ERGODIC THEORY AND DYNAMICS - NOTES, WORKSHEETS, AND PROBLEM SETS

Contents

1.	Examples and Poincare recurrence	2
2.	Ergodicity	6
3.	The Birkhoff Ergodic Theorem	10
4.	The Riesz Representation Theorem	13
5.	Banach-Alaoglu, Markov-Kakutani and Krylov-Bogolyubov	19
6.	Grassmanians, Furstenberg's Theorem on Projective Actions, and Borel Density	24
7.	Circle homeomorphisms, minimality, rotation number, and Poincare's classification	30
8.	Conditional expectation and conditional measures	35
9.	The ergodic decomposition and disintegration	39

Description of the contents. The following is the course material for Math 758 taught at the University of Wisconsin - Madison in Spring 2024. The structure of each 75 minute class was $5-10$ minutes of review, $40-45$ minutes of lecture, and 20 minutes of collaboration on an in-class worksheet, whose solutions were discussed at the end of class. Each of the sections correspond to a single class, beginning with lecture notes, and ending with the in-class worksheet. The problem sets that were assigned are included after even-numbered lectures.

1. Examples and Poincare recurrence

A *measurable space* X is a set together with a *sigma algebra* of "measurable sets" (i.e. a collection of sets containing the empty set and closed under complements and countable intersections and unions). If X is a topological space, then the *Borel sigma algebra* is the smallest sigma algebra containing all open sets. A measurable (resp. topological) *dynamical system* is a measurable (resp. topological) space X and a measurable (resp. continuous) map $f : X \longrightarrow X$ (i.e. preimages of measurable (resp. open) sets are measurable (resp. open)). A system is called *measure-preserving* if there is a measure μ on X so that, for any set U, $\mu(f^{-1}(U)) = \mu(U)$. The data (X, f, μ) is called a *measure-preserving system (m.p.s)* and, if $\mu(X) = 1$, a *p.m.p.s.*

Before examples, we recall the following (a vast generalization of which we'll see later on). Suppose that G is a *topological group*, i.e. a topological space that is also a group so that group multiplication $m : G \times G \longrightarrow G$ and group inversion $\iota : G \longrightarrow G$ given by $m(g, h) = gh$ and $\iota(g) = g^{-1}$ are continuous. A measure μ on G is called *(left) translation-invariant* if for any measurable subset A and any $g \in G$, $\mu(gA) = \mu(A)$. We will use the following.

Proposition 1. *Up to scaling, volume (i.e. Lebesgue measure) is the only translation invariant measure on* $\mathbb{R}^n$ *and* $\mathbb{T}^n := \mathbb{R}^n/\mathbb{Z}^n$.

The proposition shows that the following dynamical systems preserve Lebesgue measure.

Examples of Measure Preserving Systems.

(1) Circle rotations, i.e. $f(x) = x + \alpha \bmod 1$ where $\alpha \in \mathbb{R}$.
(2) Translations on tori, i.e. $f(x_1, \ldots, x_n) = (x_1 + \alpha_1, \ldots, x_n + \alpha_n) \bmod 1$ where $\alpha_i \in \mathbb{R}$.
(3) Translations on $\mathbb{R}^n$, i.e. $f(x) = x + v$ where $v \in \mathbb{R}^n$.

More Examples of Measure Preserving Systems.

(1) Circle doubling, e.g. $f(x) = 2x \bmod 1$.
(2) Toral Automorphisms, i.e. the self-map on $\mathbb{R}^n/\mathbb{Z}^n$ induced by a linear map $A \in \mathrm{Mat}_{n\times n}(\mathbb{Z})$.
(3) Linear maps of $\mathbb{R}^n$ to itself under matrices with determinant 1.

We will explain now why the first two maps preserve Lebesgue measure. Recall that if $f : X \longrightarrow Y$ is measurable, then the *pushforward* of a measure μ on X is defined by $f_*\mu(U) := \mu(f^{-1}(U))$ for any measurable subset U of Y.

Lemma 2. *A measure ν on Y is $f_*\mu$ if and only if for any function g in $L^1(Y, f_*\mu)$*

$$\int_X g \circ f d\mu = \int_Y g d\nu$$

Proof. The reverse direction is immediate since it holds for any indicator function. For the forward direction, by assumption, the formula holds when g is an indicator function. So it holds for finite sums of indicator functions. These are dense in $L^1(Y, f_*\mu)$ so we use dominated convergence to conclude. ☐

Note that if $f : X \longrightarrow X$ is measurable, then a measure μ is f-invariant if and only if $f_*\mu = \mu$.

Corollary 3. *If $A \in \mathrm{Mat}_{n\times n}(\mathbb{Z})$ has nonzero determinant and $f : \mathbb{T}^n \longrightarrow \mathbb{T}^n$ is the map it induces on the torus, the f preserves Lebesgue measure μ.*

Proof. We must show that $f_*\mu = \mu$, so it suffices to that $f_*\mu$ is translation-invariant. Let $v \in \mathbb{T}^n$. Since A has nonzero determinant, f is a surjection and so there is some $w \in \mathbb{T}^n$ so that $f(w) = v$. If U is any measurable subset of $\mathbb{T}^n$, then

$$f_*\mu(U + v) = \mu(f^{-1}(U + v)) = \mu(f^{-1}(U) + w) = \mu(f^{-1}(U)) = f_*\mu(U).$$

☐

Given a measurable dynamical system, (X, f) and a measurable set U a point p in U *recurs to* U if it returns to U infinitely often under iterates of f. Given a topological dynamical system (X, f), a point p is called *recurrent* if p recurs to any open set containing it.

Theorem 4 (Measurable Poincare Recurrence)**.** *If (X, f, μ) is a p.m.p.s. and U is a measurable set, then almost every point in U is U-recurrent.*

Proof. Let B be the set of all points in U that never return to U. Formally, $B = U \cap \bigcap_{n\geq 1} f^{-n}(X\backslash U)$. So $f^{-n}(B)$ consists of points p so that $f^n(p) \in U$ but $f^k(p) \notin U$ for $k > n$. So $f^{-n}(B) \cap f^{-m}(B) = \emptyset$ for $n \neq m$. Since X contains $\bigcup_{n>0} f^{-n}(B)$, each summand of which has equal measure, B has measure zero. Therefore, the set of points F_k in U that return to U under some iterate of f^k is full measure in U. Therefore, $\bigcap_{k\geq 0} F_k$ is too. This set consists of points that return to U infinitely often. ☐

Corollary 5 (Topological Poincare Recurrence)**.** *Almost every point in a second countable topological p.m.p.s. is recurrent.*

Proof. Let (U_n) be a countable basis and let B_n be the measure zero subset of U_n consisting of non-recurrent points. The union $\bigcup_{n\geq 0} B_n$ is measure zero and so its complement is full measure and consists of recurrent points. $\square$

Worksheet for Lecture 1

Problem 1. Let X be the open unit disk in $\mathbb{C}$ and let $f : X \longrightarrow X$ be given by $f(z) = z^2$. Use Poincare recurrence to show that the only f-invariant probability measure is the delta measure at 0.

Problem 2. Let α be a real number. Show that the collection of points in $[0, 1)$ given by $\{n\alpha \bmod 1\}_{n\geq 0}$ is dense if and only if α is irrational. (Hint: For α irrational, it suffices to show that the sequence gets arbitrarily close to 0. Use the fact that, for any $\epsilon > 0$, there must be two distinct integers n and m so that $n\alpha \bmod 1$ and $m\alpha \bmod 1$ are ϵ-close.)

Problem 3. Let $X = \mathbb{R}$ and let $f(x) = x+1$. Use Poincare recurrence to show that there are no finite f-invariant probability measures. Show that f induces a map $\widetilde{f}$ from the one point compacfication of $\mathbb{R}$ (i.e. the circle S^1) to itself and find all $\widetilde{f}$-invariant measures on S^1.

Problem 4. Let X be the unit circle in $\mathbb{R}^2$. Let A be a 2×2 matrix with real entries, determinant 1, and which does not have finite order. Define $f : X \longrightarrow X$ by $f(v) = \frac{A(v)}{|A(v)|}$. Find all f-invariant measures. (Hint: any 2×2 real matrix with determinant 1 is similar to one of the following: a rotation matrix, a diagonal matrix, or $\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$ for some $t \in \mathbb{R}$. In the rotation matrix case, you may use the fact (which follows from Problem 2) that any measure on the circle that is invariant under an irrational rotation is invariant under *every* rotation.)

2. Ergodicity

An m.p.s (X, T, μ) is called *ergodic* if the only T-invariant measurable sets are null or conull. Equivalently, the only almost-invariant measurable sets are null or conull (*almost-invariant* means that the symmetric difference of a set and its preimage is measure zero). Non-examples include rational rotations of the circle and the action of $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ on $\mathbb{R}^2/\mathbb{Z}^2$.

Lemma 6. *A p.m.p.s. (X, T, μ) is ergodic if and only if for any two positive measure sets A and B there is some n so that $T^{-n}(A) \cap B$ has positive measure.*

Proof. For the forward direction, we note that $\bigcup_n T^{-n}(A)$ is almost-invariant (it is contained in itself under preimages and the preimage has the same measure) and hence full measure. For the reverse direction, if A is an invariant positive measure set, then it never intersects $X - A$, so $X - A$ must be null. $\square$

Lemma 7. *Let $T : \mathbb{T}^n \longrightarrow \mathbb{T}^n$ be given by $T(x) = x + v$ where $v \in \mathbb{T}^n$. If $\{mv\}_{m \geq 0}$ is dense in $\mathbb{T}^n$, then T is ergodic with respect to Lebesgue measure.*

Proof. Let A and B be two positive measure subsets. Find density points $a \in A$ and $b \in B$, these are points so that for some $\epsilon > 0$, the ball B_a (resp. B_b) of radius ϵ around a (resp. b) has the property that 90% of the points in it are contained in A (resp. B). There is $\delta(n, \epsilon) > 0$ so that if two balls of radius ϵ in $\mathbb{R}^n$ have δ-close centers, then their intersection contains at least 99% of the points in both balls. By assumption, there is some m so that $a + mv$ is δ-close to b. So $T^m(B_a) \cap B_b$ contains 99% of points in both balls. In particular, at least 89% of its points belong to $T^m(A)$ and at least 89% belong to B. Therefore, at least 78% of points in the intersection belong to $T^m(A) \cap B$, so $\mu(T^m(A) \cap B) \geq (.78)(.99)\mu(B_a) > 0$. $\square$

Remark 8. A good exercise is showing that $\{mv\}_{m \geq 0}$ is dense in $\mathbb{T}^n$ if and only if the smallest closed subgroup of $\mathbb{T}^n$ containing v is $\mathbb{T}^n$ itself. The classification of closed subgroups of $\mathbb{T}^n$ then shows that $\{mv\}_{m \geq 0}$ is dense if and only if there is no nonzero $w \in \mathbb{Z}^n$ so that $w \cdot v = 0$ mod 1.

