Topological Models of Measure-Preserving Systems

Daniel Raban

Contents

These notes are mostly based on Chapters 4, 10, and 12 from [EFHN15], rearranging for clarity, filling in some gaps, and choosing some slightly different presentations and proofs of results for brevity.

1 Motivation: Inducing measurable dynamics from topological dynamics

In the field of dynamical systems, the idea is to study a space X with some "dynamics" occurring on it, represented by repeated action of some map T. This setup takes a few forms, including but not limited to the following:

Definition 1.1. A **topological dynamical system** is a pair (X, T), where X is a (nonempty) compact, Hausdorff topological space and $T : X \to X$ is continuous.

Definition 1.2. A **measure-preserving system** is a tuple $(X, \mathcal{B}, \mu, T)$, where $(X, \mathcal{B}, \mu)$ is a probability space (i.e. a measure space with $\mu(X) = 1$) and $T : X \to X$ is **measure-preserving**: $\mu(T^{-1}E) = \mu(E)$ for all $E \in \mathcal{B}$.

We can express the measure-preserving property as $T_*\mu = \mu$, where $T_*\mu$ denotes the push-forward measure.

Recall the Krylov-Bogoliubov theorem, which tells us that topological dynamical systems naturally give rise to measure-preserving systems:

Theorem 1.1 (Krylov-Bogoliubov)**.** *Let X be a compact metric space and $T : X \to X$ be a continuous map. Then there exists a measure μ such that $T_*\mu = \mu$.*

Thus, given a TDS (X, T), we get (possibly many) MPSs $(X, \mathcal{B}_X, \mu, T)$, where $\mathcal{B}_X$ is the Borel σ-algebra on X. Here is a proof (which cites a few high-powered results[1]):

Proof. Let $x \in X$, and consider the point-mass measure δ_x. Then $T_*\delta_x = \delta_{Tx}, T_*^2\delta_x = \delta_{T^2x}$, and so on. So consider the average measures $\mu_n = \frac{1}{n}\sum_{k=0}^{n-1} \delta_{T^kx}$. These form a sequence of probability measures in the unit ball of $\mathcal{M}(X)$, the space of measures on X, so we need to justify their convergence.

By the Riesz-Markov-Kakutani representation theorem, the space of measures is the dual space of $C(X)$, and moreover, by the Banach-Alaoglu theorem, the unit ball of this space is compact in the weak* topology (the topology of convergence in distribution). The set of probability measures is a closed and hence compact

[1] For proofs, see [Fol13] or my pillowmath Math 245B notes.

subset. The weak* topology is metrizable, so the collection of probability measures is sequentially compact. Thus, there exists a convergent subsequence of the μ_n. Let μ be the limit of such a subsequence.

We claim that μ is T-invariant. It suffices to show that $\int_X f\, dT_*\mu = \int_X f\, d\mu$ for all $f \in C(X)$. We do this by comparing μ to μ_n:

$$\left|\int_X f \circ T\, d\mu - \int_X f\, d\mu\right| \leq \left|\int_X f \circ T\, d\mu - \int_X f \circ T\, d\mu_n\right| + \left|\int_X f \circ T\, d\mu_n - \int_X f\, d\mu_n\right| + \left|\int_X f\, d\mu_n - \int_X f\, d\mu\right|$$

The middle term is $|\int f\, dT_*\mu_n - \int f\, d\mu_n|$. Since $\int f\, d\mu_n = \sum_{k=0}^{n-1} f(T^k x)$, this term telescopes:

$$= \left|\int_X f \circ T\, d\mu - \int_X f \circ T\, d\mu_n\right| + \frac{1}{n}\left|f(x) - f(T^{n-1}x)\right| + \left|\int_X f\, d\mu_n - \int_X f\, d\mu\right|$$

By the definition of weak*-convergence, the first and last terms go to 0 as $n \to \infty$. The last term is bounded by $\frac{2}{n}\|f\|_u$, so

$$\xrightarrow{n\to\infty} 0.$$

So $T_*\mu = \mu$, as claimed. $\square$

The theory of topological models allows us to answer the question: Can we go backwards? Is every measure-preserving system actually derived from some topological dynamical system? In a vague philosophical sense, we are asking if whether every probabilistic system can be modeled spatially.

Amazingly, the answer is actually "yes, sort of"! We will prove the following:

Theorem 1.2. *Every abstract measure-preserving system is isomorphic to a topological measure-preserving system.*

There will be a few technicalities, including what "isomorphic" means, but for standard probability spaces, everything will work out in the nicest possible way.

2 Properties of $C(X)$

To compare a compact space X to a measure-preserving system, we will first show that all of the information contained in X is still present in $C(X)$. This will allow us to use functional analytic methods to compare function spaces of X and a MPS.

