Unique ergodicity and the impossibility of general results on all continuous functions in a measure preserving ergodic system.

Jonathan Gontmakher | 216870782 | 03665103

December 14, 2025

Part I
Impossibility of general results

1 Impossibility of general results - no bound on general convergence speed.

In our previous discussions of ergodic theory, we have seen that it is of great interest in many cases to study the convergence of series of the type $A_N^f(x) = \frac{1}{N}\sum_{n=0}^{N-1} f(T^n x)$, and can be thought of as a "time average", that shows us that T "shuffles points well enough", such that integrals may be evaluated using averages of the function T. Birkhoff's Ergodic theorem states that

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

such as

$$\int_X \bar{f} d\mu = \int_X f d\mu$$

for each $f \in C(X)$, almost everywhere on X, assuming T is measure preserving, and in the case that T is ergodic the stronger case holds:

$$\bar{f} \equiv \int_X f d\mu$$

But, nowhere previously have we discussed how quickly the series converges. Is the convergence bounded? is it fast? Maybe the sequence even becomes constant at some point...

It turns out that no, the convergence is "as weak as it gets" - and we will illustrate this on a specific example. We do not intend to show that it is always the case - our intention is to disprove the wrong idea that it is possible to bound the convergence speed, so we will show it for a very simple and intuitive case - the real interval $[0, 1]$.

1.1 Thm (Krengel, 1977): when X is the $[0, 1]$ interval, μ is the standard Lebesgue measure, and τ is an invertible ergodic measure preserving transformation on X, for each descending null sequence $(\alpha_N)_{N=0}^{\infty}$ and $\varepsilon > 0$, there exists an $f \in C(X)$ and a subset $Y \subseteq X$ of measure more than $1 - \varepsilon$, for which

$$\forall x \in Y, \overline{\lim_{N \to \infty}} \alpha_N^{-1} \left| \frac{1}{N} \sum_{n=0}^{N-1} f(\tau^n x) - \int_X f d\mu \right| = \infty$$

To start - we need to establish a result known as Rokhlin's lemma. It will aid us in the construction of both Y and f.

1.2 Thm (Rokhlin's lemma): When τ is an ergodic measure preserving map on a probability space X, for each $\varepsilon > 0$ and $N \in \mathbb{N}$ we can find a set S for which $S, \tau^{-1}S, ..., \tau^{-N-1}S$ are pairwise disjoint, and the "error set" $(X - \bigcup_{i=0}^{N} \tau^i S)$ is of measure less than ε.

Let us assume that $C \subseteq X$ is a set of positive measure, $0 < \mu(C) < \varepsilon/N$. Now, we define the "arrival time" function - that for each $x \in X$ assigns the minimal time for its arrival in C. $f(x) = \min_n \{n \in \mathbb{N} | \tau^n(x) \in C\}$. From τ's ergodicity, we know that $\mu\left(\bigcup_{n=0}^{\infty} \{x \in X | f(x) = n\}\right) = 1$. Let us define $S = \{x \in X | f(x) \in \{N, 2N, ...\}\}$. Obviously the sets $S, \tau^{-1}S, ..., \tau^{-N+1}S$ are pairwise disjoint, and their union contains all points that do not end up in C after less than N turns. So:

$$\bigcup_{n=0}^{N-1} \tau^{-n} S = \left(\bigcup_{n=0}^{\infty} \{x \in X | f(x) = n\} \right) - \left(\bigcup_{n=0}^{N-1} \tau^{-n} C \right)$$

And as all of the sets in the equation are disjoint, we infer that

$$\mu \left(\bigcup_{n=0}^{N-1} \tau^{-n} S \right) = \mu \left(\bigcup_{n=0}^{\infty} \{x \in X | f(x) = n\} \right) - \mu \left(\bigcup_{n=0}^{\infty} \tau^{-n} C \right) > 1 - N \cdot \varepsilon / N = 1 - \varepsilon$$

