Ergodic Theory I

Cam Connor
Western University

April 22 2015

Table Of Contents

1 Abstract

Ergodic Theory refers to nice mathematical models for dynamical systems. More specifically it deals with measure preserving transformations on a measure space. We will consider ergodic theory on Hilbert spaces and take advantage of conservation of energy. Having an understanding of Hilbert spaces and other material from our Functional Analysis course is assumed. Previous knowledge of physics may help, as this theory is a direct application of physics. Additionally further experience with measure theory will be useful, but since I had limited knowledge of measure theory beforehand I avoid assuming much experience in the field.

We will mostly be following the set-up of [2], with examples from [3] interjected. In the beginning of this lecture we will set up what we are working with and explore a useful result (Koopman's Lemma) for unitary operators. Then we will define ergodic and explore an equivalent definition, which will allow us to look at a nice example of an ergodic operator. Once we understand what it means to be ergodic we will state and prove the useful von Neumann's Ergodic Theorem, which it turn reveals yet another equivalent characterization of ergodic functions. Finally, we will state the Birkhoff Ergodic Theorem and briefly mention a nice result of it.

2 The Set-Up

Following [2], we will be looking at connecting functional analysis with mathematical physics, by considering time-dependent systems' behaviour (in classical mechanics) after running for a long time.

Let's begin by looking at some phase space Γ. You can think of this as all possible states of an object, states are generally the position or momentum of an object.

For a time t, $\exists$ a map $T_t : \Gamma \to \Gamma$, where $T_t(x)$ is the state which results from looking the time evolution of a system from time t_0 to $t + t_0$. Then $T_{s+t} = T_s T_t$

While in classical mechanics we can look at observables (energy, position, temperature etc.) as functions on the phase space. Let T be the period; we want to run our system for a long time so we can consider $T \to \infty$. Note that the symbol for period T and transformation T_t are alike.

Also we will consider energy being constant in our system, which we can as energy is conserved in an isolated system, $E_{avg} = E_{int}$. Let Ω_E be our constant energy surface in our given phase space for each $w \in \Omega_E$ and for a continuous function f we get a well-defined measure $\mu(f)$[1].

$$\mu(f) = \lim_{T\to\infty} \frac{1}{T} \int_0^T f(T_t w)dt$$

We get the following nice properties:
1) $\mu(1) = 1$
Simply substitute in in $f = 1$
2) μ is linear
Shown by substituting in the following into the integration and re-arranging $\mu(af) = a\mu(f)$ and $\mu(f) + \mu(g) = \mu(f + g)$
3) $\mu(f) \geq 0$ when $f \geq 0$
As $T \geq 0$ since T is the period and therefore positive then again simply substitute in

When $F \subset \Omega_E$, let $\mu(F) := \mu(\chi_F)$

[1]Showing that this is a measure involves sufficient knowledge of measure theory on compact spaces, which is an extension of the lectures of this course, however it outside the scope of the material presented. As a result it is omitted, that being said section IV.4 of [2] deals with this concept.

The measure also has the following nice property:

For a fixed time s and a measurable set $F \subset \Omega_E$ with a characteristic function χ_F of F. Then

$$\frac{1}{T}\int_0^T \chi_{T_{s^{-1}}F}(T_t w)dt = \frac{1}{T}\int_0^T \chi_F(T_s T_t w)dt$$

So if the limit exists then $\mu(T_s^{-1}F) = \mu(F)$
We say that μ is *invariant* and T_s is *measure preserving.*

Hilbert space methods are nice enough that we can reformulate the problem in terms of $L^2(\Omega_E, d\mu_E)$. Therefore if $f \in L^2(\Omega_E, \mu_E)$ define a map:

$$U_t : f \to f \circ T_t$$

$$(U_t f)(w) = f(T_t w)$$

We can show this is unitary from the following theorem.