2.1 Separability of $C(X)$

Recall the following fact about compact metric spaces:

Lemma 2.1. *Every compact metric space X is separable.*

Proof. Fix $n \in \mathbb{N}^+$, and consider the collection of open balls $B(x, 1/n)$ with $x \in X$. Then $\{B(x, 1/n)\}_x$ forms an open cover of X, and compactness yields a finite subcover $B(x_1, 1/n), \dots, B(x_{r_n}, 1/n)$. Let $C_n = \{x_1, \dots, x_{r_n}\}$. Then $C := \bigcup_{n=1}^{\infty} C_n$ is countable and dense in X. $\square$

This separability extends to the Banach space $C(X)$, but the relationship between X and $C(X)$ is actually deeper than this. It turns out that $C(X)$ determines X up to homeomorphism, so we can obtain a lot of information about X from $C(X)$. In particular, the following is true regarding the separability of $C(X)$:

Theorem 2.1. *Let X be a compact, Hausdorff topological space. Then $C(X)$ is separable if and only if X is metrizable.*

Proof. ($\Longleftarrow$): Without loss of generality, we may assume that X is a metric space, since if $\phi : X \to Y$ is a homeomorphism with Y a metric space, then $\Phi : C(Y) \to C(X)$ sending $f \mapsto f \circ \phi$ is a homeomorphism. X is separable, so let $A \subseteq X$ be a countable dense subset; the idea is that we can approximate any continuous function using distance functions $d(\cdot, x)$ and hence by using $d(\cdot, x)$ with $x \in A$. In particular, let

$$D = \{d(\cdot, x) \in C(K) : x \in A\} \cup \{\mathbb{1}_X\},$$

where $\mathbb{1}_X$ is the constant 1 function. Then $\mathcal{D}$, the set of all finite products of elements in $\operatorname{span}_{\mathbb{Q}}(D)$, is a countable subalgebra of $C(X)$, and $\overline{\mathcal{D}}$ contains the constant functions, so $\overline{\mathcal{D}} = C(X)$ by the Stone-Weierstrass theorem.

($\Longrightarrow$): Suppose $C(X)$ is separable. We will construct a metric space homeomorphic to X. Let $\{f_0, f_1, f_2, \dots\}$ be a countable dense subset of X, and define the function

$$\varphi : X \to \prod_{n \in \mathbb{N}} \mathbb{C}, \qquad \varphi(x) = (f_0(x), f_1(x), \dots).$$

Equipping the latter space with the metric $d((z_n)_n, (w_n)_n) = \sum_{n=0}^{\infty} \frac{1}{2^n} \cdot \frac{|z_n - w_n|}{1+|z_n-w_n|}$, we claim that φ is a homeomorphism between X and $\varphi(X)$. Observe that

- φ is continuous: If $x_k \to x$, then $f_n(x_k) \to f_n(x)$ for each $n \in \mathbb{N}$ by the continuity of the f_n. So

  $$d(\varphi(x_k), \varphi(x)) = \sum_{n=0}^{\infty} \frac{1}{2^n} \cdot \frac{|f_n(x_k) - f_n(x)|}{1 + |f_n(x_k) - f_n(x)|} \xrightarrow{k\to\infty} 0$$

  (by the dominated convergence theorem). That is, $\varphi(x_k) \to \varphi(x)$.

- φ is injective: Suppose $x \neq y$. Then, since X is Hausdorff, $\{x\}$ and $\{y\}$ are closed. By Urysohn's lemma, there exists a continuous function $f : X \to [0,1]$ with $f(x) = 0$ and $f(y) = 1$. By the density of $\{f_0, f_1, f_2, \dots\}$ in $C(X)$, let $k \in \mathbb{N}$ be such that $\|f_k - f\|_u < 1/2$. Then $f_k(x) < 1/2$ and $f_k(y) > 1/2$, so $f_k(x) \neq f_k(y)$. Thus, $\varphi(x) \neq \varphi(y)$.

φ is a continuous injection from a compact space to a Hausdorff space, so its inverse $\varphi^{-1} : \varphi(X) \to X$ is automatically continuous. □

2.2 C*-algebra structure of $C(X)$

$C(X)$ does not just have the structure of a vector space. It has two other important structures: multiplication and complex conjugation.

Definition 2.1. A **Banach algebra** is a Banach space B, equipped with a continuous "multiplication" map $B \times B \to B$ such that $\|xy\| \leq \|x\|\|y\|$ for all $x, y \in B$.

Definition 2.2. A **C*-algebra** is a Banach algebra A along with an map $* : A \to A$ such that for all $x, y \in A$,

1. (Involution) $(x^*)^* = x$,
2. (Distributivity) $(x + y)^* = x^* + y^*$, $(xy)^* = y^* x^*$,
3. (Conjugation of scalars) $(\lambda x)^* = \overline{\lambda} x^*$ for $\lambda \in \mathbb{C}$.
4. (C*-algebra axiom) $\|xx^*\| = \|x\|^2$.

Example 2.1. Let H be a Hilbert space, and let $\mathcal{L}(H, H)$ be the collection of bounded linear operators from H to H. Then $\mathcal{L}(H, H)$ is a C*-algebra when equipped with the operator norm and the map $*$ sending an operator to its adjoint.

We will only use commutative C*-algebras, so the following two examples will be especially important.

Example 2.2. If X is a compact, Hausdorff space, $C(X)$ is a C*-algebra when equipped with the involution $f \mapsto \overline{f}$.

Example 2.3. Let $(X, \mathcal{B}, \mu)$ be a measure space. Then $L^\infty(X)$ is a C*-algebra when equipped with the involution $f \mapsto \overline{f}$.

2.3 Characterization of maximal ideals and homomorphisms

To understand the space $C(X)$, we will look at its **maximal ideals**, i.e. the maximal proper subspaces I with $fg \in I$ for any $f \in C(X)$ and $g \in I$.