Which is exactly the intended result. $\square$

It is important to note that in the literature, Rokhlin's lemma is usually stated as a more general result - it is not necessary for X to be a probability space, or for τ to be ergodic - weaker restrictions on τ are enough for that result, but the proof is much simpler this way and for the proof of (1.1) we do not need the more general statement of Rokhlin's Lemma. Also, the sets $S, \tau^{-1}S, ..., \tau^{-N+1}S$ are commonly referred to as a "Rokhlin tower" - the set S is known as the "base" of the tower, and the set $\tau^{-i}S$ is the i-th layer of the tower.

We intend to construct a continuous function f (for which $\int f = 1/2$), a set $Y \subseteq X$, and intervals (M_n, N_{n+1}) such as

$$\forall x \in Y, \sup_{M_n \leq K \leq N_{n+1}} \alpha_K^{-1} \left| \frac{1}{K} \sum_{i=0}^{K-1} f(\tau^i x) - 1/2 \right| > n$$

We will do this iteratively, by defining f as a limit of continuous functions $(f_n)_{n=0}^{\infty}$, $\forall n, \int f_n = 1/2$. We want for f_n to hold - for the interval of indices (M_n, N_{n+1}), on a sufficiently large set Y_n,

$$\forall x \in Y_n, \sup_{M_n \leq K \leq N_{n+1}} \alpha_K^{-1} \left| \frac{1}{K} \sum_{i=0}^{K-1} f_n(\tau^i x) - 1/2 \right| > n$$

Throughout the proof we will make use of a descending non zero sequence ε_n, which bounds the convergence of f_n - we will define the functions in a way, such as $\|f_{n+1} - f_n\|_\infty \leq \varepsilon_n$, and as the proof does not require any upper bound on the speed of ε_n's convergence, many of the arguments presented here are rationalized by making ε_n converge to 0 really quickly.
In addition, a series of integers p_n will be defined, which will be useful for definition of f_{n+1} from f_n. f_{n+1} is constructed from f_n by adding an "oscillation" g_n, and p_n may be thought of as a quantifier of this oscillation.

For the base step - let us define the function $f_1 = 1/2$ (constant), $Y_1 = X$ and $N_0 = 0, N_1 = 1$.
After the n'th step, we have a continuous function $f_n, \int f_n = 1/2$, our previous interval (M_{n-1}, N_n), and a large subset Y_n. By introducing a cleaner notation $f_{n,K}(x) = \frac{1}{K} \sum_{i=0}^{K-1} f_n(\tau^i x)$, we have that

$$\sup_{M_{n-1} \leq K \leq N_n} \alpha_K^{-1} |f_{n,K}(x) - 1/2|$$

is large on a significantly large subset of X.
In order to proceed from the n-th step, we define ε_n to be some small quantity, for which $0 < \varepsilon_n \leq \frac{1}{4N_n}$ holds. Then we pick an integer p_n for which $p_n^{-1} < \frac{\varepsilon_n}{4}$, and then pick $M_n > N_n$, for which $\alpha_{M_n}^{-1} > (n+1) / \left(\frac{\varepsilon_n}{2^4 p_n} \right)$. For a large enough

M_n (α_N tends to infinity, so we can pick an arbitrarily large M_n) Birkhoff's ergodic theorem states that

$$\sup_{K \geq M_n} |f_{n,K}(x) - 1/2| < 2^{-5}\varepsilon_n p_n^{-1}$$

almost everywhere - we pick such an M_n. Then, we define the next value $N_{n+1} = p_n M_n$. We now have our next bound - (M_n, N_{n+1}).
We now make use of Rokhlin's Lemma - The decomposition of X into a Rokhlin tower allows us to think about the action τ as moving down in the tower, and by carefully defining a function along the tower's layers, we can say a lot about its ergodic time average - as applying τ to a point just moves it down the tower. We will make use of the tower to define a function g_n, whose overall integral is 0, but it's ergodic average reaches relatively high values.
Rokhlin's lemma implies the existence of a set B_n, for which the sets $\{\tau^{-k}B_n | 0 \leq k < N_{n+1}\}$ are disjoint, and their union has a measure as close to unity as we like - in this case, we need it to be $\geq 1 - \dfrac{\varepsilon_n}{4}$. Now we define the set of the first M_n layers E_n to be