Theorem 1 (Koopman's Lemma)**.** *U_t is a unitary map of $L^2(\Omega_E, d\mu_E)$ onto $L^2(\Omega_E, d\mu_E)$.*

Proof.

$$\langle U_t f, U_t g\rangle = \int_{\Omega_E} f(T_t w)\overline{g(T_t w)}d\mu_E(w)$$

Let $y = T_t w$

$$= \int_{\Omega_E} f(y)\overline{g(y)}d\mu_E(T_t^{-1}y)$$

We have $\mu(T_t^{-1}y) = \mu(y)$ so

$$= \int_{\Omega_E} f(y)\overline{g(y)}d\mu_E(y)$$

$$= \langle f, g\rangle$$

So U is an isometry as it perseveres the inner product.
And $U_t U_{-t} = U_{t-t} = U_0$
While $(U_0 f)(w) = f(T_0 w) = f(w)$. So $U_0 = I$ and U_t is invertible.

$\Rightarrow U_t^{-1} = U_t^*$ since invertible and isometry. $U_t U_t^* = U_t^* U_t = I$
$\Rightarrow U_t$ is unitary.

□

3 Ergodic

In the continuous case we have $U_t f = f \circ T_t$, we are concerned with what functions in $L^2(\Omega_E, d\mu_E)$ will satisfy $U_t f = f$.

If f is constant, say $f(w) = c$, $\forall w \in \Omega_E$ then
$(U_t f)(w) = (f \circ T_t)(w) = f(T_t(w)) = c$ as $T_t : \Omega_E \to \Omega_E$
Definition 3.1. T_t is called **ergodic** if the constant functions are the only functions in $L^2(\Omega_E, d\mu_E)$ for which $f \circ T_t = f$ (as L^2 functions) for all t.

Proposition 1. *T_t is ergodic if and only if for all measurable sets $F \subset \Omega_E$, we have $T_t^{-1}F = F$ for all t then $\mu_E(F) = 0$ or $\mu_E(F) = 1$.*

Proof.

$\Rightarrow$ Suppose T_t is ergodic and $T_t^{-1}F = F$ $\forall t$ (so T_t^{-1} sends F to F).
Then for $f = \chi_F$, we get $T_t^{-1}\chi_F = \chi_F \Rightarrow \chi_F = T_t \circ \chi_F$ so χ_F is constant (as T_t is ergodic).
Then $\chi_F \equiv 0$ almost everywhere or $\chi_F \equiv 1$ almost everywhere.
Therefore almost everything is in F or almost nothing is in F,
which gives us $\mu_E(F) = 0$ or $\mu_E(F) = 1$ as required.

$\Leftarrow$ This direction depends on further measure theory knowledge,
however the outlining idea is as follows.
Suppose that $\mu_E(F) = 0$ or $\mu_E(F) = 1$ for measurable sets F and $T_t^{-1}F = F$ for all t.

Then $B = \{w \in F | f(w) < a\}$ is invariant under T_t, then $\mu_E(B) = 0$ and $\mu_E(F - B) = 1$, or the opposite is true; so $f(w) < a$ almost everywhere or $f(w) \geq a$ almost everywhere.
Since true for an arbitrary a, $f(w)$ can only vary on a set of measure 0. Therefore $f(w)$ is constant almost everywhere.

Therefore the statement is only true when functions f are constant, so T_t is ergodic.

□

Example 1. *Circle Rotation*

Using setup from [3], let $\mathbb{S} := [0, 1)$ with measure μ, define $\alpha \in \mathbb{S}$, $R_\alpha : \mathbb{S} \to \mathbb{S}$ by $R_\alpha(x) \equiv (x + \alpha) \bmod 1$, so $R_\alpha(x) \in \mathbb{S}$.
NOTE: The rotation on the circle map $\Theta(x) := \exp(2\pi i x)$ is isomorphic to R_α
Also R_α is measure preserving as if $F \subset \mathbb{S}$,
then $\mu(R_\alpha^{-1}F) = \mu(F)$ since it simply shifts x by $-\alpha$.