Lemma 9. *Fix $p \in [1, \infty]$. A p.m.p.s. (X, T, μ) is ergodic if and only if the only T-invariant function in $L^p(X, \mu)$ is the constant function.*

Proof. For the reverse direction, if $A \subseteq X$ is T-invariant, then so is 1_A, which must be constant a.e. and hence equal a.e. to 0 or 1. For the forward direction, if $f : X \longrightarrow \mathbb{R}$ is nonconstant and invariant then there must be some number c so that $\{x : f(x) > c\}$ and $\{x : f(x) < c\}$ have positive measure. Both must be invariant contradicting ergodicity. $\square$

Lemma 10. *Let T be the action induced by an invertible integral $n \times n$ matrix A on $\mathbb{T}^n := \mathbb{R}^n/\mathbb{Z}^n$. Let μ be Lebesgue measure on $\mathbb{T}^n$. Then T is ergodic if and only if A does not have a root of unity as an eigenvalue.*

Proof. Note that the characters of X are the continuous homomorphisms $\chi : X \longrightarrow \mathbb{C}^\times$. These are all of the form $\chi_m := \chi(x_1, \ldots, x_n) = \exp\left(2\pi i \left(m_1 x_1 + \ldots + m_n x_n\right)\right)$ where $m := (m_1, \ldots, m_n) \in \mathbb{Z}^n$. Fourier theory tells us that if f belongs to $L^2(X)$ then $f = \sum_{m \in \mathbb{Z}^n} c_m \chi_m$ where $c_m := \int_X f \overline{\chi_m}$ and that $\|f\|_2^2 = \sum_m |c_m|^2$. T is ergodic if and only if the only T-invariant function in L^2 is the constant function. Note that

$$\chi_m(Ax) = \exp(2\pi i m \cdot Ax) = \exp(2\pi i (A^T m) \cdot x) = \chi_{A^T m}(x).$$

Therefore, a function f is T-invariant means that $c_m = c_{mA^T} = c_{m(A^T)^2} = \ldots$. Therefore T is ergodic if and only if 0 is the only point in $\mathbb{Z}^n$ that is periodic under the action of A^T. This occurs if and only if $(A^T)^k - I$ has no kernel for any $k > 0$. $\square$

Worksheet for Lecture 2

Problem 1. Show that for almost every number $x \in [0,1)$ with a decimal expansion $0.a_0a_1a_2\dots$ the asymptotic fraction of the a_i that are equal to 7 is $\frac{1}{10}$. (Hint: Let $X = [0,1)$ and consider $T : X \longrightarrow X$ given by $T(x) = 10x \pmod 1$. Now apply the ergodic theorem to an appropriately chosen indicator function.)

Problem 2. The following is a version of the law of large numbers. Suppose that we flip a fair coin and record a 0 whenever heads appears and a 1 whenever tails appears. If S_n is the average of all numbers we have recorded from flip 1 to flip n, then S_n converges to $\frac{1}{2}$ almost surely. Show that this is immediate from the ergodic theorem. (Hint: Adapt the method considered in the previous problem by interpreting the space of outcomes $\{0,1\}^{\mathbb{N}}$ as the set of binary expansions of numbers in $[0,1)$.).

Problem 3. Let $T : S^1 \longrightarrow S^1$ be an irrational rotation of the circle and let μ be Lebesgue measure (i.e. length). Show that for *every* (not just almost-every) point $x \in S^1$ and every continuous function $f : S^1 \longrightarrow \mathbb{R}$, $\frac{f(x)+\ldots+f(T^{n-1}x)}{n} \longrightarrow \int_X f d\mu$.

Problem 4. Let $T : S^1 \longrightarrow S^1$ be the rotation of the circle by 90 degrees. Find an ergodic T-invariant measure (we have seen that this measure will not be Lebesgue). Similarly, find an ergodic invariant measure for the action of $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ on $\mathbb{T}^2$ and, for style points, find a such a measure that is not just a finite linear combination of delta measures.

Problem 5. Let (X, T, μ) be a pmps. Show that the only T-invariant measurable sets are null or conull if and only if the same is true of T-almost invariant measurable sets.

Problem Set 1: Due after Lecture 4

Problem 1. (Invertible extensions) Einsiedler War Exercises 2.1.7 and 2.1.8

Problem 2. (Solenoids) Einsiedler Ward Exercises 2.1.9

Problem 3. (von Neumann's mean ergodic theorem) Let V be a real *Hilbert space*, i.e. a real vector space together with a symmetric positive definite bilinear form $\langle \cdot, \cdot \rangle : V \times V \longrightarrow \mathbb{R}$ that is also complete, i.e. defining a norm $\|v\| := \sqrt{\langle v, v \rangle}$, any Cauchy sequence in V with respect to the norm converges to some vector v. Let $A : V \longrightarrow V$ be a *unitary map*, i.e. a linear map so that $\langle Av, Aw \rangle = \langle v, w \rangle$ for all $v, w \in V$. The *invariant subspace* is $V^A := \{v \in V : Av = v\}$.

(1) Consider the continuous linear map $B : V \longrightarrow V$ where $B(v) = v - Av$. Its kernel is V^A. Show that $V = V^A \oplus \overline{\mathrm{im}(B)}$ where $\oplus$ indicates that V is the direct sum of two orthogonal subspaces.
(2) Let $\pi : V \longrightarrow V^A$ be the projection onto the first summand in the decomposition $V = V^A \oplus \overline{\mathrm{im}(B)}$. Let $\mathrm{Av}_n : V \longrightarrow V$ be given by $\mathrm{Av}_n(v) := \frac{1}{n}\sum_{k=0}^{n-1} A^k v$. Show that $\mathrm{Av}_n(v) \longrightarrow \pi(v)$ for every $v \in V$. (Hint: it suffices to consider v of the form $v = w - Aw$.)
(3) Now suppose that (X, T, μ) is an ergodic pmps. Let $V = L^2(X, \mu)$. Let $A : V \longrightarrow V$ be given by $A(f) = f \circ T$. Given $f \in V$ let $c := \int_X f d\mu$. Show that $\|\frac{f(T^{n-1}) + f(T^{n-2}) + \ldots + f}{n} - c\|_{L^2} \longrightarrow 0$.

Problem 4. (Gauss map and Gauss measure) Show that the measure on $X := (0, 1)$ defined by $\mu((a, b)) = \int_a^b \frac{dx}{1+x}$ is invariant under the map $T : X \longrightarrow X$ that sends x to $\frac{1}{x}$ mod 1.

Problem 5. (Lecture 1 Worksheet Problems) Do all the problems from the worksheet from Lecture 1.

Problem 6. (Lecture 2 Worksheet Problems) Do all the problems from the worksheet from Lecture 2.

3. The Birkhoff Ergodic Theorem

Given a p.m.p.s. (X, T, μ) and $f : X \longrightarrow \mathbb{R}$ a function in L^1, set $S_0(f) := 0$,

$$S_n(f) := \sum_{k=0}^{n-1} f(T^k) \quad \text{and} \quad \mathrm{Av}_n(f) := \frac{S_n(f)}{n}.$$

Theorem 11 (The Maximal Ergodic Theorem)**.** *For $\alpha \in \mathbb{R}$, let E_α be the points in X so that $\mathrm{Av}_n(f) > \alpha$ for some n. Then $\alpha\mu(E_\alpha) \leq \int_{E_\alpha} f$.*

Proof. By replacing f with $f - \alpha$ we can suppose that $\alpha = 0$. Note

$$S_n(f) \circ T + f = S_{n+1}(f).$$

If $M_n(f) := \max_{0 \leq k \leq n}(S_n(f))$, then for $0 \leq k \leq n$,

$$M_n(f) \circ T + f \geq S_k(f) \circ T + f \geq S_{k+1}(f).$$

For x in $P_n := \{x : M_n(f)(x) > 0\}$, some $S_k(f)(x) > S_0(f)(x) = 0$, so

$$M_n(f) \circ T + f \geq M_n(f).$$

Therefore, since $M_n(f) = 0$ on $X - P_n$ and since $M_n(f) \geq 0$.

$$\int_{P_n} f \geq \int_{P_n} M_n(f) d\mu - \int_{P_n} M_n(f) \circ T d\mu \geq \int_X M_n(f) d\mu - \int_X M_n(f) \circ T d\mu = 0.$$

Notice that (P_n) is an ascending chain of sets whose union is E_0. By the dominated convergence theorem,

$$\int_{E_0} f d\mu = \lim_{n \longrightarrow \infty} \int_X f \chi_{P_n} d\mu \geq 0.$$

$\square$

Remark 12. By replacing f with $-f$, the maximal ergodic theorem also implies that for any $\beta \in \mathbb{R}$, if F_β is the set of points in X so that $\mathrm{Av}_n(f) < \beta$ for some n, then $\beta\mu(F_\beta) \geq \int_{F_\beta} f$.

Theorem 13 (The Birkhoff Ergodic Theorem, 1931)**.** *If $f^*(x) = \limsup_n \mathrm{Av}_n(f)$ and $f_*(x) = \liminf_n \mathrm{Av}_n(f)$, then $f_* = f^*$, these functions are invariant, and $\int_X f^* d\mu = \int_X f d\mu$. In particular, if (X, μ, T) is ergodic, $\mathrm{Av}_n(f)$ converges pointwise almost everywhere to $\int_X f d\mu$.*

Proof. The following idea shows invariance for f^* and f_*. Note that

$$\frac{n}{n+1}\mathrm{Av}_n(f)(T(x)) + \frac{1}{n+1}f(x) = \mathrm{Av}_{n+1}(f)(x)$$

Fixing x and applying limsup to both sides gives $f^* \circ T = f^*$,

Let $a < b$ be two rational numbers and let $E(a, b)$ be the set of x so that $f_*(x) \leq a < b \leq f^*(x)$. This is a T-invariant set and we claim that it has measure zero. If not, then after replacing X with $E(a,b)$ and μ with $\frac{\mu}{\mu(E(a,b))}$, the maximal ergodic theorem yields $b < \int_{E(a,b)} f \frac{d\mu}{\mu(E(a,b))} < a$, a contradiction. The set where $f^* \neq f_*$ is $\bigcup_{(a,b)\in\mathbb{Q}} E(a,b)$, which is null. This shows that $\mathrm{Av}_n(f)$ converges pointwise almost everywhere to f^*.

Now suppose that f is bounded. It follows that $(\mathrm{Av}_n(f))$ are all bounded by the same constant and so the dominated convergence theorem yields that

$$\int_X f^* d\mu = \lim_{n\longrightarrow\infty} \int_X \mathrm{Av}_n(f) d\mu = \int_X f d\mu$$

where the second equality follows from the T-invariance of μ.

Now suppose that f is unbounded. It suffices to show that $(\mathrm{Av}_n(f))$ converges to f^* in $L^1(X,\mu)$. Since this sequence converges pointwise a.e. to f^*, it suffices to show that $(\mathrm{Av}_n(f))$ is Cauchy in $L^1(X,\mu)$. Let $\epsilon > 0$. Choose a bounded function $g : X \longrightarrow \mathbb{R}$ so that $\|f - g\|_1 < \frac{\epsilon}{3}$ (the existence of g follows by dominated convergence) and note that we have already seen that $(\mathrm{Av}_k(g))_{k\geq 0}$ converges pointwise and hence in L^1 to g^* (going from pointwise to L^1 convergence uses the boundedness $(\mathrm{Av}_n(g))$). Choose N so that $k,m > N$ implies that $\|\mathrm{Av}_k(g) - \mathrm{Av}_m(g)\|_1 < \frac{\epsilon}{3}$. Then for $k,m > N$, by the triangle inequality, we bound $\|\mathrm{Av}_k(f) - \mathrm{Av}_m(f)\|_1$ by

$$\|\mathrm{Av}_k(f) - \mathrm{Av}_k(g)\|_1 + \|\mathrm{Av}_k(g) - \mathrm{Av}_m(g)\|_1 + \|\mathrm{Av}_m(g) - \mathrm{Av}_m(f)\|_1 < \epsilon.$$

□

Worksheet for Lecture 3

Problem 1. Suppose that (X, T, μ) is an ergodic topological p.m.p.s on a compact metric space X. A point x is called *μ-generic* if for every continuous function $f : X \longrightarrow \mathbb{R}$, $\frac{f(x)+f(Tx)...+f(T^{n-1}x)}{n} \longrightarrow \int_X f d\mu$. Show that μ-a.e. point in X is generic. (Hint: Use the fact that if X is a compact metric space, then a consequence of the Stone-Weierstrass theorem is that there is a countable dense set of functions in $C(X)$.)

