Monstrous Menagerie with Vandehey 7/4

Jason Schuchardt

July 4, 2019

1 The ergodic theorem

Definition 1.1. $L^1(\mu, X)$ is the set of equivalence classes of functions $f : X \to \mathbb{R}$ such that

$$\int_X |f| \, d\mu$$

is meaningful and finite modulo the equivalence relation where $f \sim g$ if $f = g$ a.e.

There are many ergodic theorems. There are the ratio ergodic theorem, the maximal ergodic theorem, and the mean ergodic theorem. However, we're going to talk about the pointwise ergodic theorem.

Theorem 1.1 (Birkhoff's Pointwise Ergodic Theorem)**.** *Let (X, μ, T) be a measure-preserving dynamical system, with $\mu(X) = 1$. Then for any $f \in L^1(X, \mu)$, we have*

$$f^*(x) := \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x)$$

exists almost everywhere, satisfies

$$\int_X f \, d\mu = \int_X f^* \, d\mu,$$

and is T-invariant, i.e. $f^* \circ T = f^*$.

Moreover, if T is ergodic, then f^* equals the constant value

$$\int_X f \, d\mu$$

almost everywhere.

Example 1.1. Let $f = 1_{C_s}$ be the indicator function for a cylinder set and let T be ergodic.

Then the theorem says that for almost all $x \in X$,

$$f^*(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) = \int_X f \, d\mu.$$

That is, when we iterate the function T on x, the proportion of time x spends in the cylinder set C_s is equal to the measure of C_s.

This is much stronger than recurrence. Recurrence only says that if you start in C_s, you should expect to return to C_s. This result tells you with what frequency you should expect to return to C_s.

Example 1.2. RCF is ergodic with the Gauss measure,

$$\mu(A) = \int_A \frac{1}{() \log 2)(1+x)}\, dx$$

So how often does a 1 appear in a typical RCF expansion?

The answer should be

$$\mu(C_1) = \mu(1/2, 1) = \int_{1/2}^{1} \frac{1}{(\log 2)(1+x)}\, dx = \left.\frac{\log(1+x)}{\log 2}\right|_{1/2}^{1} = \frac{\log 2 - \log(3/2)}{\log 2} = \frac{\log(4/3)}{\log 2} \approx 41.5\%.$$

Question: Does this answer depend on the measure? Yes.

However, Gauss measure is continuous with respect to Lebesgue measure, so if something is true for Gauss almost all points, then it is true for Lebesgue almost all points.

Now we'll prove the proof. This will be the most technical proof that we will give.

There are easy proofs, that proceed from a more abstract framework. This is a more elementary proof, but it will be more technical as a result. We'll need a lemma first.

Lemma 1.1. *For a probability space (X, μ), a measure preserving transformation T is ergodic if and only if every f that is T-invariant almost everywhere is actually a constant a.e.*

Proof. The if direction will be an exercise. Use $f = 1_A$, where A is some set.

Thus suppose T is ergodic and f is a T-invariant function. Define for any $r \in \mathbb{R}$,

$$A_r = \{x \in X : f(x) > r\}.$$

Since f is T-invariant a.e., A_r is T-invariant a.e. Thus $\mu(T^{-1}A_r \Delta A_r) = 0$. Hence by ergodicity, $\mu(A_r) = 0$ or $\mu(A_r) = 1$. If f were not a constant almost everywhere, then we could find r such that $0 < \mu(A_r) < 1$, which is a contradiction. Hence f is a constant a.e. ∎

Now we can prove the theorem.

Proof of the Ergodic Theorem. We only need to prove the theorem for indicator functions. The general case follows in the typical way by building up L^1 out of indicator functions.

Note $f(T^i x)$ measures if $T^i x$ is in A, where $f = 1_A$.

We define

$$\overline{f}(x) = \limsup_n \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x),$$

and

$$\underline{f}(x) = \liminf_n \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x).$$

The lim sup and lim inf always exist, and the limit exists if and only if lim sup = lim inf.

Clearly, $\underline{f}(x) \le \overline{f}(x)$. Also, both of these are T-invariant. If we consider $\overline{f}(x)$, we get

$$\overline{f}(Tx) = \limsup_n \frac{1}{n} \sum_{i=0}^{n-1} f(T^{i+1}x) = \limsup_n \frac{n+1}{n} \cdot \frac{1}{n+1} \sum_{i=0}^{n} f(T^i x) - \frac{1_A(x)}{n} = \overline{f}(x).$$

Similarly, $\underline{f}(x)$ is also T-invariant.