$$E_n = \bigcup_{k=0}^{M_n - 1} \tau^{-k} B_n$$

and the new function g_n^* to be

$$g_n^* = 1_{E_n} - 1_{\tau^{-M_n} \circ E_n}$$

More intuitively, it means that $g_n^* = 1$ on the first M_n levels of the tower, -1 on the next M_n,and 0 for the rest of the levels.
We may notice that it's integral is obviously zero - it has values of -1 and 1 on sets with similar measures.
Analogously to the definition of $f_{n,K}$, for a clearer notation we define

$$g_{n,K}^* = \frac{1}{K} \sum_{i=0}^{K-1} g_n^* \circ \tau^i$$

Now, we climb mount Rokhlin.
The big set we are interested in is $Y_n = \bigcup_{k=2M_n}^{N_{n+1}-1} \tau^{-k} B_n$ - so, the entirety of the tower, except the first $2M_n$ layers. The measure of this set is, of course,

$$\mu(Y_n) = \mu(B_n) \cdot (N_{n+1} - 2M_n) = \left(1 - \frac{\varepsilon_n}{4}\right) \cdot \frac{N_{n+1} - 2M_n}{N_{n+1}} = \left(1 - \frac{\varepsilon_n}{4}\right) \cdot \left(1 - \frac{2}{p_n}\right) > \left(1 - \frac{\varepsilon_n}{4}\right) \cdot \left(1 - \frac{\varepsilon_n}{2}\right)$$

The measure tends to 1 as ε_n tends to 0 - this is exactly the property that we need, as ε_n can be arbitrarily small.
Let us suppose that $x \in Y_n$, and denote the layer it is located in as L. If we define L' to be the number of layers between x and the M_n'th layer ($L' = L - M_n$) we

will get that $g^*_{n,L'}(x) = \dfrac{M_n}{L'}$. As $L > 2M_n$, $L' \in [M_n, N_{n+1}]$ - and as such

$$\sup_{M_n \le K \le N_{n+1}} |g^*_{n,K}(x)| \ge |g^*_{n,L'}(x)| = \frac{M_n}{L'} = \frac{M_n}{L - M_n} \ge \frac{M_n}{N_{n+1} - M_n}$$
$$= \frac{M_n}{(p_n - 1)M_n} = \frac{1}{p_n - 1} \underset{p_n \text{is large enough}}{\ge} \frac{1}{2p_n}$$

This holds, of course, for each $x \in Y_n$ - the final bound does not depend on L, and from p_n^{-1}.
A result from real analysis implies that g^*_n can be approximated by a continuous function g_n, such as $\int g_n = 0$, $|g_n| \le 1$, and the approximation is close enough such as $g_{n,K} = \dfrac{1}{K} \sum_{i=0}^{K-1} g_n \circ \tau^i$ satisfies a slightly weaker bound:

$$\sup_{M_n \le K \le N_{n+1}} |g_{n,K}| \ge \frac{1}{4p_n}$$

The fact that such an approximation exists is non trivial, and is left as an exercise for the curious reader.
Now we define $f_{n+1} = f_n + \varepsilon_n g_n$, and we see that for each $x \in Y_n$

$$\sup_{M_n \le K \le N_{n+1}} |f_{n+1,K}(x) - 1/2| = \sup_{M_n \le K \le N_{n+1}} |f_{n,K}(x) + \varepsilon_n g_{n,K}(x) - 1/2|$$