Problem 2. Suppose that (X, T, μ) is a p.m.p.s. and that T is invertible. Show that $\lim_n \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) = \frac{1}{n} \lim_n \sum_{k=0}^{n-1} f(T^{-k} x)$ almost surely. (Hint: Consider the invariant set where one limit is bigger than the other. Integrate both functions over this set.)

4. The Riesz Representation Theorem

Let X be a compact metric space. Let $C(X)$ be the space of continuous functions from X to $\mathbb{R}$. Let $\mathcal{M}(X)$ be the space of finite Borel probability measures. A *(real) Banach space* $(V, \|\cdot\|)$ is a (real) vector space V with a *norm* $\|\cdot\|$ (i.e. a definite function $\|\cdot\| : V \longrightarrow [0, \infty)$ satisfying the triangle inequality and so $\|av\| = |a|\|v\|$ for $a \in \mathbb{R}$ and $v \in V$) that is *complete*, i.e. any Cauchy sequence converges. The space $C(X)$ is a Banach space with $\|f\| := \max_{x \in X} |f(x)|$. The *dual space* V^* is the set of *linear functionals* $\phi : V \longrightarrow \mathbb{R}$, i.e. continuous linear maps.

Lemma 14. *Suppose that X is a compact metric space. If K is a closed subset and μ is a finite measure, then*

$$\mu(K) = \inf\{\int_X f d\mu : 1_K \leq f \in C(X)\}.$$

Proof. The lefthand side is less than the righthand side. Let U_n be the points of distance at most $\frac{1}{n}$ from K. This is an open set and $\bigcap_n U_n = K$. By Ursyohn's lemma, we can find $f_n : X \longrightarrow [0, 1]$ so that $1_K \leq f_n \leq 1_{U_n}$. By dominated convergence, $\mu(K) = \int_X 1_K d\mu = \lim_{n \longrightarrow \infty} \int_X f_n d\mu$. □

Lemma 15. *If X is a compact metric space, then $\mathcal{M}(X)$ injects into $C(X)^*$.*

Proof. If $\|f - g\| < \epsilon$, then

$$|\mu(f) - \mu(g)| = \left| \int_X (f - g) d\mu \right| \leq \int_X |f - g| d\mu \leq \epsilon \int_X d\mu = \epsilon \mu(X).$$

So we have continuity as desired. The injectivity of the map from $\mathcal{M}(X)$ to $C(X)^*$ follows since the measure of any closed subset is determined by the image of μ in $C(X)^*$. Any finite Borel measure is determined by its values on closed sets. □

Recall that the *Cantor set* is homeomorphic to $\{0, 1\}^{\mathbb{N}}$.

Lemma 16. *If X is a compact metric space, then there is a continuous surjection from the Cantor set C to X.*

Proof. Cover X be a collection of n_1 balls U_i of diameter at most 1. For each i, cover U_i by a collection of n_2 balls U_{ij} of diameter at most $\frac{1}{2}$. Repeat. Elements of $D := \prod_i \{1 \ldots n_i\}$ determine a nested sequence of compact sets whose intersection is a singleton. The map sending D to X given by sending a sequence to this singleton is continuous (for any sequence d and any $\epsilon > \frac{1}{m}$ the (open) set of sequences in D that agree with d in the first m places are sent to points that are within $\frac{1}{m}$ of the image of d). Suppose without loss of generality that $n_i = 2^{q_i}$ for some integer q_i. Then D can be expressed as $\prod_{i\geq 0} \mathrm{Maps}(\{1, \ldots, q_i\} \longrightarrow \{0,1\})$. This is homeomorphic to the Cantor, i.e. functions from $\mathbb{N}$ to $\{0,1\}$, by thinking of these as functions from $\{1, \ldots, q_1 | q_1+1, \ldots, q_1+q_2 | q_1+q_2+1, \ldots q_1+q_2+q_3 | \ldots\}$ to $\{0,1\}$. □

We will say that a functional $\mu : C(X) \longrightarrow \mathbb{R}$ is *positive* if $\mu(f) \geq 0$ for any $f : X \longrightarrow (0,\infty)$. This forms a *cone*, i.e. a subset of a vector space that is closed under addition and positive scalar multiplication. Note that if $\phi \in C(X)^*$ is positive and $f \geq g \in C(X)$, then $\phi(f) \geq \phi(g)$.

Lemma 17. *Suppose that X is the Cantor set. Then the cone of positive linear functionals in $C(X)^*$ can be identified with $\mathcal{M}(X)$.*

Proof. Let $\phi \in C(X)^*$ be a positive linear functional. Let $\mathcal{B}$ be the subsets of $\{0,1\}^{\mathbb{N}}$ that are preimages of subsets of $\{0,1\}^n$ under the restriction maps $\pi_n : \{0,1\}^{\mathbb{N}} \longrightarrow \{0,1\}^n$. Any subset of $\{0,1\}^n$ is closed and open so the same holds for the subsets of $\mathcal{B}$. In particular, the indicator functions $(\chi_B)_{B\in\mathbb{B}}$ are continuous. Moreover, $\mathcal{B}$ forms a *ring of sets*, i.e. it is closed under finite union and relative complement. By compactness, any countable disjoint union of nonempty elements of $\mathcal{B}$ belonging to $\mathcal{B}$ is actually a finite disjoint union. Therefore, ϕ determines a countably additive function $\mu : \mathcal{B} \longrightarrow [0,\infty)$ by $\mu(B) := \phi(\chi_B)$. By the Caratheodory extension theorem, μ may be extended to a measure on X. Notice that for finite linear combinations of elements of $(\chi_B)_{B\in\mathcal{B}}$, μ and ϕ agree, i.e. for constants a_B

$$\mu(\sum_{B\in\mathcal{B}} a_B \chi_B) = \sum_{B\in\mathcal{B}} a_B \mu(\chi_B) = \sum_{B\in\mathcal{B}} a_B \phi(\chi_B) = \phi(\sum_{B\in\mathcal{B}} a_B \chi_B).$$

Such functions are dense in $C(X)$ so since μ and ϕ are continuous and agree on a dense set, they are equal. □

Recall that the *(dual) norm* $\|\mu\|$ of a functional $\mu \in C(X)^*$ is the smallest constant C so that $|\mu(f)| \leq C\|f\|$ for all $f \in C(X)$.

Lemma 18. *A nonzero linear functional $\mu \in C(X)^*$ is positive if and only if $\mu(\chi_X) = \|\mu\|$.*

Proof. For the forward direction note that $\langle f, g\rangle := \mu(fg)$ defines a nonnegative semidefinite bilinear form and hence Cauchy-Schwarz applies. For any $f \in C(X)$,

$$|\mu(f)|^2 = |\mu(f\cdot\chi_X)|^2 \leq \mu(f^2)\mu(\chi_X) \leq \mu(\|f\|^2\cdot\chi_X)\mu(\chi)_X) = \|f\|^2\mu(\chi_X)^2.$$

Since equality holds when $f = \chi_X$, $\|\mu\| = \mu(\chi_X)$.

For the reverse direction suppose without loss of generality that $\mu(\chi_X) = \|\mu\| = 1$. Let $f : X \longrightarrow [m, 1]$ be continuous with $m > 0$. Noting that $\frac{1+m}{2}$ is the midpoint of this interval,

$$|\mu(f) - \frac{1+m}{2}| = |\mu(f) - \mu(\frac{1+m}{2}\chi_X)| \leq \|f - \frac{1+m}{2}\| \leq \frac{1-m}{2},$$

i.e. $\mu(f) \in [m, 1]$. $\square$

Theorem 19 (Riesz Representation Theorem, 1941)**.** *If X is any compact metric space, then the cone of positive linear functionals in $C(X)^*$ can be identified with $\mathcal{M}(X)$.*

Proof. Let C be the Cantor set and let $p : C \longrightarrow X$ be a continuous surjection. Since p is a surjection, $p^* : C(X) \longrightarrow C(C)$ given by $p^*(f) = f \circ p$ is a linear isometry (and hence an injection). Let $\phi \in C(X)^*$ be a positive linear functional and define $p^*\phi : p^*(C(X)) \longrightarrow \mathbb{R}$ to be the function that sends $f \circ p$ to $\phi(f)$. By Hahn-Banach, $p^*\phi$ can be extended to a functional $\psi : C(C) \longrightarrow \mathbb{R}$ of the same norm. The previous two lemmas imply that the extension is positive and hence given by integration against a measure μ on C. For $f \in C(X)$,

$$\phi(f) = \psi(p \circ f) = \int_C f(p) d\mu = \int_X f dp_*\mu.$$

$\square$

Worksheet for Lecture 4

Problem 1. Let X be a compact metric space. Suppose that $T_n : X \longrightarrow X$ is a sequence of continuous functions that uniformly converge to $T : X \longrightarrow X$. Suppose that μ is a Borel probability measure on X that is invariant under T_n for each n. Show that μ is also T-invariant. (Hint: Apply the dominated convergence theorem and the Riesz representation theorem). Conclude that the only measure on the circle that is invariant under an irrational rotation is Lebesgue. (Hint: recall that if θ is irrational, then $\{n\theta \bmod 1\}_{n\geq 0}$ is dense in $[0,1)$.)

Problem 2. Let $\pi_n : \{0,1\}^{\mathbb{N}} \longrightarrow \{0,1\}^n$ be the restriction map. Show that functions of the form $f \circ \pi_n$ (where f and n are arbitrary) are dense in $C(\{0,1\}^{\mathbb{N}})$.

Problem set: Due after Lecture 6

Problem 1 (Another L^2 characterization of ergodicity): Einsiedler-Ward Exercise 2.5.1.

Problem 2 (Souped-Up Poincare Recurrence): Einsiedler-Ward Exercise 2.5.4 and 2.5.5.

Problem 3 (Metrizability of $\mathcal{M}(X)$): Let X be a compact metric space. Show that the set of contracting linear functionals in $C(X)^*$, i.e. functionals $\phi : C(X) \longrightarrow \mathbb{R}$ so that $|\phi(f)| \leq \|f\|$ for all $f \in C(X)$, is a compact metric space with respect to the weak* topology. (Hint: By the Stone-Weierstrass theorem, there is a countable collection of continuous functions $f_n : X \longrightarrow \mathbb{R}$ that are dense in $C(X)$. Define $d(\mu_1, \mu_2) := \sum_{n\geq 0} 2^{-n}|\mu_1(f_n) - \mu_2(f_n)|$.)

Problem 4 (The Jordan Decomposition): Let X be a compact metric space. Let $\phi : C(X) \longrightarrow \mathbb{R}$ be a bounded linear functional. Recall that given a continuous function $f : X \longrightarrow \mathbb{R}$ we define its positive part $f^+ := \max(0, f)$ and its negative part $f^- := \max(0, -f)$ so that $f = f^+ - f^-$. Let $N(X) \subseteq C(X)$ be the collection of continuous maps from X to $[0, \infty)$. Define $\phi^+ : N(X) \longrightarrow \mathbb{R}$ by $\phi^+(f) := \sup\{\phi(g) : g \leq f, g \in N(X)\}$.