R_α is ergodic if and only if $\alpha \notin \mathbb{Q}$

First if $\alpha \in \mathbb{Q}$
Suppose $\alpha = \frac{p}{q}$, when $p, q \in \mathbb{Z}, q \neq 0$.
Then $(R_\alpha(x))^q = R_\alpha(R_\alpha(\dots R_\alpha(x))) = x + \frac{p}{q} + \frac{p}{q} \dots \frac{p}{q} = x + q\left(\frac{p}{q}\right) = x + p \equiv x \bmod 1$.
Now fix some $x \in \mathbb{S}$ and ϵ sufficiently small that the ϵ-neighbourhoods of $x + k\alpha$ for $k = 0, 1, \dots q-1$ are disjoint. The union of these neighbourhoods (call it F) is an invariant set of positive measure, and we choose ϵ small enough so that $F \neq \mathbb{S}$. Thus $0 < \mu(F) < 1$ so R_α is not ergodic.

Now if $\alpha \notin \mathbb{Q}$
For a measurable set F such that $T_t^{-1}F = F$,choose f s.t. $f(x) = x \; \forall x \in F$.
We can write $f = \sum_{n\in\mathbb{Z}} \hat{f}(n) \exp(2\pi int)$ (Fourier Series expansion).
But since R_α is measure preserving over F, $f = f \circ R_\alpha = \sum_{n\in\mathbb{Z}} \exp(2\pi in\alpha)\hat{f}(n) \exp(2\pi int)$

$\therefore \hat{f}(n) = \hat{f}(n) \exp(2\pi in\alpha)$ but $\alpha \notin \mathbb{Q} \Rightarrow \exp(2\pi in\alpha) \neq 1$, so $\hat{f}(n) = 0$, $n \neq 0$.
So $f = \hat{f}(0)$ almost everywhere, and $\hat{f}(0) = \int_F f(x)dx = \mu(F)$.
Then when $\mu(F) > 0$, $f(x) = x = \mu(F) \times x \Rightarrow \mu(F) = 1$.
Therefore for all measurable sets F we get $\mu(F) = 0$ or $\mu(F) = 1$ Therefore R_α is ergodic.

Example 2. *Angle Doubling*

There is also a similar example is for angle doubling. Again from [3], let $\mathbb{S} = [0, 1)$, and $\mathcal{D} : \mathbb{S} \to \mathbb{S}$ s.t. $\mathcal{D}(x) = 2x \bmod 1$. $\mathcal{D}$ is an angle doubling map as $\Theta(x) := \exp(2\pi ix)$ is an isomorphism between $\mathcal{D}$ and the angle doubling map $\exp(x) \to \exp(2x)$ on the circle.
However to show that $\mathcal{D}$ is ergodic requires more advanced knowledge (i.e. mixing). So it will not be shown.[2]

4 von Neumann's Ergodic Theorem

Continuing from [2], we want to look at:

$$\frac{1}{T}\int_0^T (U_t f)(w) d\mu(w)$$

In order to simplify some calculations we will look at the discrete analogue:

$$\frac{1}{N}\sum_{n=0}^{N-1} (U^n f)(w)$$

This leads to the following very powerful result.

[2][3] Chapter 2 deals with mixing and proves that Angle Doubling is ergodic.

Theorem 2 (von Neumann's Ergodic Theorem)**.** *Let U be a unitary operator on a Hilbert space $\mathcal{H}$, with P being the orthogonal projection onto $\{\psi \in \mathcal{H} | U\psi = \psi\}$. Then, for any $f \in \mathcal{H}$.*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} U^n f = Pf$$

However, before we prove this, let's prove some easy results to help us.