But, we know that

$$\sup_{M_n \le K} |f_{n,K}(x) - 1/2| < 2^{-5} \varepsilon_n p_n^{-1}$$

and as such:

$$\sup_{M_n \le K \le N_{n+1}} |f_{n,K}(x) + \varepsilon_n g_{n,K}(x) - 1/2| \ge \sup_{M_n \le K \le N_{n+1}} |g_{n,K}| - \sup_{M_n \le K \le N_{n+1}} |f_{n,K}(x) - 1/2| \ge$$
$$\ge \frac{\varepsilon_n}{4p_n} - \frac{\varepsilon_n}{2^5 p_n} \ge \frac{\varepsilon_n}{2^4 p_n}$$

on a sufficiently large measure set - Y_n
The desired result is thus obtained. $2^{-4} \varepsilon_n p_n^{-1}$ is the sufficiently large value that we require.
As we know that $\|f_{n+1} - f_n\| < \varepsilon_n$, and $\sum \varepsilon_n$ converges, and as slowly as we like, we know that there is a continuous function f, such as $f_n \underset{n \to \infty}{\to} f$.
We may notice that, because of our selection of M_n - we selected it such as $\alpha^{-1}_{M_n} > (n+1) / \left(\dfrac{\varepsilon_n}{2^4 p_n} \right)$, it holds that

$$\sup_{M_n \le K \le N_{n+1}} \alpha^{-1}_{M_n} |f_{n+1,K}(x) - 1/2| \ge n + 1$$

and as $f_n \underset{n \to \infty}{\to} f$ - and we may make it converge as quickly as we like by picking small ε_n's, we may see that for the limit function f, those results hold, and as

such, the sequence $\alpha_N^{-1}\left|\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x) - \int_X f d\mu\right|$ is indeed unbounded, on the set that is the intersection of all Y_n's - which we can force to be of a measure larger than $1-\varepsilon$ by picking ε_n's to be very small!
So, the proof is complete. A corollary exists - the proof may be modified to yield a function f for which this result is true almost everywhere, but this is outside the scope of this work and is left as an exercise to the interested reader.

Part II
Unique ergodicity

2 On T invariant measures

It is of great interest for us to discuss and identify Borel measures, in the study of dynamical systems. In the second part of the lecture, we will do just that.

2.1 Def: $\mathcal{M}(X)$ is the space of Borel probability measures on a compact metric space (X, d).

2.2 Def: $\mathcal{M}^T(X)$ is the subspace of T invariant Borel probability measures on (X, d).

Every continuous $T : X \to X$ induces a continuous map (in respect to the weak star topology!) $T_* : \mathcal{M}(X) \to \mathcal{M}(X)$, by defining $T_*(\mu)(A) = \mu(T^{-1}A)$.
An obvious measure we can look at is $\delta_x(A)$ which equals to 1 if $x \in A$, and 0 if $x \notin A$. We may notice that:$T_*\delta_x(A) = \delta_x(T^{-1}A) = \delta_{T(x)}(A)$, and as such $T_*\delta_x = \delta_{T(x)}$. This is interesting - in some sense, $\mathcal{M}(X)$ "contains" all the points of X in itself, as the δ_x measures - and as such, T_* may be viewed as an extension of the mapping $T : X \to X$ to the broader space $\mathcal{M}(X)$. An Important property of T_*, which is an extension of the previously described one, is that for each $f \in C(X), \mu \in \mathcal{M}(X)$:

$$\int_X f d(T_*\mu) = \int_X f \circ T d\mu$$

It may be seen that $\mathcal{M}^T(X)$ is the set of all measures for which $T_*(\mu) = \mu$, and as T_* is affine (obviously) and continuous, the fixed set $\mathcal{M}(X)$ is closed and convex. (closure can be shown by constructing a sequence of invariant measures).

Cool! but we haven't shown that $\mathcal{M}^{\mathcal{T}}(X)$ is non empty...