Proposition 2.1. *Let X be a compact, Hausdorff space. The closed ideals I of $C(X)$ are precisely the sets of the form $I_F := \{f \in C(X) : f = 0 \text{ on } F\}$, where $F \subseteq X$ is closed.*

Here is a proof sketch. For the whole proof, see Section 4.2 of [EFHN15].

Proof. First, to check that I_F is a (closed) ideal, note that $I_F = \ker(\mathrm{res}_F)$, where restriction to F, $\mathrm{res}_F : C(X) \to C(F)$, is an algebra homomorphism.

Conversely, given I, define $F := \{x \in X : f(x) = 0 \,\forall f \in I\}$. The set F is closed, as $F = \bigcap_{f \in I} f^{-1}(\{0\})$, an intersection of closed sets. This construction gives $I \subseteq I_F$, and an approximation argument gives the reverse containment. □

Corollary 2.1. *Let X be a compact, Hausdorff space. The maximal ideals of $C(X)$ are $I_{\{x\}}$ for $x \in X$.*

Proof. To check that each $I_{\{x\}}$ is maximal, let $J \supseteq I_{\{x\}}$ be a proper ideal of $C(X)$. Then $\overline{J}$ is an ideal which is not all of $C(X)$ because the set of invertible elements of $C(X)$ is open. By the proposition, $\overline{J} = I_F$ for some closed set $F \subseteq X$. But since $I_F \supseteq I_{\{x\}}$ iff $F \subseteq \{x\}$, we get $I_{\{x\}} = \overline{J} = J$.

Conversely, suppose that I is a maximal (proper) ideal of $C(X)$. Then $\overline{I}$ is a closed ideal of $C(X)$ which is not all of $C(X)$ because the set of invertible elements of $C(X)$ is open. So $I = \overline{I} = I_F$ for some closed $F \subseteq X$, and maximality implies that F is a singleton. □

What we have shown is that by looking at the maximal ideals of $C(X)$, we can recover all the points in the space X. It now remains to show that we can recover the topological structure of X. To do this, we will step away from the maximal ideal characterization of the points and instead think of them as point-mass measures.

Let δ_x denote the point-mass probability measure at $x \in X$. We can think of δ_x as a linear functional on $C(X)$, namely via $\delta_x(f) = f(x)$. Moreover, observe that $I_{\{x\}} = \ker \delta_x$. So instead of relating points in X to maximal ideals, we will relate them to particular linear functionals on $C(X)$, which have a topology. In particular, we will relate them to **algebra homomorphisms** $C(X) \to \mathbb{C}$, i.e. linear functionals satisfying $\psi(fg) = \psi(f)\psi(g)$ and $\psi(\mathbb{1}_X) = 1$.

Lemma 2.2. *Let X be a compact, Hausdorff space. A linear functional $\psi : C(X) \to \mathbb{C}$ is an algebra homomorphism if and only if $\psi = \delta_x$ for some $x \in X$.*

Proof. First, observe that δ_x is multiplicative with $\delta_x(\mathbb{1}_X) = 1$. Conversely, suppose ψ is an algebra homomorphism. Then $\ker \psi$ is an ideal of $C(X)$, and it is maximal because $\dim C(X)/\ker \psi = \dim \operatorname{im} \psi = 1$. Thus, $\ker \psi = I_{\{x\}}$ for some $x \in X$. Now, for any $f \in C(X)$, $f - \psi(f)\mathbb{1}_X \in \ker \psi = I_{\{x\}}$, so

$$0 = f(x) - \psi(f)\mathbb{1}_X(x) = f(x) - \psi(f).$$

Thus, $\psi(f) = f(x) = \delta_x(f)$ for all $f \in C(X)$. □

2.4 The Gelfand-Naimark theorem

This gives us our characterization of X from $C(X)$.

Theorem 2.2. *Let X be a compact, Hausdorff space, and let the* ***Gelfand space*** *of $C(X)$ be*

$$\Gamma(C(X)) := \{\psi \in C(X)^* : \psi \text{ is an algebra homomorphism}\}.$$

Then the map $\delta : X \to \Gamma(C(X))$ sending $x \mapsto \delta_x$ is a homeomorphism, where $\Gamma(C(K))$ has the weak topology inherited from $C(X)^*$.*

Proof. By the lemma, the map δ is surjective. The map δ is injective, as if $x \neq y$, then since X is Hausdorff, $\{x\}$ and $\{y\}$ are closed. By Urysohn's lemma, there exists a continuous function $f : X \to [0, 1]$ with $f(x) = 0$ and $f(y) = 1$, which shows $\delta_x(f) = f(x) \neq f(y) \neq \delta_y(f)$. To show that the map is continuous, suppose $x_n \to x$. To show that $\delta_{x_n} \to \delta_x$ in the weak* topology, we test these against any continuous function f:

$$\delta_{x_n}(f) = f(x_n) \xrightarrow{n\to\infty} f(x) = \delta_x(f),$$

by the continuity of f. Finally, since δ is a continuous bijection from a compact space to a Hausdorff space, its inverse is automatically continuous. □

Remark 2.1. At first glance, you might think that this whole process of characterizing maximal ideals and algebra homomorphisms was a total waste of time, since the map $x \mapsto \delta_x$ can be defined without knowing these things. If you look carefully at the proof, the only use of these characterizations came in play to show that the map δ was surjective, so it seems like we could have bypassed all this work by just concluding that X is homeomorphic to $\delta(X)$. However, this is not sufficient for our purposes because $\delta(X)$ needs to be characterized intrinsically via $C(X)$ without knowledge of the space X. Otherwise, we would not be able to show that $C(X)$ determines X, as we want.

