E$. Thus μ-a.s. we have

$$(1_E f)'(y) = \int 1_E f\,d\mu'_y = 1_E(y) \int f\,d\mu'_y = 1_E \cdot f'$$

By a well-known characterization of conditional expectation, $f' = \mathbb{E}(f|\mathcal{E}) = \overline{f}$. □