Lemma 1. *If U is unitary, $Uf = f \iff U^* f = f$*

Proof.

$$Uf = f \iff f = U^{-1} f \iff f = U^* f$$

□

Lemma 2. *If U is unitary, $\|U\| = 1$*

(This was discussed in class)

Proof.

$$\|U\| = \sup_{x\in\mathcal{H}} \frac{||Ux||}{||x||} = \sup_{x\in\mathcal{H}} \frac{(\langle Ux, Ux\rangle)^{\frac{1}{2}}}{(\langle x, x\rangle)^{\frac{1}{2}}} = \sup_{x\in\mathcal{H}} \left(\frac{\langle x, U^*Ux\rangle}{\langle x, x\rangle}\right)^{\frac{1}{2}} = \sup_{x\in\mathcal{H}} \left(\frac{\langle x, x\rangle}{\langle x, x\rangle}\right)^{\frac{1}{2}} = 1$$

□

Lemma 3. *For any operator A on a Hilbert Space $\mathcal{H}$, $(imA)^\perp = kerA^*$*

(This was done during an assignment)

Proof.

$$\psi \in kerA^* \text{ then } \langle \phi, A^*\psi\rangle = \langle \phi, 0\rangle = 0, \forall \phi \in \mathcal{H}$$

$$\text{While } \psi \in (imA)^\perp \text{gives } \langle A\varphi, \psi\rangle = 0, \forall \varphi \in \mathcal{H}$$

$$\therefore \psi \in (imA)^\perp \iff \langle A\varphi, \psi\rangle = 0 \forall \varphi \in \mathcal{H} \iff \langle \varphi, A^*\psi\rangle = 0 \forall \varphi \in \mathcal{H} \iff \psi \in kerA^*$$

□

Now we can prove the important theorem stated above.

Proof of von Neumann's Ergodic Theorem.

Let $f = g - Ug \Rightarrow f \in im(I-U)$ then

$$\left\| \frac{1}{N}\sum_{n=0}^{N-1} U^n f \right\| = \left\| \frac{1}{N}\sum_{n=0}^{N-1} U^n (g - Ug) \right\| = \left\| \frac{1}{N}(g - U^N g) \right\|$$

Also we can use the previous lemma 2 to get $\|U^N\| = 1$ as U^N is unitary when U is.
Then $\|g - U^N g\| \leq \|g\| + \|U^N\| \, \|g\| = \|g\|(1 + \|U^N\|) = \|g\|(2) = 2\|g\|$

$$\therefore \left\| \frac{1}{N}\sum_{n=0}^{N-1} U^n f \right\| = \left\| \frac{1}{N}(g - U^N g) \right\| \leq \frac{2\,\|g\|}{N}$$

Which goes to 0 as $N \to \infty$

$$\Rightarrow \left\| \frac{1}{N}\sum_{n=0}^{N-1} U^n f \right\| \to 0 \quad \forall f \in im(I-U)$$

Additionally we show this holds for $\overline{im(I-U)}$, consider $f \in \overline{im(I-U)}$ then
$\exists \tilde{f} \in im(I-U)$ s.t. $\|f - \tilde{f}\| < \frac{\epsilon}{2}$. And N such that

$$\left\| \frac{1}{N}\sum_{n=0}^{N-1} U^n \tilde{f} \right\| < \frac{\epsilon}{2}$$

$$\therefore \left\| \frac{1}{N}\sum_{n=0}^{N-1} U^n f \right\| = \left\| \frac{1}{N}\sum_{n=0}^{N-1} U^n ((f - \tilde{f}) + \tilde{f}) \right\| \leq \left\| \frac{1}{N}\sum_{n=0}^{N-1} U^n (f - \tilde{f}) \right\| + \left\| \frac{1}{N}\sum_{n=0}^{N-1} U^n \tilde{f} \right\|$$

$$< \frac{\epsilon}{2}\left(\frac{1}{N}\sum_{n=0}^{N-1} \|U^n\| \right) + \frac{\epsilon}{2} \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$ Using that U^n is unitary

$$\therefore \left\| \frac{1}{N}\sum_{n=0}^{N-1} U^n f \right\| < \epsilon$$

From our lemma 1 and 3
$(im(I-U))^\perp = ker(I - U^*) = \{\psi \in \mathcal{H} | U^*\psi - \psi = 0\} = \{\psi \in \mathcal{H} | U^*\psi = \psi\} = \{\psi \in \mathcal{H} | U\psi = \psi\}$

$$\therefore Pf = 0 \iff f \in im(P)^\perp \iff f \in (im(I-U)^\perp)^\perp \iff f \in \overline{im(I-U)}$$