Corollary 2.2. *Let X, Y be compact Hausdorff spaces. Then X, Y are homeomorphic if and only if the algebras $C(X), C(Y)$ are isomorphic.*

Proof. ($\implies$): If $\phi : X \to Y$ is a homeomorphism, then $\Phi : C(Y) \to C(X)$ sending $f \mapsto f \circ \phi$ is an algebra isomorphism.

($\impliedby$): If $C(X) \cong C(Y)$ as algebras, then $\Gamma(C(X))$ is homeomorphic to $\Gamma(C(Y))$. By the theorem, X is homeomorphic to Y. $\square$

The Gelfand map Γ plays a very important role in the theory of all commutative C*-algebras, not just $C(X)$. The following theorem will be instrumental in our proof of the existence of topological models.

Theorem 2.3 (Gelfand-Naimark)**.** *Let A be a commutative C*-algebra. Then there is a compact, Hausdorff space X and an isometric isomorphism $\Phi : A \to C(X)$ that commutes with $*$. The space X is unique up to homeomorphism.*

We will not prove this, but the construction is to set $X = \Gamma(A)$, the set of linear functionals which are algebra homomorphisms. For the full proof, see Chapter 4 of [EFHN15], [Dix82], or my pillowmath Math 259A notes.

Remark 2.2. This may be surprising, given that L^∞ of a measure space is a C*-algebra. You can reassure yourself with the notion that if $C(X) \cong L^\infty(Y)$ as C*-algebras, then X may not look too similar to Y. We will later see what the compact space X may look like.

3 Studying dynamics via Koopman operators

The purpose of this section is to show that instead of the studying the topological or measure-preserving systems themselves, it is sufficient to study their Koopman operators.

3.1 Koopman operators for topological dynamical systems

We can study a topological dynamical system (X, T) by looking at the action of T on continuous functions by pre-composition:

Definition 3.1. Let (X, T) be a topological dynamical system. The **Koopman operator** is the operator $U_T : C(X) \to C(X)$ sending $f \mapsto f \circ T$.

The following theorem tells us that if we can find an operator which looks like the Koopman operator, we can recover the dynamics on the space X. The key property is that U_T is an **algebra homomorphism**, a linear map with $U_T(fg) = U_T(f)U_T(g)$ and $U_T(\mathbb{1}_X) = \mathbb{1}_X$.

Theorem 3.1. *Let X be a compact, Hausdorff space, and let $U : C(X) \to C(X)$ be an algebra homomorphism. Then there exists a unique $T \in C(X)$ such that $U = U_T$, i.e. $U(f) = f \circ T$ for all $f \in C(X)$.*

The first step is proving that if we can find such a T, then it will be continuous.

Lemma 3.1. *Let X be a compact, Hausdorff space. $T : X \to X$ is continuous if and only if $f \circ T$ is continuous for all $f \in C(X)$.*

Let's prove the lemma.

Proof. If T is continuous, then $f \circ T$ is continuous for $f \in C(X)$ as compositions of continuous functions are continuous. Conversely, suppose $f \circ T$ is continuous for all $f \in C(X)$. We want to show that $T^{-1}(V)$ is open for all open $V \subseteq X$, and we have that $T^{-1}(f^{-1}(W)) = (f \circ T)^{-1}(W)$ is open for each $f \in C(X)$ and open $W \subseteq \mathbb{C}$. So it suffices to show that every open $V \subseteq X$ can be expressed as a union of $f^{-1}(W)$ for open $W \subseteq X$.

For a nonempty open $V \subsetneq X$, let $x \in V$. Then, as X is Hausdorff, $\{x\}$ is closed, so by Urysohn's lemma, there exists an $f_x \in C(X)$ such that $f_x(x) = 0$ and $f_x(y) = 1$ for all $y \in X \setminus V$. Thus, $f_x^{-1}(\mathbb{C} \setminus \{1\})$ is an open set containing x which is contained in V, and we thus have $V = \bigcup_{x \in V} f_x^{-1}(\mathbb{C} \setminus \{1\})$. $\square$

Now we can prove the theorem:

Proof. Consider the adjoint map $U^* : C(X)^* \to C(X)^*$, which satisfies $[U^*F](f) = F(Uf)$ for each $f \in C(X)$ and $F \in C(X)^*$. Then, letting δ_x be the point-mass measure at $x \in X$, viewed as the linear functional $\delta_x(f) = f(x)$, observe that $U^*(\delta_x) : C(X) \to \mathbb{C}$ is an algebra homomorphism:

$$[U^*(\delta_x)](fg) = \delta_x(U(fg)) = \delta_x(UfUg) = \delta_x(Uf)\delta_x(Ug) = [U^*(\delta_x)](f) \cdot [U^*(\delta_x)](g),$$

$$[U^*(\delta_x)](\mathbb{1}_X) = \delta_x(U\mathbb{1}_X) = \delta_x(\mathbb{1}_X) = 1.$$