(1) Show that ϕ^+ is a continuous linear map. (Hint: To show linearity note that $\phi^+(f_1 + f_2) \geq \phi^+(f_1) + \phi^+(f_2)$ is clear. For the reverse inequality, suppose that $0 \leq g \leq f_1 + f_2$ and define $g_1 = \min(g, f_1)$ and $g_2 = g - g_1$.)
(2) Extend ϕ^+ to $C(X)$ by defining $\mu_1 : C(X) \longrightarrow \mathbb{R}$ so that $\mu_1(f) := \phi^+(f^+) - \phi^+(f^-)$. Show that μ_1 is a positive continuous linear map and hence defines a measure. (Hint: To show linearity, it suffices to show that if $f = g - h$ where $g, h \geq 0$, then $\mu_1(f) = \phi^+(g) - \phi^+(h)$. Write $f^+ - f^- = g - h$, rearrange, and use the linearity of ϕ^+.)
(3) Conclude that there are two finite Borel measures μ_1 and μ_2 so that $\phi(f) = \mu_1(f) - \mu_2(f)$ for any $f \in C(X)$.

Problem 5 (Choquet's theory of barycenters): Let V be a locally convex topological vector space (this just means that we can apply the Hahn-Banach theorem, which says that for any two distinct vectors $x, y \in V$ there is a linear functional $f \in V^*$ so that $f(x) \neq f(y)$). Let $C \subseteq V$ be a convex compact subset. Let μ be a probability measure on C. A point $c \in C$ is a *barycenter* if for every continuous linear map

$f : C \longrightarrow \mathbb{R}$, $f(c) = \mu(f)$, i.e. the value of f on c is the μ-average of f over C.

(1) Show that there is a barycenter if and only if for every finite collection $T = (f_1, \dots, f_n)$ of continuous maps from V to $\mathbb{R}^n$ the point $p := (\mu(f_1), \dots, \mu(f_n))$ is contained in $T(C)$.
(2) Show that C has a barycenter. (Hint: If not, then find a hyperplane in $\mathbb{R}^n$ that separates p from $T(C)$.)
(3) Use the Hahn-Banach theorem to show that barycenters are unique.
(4) Let $b : \mathcal{M}(C) \longrightarrow C$ be the map that associates a probability measure on C to its barycenter. Show that when $\mathcal{M}(C)$ is equipped with the weak* topology that this map is continuous and G-equivariant.

Problem 6: Do Lecture 3 worksheet Problems 1 and 2 and Lecture 4 worksheet Problem 1.

5. Banach-Alaoglu, Markov-Kakutani and Krylov-Bogolyubov

Let X be a compact metric space. Let $\mathcal{M}^1(X)$ be the space of Borel probability measures. The space of all functions from $C(X)$ to $\mathbb{R}$ equipped with the topology of pointwise convergence is $\prod_{f \in C(X)} \mathbb{R}$ equipped with the product topology. Since $C(X)^*$ is a subspace the induced topology on it is called the *weak** topology.

Theorem 20 (Banach-Alaoglu, 1940)**.** *The set of contracting functionals in $C(X)^*$, i.e. those for which $|\mu(f)| \leq \|f\|$ for all $f \in C(X)$ is compact in the weak* topology.*

Proof. The contracting linear functionals are a subset of $C := \prod_{f \in C(X)} [-\|f\|, \|f\|]$, which is compact by Tychonoff's theorem. More exactly, it is the intersection of C and $C(X)$, so it is a closed subset of a compact set and hence compact. ☐

Corollary 21. *$\mathcal{M}^1(X)$ is compact and convex in the weak* topology.*

Proof. These are precisely the contracting linear functionals in $C(X)^*$ that (1) are positive and (2) send χ_X to 1. The two numbered conditions are closed conditions. ☐

For any continuous map $g : X \longrightarrow X$ we note that $g_* : \mathcal{M}^1(X) \longrightarrow \mathcal{M}^1(X)$ is continuous since for any sequence (μ_n) that weak* converges to μ, $(g_*\mu_n)$ weak* converges to $g_*\mu$ since for any $f \in C(X)$,

$$g_*\mu_n(f) = \mu_n(f \circ g) \longrightarrow \mu(f \circ g) = g_*\mu(f).$$

A *topological semigroup* is a group G together with a topology so that multiplication $m : G \times G \longrightarrow G$ given by $m(g, h) = gh$ is continuous. The main examples for us are $\mathbb{N}$ and $\mathbb{Z}$ equipped with the discrete topology and $[0, \infty)$ and $\mathbb{R}$ equipped with the usual topology. An action of G on a topological space X is a continuous (semi)group action $G \times X \longrightarrow X$. In the sequel we will suppress the words "topological" and "continuous". An $\mathbb{N}$-action on a space, is just a continuous map $f : X \longrightarrow X$. An $\mathbb{R}$-action, is just a flow $\phi_t : X \longrightarrow X$. A semigroup G is *amenable* if every continuous action of G on a compact metric space X admits a G-invariant measure, i.e. a measure μ on X so that $g_*\mu = \mu$ for all $g \in G$.

Lemma 22 (Markov-Kakutani, 1936)**.** *If G is abelian then it is amenable.*

Proof. Let G act continuously on a compact metric space X. Set $\mathcal{M} := \mathcal{M}^1(X)$. For each $g \in G$, let $A_{n,g}(\mu) = \frac{1}{n} \sum_{i=1}^{n} (g^i)_*\mu$. Let S be the set of $\{A_{n,g}\}_{n \geq 0; g \in G}$ after closing under composition. This is an abelian semigroup since G is abelian. We note that if $\phi_1, \dots, \phi_n$ are transformations in S then $\bigcap_i \phi_i(\mathcal{M})$ is nonempty since it contains $\phi(\mathcal{M})$ where $\phi = \phi_1 \dots \phi_n$ (since S is abelian). Since $\mathcal{M}$ is compact, the finite-intersection property implies that $\bigcap_{s\in S} s(\mathcal{M})$ is nonempty. Let μ be an element of this set. For each $g \in G$, there is some μ_n so $A_{n,g}(\mu_n) = \mu$. Therefore,

$$\|\mu - g_*\mu\| = \frac{1}{n}\left\|\sum_{i=1}^{n}(g_i)_*\mu_n - \sum_{j=2}^{n+1}(g^j)_*\mu_n\right\| \leq \frac{2}{n} \longrightarrow 0$$

□

Corollary 23 (Krylov-Bogolyubov, 1937)**.** *If X is a compact metric space and $T : X \longrightarrow X$ is continuous then there is an T-invariant measure.*

Later we will see that every compact Hausdorff group G admits a G-invariant probability measure called *Haar measure.*

Corollary 24. *Compact groups are amenable.*

Proof. Let G act on a compact space X and let μ be the Haar measure on G. Let $x \in X$ be any point and let $\phi : G \longrightarrow X$ be the map that sends g to $g \cdot x$. Note $\phi(g \cdot h) = g \cdot \phi(h)$. Let $\nu := \phi_*\mu$. Then for any measurable set $A \subseteq X$

$$\nu(g^{-1}A) = \phi_*\mu(g^{-1}A) = \mu(\phi^{-1}g^{-1}A) = \mu(g^{-1}\phi^{-1}A) = \mu(\phi^{-1}A) = \nu(A).$$

□

Other than tori the quintessential examples of compact groups are the *orthogonal groups* $O(n)$ and the *unitary groups* $U(n)$, which consist of matrices A in $\mathrm{GL}(n,\mathbb{R})$ (resp. $\mathrm{GL}(n,\mathbb{C})$) that preserve the dot product, i.e. $Av \cdot Aw = v \cdot w$ (resp. Hermitian product, i.e. $Av \cdot \overline{Aw} = v \cdot \overline{w}$) for all v and w. There is also the *compact symplectic group* $\mathrm{Sp}(n)$ which consists of matrices in $\mathrm{GL}(n,\mathbb{H})$ that preserve the standard Hermitian product on $\mathbb{H}^n$. Up to taking finite products, finite covers, finite index subgroups, and finite extensions (and adding in five exceptional groups) $O(n), U(n)$, and $\mathrm{Sp}(n)$ (for $n \geq 1$) account for all compact groups that are also manifolds.

5.1. **Bonus Section: Continuous Semigroup Actions and their induced representations.** The *compact-open topology* on the space $C(X,X)$ of continuous self-maps of X is that of uniform convergence, i.e. the one metrized by

$$d_{C(X,X)}(f,g) := \max_{x\in X} d_X(f(x),g(x)).$$

Lemma 25. *If G acts continuously on X, then the homomorphism $G \longrightarrow C(X,X)$ is a continuous.*

Proof. Fix $\epsilon > 0$ and $h \in G$. For each $x \in X$, there is neighborhood U_x of x and a neighborhood W_x of h in G so that $W_x \cdot U_x \subseteq B(h(x), \frac{\epsilon}{2})$. By compactness, we can find a finite collection (x_i) of points so that $\bigcup_i U_{x_i}$ covers X. Let $W := \bigcap_i W_{x_i}$. If $g \in W$ and $x \in X$, then $x \in U_{x_i}$, so $|g(x) - h(x)| < \epsilon$ since both points are $\frac{\epsilon}{2}$-close to $h(x_i)$. □

Given a Banach space V, the vector space $B(V)$ of bounded linear maps from V to itself is topologized with the topology of pointwise convergence, called the *strong operator topology*. Given a topological semigroup G acting continuously on a metric space X, each element $g \in G$ defines a linear contraction $g^* : C(X) \longrightarrow C(X)$ where $g^*(f) = f \circ g$. This is homomorphism from the opposite semigroup G^{op} to $B(C(X))$.

Lemma 26. *The homomorphism from G^{op} to $B(C(X))$ that sends g to g^* is continuous. Moreover, G acts continuously on $\mathcal{M}^1(X)$*

Proof. We must show that for any continuous function $f : X \longrightarrow \mathbb{R}$, any $\epsilon > 0$, and any $g \in G$, there is a neighborhood U of g so that $h \in U$ implies that $\|f \circ h - f \circ g\| < \epsilon$. Since f is uniformly continuous, there is a $\delta > 0$ so that if $d(x,y) < \delta$ then $|f(x) - f(y)| < \epsilon$. By the previous lemma, there is a neighborhood U of g so that if $h \in U$, then $d(h(x), g(x)) < \delta$ for all $x \in X$, which implies that $\|f \circ h - f \circ g\| < \epsilon$. The second claim is an exercise. □

Worksheet for Lecture 5

Problem 1. (Nonabelian free groups are not amenable) Consider the action of $\mathrm{GL}(2,\mathbb{R})$ on the unit circle X where a matrix A sends v to $\frac{A(v)}{|A(v)|}$. Show that there are no invariant measures on X for the action of the group generated by $A = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$ and $B := r_{-\frac{\pi}{4}} A r_{\frac{\pi}{4}}$ where r_θ is the rotation by θ matrix. Conclude that any nonabelian free group is not amenable. (Adapting this argument will show that $\mathrm{SL}(2,\mathbb{Z})$ is also not amenable.)

Problem 2. Consider the measure μ_r on $\mathbb{R}^2$ that averages a function over the circle of radius r. Show that (μ_r) weak* converges to the delta measure supported at the origin as $r \longrightarrow 0$.

Problem 3. Suppose that (μ_n) are Borel probability measures on a compact metric space X that weak* converge to μ. Show that for any compact subset K, $\limsup_n \mu_n(K) \leq \mu(K)$. (Hint: recall that there is a collection of continuous functions $f_n : X \longrightarrow [0,1]$ so that $1_K \leq f_n \leq f_{n-1}$ and so that, for any Borel probability measure ν, $\nu(K) = \lim_n \int_X f_n d\nu$.) Use the previous problem to show that equality need not always hold.

