



THE ERGODIC THEOREM

Emmanuel Lesigne

► To cite this version:

| Emmanuel Lesigne. THE ERGODIC THEOREM. Master. France. 2023, pp.8. hal-04012715

HAL Id: hal-04012715

<https://hal.science/hal-04012715>

Submitted on 3 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THE ERGODIC THEOREM

EMMANUEL LESIGNE

We give here a self-contained proof of the ergodic theorem for a measure preserving action of the additive group of real numbers. There are two versions of this theorem : von Neumann ergodic theorem gives a convergence in the mean ; Birkhoff ergodic theorem gives a pointwise convergence. Nothing original here, but a direct access to a complete proof which can be useful to non specialists of the subject.

1. VON NEUMANN ERGODIC THEOREM

Theorem. *Let $(T_t)_{t \in \mathbb{R}}$ be a one-parameter group of measure preserving transformations of the finite measure space (X, m) . For any $p \in [1, +\infty)$ and any $F \in L^p(m)$,*

$$(1) \quad \lim_{s \rightarrow +\infty} \frac{1}{s} \int_0^s F(T_t x) dt$$

exists in $L^p(m)$.

Remarks.

- (i) A one-parameter group of measure transformations of (X, m) is a family $(T_t)_{t \in \mathbb{R}}$ of measurable maps from X into itself satisfying the four following properties :

$$T_0 = \text{id}_X, \quad T_t \circ T_s = T_{t+s} \text{ for all reals } t \text{ and } s,$$

the map $(t, x) \mapsto T_t(x)$ is measurable from $\mathbb{R} \times X$ into X ,

and, of course, for all measurable subset A of X and all t , $m(T_t(A)) = m(A)$.

(This last property is equivalent to the fact that, for all t , the map $F \mapsto F \circ T_t$ is an isometry in any of the $L^p(m)$ spaces.)

- (ii) In the Hilbertian situation ($p = 2$), the limit (1) is the orthogonal projection of F onto the subspace of invariant functions

$$\mathcal{I} := \{ F \in L^2(m) \mid F \circ T_t = F \text{ for all } t \}.$$

(In case when m is a probability, this projection is a conditional expectation.)

- (iii) The m -almost everywhere existence of the limit (1) is true, but it is not a direct consequence of the convergence in norm. The pointwise convergence theorem is

known as Birkhoff ergodic theorem, and its proof is more subtle than the present one. See Section 2.

First step of the proof.

We claim that the set of $F \in L^p(m)$ such that the limit (1) exists in $L^p(m)$ is closed in $L^p(m)$. Indeed, if $F, G \in L^p(m)$ then, for all $S > 0$,

$$\begin{aligned} \left\| \frac{1}{S} \int_0^S F(T_t x) dt - \frac{1}{S} \int_0^S G(T_t x) dt \right\|_{L^p(m)} &= \left(\int_X \left| \frac{1}{S} \int_0^S (F - G)(T_t x) dt \right|^p dm(x) \right)^{1/p} \\ &\leq \left(\int_X \left(\frac{1}{S} \int_0^S |F - G|^p(T_t x) dt \right) dm(x) \right)^{1/p} = \left(\frac{1}{S} \int_0^S \left(\int_X |F - G|^p(T_t x) dm(x) \right) dt \right)^{1/p} \\ &= \left(\frac{1}{S} \int_0^S \left(\int_X |F - G|^p(x) dm(x) \right) dt \right)^{1/p} = \|F - G\|_{L^p(m)}, \end{aligned}$$

and from this contraction property we deduce easily that, if $F \in L^p(m)$ can be approximated by functions G such that $\frac{1}{S} \int_0^S G(T_t x) dt$ has the Cauchy property in L^p -norm when S goes to infinity, then the same is true for $\frac{1}{S} \int_0^S F(T_t x) dt$.

Second step of the proof : case $p = 2$.