By our previous characterization of algebra homomorphisms $C(X) \to \mathbb{C}$, there is a unique $y =: T(x)$ such that $U^*(\delta_x) = \delta_{T(x)}$. Thus,

$$[Uf](x) = \delta_x(Uf) = [U^*(\delta_x)](f) = \delta_{T(x)}(f) = f(T(x)),$$

yielding $Uf = f \circ T$. By the lemma, T is continuous, so we are done. □

3.2 Replacing measure-preserving systems by their Koopman operators

Similarly to the topological case, a measure-preserving system $(X, \mathcal{B}, \mu, T)$ also has an associated Koopman operator:

Definition 3.2. Let $(X, \mathcal{B}, \mu, T)$ be a measure-preserving system. The **Koopman operator** is the operator $U_T : L^1(X) \to L^1(X)$ sending $f \mapsto f \circ T$.

To use functional analytic techniques to compare measure-preserving systems, we will be comparing their Koopman operators. Just as algebra homomorphisms $C(X) \to C(X)$ correspond to Koopman operators for topological dynamical systems, Koopman operators have an analogue in the $L^1(X)$ setting.

Definition 3.3. Let $(X, \mathcal{B}_X, \mu)$ and $(Y, \mathcal{B}_Y, \nu)$ be measure spaces. An operator $U : L^1(Y) \to L^1(X)$ is called a **Markov embedding** if

(i) (Positivity) $Uf \geq 0$ when $f \geq 0$,

(ii) (Preserves identity) $U\mathbb{1}_Y = \mathbb{1}_X$,

(iii) (Preserves integration) $\int_X Uf\, d\mu = \int_Y f\, d\nu$ for all $f \in L^1(Y)$,

(iv) (Embedding condition) $|Uf| = U|f|$ for all $f \in L^1(Y)$.

In the case $X = Y$, the pair (X, U) is called an **abstract measure-preserving system**.

Observe that a Koopman operator $U_T : L^1(X) \to L^1(X)$ is a Markov embedding, so every measure-preserving system gives rise to an abstract measure-preserving system.

You may be wondering why this definition allows for a different domain and codomain. This is because Markov embeddings describe both Koopman operators and the maps that relate Koopman operators to each other. First, recall how we usually compare measure-preserving systems:

Definition 3.4. Let $(X, \mathcal{B}_X, \mu, T)$ and $(Y, \mathcal{B}_Y, \nu, S)$ be measure-preserving systems. A **factor map** is a measurable map $\phi : X \to Y$ satisfying the following (replacing X and Y by full measure subsets $X' \subseteq X$ and $Y' \subseteq Y$, respectively, if needed):

(i) ϕ is measure preserving: $\mu(\phi^{-1}A) = \nu(A)$ (or equivalently, $\phi_*\mu = \nu$).

(ii) ϕ *converts the dynamics* of X into the dynamics of Y: $\phi \circ T(x) = S \circ \phi(x)$ for every $x \in X$.

$$\begin{tikzcd} X \arrow[r, "T"] \arrow[d, "\phi"'] & X \arrow[d, "\phi"] \\ Y \arrow[r, "S"'] & Y \end{tikzcd}$$

Note that if we replace X with X' and Y with Y', then we still need $TX' \subseteq X'$ and $SY' \subseteq Y'$ for the dynamics to make sense.

Here are the Markov operators that act as factor maps to compare Koopman operators:

Definition 3.5. Let $U : L^1(X) \to L^1(X)$ and $V : L^1(Y) \to L^1(Y)$ be Markov operators. A Markov embedding $\Phi : L^1(Y) \to L^1(X)$ is **intertwining** for U and V if $\Phi \circ V = U \circ \Phi$.

Definition 3.6. A **Markov isomorphism** is a surjective (and hence invertible) Markov embedding. Two abstract measure-preserving systems $(X, U), (Y, V)$ are **isomorphic** if there exists an intertwining Markov isomorphism $\Phi : L^1(Y) \to L^1(X)$.

Proposition 3.1. *Let* $(X, \mathcal{B}_X, \mu, T)$ *and* $(Y, \mathcal{B}_Y, \nu, S)$ *be isomorphic as measure-preserving systems. Then the abstract measure-preserving systems* $(X, U_T), (Y, U_S)$ *are isomorphic.*

Proof. Let $\phi : X \to Y$ be an isomorphism, and define $\Phi : L^1(Y) \to L^1(X)$ by $\Phi f = f \circ \phi$. Then Φ is an intertwining Markov embedding, and it is invertible via Φ^{-1} sending $g \mapsto g \circ \phi^{-1}$. $\square$

Unfortunately, we come now to the main caveat of our theory: The correspondence does not always go the other way. The Koopman operator $U_T : L^1(X) \to L^1(X)$ gives us information about how T acts on sets in $\mathcal{B}_X$ because $\mathcal{B}_X/\sim \subseteq L^1(X)$, where $A \in \mathcal{B}_X$ is identified $\mathbb{1}_A \in L^1(X)$ and $A \sim B$ iff $\mu(A \triangle B) = 0$.[2] Indeed an isomorphism of abstract measure-preserving systems induces an isomorphism of $\mathcal{B}_X/\sim$.[3] But, as the following example shows, if the σ-algebras involved are not rich enough to give good resolution of measurable subsets of our space, we may not get isomorphism of the underlying measure-preserving systems.