Problem 4. (Equidistribution of generic points) Under the same assumptions as the previous problem, show that for any open set U with μ-measure zero boundary $\lim_n \mu_n(U) = \mu(U)$. (Hint: Note that the previous problem implies that $\liminf_n \mu_n(U) \geq \mu(U)$ and use that $\nu(U) \leq \nu(\overline{U})$ for any measure ν). Finally, show that if (X, T, μ) is an ergodic topological pmps and $x \in X$ is generic, then for any open set U with measure zero boundary

$$\lim_{n \longrightarrow \infty} \frac{\#\{1 \leq k \leq n : T^k(x) \in U\}}{n} = \mu(U).$$

(Hint: Note that genericity of x is equivalent to the statement that $\frac{1}{n}\sum_{k=0}^{n-1} \delta_{T^k x}$ weak* converges to μ.)

Problem 5. (Compactness of Orthogonal Groups) Recall that an $n \times n$ matrix A with real entries is orthogonal if and only if $A^T A = I$ or, equivalently, if its columns form an orthonormal basis of $\mathbb{R}^n$. Use these two characterizations to explain why $O(n)$ is compact.

Problem 6. ($O(n)$-invariant measures on spheres) Note that $O(n)$ acts continuously on the unit sphere $S^{n-1} \subseteq \mathbb{R}^n$. Show that the standard (normalized) $(n-1)$-dimensional volume (called simply

Lebesgue measure m) is the only $O(n)$-invariant probability measure. (Hint: If μ is another invariant probability measure, then argue that it is absolutely continuous with respect to Lebesgue and use the Radon-Nikodym theorem to write $d\mu = f dm$. Conclude that f must be invariant and hence constant.)

6. Grassmanians, Furstenberg's Theorem on Projective Actions, and Borel Density

Given a topological space X and an equivalence relation $\sim$, the *quotient topology* on $X/\sim$ is the one for which $A \subseteq X/\sim$ is open if and only if $p^{-1}(A)$ is open where $p : X \longrightarrow X/\sim$ is the *projection map*. In other words, $f : X/\sim \longrightarrow Y$ is continuous if and only if $f \circ p$ is. If a group G acts on X, then X/G is the space $X/\sim$ where $x \sim y$ if there is an element $g \in G$ so that $g \cdot x = y$.

Lemma 27. *Let G be a compact group. If G acts continuously and transitively on a Hausdorff space X with point stabilizer H, then X is homeomorphic to G/H.*

Proof. Suppose that H stabilizes $x \in X$. Then the map sending G to X by sending g to $g \cdot x$ factors through a map $\phi : G/H \longrightarrow X$, which is a continuous surjection by assumption. It is injective since $gH \cdot x = g'H \cdot x$ if and only if $g^{-1}g' \in H$, i.e. $g' \in gH$. Continuous bijections from a compact space to a Hausdorff one are homeomorphisms. □

In fact the conclusion holds if G is a locally compact Hausdorff group that is σ-compact and X is a locally compact Hausdorff space. Recall that the *Grassmannian* $\mathrm{Gr}_d(\mathbb{R}^n)$ *of d-dimensional subspaces in* $\mathbb{R}^n$ is the set of all d-dimensional subspaces of $\mathbb{R}^n$. If V_n is a sequence of subspaces, then we say that V_n converges to V if a basis of V_n converges to a basis of B. We remark that the universal Grassmannian $\mathrm{Gr}_d(\mathbb{R}^\infty)$ is the classifying space for rank d vector bundles.

Corollary 28. *Since $O(n)$ is compact and acts transitively on the Grassmannian, $\mathrm{Gr}_d(\mathbb{R}^n)$ is compact for all d and n. In particular, it is homeomorphic to $O(n)/O(d) \times O(n-d)$.*

Proof. If $(e_1, \dots, e_n)$ is the standard basis, then the stabilizer of the span of $(e_1, \dots, e_d)$ is $O(d) \times O(n-d)$. □

When $d = 1$, the Grassmannian is called *projective space* $\mathbb{P}^{n-1} := \mathbb{P}(\mathbb{R}^n)$, which is the space of lines through the origin of $\mathbb{R}^n$. We note that $\mathrm{GL}(n, \mathbb{R})$ acts on this space with the matrices of the form λI acting trivially. The resulting quotient group is denoted $\mathrm{PGL}(n, \mathbb{R})$. We are interested in the dynamics of matrices in $\mathrm{GL}(n, \mathbb{R})$ acting on the space of lines (more or less equivalently, after taking a $\mathbb{Z}/2$ cover, on spheres). Since the Plücker embedding sends $\mathrm{Gr}_d(\mathbb{R}^n)$ into $\mathbb{P}(\Lambda^d\mathbb{R}^n)$ the dynamics of matrix groups on Grassmannians is included in the study of their dynamics on projective space.

Lemma 29 (Furstenberg's Lemma, 1963)**.** *Suppose that (g_m) is a sequence of matrices in $\mathrm{SL}(n,\mathbb{R})$ with unbounded entries. Suppose that for $\mu, \nu \in \mathcal{M}^1(\mathbb{P}(\mathbb{R}^n))$, $(g_m)_*\mu$ weak* converges to ν. Then there are proper subspaces R and V of $\mathbb{R}^n$ so that ν is supported on $\mathbb{P}(R) \cup \mathbb{P}(V)$.*

As an example if $g_m = \begin{pmatrix} m & \\ & \frac{1}{m} \end{pmatrix}$ then ν will be supported on the projectivization of the x and y axes.

Proof. Let $\|g_m\|_\infty$ be the maximum absolute value of an entry in g_m. Passing to a subsequence, suppose that $\frac{g_m}{\|g_m\|_\infty}$ converges to a matrix g. Then

$$\det(g) = \lim_m \frac{\det g_m}{\|g_m\|_\infty^n} = 0.$$

Let N and R be the kernel and image of g respectively. Passing to a subsequence, suppose that $g_m \cdot N$ converges in the Grassmannian to a subspace V. If ℓ is a line in $\mathbb{R}^n$, then $g_m \ell$ converges to a line in V (resp. R) provided that ℓ is contained in N (resp. not contained in N). Write $\mu = \mu_1 + \mu_2$ where $\mu_1(A) := \mu(A \cap \mathbb{P}(N))$ and $\mu_2(A) = \mu(A \backslash \mathbb{P}(N))$ for any measurable set A. Passing to a subsequence, suppose that $(g_m)_*\mu_1$ and $(g_m)_*\mu_2$ weak* converge to measures ν_1 and ν_2 respectively. We claim that ν_1 is supported on $\mathbb{P}(V)$ and ν_2 on $\mathbb{P}(R)$. Since the arguments are identical we will just prove the first claim. We want to show that if $f : \mathbb{P}(\mathbb{R}^n) \longrightarrow \mathbb{R}$ is continuous and vanishes on $\mathbb{P}(V)$, then $\nu_1(f) = 0$. We have seen that $f \circ g_m \restriction_{\mathbb{P}(N)}$ pointwise converges to 0. By the dominated convergence theorem and the fact that μ_1 is supported on $\mathbb{P}(N)$ we have the following.

$$\int_{\mathbb{P}(\mathbb{R}^n)} f d\nu_1 := \lim_m \int_{\mathbb{P}(\mathbb{R}^n)} f(g_m) d\mu_1 = \lim_m \int_{\mathbb{P}(N)} f(g_m) d\mu_1 = \int_{\mathbb{P}(N)} \lim_m f(g_m) d\mu_1 = 0.$$

□

Corollary 30. *Let G be a noncompact subgroup of $\mathrm{SL}(n,\mathbb{R})$ for which there is a G-invariant ergodic probability measure μ on $\mathbb{P}(\mathbb{R}^n)$. Then there is a finite collection $V_1, \ldots, V_m$ of equidimensional subspaces that G permutes transitively and so μ is supported on $\mathbb{P}(V_1) \cup \ldots \cup \mathbb{P}(V_m)$. Moreover, if G_i is the subgroup of matrices that fix V_i, then G_i acts on $\mathbb{P}(V_i)$ via a compact group, i.e. one that conjugates into $O(V_i) \times \mathbb{R}^\times$.*

Proof. Since G is not compact there is a sequence (g_m) with unbounded entries. Since these fix μ, μ is supported on the union of two invariant subspaces by Furstenberg's Lemma. Let V be a smallest dimensional subspace to which μ assigns positive measure. If $g \in G$ and $\mu((g \cdot V) \cap V) > 0$, then the minimality assumption implies that $g \cdot V = V$. So up to null sets $\{g \cdot V\}_{g \in G}$ is a disjoint union of subspaces, each of which has measure $\mu(V)$. Since μ is a probability measure, this set is finite. Since μ is ergodic, G permutes these subspaces transitively. If H is the subgroup that fixes V, then $\mu \restriction_{\mathbb{P}(V)}$ is fully supported and so H must act by a compact group, i.e. one conjugated into $O(V)$. □

6.1. **Bonus Material: Borel Density.** A closed subgroup H of G is said to have *cofinite volume* if there is a G-invariant probability measure on G/H. An example is $\mathrm{SL}(n, \mathbb{Z})$ in $\mathrm{SL}(n, \mathbb{R})$. The *Zariski closure* of H is the largest subgroup L of G so that if $p(A) := p(a_{11}, \ldots, a_{nn})$ is a polynomial in the entries of a matrix $A = (a_{ij})$ and $p(A) = 0$ for all $A \in H$, then $p(A) = 0$ for all $A \in L$. When $G = \mathrm{SL}(n, \mathbb{R})$, Chevalley showed that one can always find a vector space V that G acts irreducibly on and so that L is the stabilizer of a line.

Corollary 31 (Borel Density)**.** *If H is a closed cofinite volume subgroup of $G := \mathrm{SL}(n, \mathbb{R})$, then its Zariski closure L is $\mathrm{SL}(n, \mathbb{R})$ itself.*

Proof. Push forward the invariant measure on G/H to $G/L \subseteq \mathbb{P}(V)$. Since G acts irreducibly, the measure is not supported on a proper subspace and so G acts via a compact group by Furstenberg's theorem. But since $\mathrm{SL}(n, \mathbb{R})$ is simple it does not map to compact Lie groups except for the trivial group. So G acts trivially on V and hence G/L is a point, i.e. $G = L$. □

Worksheet for Lecture 6

Problem 1: (Upper triangular matrices acting on S^2) Let P be the group of 3×3 upper triangular matrices. Let $g \in P$ act on the sphere S^2 by sending v to $\frac{gv}{|gv|}$. Find all P-invariant measures on S^2.

Problem 2: (Flag manifolds) Let $0 < d_1 < \ldots < d_k$ be any increasing sequence of positive integers. Let $d = (d_1, \ldots, d_k)$. The *flag manifold* $F_d(\mathbb{R}^n)$ is the set of subspaces $V_1 \subseteq \ldots \subseteq V_k \subseteq \mathbb{R}^n$ so that $\dim V_i = d_i$. This space is topologized similarly to the Grassmannian. Prove that $O(n)$ acts transitively on F_d find a closed subgroup H so that $O(n)/H$ is homeomorphic to F_d. In particular, this shows that flag manifolds are compact.

Problem 3: (Solvable Matrix Groups and Representation Theory) Let G be a solvable subgroup of $\mathrm{SL}_n(\mathbb{R})$ for $n > 1$ that contains some sequence of matrices with unbounded entries. (The group in the previous problem is an example). Show that there is some finite index subgroup H of G whose action on $\mathbb{R}^n$ is not irreducible (i.e. for which there is a proper nonzero H-invariant subspace). Recall that solvable groups are amenable.