It now suffices to prove

$$\int_X \overline{f}\, d\mu \le \mu(A) \le \int_X \underline{f}\, d\mu. \tag{*}$$

To see this, recall

$$\overline{f} - \underline{f} \ge 0,$$

but this inequality says that

$$\int_X \overline{f} - \underline{f}\, d\mu \le 0.$$

This is only possible if $\overline{f} = \underline{f}$ a.e., which is equivalent to f^* existing a.e.

Then the inequality tells us that

$$\int_X f^*\, d\mu \le \mu(A) \le \int_X f^*\, d\mu,$$

so

$$\int_X f^*\, d\mu = \mu(A) = \int_X f\, d\mu.$$

This function f^* is T-invariant a.e., so if T is ergodic, then by the lemma, f^* is a constant almost everywhere.

Since $\mu(X) = 1$, we can then conclude that if T is ergodic, then $f^* = \int_X f\, d\mu$ a.e.

Thus all we need to do is prove the inequality. By symmetry, it suffices to prove half of the inequality. We will prove that

$$\int_X \overline{f}\, d\mu \le \mu(A).$$

Let $\epsilon > 0$. Define

$$S_n(x, B) = \sum_{i=0}^{n-1} 1_B(T^i x).$$

Let

$$N(x) = \min\{n \ge 1 : n \in \mathbb{N}, S_n(x, A) \ge (\overline{f}(x) - \epsilon)n\}.$$

By definition,

$$S_{N(x)}(x, A) \ge (\overline{f}(x) - \epsilon)N(x).$$

For $M > 0$, let

$$A_M := \{x \in X : N(x) > M\}.$$

Since $N(x)$ is finite, we can find M such that $\mu(A_M) < \epsilon$.

For this M, let $A' = A \cup A_M$. Define

$$N'(x) := \begin{cases} N(x) & N(x) \le M \\ 1 & N(x) > M. \end{cases}$$

We see that

$$S_{N'(x)}(x, A') \ge (\overline{f}(x) - \epsilon)N'(x)$$

This is true because, if $N(x) > M$, then $x \in A'$, and $N'(x) = 1$, so $S_{N'(x)}(x, A') = 1$, which is larger than

$$(\overline{f}(x) - \epsilon)N'(x) \le (1 - \epsilon)1 < 1.$$

On the other hand, if $N(x) \le M$, then $N'(x) = N(x)$, and $A \subseteq A'$, so

$$S_{N'(x)}(x, A') = \sum_{i=0}^{N'(x)-1} 1_{A'}(T^i x) \ge \sum_{i=0}^{N(x)-1} 1_A(T^i x) \ge (\overline{f}(x) - \epsilon)N(x).$$

Now we want to look at $S_n(x, A')$, and somehow replace it with nicer N' stuff.

Let $n_0(x) = 0$. Let

$$n_k(x) = n_{k-1}(x) + N'(T^{n_{k-1}(x)}x).$$

Choose $n > M$, let

$$\ell = \ell(n, x) = \max\{k \ge 1 : n_k(x) \le n - 1\}$$

Using the T-invariance of $\overline{f}$, we have

$$\begin{aligned}
S_n(x, A') &\ge \sum_{i=0}^{n_\ell(x)-1} 1_{A'}(T^i x) \\
&= \sum_{i=0}^{\ell-1} \sum_{j=n_i(x)}^{n_{i+1}(x)} 1_{A'}(T^i x) \\
&= \sum_{i=0}^{\ell-1} S_{N'(T^{n_i(x)}x)}\Big(T^{n_i(x)}x, A'\Big) \\
&\ge \sum_{i=0}^{\ell-1} \Big(\overline{f}(T^{n_i(x)}x) - \epsilon\Big) N'\Big(T^{n_i(x)}x\Big) \\
&= \Big(\overline{f}(x) - \epsilon\Big) \sum_{i=0}^{\ell-1} N'\Big(T^{n_i(x)}x\Big) \\
&= \Big(\overline{f}(x) - \epsilon\Big) \sum_{i=0}^{\ell-1} (n_{i+1}(x) - n_i(x)) \\
&\ge (\overline{f}(x) - \epsilon)(n - M)
\end{aligned}$$

Now for all B,

$$\mu(B) = \mu(T^{-i}B) = \int_X 1_{T^{-i}B}\, d\mu = \int_X 1_B \circ T^i\, d\mu,$$

so in particular,

$$\mu(A') = \frac{1}{n} \int_X S_n(x, A')\, d\mu.$$

Combining this fact, with the previous long inequality, we derive

$$\mu(A') \ge \frac{n - M}{n} \bigg(\int_X \overline{f}\, d\mu - \epsilon \bigg).$$

Finally, we can let $n \to \infty$, giving

$$\mu(A') \ge \int_X \overline{f}\, d\mu - \epsilon,$$

and then we can let ϵ go to 0 giving us the inequality we wanted.

■