Now suppose $Pf = f$. Then $Uf = f$, where $f \in \{\psi \in \mathcal{H} | U\psi = \psi\} \Rightarrow f \in im(I-U)^{\perp}$

$$\therefore \frac{1}{N}\sum_{n=0}^{N-1} U^n f = \frac{Nf}{N} = f, \text{ so it converges to } f = Pf$$

So the following holds $\forall f \in \overline{im(I-U)} \oplus (im(I-U))^{\perp} = \mathcal{H}$

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} U^n f = Pf$$

□

In the continuous case we then get[3]

$$\frac{1}{T}\int_0^T (U_t f)(w) d\mu = Pf$$

Corollary 1. *For T_t ergodic then for any $f \in L^2(\Omega_E, d\mu_E)$*

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T f(T_t w) dt = \int_{\Omega_E} f(y) d\mu_E(y)$$

Proof.

Since T_t is ergodic, $\{f \in L^2(\Omega_E, d\mu_E) | Uf = f\}$ is a collection of constant maps, so $Pf \equiv c$, where $c \in \mathbb{R}$

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T f(T_t w) dt = \lim_{T\to\infty} \frac{1}{T}\int_0^T (U_t f)(w) dt = Pf = c = \langle f, 1\rangle = \int_{\Omega_E} f(y) d\mu_E(y)$$

As $\langle f - Pf, Pf\rangle = 0 \Rightarrow \langle f - Pf, Pf\rangle = \langle f - c, c\rangle = \langle f, c\rangle - \langle c, c\rangle = 0$
So $\langle f, c\rangle = \langle c, c\rangle \Rightarrow c\langle f, 1\rangle = c^2 \Rightarrow \langle f, 1\rangle = c$

□

The converse can be shown similarly, so T_t is ergodic if and only if the corollary holds.

[3] This is omitted as it is similar but with more work (as it involves integration as compared to summation). However section 35.2 of [1] proves the continuous case.

5 Conclusion

Finally, I would like to introduce the following theorem.
Theorem 3 (Birkhoff Ergodic Theorem)**.** *Let T be a measure preserving transformation on a measure space (Ω, μ). Then, for any $f \in L^1(\Omega, \mu)$, we get*

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

which exists pointwise almost everywhere and is some function $\tilde{f} \in L^1(\Omega, \mu)$ s.t. $\tilde{f}(Tx) = \tilde{f}(x)$.

If $\mu(\Omega) < \infty$, then

$$\int_\Omega \tilde{f}(x) d\mu(x) = \int_\Omega f(x) d\mu(x)$$

Furthermore, if μ is ergodic and $\mu(\Omega) = 1$, then as $N \to \infty$, the space average of f becomes the time average f.

$$\lim_{N\to 0} \frac{1}{N} \sum_{n=0}^{N-1} f(T^n x) \longrightarrow \int_\Omega f(y) d\mu(y)$$

for almost all x.

In general, the space and time averages of a function are different, so this can be a very useful result in physics. Additionally, this theorem has better connections with statistical mechanics than the von Neumann theorem.

6 References

1) Lax, Peter D. “Groups of Unitary Operators, Ergodic Theory.” *Functional Analysis.* New York: Wiley, 2002. Print.

2) Reed, Michael, and Barry Simon. “Ergodic Theory: An Introduction.” *Functional Analysis.* Rev. and Enl. ed. London: Academic, 1980. Print.

3) Sarig, Omri. “Lecture Notes on Ergodic Theory.” Penn State, University Park, Pennsylvania. 12 Dec. 2008. Lecture.