Von Neumann's Ergodic Theorem

Theorem (Von Neumann's Ergodic Theorem; 1932)**.** *Let μ be a probability measure on a measurable space $\mathcal{X}$ and $T\colon \mathcal{X} \to \mathcal{X}$ a measurable map that preserves μ. For every $f \in L^2_\mu(\mathcal{X})$,*

$$\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k \to \widetilde{f} \qquad \textit{in } L^2_\mu(\mathcal{X}) \textit{ as } n \to \infty$$

for some $\widetilde{f} \in L^2_\mu(\mathcal{X})$ that satisfies $\widetilde{f} \circ T = \widetilde{f}$ μ-almost everywhere.

Remark (Characterization of the limit)**.** The proof below will characterize $\widetilde{f}$ as the orthogonal projection of f onto the closed linear subspace

$$\{g \in L^2_\mu(\mathcal{X}) : g \circ T = g \ \mu\text{-a.e.}\}$$

of $L^2_\mu(\mathcal{X})$. ◇

Remark (Induced isometry)**.** The map $f \mapsto f \circ T$ is a positive linear isometry of $L^2_\mu(\mathcal{X})$. To see that it is an isometry (i.e., it preserves distances in L^2_μ), note that

$$\|f_2 \circ T - f_1 \circ T\|^2 = \int |f_2 \circ T - f_1 \circ T|^2 \,\mathrm{d}\mu = \int |f_2 - f_1|^2 \,\mathrm{d}(T\mu) = \int |f_2 - f_1|^2 \,\mathrm{d}\mu = \|f_2 - f_1\|^2 \,.$$

The linearity and the positivity ($f \geq 0$ implies $f \circ T \geq 0$) are clear. ◇

Proof of von Neumann's theorem. We start with some special cases:

- If $f \circ T = f$ μ-a.e., then the claim trivially holds and $\widetilde{f} = f$.
- If $f = h \circ T - h$ for some $h \in L^2_\mu(\mathcal{X})$, then $\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k \to 0$ in $L^2_\mu(\mathcal{X})$, hence $\widetilde{f} = 0$.

  Argument. Note that

  $$\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k = \frac{1}{n}\sum_{k=0}^{n-1} (h \circ T - h) \circ T^k = \frac{1}{n}\sum_{k=0}^{n-1} (h \circ T^{k+1} - h \circ T^k) = \frac{h \circ T^n - h}{n} \,.$$

  Therefore,

  $$\left\|\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k\right\| = \frac{1}{n}\|h \circ T^n - h\| \leq \frac{1}{n}\big(\|h \circ T^n\| + \|h\|\big) = \frac{2}{n}\|h\| \,,$$

  which goes to 0 as $n \to \infty$. (For the last equality, we have used the fact that $h \mapsto h \circ T$ is an isometry.)
- If $f = c_1 f_1 + c_2 f_2$ for some $f_1, f_2 \in L^2_\mu(\mathcal{X})$ and $c_1, c_2 \in \mathbb{R}$, and the claim holds for f_1 and f_2, then the claim clearly also holds for f with $\widetilde{f} = c_1 \widetilde{f}_1 + c_2 \widetilde{f}_2$.

Let

$$B := \big\langle h \circ T - h : h \in L^2_\mu(\mathcal{X}) \big\rangle$$

be the linear span of all functions of the form $h \circ T - h$. Let us consider one more special case:

Notes for Introduction to Ergodic Theory (MATH 307K), Siamak Taati
Last update: February 22, 2022

- If $f \in \overline{B}$, then $\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k \to 0$ in $L^2_\mu(\mathcal{X})$, hence $\widetilde{f} = 0$.

  (Here, $\overline{B}$ stands for the closure of B in $L^2_\mu(\mathcal{X})$.)

Argument. The case in which $f \in B$ is already handled. Let $f \in \overline{B}$ and suppose there exist a sequence $f_1, f_2, \ldots \in B$ such that $f_m \to f$ as $m \to \infty$.

Observe that for a fixed $m \geq 1$,

$$\begin{aligned}\left\|\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k\right\| &= \left\|\frac{1}{n}\sum_{k=0}^{n-1} f_m \circ T^k + \frac{1}{n}\sum_{k=0}^{n-1}(f - f_m) \circ T^k\right\| \\ &\leq \left\|\frac{1}{n}\sum_{k=0}^{n-1} f_m \circ T^k\right\| + \frac{1}{n}\sum_{k=0}^{n-1}\underbrace{\left\|(f - f_m) \circ T^k\right\|}_{\|f-f_m\|} \\ &= \left\|\frac{1}{n}\sum_{k=0}^{n-1} f_m \circ T^k\right\| + \|f - f_m\| \ .\end{aligned}$$

Letting $n \to \infty$, we obtain

$$\limsup_{n\to\infty}\left\|\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k\right\| \leq \lim_{n\to\infty}\left\|\frac{1}{n}\sum_{k=0}^{n-1} f_m \circ T^k\right\| + \|f - f_m\| = \|f - f_m\| \ .$$

Finally, letting $m \to \infty$, we find that

$$\lim_{n\to\infty}\left\|\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k\right\| = 0 \ .$$

Now, let

$$I := \{g \in L^2_\mu(\mathcal{X}) : g \circ T = g \ \mu\text{-a.e.}\} \ ,$$

and observe that I is a closed linear subspace of $L^2_\mu(\mathcal{X})$.