Example 3.1. Let $X = \{0\}$, $\mathcal{B}_X = \{\varnothing, X\}$, $\mu(X) = 1$, and $T = \mathrm{id}_X$, and let $Y = \{0, 1\}$, $\mathcal{B}_Y = \{\varnothing, Y\}$, $\nu(Y) = 1$, and $S = \mathrm{id}_Y$. These are not isomorphic as measure-preserving systems, but $L^1(X)$ and $L^1(Y)$ are the constant functions on X and Y, respectively, so $(X, U_T), (Y, U_S)$ are isomorphic via the intertwining Markov isomorphism $c\mathbb{1}_Y \mapsto c\mathbb{1}_X$.

For nice spaces, these notions agree!

Theorem 3.2. *Let $(X, \mathcal{B}_X, \mu), (Y, \mathcal{B}_Y, \nu)$ be standard probability spaces. An isomorphism on (X, U_T) and (Y, U_S) induces an a.e.-uniquely determined isomorphism beween X and Y as measure-preserving systems.*

Lemma 3.2 (von Neumann)**.** *Let $(X, \mathcal{B}_X, \mu), (Y, \mathcal{B}_Y, \nu)$ be standard probability spaces, and let $U : L^1(Y) \to L^1(X)$ be a Markov embedding. Then there is a μ-almost everywhere unique measure-preserving map $f : X \to Y$ such that $U = U_f$ (i.e. U sends $g \mapsto g \circ f$).*

We omit the proof of the lemma. For the proof (which is not so long), see Appendix F of [EFHN15].

Proof. Using the lemma with U, V, we get measure-preserving maps $T : X \to X$ and $S : Y \to Y$ with $U = U_T$ and $V = U_S$. Using the lemma with an intertwining Markov isomorphism $\Phi : L^1(Y) \to L^1(X)$, we get measure-preserving maps $\phi : X \to Y$ and $\phi^{-1} : Y \to X$ with $\Phi = U_\phi$ and $\Phi^{-1} = U_{\phi^{-1}}$. To show these are actually (measurable) inverses, observe that

$$U_{\phi^{-1} \circ \phi} = U_\phi \circ U_{\phi^{-1}} = \Phi \circ \Phi^{-1} = \mathrm{id}_{L(Y)} = U_{\mathrm{id}_Y},$$

[2] The structure $\mathcal{B}_X/\sim$ is sometimes referred to as the **measure algebra** of the measure-preserving system.

[3] For a proof of this, see Theorem 12.10 of [EFHN15].

so uniqueness in the lemma gives $\phi^{-1} \circ \phi = \mathrm{id}_Y$ ν-a.e. The same argument applies to $\phi \circ \phi^{-1}$. To show that these are μ-a.e. intertwining, we use the same argument with

$$U_{\phi \circ T} = U_T \circ U_\phi = U \circ \Phi = \Phi \circ V = U_\phi \circ U_S = U_{S \circ \phi}$$

and apply the a.e. uniqueness once more. $\square$

Our construction of topological models will apply to abstract measure-preserving systems and use this looser notion of isomorphism, so it may not completely satisfy your philosophical broodings about the nature of measure-preserving systems. However, often in applications, dealing with the Koopman operator of a measure-preserving system (and more generally Markov operators) is enough to understand the systems at play, so we cheerfully sweep this philosophical discrepancy under the rug.[4]

3.3 Properties of Markov embeddings

For our proof, we will need to understand Markov embeddings a bit better, so we'll prove a few properties here.

Lemma 3.3. *If $U : L^1(Y) \to L^1(X)$ is a Markov embedding, then $U\mathbb{1}_A$ is an indicator which we denote by $\mathbb{1}_{UA}$. Moreover, $U(A \cup B) = UA \cup UB$ and $U(A \cap B) = UA \cap UB$.*

Proof. For the first claim, it suffices to show that $U\mathbb{1}_A$ takes values in $\{0, 1\}$. The embedding property gives

$$\left| U\mathbb{1}_A - \frac{1}{2} \right| = \left| U\left(\mathbb{1}_A - \frac{1}{2}\mathbb{1}_Y \right) \right| = U\left| \mathbb{1}_A - \frac{1}{2}\mathbb{1}_Y \right| = U\left(\frac{1}{2}\mathbb{1}_Y \right) = \frac{1}{2}\mathbb{1}_X,$$

so $U\mathbb{1}_A$ is always distance $1/2$ from $1/2$.

For the union property, the embedding property gives

$$\begin{aligned} U\mathbb{1}_{A\cup B} &= U \max\{\mathbb{1}_A, \mathbb{1}_B\} \\ &= U\left(\frac{\mathbb{1}_A + \mathbb{1}_B + |\mathbb{1}_A - \mathbb{1}_B|}{2} \right) \\ &= \frac{\mathbb{1}_{UA} + \mathbb{1}_{UB} + |\mathbb{1}_{UA} - \mathbb{1}_{UB}|}{2} \end{aligned}$$

[4] The real answer is "topological models are cool, so I tricked you into reading these notes with a vague promise of philosophy."

$$
\begin{aligned}
&= \max\{\mathbb{1}_{UA}, \mathbb{1}_{UB}\} \\
&= \mathbb{1}_{UA\cup UB}.
\end{aligned}
$$