Problem 4: (Matrix Groups acting on S^2) Let G be a subgroup of 3×3 invertible matrices. Let $g \in G$ act on the sphere S^2 by sending v to $\frac{gv}{|gv|}$. Suppose that G preserves a probability measure μ on the sphere. Show that one of the following occurs:

(1) A finite index subgroup of G fixes a point or stabilizes a great circle.
(2) There is some $A \in \mathrm{GL}(3, \mathbb{R})$ so that AGA^{-1} acts on the sphere via the (projective) orthogonal group $PO(3)$ (in which case G preserves $A_* m$ where m is the area measure on the sphere).

Problem 5: (Amenable Matrix Groups) Let G be a closed subgroup of $\mathrm{GL}(n, \mathbb{R})$. Let N be the maximal closed normal solvable subgroup of G. Levi's Theorem states that G/N is either compact or admits an irreducible linear action by a noncompact group on $\mathbb{R}^m$ for some m. Show that G is amenable if and only if G/N is compact. (Hint: On the next homework you will show that a group G with a normal subgroup N is amenable if N and G/N both are).

Problem 6: (Refresher on group quotients and topology) Let G be a locally compact σ-compact Hausdorff topological group acting transitively on a locally compact Hausdorff space X. If H is the stabilizer of $x_0 \in X$, then show that X is homeomorphic to G/H.

(Hint: (1) Show that it suffices to show that the continuous bijection $\phi : G/H \longrightarrow X$ (equivalently from G to X) is open where $\phi(gH) = g{\cdot}x$, i.e. that open sets are sent to open sets. (2) Then show that it suffices to show that open subsets of the identity in G with compact closure are sent to open subsets of X. (3) If U is such a subset, then use σ-compactness to show that there is a countable set (g_n) of points in G so that $\bigcup_n g_n U$ covers G. Therefore, $X = \bigcup_n g_n \phi(\overline{U})$. Use the Baire category theorem to conclude that $\phi(U)$ has interior. (4) If $\phi(h)$ is the interior point, then $\phi(e)$ is an interior point of $h^{-1}\phi(U)$. Use this to conclude.)

Problem Set 3: Due after Lecture 8

Problem 1 (Solvable groups are amenable): Prove the following:

(1) Show that if G is an amenable group that acts continuously and linearly on a compact convex subset C of a locally convex topological vector space, then G has a fixed point. (Hint: Use barycenters (see last week's problem set).)
(2) Show that if N is an amenable normal subgroup of G and G/N is amenable, then so is G. (Hint: If X is a compact metric space on which G acts, consider the compact convex subset of $\mathcal{M}^1(X)$ consisting of N-invariant measures. Show that this set has a G/N-fixed point, which must be a G-invariant measure).
(3) Conclude that solvable groups are amenable.

Problem 2. (Unique ergodicity) Let $T : X \longrightarrow X$ be a continuous self-map of a compact metric space. T is called *uniquely ergodic* if there is only one T-invariant Borel probability measure μ on X. We have seen that irrational rotations on the circle are examples.

(1) Show that μ is ergodic.
(2) Show that *every* (not just almost every) point is μ-generic.
(3) Conversely, show that if every point is μ-generic for some T-invariant measure μ, then μ is uniquely ergodic.

Problem 3: (Amenability and Folner Sequences) Let G be a σ-compact locally compact Hausdorff topological group and let m be its (left) Haar measure. For any compact subset K of G and any $\epsilon > 0$ say that a finite-measure set F is (K, ϵ)- invariant if $m(F \Delta KF) < \epsilon m(F)$. Suppose that (F_n) is a sequence of compact finite measure subsets so that for any compact K and any ϵ there is an N so that $n > N$ implies that F_n is (K, ϵ)-invariant. Show that G is amenable. (Hint: If G acts on a compact metric space X let ν be any Borel probability measure on X and, for any measurable subset A of X, define $\mu_n(A) := \frac{1}{m(F_n)} \int_{F_n} \nu(g^{-1}A) dm(g)$. Let μ be any weak* limit of μ_n and show that μ is G-invariant.) A set (F_n) is called a *Folner sequence*. Explain why $\mathbb{Z}^d$ has a Folner sequence.

Problem 4: Do problems 3 and 4 from Worksheet 5.

Problem 5: Do problems 1 and 3 from Worksheet 6.

Problem 6: Do problems 4 and 5 from Worksheet 6.

7. Circle homeomorphisms, minimality, rotation number, and Poincare's classification

The action of G on a topological space X is called *minimal* if the smallest closed nonempty G-invariant set is X itself. (A *minimal set*, i.e. a closed invariant set with no proper closed invariant subsets always exists by Zorn's lemma). Equivalently the action is minimal if, for any $x \in X$, $G \cdot x$ is dense in X. (Birkhoff defined "minimality" for the first time in 1912).

Examples:

(1) For a rational circle rotation, minimal sets all have cardinality the order of the rotation. In contrast, irrational circle rotations are minimal.
(2) For the map on the open unit disk in $\mathbb{C}$ given by $T(z) = z^2$, the only minimal set is $\{0\}$.

Lemma 32 (Classification of minimal sets on the circle)**.** *If G does not act minimally then it has a minimal set that is either finite or a Cantor set (i.e. a perfect nowhere dense set).*

Proof. Suppose that G has no finite invariant sets. Let C be a minimal set. It must be perfect since its set of limit points is G-invariant and closed. We will show that C must also be nowhere dense. If not, then C contains a closed interval and the union of all closed intervals contained in C is invariant. The closure of the endpoints of these intervals is a proper nonempty invariant subset contained in C contradicting minimality. □

On the homework you will show Denjoy's theorem (1932) which says that the Cantor set is not a minimal set if the homeomorphism is C^2.

Lemma 33 (Fekete's Lemma, 1923)**.** *A sequence of positive real numbers is subadditive if $a_{n+m} \leq a_n + a_m$ for all n, m. If (a_n) is such a sequence then $\lim_n \frac{a_n}{n}$ exists and equals $\inf_{n \geq 0} \frac{a_n}{n}$.*

Proof. Fix k and for any $n > k$ write $n = mk + r$ for $r \in \{0, \ldots, k-1\}$. Note that $a_{mk} \leq m a_k$. Therefore, $\frac{a_n}{n} \leq \frac{m a_k}{mk+r} + \frac{\max(a_0, \ldots, a_r)}{n}$. Therefore, $\limsup_n \frac{a_n}{n} \leq \frac{a_k}{k}$ for all k. So $\limsup_n \frac{a_n}{n} \leq \inf_n \frac{a_n}{n} \leq \liminf_n \frac{a_n}{n}$. □

Let $f : S^1 \longrightarrow S^1$ be a homeomorphism. Lift the homeomorphism to a map $F : \mathbb{R} \longrightarrow \mathbb{R}$ so that $F(x+1) = F(x) + 1$. The *escape rate of a point x_0 (aka rotation number $\tau(f)$)* is $\lim_{n \longrightarrow \infty} \frac{d(F^n x_0, x_0)}{n}$ (we will show now that limit exists). Given a lift F, $G(x) = F(x) + m$ is another lift for any $m \in \mathbb{Z}$. Note that $G^k(x) = F^k(x) + km$. Therefore, rotation number is defined in $\mathbb{R}/\mathbb{Z}$. Alternatively, we could choose F so that $d(F(x), x) < 1$ for all $x \in \mathbb{R}$.

Lemma 34. *Rotation number is well-defined and independent of the point x_0.*

Proof. Let $G_n := \max_{x\in[0,1]} d(F^n x, x)$. Then $G_{n+m} \leq G_n + G_m$. We note too that $d(F^n x, F^n y) < 1$ if $d(x, y) < 1$. So if $g_n := \min_{x\in[0,1]} d(F^n x, x)$, then $g_n \leq G_n \leq g_n + 2$. We're done now by Fekete. $\square$

Corollary 35 (The Monotonicity Lemma)**.** *Let $F : \mathbb{R} \longrightarrow \mathbb{R}$ be any lift of a circle homeomorphism f with rotation number τ. Let x_0 be any point in $\mathbb{R}$. Then $F^{n_1}(x_0) + m_1 < F^{n_2}(x_0) + m_2$ implies that $n_1\tau + m_1 < n_2\tau + m_2$.*

Proof. Set $y_0 := F^{n_2}(x_0)$. By Fekete's Lemma,

$$\tau = \inf_{n\geq 0} \frac{F^n(y_0) - y_0}{n} \leq \frac{F^{n_1-n_2}(y_0) - y_0}{n_1 - n_2} \leq \frac{m_2 - m_1}{n_1 - n_2}$$

$\square$

Lemma 36. *Rotation number is* 0 *if and only if f has a fixed point.*

Proof. If f has a fixed point, then there is a lift F that does as well and hence the rotation number is zero by definition. Conversely, suppose that the rotation number is zero and choose a lift so that $d(Fx, x) < 1$ for all x. Set $\delta = \min_{x\in[0,1]} d(F(x), x)$. So $n > d(F^n(x), x) \geq n\delta$. So rotation number vanishes in $\mathbb{R}/\mathbb{Z}$ if and only if $\delta = 0$. $\square$

Corollary 37. *Rotation number is rational if and only if f has a periodic point.*

Proof. We note that $\tau(f^n) = n\tau(f)$ mod 1 and that f has a periodic point with period n if and only if f^n has a fixed point. $\square$

A measurable (resp. topological) dynamical system $T : X \longrightarrow X$ is said to be *semiconjugate* to a dynamical system $G : Y \longrightarrow Y$ if there is a measurable (resp. continuous) surjection h so that the following diagram commutes,

$$\begin{tikzcd} X \arrow[r, "T"] \arrow[d, "h"] & X \arrow[d, "h"] \\ Y \arrow[r, "G"] & Y \end{tikzcd}$$

The systems are *conjugate* if h is invertible and the inverse is measurable (resp. continuous).

Theorem 38 (Poincare's classification of circle homeomorphisms, 1885)**.** *If $f : S^1 \longrightarrow S^1$ is a homeomorphism that has irrational rotation number, then f is conjugate to a rotation by $\tau(f)$ if it is minimal. Otherwise it is semiconjugate to one.*

Proof. Let x_0 be a point on the circle that lifts to 0. Let $\mathcal{O}$ be the lift of $\{f^n x_0\}_{n\geq 0}$ to $\mathbb{R}$. Let $F : \mathbb{R} \longrightarrow \mathbb{R}$ be a lift of f. Define a map $H : \mathcal{O} \longrightarrow \mathbb{R}$ that sends $F^n(0) + m$ to $n\tau + m$ for any $n, m \in \mathbb{N}$. (Since τ is irrational, f has no periodic orbits so each point in $\mathcal{O}$ can be written uniquely as $F^n(0) + m$.) By the monotonicity lemma, H is increasing and its image is dense in $\mathbb{R}$. Therefore H extends to a continuous increasing bijection from $\overline{\mathcal{O}}$ to $\mathbb{R}$ (note that $\{f^n x_0\}$ and hence $\mathcal{O}$ is perfect and the limit of a sequence of points approaching a number from above and below must agree by density of the image of H). H therefore extends to a surjection from $\mathbb{R}$ to $\mathbb{R}$ and fits into a commutative diagram

$$\begin{tikzcd}
\mathbb{R} \arrow[r, "F"] \arrow[d, "H"] & \mathbb{R} \arrow[d, "H"] \\
\mathbb{R} \arrow[r, "x \longrightarrow x+\tau"] & \mathbb{R}
\end{tikzcd}$$

Note that $F(x+1) = F(x) + 1$ and similarly for the bottom horizontal arrow. So taking a quotient by $\mathbb{Z}$ gives us the desired semiconjugacy. $\square$

7.1. Bonus Section: Semiconjugacy for group actions, Compact Subgroups of $\mathrm{Homeo}(S^1)$, and the Tits alternative.