So far, we have verified that the conclusion of the theorem holds if f belongs to either of the two closed subspaces I and $\overline{B}$. Remarkably, every element of $L^2_\mu(\mathcal{X})$ can be decomposed (in a unique fashion) as a sum of two elements one belonging to I and the other belonging to $\overline{B}$. More specifically:

Lemma (Decomposition lemma). *Every $f \in L^2_\mu(\mathcal{X})$ has a unique decomposition $f = \widetilde{f} + f_0$ where*

$$\widetilde{f} \in I \ , \qquad f_0 \in \overline{B} \ , \qquad \widetilde{f} \perp f_0 \ .$$

In other words, $L^2_\mu(\mathcal{X}) = I \oplus \overline{B}$.

(The proof of the lemma comes below.)

It follows that for every $f \in L^2_\mu(\mathcal{X})$,

$$\frac{1}{n}\sum_{k=0}^{n-1} f \circ T^k = \underbrace{\frac{1}{n}\sum_{k=0}^{n-1} \widetilde{f} \circ T^k}_{\widetilde{f}} + \underbrace{\frac{1}{n}\sum_{k=0}^{n-1} f_0 \circ T^k}_{\to 0} \to \widetilde{f} \qquad \text{in } L^2_\mu(\mathcal{X}) \text{ as } n \to \infty,$$

and this concludes the proof. □

Proof of the decomposition lemma. We need to show that I is the orthogonal complement of $\overline{B}$.

First, let $g \in I$. Then, for every $h \in L^2_\mu(\mathcal{X})$, we have

$$\begin{aligned}\int (h \circ T - h)\overline{g}\,\mathrm{d}\mu &= \int (h \circ T)\overline{g}\,\mathrm{d}\mu - \int h\overline{g}\,\mathrm{d}\mu \\ &= \int (h \circ T)(\overline{g} \circ T)\,\mathrm{d}\mu - \int h\overline{g}\,\mathrm{d}\mu && \text{(because } g \circ T = g \ \mu\text{-a.e.)} \\ &= \int h\overline{g}\,\mathrm{d}(T\mu) - \int h\overline{g}\,\mathrm{d}\mu \\ &= 0 && \text{(because } T\mu = \mu\text{)}\end{aligned}$$

hence $g \perp (h \circ T - h)$. It follows that $g \perp \overline{B}$, and in particular $I \subseteq \overline{B}^{\perp}$.

Conversely, let $g \in \overline{B}^{\perp}$. Then, for every $h \in L^2_\mu(\mathcal{X})$, we have $\int (h \circ T - h)\overline{g}\,\mathrm{d}\mu = 0$, which implies

$$\int (h \circ T)\overline{g}\,\mathrm{d}\mu = \int h\overline{g}\,\mathrm{d}\mu\ . \tag{$\clubsuit$}$$

Now, we can write

$$\begin{aligned}\int |g \circ T - g|^2\,\mathrm{d}\mu &= \int (g \circ T - g)(\overline{g \circ T - g})\,\mathrm{d}\mu \\ &= \int (g \circ T)(\overline{g} \circ T)\,\mathrm{d}\mu - \int (g \circ T)\overline{g}\,\mathrm{d}\mu + \int g\overline{g}\,\mathrm{d}\mu - \int g(\overline{g} \circ T)\,\mathrm{d}\mu\ .\end{aligned}$$

Now observe that all the four terms on the right-hand side are equal to $\int g\overline{g}\,\mathrm{d}\mu$: the first term because of the invariance of μ, and the 2nd and the 4th by ($\clubsuit$). Therefore, $\int |g \circ T - g|^2\,\mathrm{d}\mu = 0$, which implies $g \in I$. It follows that $\overline{B}^{\perp} \subseteq I$.

We conclude that $\overline{B}^{\perp} = I$. □

Remark (Extension to higher dimensions)**.** The above proof can be readily generalized to higher dimensional dynamics, where the transformation T is replaced with a measure-preserving action of $\mathbb{Z}^d$ on $\mathcal{X}$. More generally, the same argument works if $\mathbb{Z}^d$ is replaced with any countable amenable group. ◇

Remark (Hilbert space variant)**.** The theorem and its proof rely only on the fact that $L^2_\mu(\mathcal{X})$ is a Hilbert space and $f \mapsto f \circ T$ is a linear isometry.

The theorem extends to an abstract setting where $L^2_\mu(\mathcal{X})$ is replaced with a Hilbert space and $\mathcal{H}$ and $f \mapsto f \circ T$ is replaced with a linear contraction $U : \mathcal{H} \to \mathcal{H}$. ◇

Exercise. Formulate and prove the abstract version of von Neumann's theorem mentioned in the latter remark.