For the intersection property, we can use

$$
\begin{aligned}
U\mathbb{1}_{A\cap B} &= U\min\{\mathbb{1}_A, \mathbb{1}_B\} \\
&= U(\mathbb{1}_A + \mathbb{1}_B - \max\{\mathbb{1}_A, \mathbb{1}_B\}) \\
&= \mathbb{1}_{UA} + \mathbb{1}_{UB} - \max\{\mathbb{1}_{UA}, \mathbb{1}_{UB}\} \\
&= \min\{\mathbb{1}_{UA}, \mathbb{1}_{UB}\} \\
&= \mathbb{1}_{UA\cap UB}.
\end{aligned}
$$

□

Remark 3.1. This property can be extended to show that U induces a homomorphism on $\mathcal{B}_X/\sim$ and $\mathcal{B}_Y/\sim$, but we will not show that here.

Proposition 3.2. *If $U : L^1(X) \to L^1(X)$ is a Markov embedding, it is an algebra homomorphism when restricted to L^∞.*

Proof. Since U is linear and preserves the identity, we need only prove that it preserves multiplication of $f, g \in L^\infty$. By linearity, it also suffices to prove this when f, g are real-valued. First, we can show this when $f = g$: Let $\phi_j = \sum_{i=1}^{n_j} c_{i,j}\mathbb{1}_{A_{i,j}}$ be an L^1 approximation to f by simple functions (with indicators on disjoint sets). Then

$$
U\phi_j^2 = U\sum_{i=1}^{n_j} c_{i.j}^2\mathbb{1}_{A_{i,j}} = \sum_{i=1}^{n_j} c_{i.j}^2\mathbb{1}_{UA_{i,j}} = (U\phi_j)^2,
$$

So

$$
\begin{aligned}
\|(Uf)^2 - Uf^2\|_1 &\le \|(Uf)^2 - (U\phi_j)^2\|_1 + \|(U\phi_j)^2 - U\phi_j^2\|_1 + \|U\phi_j^2 - Uf^2\|_1 \\
&\le \|Uf + U\phi_j\|_\infty \|U(f-\phi_j)\|_1 + \|\phi_j^2 - f^2\|_1 \\
&= \|Uf + U\phi_j\|_\infty \|f-\phi_j\|_1 + \|\phi_j + f\|_\infty \|\phi_j - f\|_1
\end{aligned}
$$

For large enough j,

$$
\begin{aligned}
&\le (2\|Uf\|_\infty + 1)\|(f-\phi_j)\|_1 + (2\|f\|_\infty + 1)\|\phi_j - f\|_1 \\
&\xrightarrow{j\to\infty} 0.
\end{aligned}
$$

We now apply $Uf^2 = (Uf)^2$ to the polarization identity $2fg = (f+g)^2 - f^2 - g^2$ to get

$$
U(fg) = \frac{1}{2}U((f+g)^2 - f^2 - g^2) = \frac{1}{2}\left((Uf + Ug))^2 - (Uf)^2 - (Ug)^2\right) = (Uf)(Ug),
$$

which completes the proof. □

4 Topological models

4.1 Construction of topological models

We are now prepared to construct topological models.

Definition 4.1. Let (X, U) be an abstract measure-preserving system. A **topological model** of X is a measure-preserving system $(K, \mathcal{B}, \nu, T)$ such that K is a compact, Hausdorff space, $T : K \to K$ is continuous, and there is an intertwining Markov isomorphism

$$\Phi : (K, U_T) \to (X, U).$$

Theorem 4.1. *Every abstract measure-preserving system admits at least one topological model.*

Remark 4.1. Topological models are not in general unique. The proof will actually produce machinery to construct topological models using *U-invariant C*-subalgebras of* $L^\infty(X)$ (which are dense in $L^1(X)$). Different subalgebras may result in different topological models, which may have different properties such as ergodicity.

Proof. Let (X, U) be an abstract measure-preserving system, and let A be a U-invariant C*-subalgebra of $L^\infty(X)$ which is dense in $L^1(X)$ (for example, we could take $A = L^\infty(X)$ itself).[5] By the Gelfand-Naimark theorem, there exist a compact, Hausdorff space K an a C*-algebra isomorphism $\Phi : C(K) \to A$.

Having constructed the space K, we now construct the probability measure on K. Consider the linear functional $L : C(K) \to \mathbb{C}$ sending $f \mapsto \int_X \Phi f \, d\mu$. L is bounded, as $\|L\| \leq 1$, so by the Riesz-Markov-Kakutani representation theorem, there exists a measure ν on K such that $\int_K f \, d\nu = \int_X \Phi f \, d\mu$ for all $f \in C(X)$. To check that ν is a positive measure, note that $\int_K f \, d\nu = \int_X \Phi f \, d\mu \geq 0$ whenever $f \geq 0$; approximating an indicator by nonnegative continuous functions in L^1, we get $\nu(E) = \int_K \mathbb{1}_E, d\nu \geq 0$ for all measurable $E \subseteq K$. Moreover,

$$\nu(K) = \int_K \mathbb{1}_K \, d\nu = \int_X \Phi \mathbb{1}_K \, d\mu = \int_X \mathbb{1}_X \, d\mu = \mu(X) = 1,$$

so ν is a probability measure.