Lemma 39 (Semiconjugacy of G-actions to minimal actions)**.** *If G leaves a minimal Cantor set C invariant, then there is a continuous surjection $\phi : S^1 \longrightarrow S^1$ so that $\phi(C) = S^1$ and so there is a minimal G action on S^1 so that ϕ is equivariant. Moreover, there is an G-invariant measure on C if and only if there is one on $\phi(S^1)$.*

Proof. We have already seen that there is a continuous surjection from C to S^1. We extend this over the complement of C by demanding that the function be constant on intervals in the complement of C. As in the construction of the Cantor function, we may demand that this map is "non-decreasing", which shows us that $S^1/\sim$ is homeomorphic to S^1, where $a \sim b$ if a and b are in the closure of the same interval in the complement of the Cantor set. Therefore, the action of G descends to an action on $S^1/\sim$, which is a circle. For the second claim, let $E \subseteq C$ be the countable subset of interval endpoints. Then $\phi : C - E \longrightarrow$

$S^1 - \phi(E)$ is a G-equivariant bijection. So there is an invariant measure on one if and only if there is an invariant measure on the other. $\square$

Theorem 40 (Tits alternative for $\mathrm{Homeo}(S^1)$, Margulis 2000)**.** *If G is a subgroup of* $\mathrm{Homeo}(S^1)$ *then one of the following occurs:*

(1) There is a finite G-invariant subset of the circle
(2) G is isomorphic to a subgroup of $O(2)$.
(3) G contains a nonabelian free group.

In particular, G satisfies a version of the *Tits alternative*, i.e. either G has an invariant measure or contains a nonabelian free subgroup. The usual Tits alternative says that a finitely generated matrix group is either virtually solvable (and hence is amenable) or contains a nonabelian free subgroup (and hence is not amenable). We have already seen that we may suppose that G acts minimally on the circle.

The topology on $\mathrm{Homeo}(S^1)$ is the *compact-open topology*, i.e. the one metrized by $d(f, g) = \max_{x \in S^1} d(f(x), g(x))$.

Lemma 41. *If G is a subgroup of* $\mathrm{Homeo}(S^1)$ *with compact closure, then G is conjugate to a subgroup of the rotation group $O(2)$.*

Proof. If G is not compact, then we replace it with its compact closure and let m be its Haar measure. Let λ be the usual length measure on the circle. Define a new measure on the circle by

$$\mu(A) := \int_G \lambda(g^{-1}A) dm(g)$$

Note that if $h \in G$, then

$$\mu(h^{-1}A) = \int_G \lambda(g^{-1}h^{-1}A) dm(g) = \int_G \lambda(g^{-1}A) d(h_* m) = \int_G \lambda(g^{-1}A) dm = \mu(A).$$

This measure is nonatomic and nonzero on any open set (take a small neighborhood of the origin, this neighborhood has positive measure and only changes the length of any interval by ϵ). These properties imply that $h(x) = \mu([0, x])$ is a strictly increasing function from 0 to 1 and hence defines a circle homeomorphism. Its inverse is then the function that sends $x \in [0, 1)$ to y so that $\mu([0, y]) = x$. So

$$h_*^{-1} \mu([0, x]) = \mu([0, y]) = x$$

which shows us that $h_*^{-1}\mu = \lambda$. Therefore, $h^{-1}Gh$ preserves length and hence belongs to $O(2)$. $\square$

Worksheet for Lecture 7

Problem 1: Show that if $f : S^1 \longrightarrow S^1$ is a homeomorphism with rational rotation number $\frac{p}{q}$ where p and q are coprime positive integers, then any periodic point has period q. (Hint: Show that any periodic point has period a multiple of q, then explain why the only periodic points of f^q are fixed points).

Problem 2: Show that if two circle homeomorphisms are in the same conjugacy class then they have the same rotation number.

Problem 3: Explain why $G := \mathrm{Homeo}(S^1)$ is a topological group, i.e. explain why $m : G \times G \longrightarrow G$ so that $m(g,h) = gh$ and $\iota(g) = g^{-1}$ are continuous. The topology on $\mathrm{Homeo}(S^1)$ is the *compact-open topology*, i.e. the one metrized by $d(f,g) = \max_{x \in S^1} d(f(x), g(x))$.

Problem 4: Show that the map from $\mathrm{Homeo}(S^1)$ to $\mathbb{R}/\mathbb{Z}$ given by rotation number is continuous.

8. Conditional expectation and conditional measures

Let $(X, \mathcal{B}, \mu)$ be a probability measure space. Let $i : \mathcal{A} \longrightarrow \mathcal{B}$ be an inclusion of σ-algebras. The typical $\mathcal{B}$-measurable function won't be $\mathcal{A}$-measurable, but what's the best approximation of a function in $L^1(X, \mathcal{B}, \mu)$ by a function in $L^1(X, \mathcal{A}, \mu)$, i.e. can we find a "reasonable map" $i^* : L^1(X, \mathcal{B}, \mu) \longrightarrow L^1(X, \mathcal{A}, \mu)$? Among other things it should be a one-sided inverse to the inclusion map $\iota : L^1(X, \mathcal{A}, \mu) \longrightarrow L^1(X, \mathcal{B}, \mu)$ that sends an $\mathcal{A}$-measurable function f to f. First we develop some intuition.

Examples of i^*.

(1) The smallest σ-algebra is $\{\emptyset, X\}$. Given $f \in L^1(X, \mathcal{B}, \mu)$, $i^*(f)$ must be constant and we choose $i^*(f) := \int f d\mu$.
(2) The next smallest σ-algebra is $\{\emptyset, A, X \backslash A, X\}$ for some nonempty subset A. Given $f \in L^1(X, \mathcal{B}, \mu)$, $i^*(f)$ should return the average over A if $x \in A$ and the average over $X \setminus A$ for $x \notin A$.

Theorem 42 (Conditional Expectation)**.** *There is a linear contraction $i^* : L^1(X, \mathcal{B}, \mu) \longrightarrow L^1(X, \mathcal{A}, \mu)$ that is a one-sided inverse to $\iota : L^1(X, \mathcal{A}, \mu) \longrightarrow L^1(X, \mathcal{B}, \mu)$, i.e. $i^* \circ \iota = \mathrm{id}$. It is functorial in the sense that if $j : \mathcal{B} \longrightarrow \mathcal{C}$ is an inclusion of σ-algebras then $(j \circ i)^* = i^* \circ j^*$ and $(\mathrm{id})^* = \mathrm{id}$. Moreover, $i^*(f)$ is the uniquely characterized (up to its definition on sets of measure zero) as the $\mathcal{A}$-measurable function such that $\int_A i^*(f) d\mu = \int_A f d\mu$ for all $A \in \mathcal{A}$.*

The standard notation is $E[f|\mathcal{A}] := i^*(f)$.

Proof. Start with a nonnegative function $f \in L^1(X, \mathcal{B}, \mu)$. Define a measure $\mu_f : \mathcal{A} \longrightarrow [0, \infty)$ by $\mu_f(A) = \int_A f d\mu$. This is absolutely continuous with respect to μ so the Radon-Nikodym theorem provides a unique (up to its definition on sets of measure zero) $\mathcal{A}$-measurable function $i^*(f)$ so that $\mu_f = i^*(f)\mu$ which is characterized by the property that $\int_A f d\mu = \int_A i^*(f) d\mu$ for any $A \in \mathcal{A}$. With this definition define $i^*(f) = i^*(f^+) - i^*(f^-)$ where f^+ and f^- are functions so that $f = f^+ - f^-$. Functoriality and linearity follow from the characteristic property. □

Corollary 43 (Restatement of the Ergodic Theorem)**.** *Let $(X, \mathcal{B}, \mu, T)$ be a pmps. Let $f \in L^1$. Let $\mathcal{A}$ be the σ-subalgebra of T-invariant sets. The functions $\mathrm{Av}_n(f) := \frac{f(x)+f(Tx)+\ldots+f(T^{n-1}x)}{n}$ converge pointwise almost everywhere and in L^1 to $E[f|\mathcal{A}]$.*

Proof. By the ergodic theorem, the averages converge pointwise almost everywhere and in L^1 to f^*, which is T-invariant and satisfies $\int_A f^* d\mu =$

$\int_A f d\mu$ for any T-invariant set. This is the characterization of $E[f|\mathcal{A}]$. ☐

Theorem 44 (Disintegration of measures - continuous version)**.** *Let $(X, \mathcal{B}, \mu)$ be a Borel probability measure space with X a compact metric space. Let $\mathcal{A}$ be a σ subalgebra of $\mathcal{B}$. For almost every $x \in X$ there is a probability measure $\mu_x^{\mathcal{A}}$ which is characterized by the property that for any $f \in C(X)$, $E[f|\mathcal{A}](x) = \int_X f d\mu_x^{\mathcal{A}}$.*

In the context of the ergodic theorem, $\mu_x^{\mathcal{A}}$, with $\mathcal{A}$ the σ-algebra of invariant sets, will be a choice of a weak* limit of $\frac{\delta_x + \delta_{Tx} + \ldots + \delta_{T^{n-1}x}}{n}$

Proof. The map $i^* : C(X) \longrightarrow L^1(X, \mathcal{A}, \mu)$ that sends f to $E[f|\mathcal{A}]$ is a linear contraction with the property that $E[f|\mathcal{A}] \geq 0$ if $f \geq 0$. So $(i^*)_x : C(X) \longrightarrow \mathbb{R}$ that sends f to $(i^*(f))(x)$ is a positive linear functional and hence determines a measure $\mu_x^{\mathcal{A}}$ by the Riesz representation theorem. This measure is characterized by the fact that $\mu_x^{\mathcal{A}}(f) = E[f|\mathcal{A}](x)$ for all $f \in C(X)$ and all $x \in X$. ☐

Remark 45. In fact the above proof is incomplete in that $E[f|\mathcal{A}]$ is only uniquely determined up to its values on sets of measure zero. To fix this, one should start with a countable dense subspace $V \subseteq C(X)$ and define $E[f|\mathcal{A}]$ for $f \in V$ as outputting an actual function (not an equivalence class of functions). This will require deleting a set of measure zero to ensure that the desired linearity, contraction, and positivity properties hold on the nose and not just up to sets of measure zero. Then the functionals $(i^*(f))(x)$ can be extended by linearity from V to the subspace W of real linear combinations of elements of V and then to $C(X)$ by Hahn-Banach.

Theorem 46 (Disintegration of measures - L^1 version)**.** *The conclusion of the previous theorem holds if f is in $L^1(X, \mathcal{B}, \mu)$.*

Proof. We must show that for any $f \in L^1(X, \mathcal{B}, \mu)$, $\int_X f d\mu_x$ is $\mathcal{A}$-measurable and that for any $A \in \mathcal{A}$ $\int_A \int_X f d\mu_x d\mu = \int_A f d\mu$. It suffices to show this fact for nonnegative functions. By taking monotone increasing limits of simple functions, it suffices to prove the claim for simple functions and hence for indicator functions. By dominated convergence, the claim for continuous functions implies the claim for indicator functions of open (and hence closed) sets. By monotone convergence it also holds for countable increasing unions of sets. Therefore, the sets for which the claim holds are closed under countable increasing union, countable nested intersection, and complements. This is a monotone class including all the open sets and hence coincides with the Borel σ-algebra by the monotone class theorem. ☐

Homework 4: Due after lecture 10

Problem 1: (Flows on the circle) A *topological flow* is a continuous $\mathbb{R}$-action on a topological space X. Show that every topological flow on the circle either has a fixed point or is conjugate to a flow where, for some fixed λ, and for all $t \in \mathbb{R}$, t acts by rotating the circle by $t\lambda$ radians. (Hint: Distinguish between the cases where the action is transitive and where it is not.)