3 Kryloff - Bogoliouboff theorem

As usual, we let $T : X \to X$ be a continuous map of a compact metric space, and ν_n to be a sequence in $\mathcal{M}(X)$. The sequence $\mu_n = \frac{1}{n}\sum_{j=0}^{n-1} T_*^j \nu_n$ is induced by the sequence ν_n, and as $\mathcal{M}(X)$ is weak * compact, We know that it has a converging subsequence $\mu_{n_k} \to \mu$. Let us take a look at the value $\left|\int_X f \circ T d\mu_{n_k} - \int_X f d\mu_{n_k}\right|$, for some $f \in C(X)$:

$$\left|\int_X f \circ T d\mu_{n_k} - \int_X f d\mu_{n_k}\right| = \frac{1}{n_k}\left|\int_X \sum_{i=0}^{n_k-1} \left(f \circ T^{i+1} - f \circ T^i\right) d\nu_{n_K}\right|$$
$$= \frac{1}{n_k}\left|\int_X \left(f \circ T^{n_k} - f\right) d\nu_{n_K}\right| \underset{\text{probability space}}{\leq} \frac{1}{n_k} \cdot 2 \cdot \sup\{|f(x)||x \in X\} \to 0$$

As such, $\int f \circ T d\mu = \int f d\mu$ for all $f \in C(X)$, and as such μ is a member of $\mathcal{M}^T(X)$ (recall previous lecture).
And so, the statement of thew Kryloff - Bogolioboff theorem may be inferred:

3.1 Thm (Kryloff Bogolioboff): $\mathcal{M}^T(X)$ is non empty (for a compact space).

It may be noted that the theorem is obviously false for many non compact (and possibly locally compact) spaces, such as the real line.

4 Unique ergodicity

An interesting class of transformations is the class of those that have a unique T preserving borel measure.

4.1 Def: $T : X \to X$ is uniquely ergodic if $|\mathcal{M}^T(X)| = 1$.

The uniquely ergodic measure μ is automatically ergodic, because if we have a set $A \subset X$,such as $\mu(A) \notin \{0,1\}$ and $T^{-1}A = A$ We may define a measure $\mu'(B) = \dfrac{1}{\mu(A)}\mu(A \cap B)$, and we get that $\mu'(B) = \dfrac{\mu(A \cap B)}{\mu(A)} = \dfrac{\mu(T^{-1}(A \cap B))}{\mu(A)} = \dfrac{\mu(A \cap T^{-1}B)}{\mu(A)} = \mu'(T^{-1}B)$, and as such we got another measure in $\mathcal{M}^T(X)$.

Now, let us introduce an important result:

4.2 Thm: when T is uniquely ergodic, $\forall f \in C(X)$, the following statement holds:

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

without dependence of C_f on x, and the convergence is uniform.

Let μ be the unique member of $\mathcal{M}^T(X)$. We may notice that, according to the prior result, the following relation must be true:

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

And as such, because f is continuous in respect to the weak star topology, it holds that

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

Now, let us assume that the convergence is non uniform. Non uniformity of the convergence implies the existence of a function $g \in C(X)$, and $\varepsilon > 0$, s.a. for every N_0, there is $N > N_0$ and $x_j \in X$, for which

$$\left|\frac{1}{N}\sum_{n=0}^{N-1} g(T^n x_j) - C_g\right| \geq \varepsilon$$

By defining $\mu_N = \frac{1}{N}\sum_{n=0}^{N-1}\delta_{T^n x_j}$, we may express the sum as an integral

$$\left|\int_X g d\mu_N - C_g\right| \geq \varepsilon$$

compactness implies existence of a converging sequence $\mu_{N_k} \to \nu$. By the earlier result we see that $\nu \in \mathcal{M}^T(X)$, and

$$\left|\int_X g d\nu - C_g\right| \geq \varepsilon$$

but as $C_g = \int_X g d\mu$, this contradicts $\mu = \nu$.
As such, we have established our result - the convergence is uniform!