We now upgrade the C*-algebra isomorphism Φ into the desired intertwining Markov isomorphism. To extend Φ to all of $L^1(K)$, we will first show that it is an isometry in the L^1 norm. Using the properties of Φ as an algebra homomorphism,

$$(\Phi|f|)^2 = \Phi|f|^2 = \Phi(f\overline{f}) = (\Phi f)(\Phi\overline{f}) = (\Phi f)(\overline{\Phi f}) = |\Phi f|^2,$$

[5] A C*-subalgebra is by definition closed in the norm topology of $L^\infty(X)$.

which gives $\Phi|f| = |\Phi f|$. Thus, Φ is an isometry:

$$\|\Phi f\|_{L^1(X)} = \int_X |\Phi f|\, d\mu = \int_X \Phi|f|\, d\mu = \int_K |f|\, d\nu = \|f\|_{L^1(K)}.$$

So Φ extends uniquely to an isometry $L^1(K) \to L^1(X)$ by defining $\Phi(\lim f_n) = \lim \Phi f_n$, where the limits are in the L^1 sense. Moreover, Φ is a Markov isomorphism:

(i) (Embedding condition) For $f \in C(K)$, we already have $|\Phi f| = \Phi|f|$. The property extends to all $f \in L^1(K)$ by approximation, due to the continuity of Φ and $|\cdot|$.

(ii) (Positivity) If $f \geq 0$, then $|\Phi f| = \Phi|f| = \Phi f$, so $\Phi f \geq 0$.

(iii) (Preserves identity) This follows from the algebra homomorphism property.

(iv) (Preserves integration) $\int_X \Phi f\, d\mu = \int_K f\, d\nu$ for all $f \in C(K)$ by the definition of ν, and this equality extends to all $f \in L^1(K)$ by approximation.

(v) (Surjective): Since Φ is an isometry, it is an open map. Thus, the range of Φ is closed in $L^1(X)$, and since A is dense in $L^1(X)$, Φ is surjective.

We now construct the dynamics on K. Consider the map $\Phi^{-1}U\Phi : C(K) \to C(K)$. This is an algebra homomorphism, as Φ, Φ^{-1}, and U are. So there exists a unique continuous $T : K \to K$ such that $U_T = \Phi^{-1}U\Phi$. This implies $\Phi \circ U_T = U \circ \Phi$, so Φ is intertwining. Finally, ν is T-preserving, as

$$\int_K f \circ T\, d\nu = \int_K \Phi^{-1}U\Phi f\, d\nu = \int_X U\Phi f\, d\mu = \int_X \Phi f\, d\mu = \int_K f\, d\nu$$

for all $f \in C(K)$. □

4.2 Properties of topological models

Now that we have constructed topological models from abstract measure-preserving systems, let's endeavor to understand these topological models better.

Proposition 4.1. *Let $(K, \mathcal{B}, \nu, T)$ be a topological model for the abstract measure-preserving system (X, T), constructed as above. Then K is* ***faithful****; i.e.* $\operatorname{supp} \nu = K$.

Proof. □

We may also ask whether the compact space K is a metric space or not, i.e. whether there exists a **metric model** for (X, U). Fortunately, there is precisely a characterization of this!

Proposition 4.2. *An abstract measure-preserving system* (X, U) *admits a metric model* $(K, \mathcal{B}, \nu, T)$ *if and only if* $L^1(X)$ *is separable.*

The idea is to leverage our characterization of which compact, Hausdorff spaces are metrizable: K is metrizable iff $C(K)$ is separable.

Proof. ($\implies$): Suppose K is a metric space. Then $C(K)$ is separable, so we get that $L^1(K)$ is separable (since uniform convergence implies L^1 convergence). The isometric isomorphism $\Phi : L^1(K) \to L^1(X)$ thus provides a countable dense subset of $L^1(X)$.

($\impliedby$): If $L^1(X)$ is separable, then so is $L^\infty(X)$ (with the L^1 norm), as we can replace any countable dense subset of $L^1(X)$ by a countable dense subset in $L^\infty(X)$ using the density of $L^\infty(X)$ in $L^1(X)$. Moreover, we can assume that the countable subset $D \subseteq L^\infty(X)$ contains $\mathbb{1}_X$. By adding $Uf, U^2 f, \dots$ for each $f \in D$ (still keeping D countable) and by adding in conjugates, we can assume that D is closed under U and under conjugation. Now let A be the closure of D in L^∞ (with respect to the $\|\cdot\|_\infty$ topology). Then A is a C^*-subalgebra of $L^\infty(X)$ which is separable in the $L^1(X)$ topology (as L^∞ convergence implies L^1 convergence) and dense in $L^1(X)$. Our construction of topological models gives an isometric isomorphism $\Phi : C(K) \to A$, and we can use Φ^{-1} on D to obtain a countable, dense subset of $C(K)$. Thus, $C(K)$ is separable, so K is metrizable. $\square$

References

[Dix82] J. Dixmier. *C^*-algebras*. North-Holland mathematical library. North-Holland, 1982.

[EFHN15] Tanja Eisner, Bálint Farkas, Markus Haase, and Rainer Nagel. *Operator theoretic aspects of ergodic theory*, volume 272. Springer, 2015.

[Fol13] G.B. Folland. *Real Analysis: Modern Techniques and Their Applications*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2013.