Problem 2: (Homeo$^+(S^1)$ is perfect) Let Homeo$^+(S^1)$ be the group of orientation preserving homeomorphisms of the circle. The goal of this problem is to show that the only homomorphism from this group to an abelian group is the trivial one.

(1) Consider the group G of increasing bijections from $[0, 1]$ to itself. Show that if $f, g \in G$ and $f(x) > x$ and $g(x) > x$ for all $x \in (0, 1)$ then f and g are conjugate in G.
(2) Now suppose that $f : S^1 \longrightarrow S^1$ is an orientation preserving homeomorphism that has a fixed point. Using the previous problem, show that f and f^2 are conjugate in Homeo$^+(S^1)$. Conclude that f can be written as a commutator.
(3) Recall that any 2×2 rotation matrix can be written as a commutator of elements in $\mathrm{SL}(2, \mathbb{R})$. Conclude that any element of Homeo$^+(S^1)$ can be written as a product of two commutators.

Problem 3: (Rotation number is a class function) Show that rotation number is a class function on Homeo(S^1), i.e. two homeomorphisms have the same rotation number if they are in the same conjugacy class.

Problem 4: (Homeo$^+(S^1)$ is topologically a circle) Show that Homeo(S^1) deformation retracts to $O(2)$. (Hint: It is easier to work with lifts of (orientation-preserving) circle homeomorphisms to $\mathbb{R}$ since these are just strictly increasing functions $F : \mathbb{R} \longrightarrow \mathbb{R}$ that are 1-periodic, i.e. so that $F(x+1) = F(x)+1$. The lift of a rotation by α is just $F(x) = x + \alpha$. So you need to exhibit a deformation retract from the space of increasing 1-periodic functions to the space of functions of the form $x + \alpha$. A "straight-line homotopy" should do the trick.)

Problem 5: (Rotation number is a continuous map) Problems 3 and 4 from Worksheet 7.

Problem 6: (Expansive circle maps) Let G be a group of orientation preserving circle homeomorphisms. Show that either G is abelian or G is *expansive*, i.e. there is a sequence of intervals I_n on the circle whose length goes to 0 and a sequence of elements $g_n \in G$ so that $g_n(I_n)$ has length bounded away from zero.

9. The ergodic decomposition and disintegration

Let X be a compact metric space with Borel σ-algebra $\mathcal{B}$ with a σ-subalgebra $\mathcal{A}$ and a probability measure μ. Let $T : X \longrightarrow X$ be a measurable map that preserves μ. Suppose that $\mathcal{E}$ is the σ subalgebra of T-almost-invariant sets.

Lemma 47. *The map $X \longrightarrow \mathcal{M}(X)$ that sends x to $\mu_x^{\mathcal{A}}$ is $\mathcal{A}$-measurable.*

Proof. A basis of the σ-algebra is given by sets of the form $U_{f,r,\epsilon} := \{\mu : |\mu(f) - r| < \epsilon\}$ where $f \in C(X)$ is a fixed continuous function and r and ϵ are fixed real numbers. The preimage is

$$\{x \in X : |\mu_x(f) - r| < \epsilon\} = \{x \in X : E[f|\mathcal{A}](x) \in (r - \epsilon, r + \epsilon)\},$$

which is just $E[f|\mathcal{A}]^{-1}((r - \epsilon, r + \epsilon))$. $\square$

Lemma 48. *There is a countably generated σ-algebra that agrees with $\mathcal{A}$ up to sets of measure zero.*

Proof. Since $C(X)$ is separable and dense in $L^1(X, \mu)$, $L^1(X, \mu)$ is also separable. Therefore, there is a countable collection (A_i) of sets in $\mathcal{A}$ whose indicator functions are dense in $\{\chi_A\}_{A\in\mathcal{A}}$. Let $\widetilde{\mathcal{A}} \subseteq \mathcal{A}$ be the σ-algebra they generate. Let $A \in \mathcal{A}$. There is a sequence $\chi_{A_{n_i}}$ that converges to χ_A in $L^1(X, \mathcal{A}, \mu)$. This sequence is Cauchy in $L^1(X, \widetilde{\mathcal{A}}, \mu)$ and hence has a limit f which is $\widetilde{\mathcal{A}}$-measurable and so $\|f - \chi_A\|_1 = 0$. So $f^{-1}(1)$ is in $\widetilde{\mathcal{A}}$ and agrees with A up to sets of measure zero. $\square$

If $\mathcal{A}$ is countably generated by sets $(A_i)_{i=1}^{\infty}$, then the *atom (in $\mathcal{A}$)* of $x \in X$ is the smallest subset of $\mathcal{A}$ containing x. More precisely, for each i define B_i to be A_i (resp. $X \setminus A_i$) if x is in (resp. not in) A_i. The atom is $[x]_{\mathcal{A}} := \bigcap_{i=1}^{n} B_i$. To justify this, let $\mathcal{S}$ be the closure of the set $\{\chi_{A_i}\}_{i=1}^{\infty}$ under products, increasing limits, the operation of sending f to $1 - f$, and addition of functions f and g provided that $\{x : f \neq 0\}$ and $\{x : g \neq 0\}$ are disjoint. If x and y are two points in an atom then they are sent to the same value under the generating functions and hence under every function in $\mathcal{S}$ as desired. We have therefore also shown that $y \in [x]_{\mathcal{A}}$ implies that $x \in [y]_{\mathcal{A}}$.

Corollary 49. *If $\mathcal{A}$ is countably generated, $[x]_{\mathcal{A}}$ has full $\mu_x^{\mathcal{A}}$-measure almost surely. It follows that $\mu_x^{\mathcal{A}} = \mu_y^{\mathcal{A}}$ if $[x]_{\mathcal{A}} = [y]_{\mathcal{A}}$.*

Proof. If x is in the full measure set where $\mu_x^{\mathcal{A}}(A_i) = E[\chi_{A_i}|\mathcal{A}] = \chi_{A_i}(x)$, then μ_x assigns full measure to each B_i so the atom has full measure. $\square$

Lemma 50. *Let $\mathcal{F}$ be a dense collection of functions in $C(X)$. Then (X, T, μ) is ergodic if and only if for each $f \in \mathcal{F}$, $\mathrm{Av}_n(f)$ converges to $\int_X f d\mu$ pointwise almost-everywhere for each $f \in \mathcal{F}$.*

Proof. The forward direction is Birkhoff. For the reverse direction, von Neumann's ergodic theorem implies that $\mathrm{Av}_n(f)$ converges in L^2 to the projection of f onto the subspace of T-invariant functions. The image of $\mathcal{F}$ is onto the subspace of constant functions, so the only T-invariant functions are constant and the system is ergodic. $\square$

Lemma 51. *Given a pmps $(X, T, \mathcal{B}, \mu)$ of a compact metric space, let $\mathcal{E}$ be the σ-algebra of almost-invariant sets. Then $\mu_x^{\mathcal{E}}$ is almost surely an ergodic invariant probability measure.*

Proof. Fix $f \in L^1$. Note that $E[f \circ T|\mathcal{E}] = E[f|\mathcal{E}]$ since $\int_A f \circ T d\mu = \int_A f d\mu$ for any T-invariant A. Therefore,

$$\int_X f dT_* \mu_x^{\mathcal{E}} = \int_X f \circ T d\mu_x^{\mathcal{E}} = E[f \circ T|\mathcal{E}](x) = E[f|\mathcal{E}](x) = \int_X f d\mu_X^{\mathcal{E}}$$

for almost every x. So there is a full measure set where this equality holds on a countable dense set of $C(X)$ and hence on all of $C(X)$ and hence the desired invariance follows from the Riesz representation theorem. For ergodicity, choose a countable dense collection of function (f_n) in $C(X)$. For almost every $x \in X$ and every m, $\mathrm{Av}_n(f_m)(x) \longrightarrow E[f_m|\mathcal{E}](x) = E[f_m|\mathcal{E}](x)$. The limit is unchanged if we replaced x with $y \in [x]_{\mathcal{E}}$. This implies that $\mu_x^{\mathcal{E}}$ is ergodic. $\square$

Theorem 52 (The ergodic decomposition, 1959)**.** *For any T-invariant $\mu \in \mathcal{M}^1(X)$, there is a measure ν supported on the T-invariant ergodic measures of $\mathcal{M}^1(X)$ so that for any $f \in L^1$, $\int_X f d\mu = \int_{\mathcal{M}^1(X)} \int_X f d\tau d\nu(\tau)$.*

This is usually just written as $\mu = \int_{\mathcal{M}^1(X)} d\nu$. As a historical note, this is a consequence of work of Choquet-Bishop-de-Leeuw in 1959.

Proof. Pushforward the measure μ to a measure ν on $\mathcal{M}(X)$ under the map sending x to μ_x. Then for any $f \in L^1(X, \mu)$

$$\int_X f d\mu = \int_X E[f|\mathcal{E}](x) d\mu = \int_X \int_X f d\mu_x^{\mathcal{E}} d\mu = \int_{\mathcal{M}^1(X)} \int_X f d\tau d\nu(\tau).$$

$\square$

Remark 53. A more intricate proof that follows a similar argument is available for G-invariant probability measures where G is any σ-compact metrizable group. Note that such a measure is *ergodic* if the only G-almost-invariant sets are null or conull.

Worksheet for Lecture 9

Problem 1. (The Rokhlin Disintegration Theorem, 1952). Let $T : X \longrightarrow Y$ be a continuous map between compact metric spaces. Let $\mathcal{A}$ (resp. $\mathcal{B}$) be the Borel σ-algebra on X (resp. Y). Let μ be a probability measure on X and let ν be its pushforward to Y. Show that for almost every $y \in Y$ there is a probability measure μ_y supported on $T^{-1}(y)$ with the property that for any map $f \in L^1(X, \mu)$,

$$\int_X f d\mu = \int_Y d\nu(y) \int_{T^{-1}(y)} f d\mu_y$$

(Hint: Let $\mathcal{C}$ be the pullback of $\mathcal{B}$. Show that $\mu_x^{\mathcal{C}} = \mu_y^{\mathcal{C}}$ almost surely if and only if $T(x) = T(y)$).

Problem 2. (An example with disintegration) If $T : [0,1]^2 \longrightarrow [0,1]$ is the projection onto the first factor then find μ_x when μ is Lebesgue measure and when $\mu = f(x,y)dxdy$ for some bounded measurable function f.

Problem 3. In the notation of the last lecture, argue that $\mu_x^{\mathcal{E}} = \mu_{Tx}^{\mathcal{E}}$ almost surely. (Hint: Show that $E[f|\mathcal{E}] \circ T = E[f|\mathcal{E}]$ for any function f in L^1.)

Problem 4 (Lifts) Let B and X be compact metric spaces and suppose that we have the following commutative diagram,

$$\begin{tikzcd} B \times X \arrow[r, "S"] \arrow[d, "\pi"] & B \times X \arrow[d, "\pi"] \\ B \arrow[r, "T"] & B \end{tikzcd}$$

Suppose that S and T are measurable, invertible, and that μ is a T-invariant probability measure. Suppose that λ is a measure on $B \times X$ whose pushforward to B is μ. Show that λ is S-invariant if and only if the disintegrations $(\mu_b)_{b \in B}$ have the property that $S_* \mu_b = \mu_{S(b)}$ almost surely. (Hint: For the forward direction, in the notation of Problem 1, imitate the proof that $\mu_x^{\mathcal{E}}$ is T-invariant that we saw in class)