Let us denote by \mathcal{I} the space of square integrable invariant functions and by \mathcal{C} the closed linear subspace of $L^2(m)$ generated by the set of all functions of the form $G \circ T_t - G$ with $G \in L^\infty(m)$ and $t \in \mathbb{R}$. We claim that

$$L^2(m) = \mathcal{I} \oplus \mathcal{C}.$$

Indeed, on one hand, if $F \in \mathcal{I}$, $G \in L^\infty(m)$ and $t \in \mathbb{R}$, then

$$\int F \cdot (G \circ T_t - G) dm = \int F \cdot G \circ T_t dm - \int F \cdot G dm = \int F \circ T_t \cdot G \circ T_t dm - \int F \cdot G dm = 0,$$

so that $\mathcal{I} \subset \mathcal{C}^\perp$,

and on the other hand, if $F \in \mathcal{C}^\perp$ then for any $G \in L^\infty(m)$ and any $t \in \mathbb{R}$,

$$\begin{aligned} \int (F \circ T_t - F) \cdot G dm &= \int F \circ T_t \cdot G dm - \int F \cdot G dm = \int F \cdot G \circ T_{-t} dm - \int F \cdot G dm \\ &= \int F \cdot (G \circ T_{-t} - G) dm = 0 \end{aligned}$$

which shows that $F \circ T_t - F = 0$, hence $F \in \mathcal{I}$, so that $\mathcal{C}^\perp \subset \mathcal{I}$. We conclude that $\mathcal{C}^\perp = \mathcal{I}$, which gives the claim.

In order to prove the existence of (1) for all $F \in L^2(m)$ it is enough to consider the case when $F \in \mathcal{I}$ and the case when $F \in \mathcal{C}$. The first case is trivial and in the second case we have $F = G \circ T_s - G$ and

$$\begin{aligned} \frac{1}{S} \int_0^S F(T_t x) dt &= \frac{1}{S} \left(\int_0^S G(T_{s+t} x) dt - \int_0^S G(T_t x) dt \right) \\ &= \frac{1}{S} \left(\int_s^{S+s} G(T_t x) dt - \int_0^S G(T_t x) dt \right) = \frac{1}{S} \left(\int_s^{S+s} G(T_t x) dt - \int_0^s G(T_t x) dt \right), \end{aligned}$$

which goes uniformly to zero when S goes to infinity if G is bounded.

Last step of the proof : any exponent $p \in [1, +\infty)$.

Let F be a bounded measurable function on X . The averages $\frac{1}{S} \int_0^S F(T_t x) dt$ stay uniformly bounded and are convergent in $L^2(m)$. We claim that these averages converge also in $L^p(m)$ and since bounded functions form a dense subspace of $L^p(m)$ this claim concludes the proof. The claim can be based on one of the following inequalities :

— if $p \in [1, 2]$ and $G \in L^2(m)$, then

$$\|G\|_{L^p(m)} \leq (m(X))^{\frac{1}{p}-\frac{1}{2}} \|G\|_{L^2(m)};$$

— if $p \geq 2$ and $G \in L^\infty(m)$, then

$$\|G\|_{L^p(m)} \leq \|G\|_{L^\infty(m)}^{1-\frac{2}{p}} \|G\|_{L^2(m)}^{\frac{2}{p}}.$$

2. BIRKHOFF ERGODIC THEOREM

Theorem. *Let $(T_t)_{t \in \mathbb{R}}$ be a one-parameter group of measure preserving transformations of the finite measure space (X, m) . For any $F \in L^1(m)$,*

$$(2) \quad \lim_{S \rightarrow +\infty} \frac{1}{S} \int_0^S F(T_t x) dt \quad \text{exists for } m\text{-almost all } x.$$

Remark. Thanks to Fubini theorem, for any $F \in L^1(m)$ and any $S > 0$, the expression $\int_0^S F(T_t x) dt$ makes sense for almost all x . In the following, we'll define and manipulate some maximal functions. The fact that they are well defined and measurable can be established in details.

Lemma (A maximal inequality). *Let F be a positive integrable function on X . For any $\alpha > 0$,*

$$m \left(\left\{ x \in X \mid \sup_{S > 0} \frac{1}{S} \int_0^S F(T_t x) dt > \alpha \right\} \right) \leq \frac{6}{\alpha} \int_X F(x) dm(x).$$

Proof of the theorem, knowing the lemma. We claim that the set of functions F satisfying the conclusion of the theorem is closed in $L^1(m)$. This will conclude the proof of the theorem since

- the result is true for the functions which are invariant under the flow (T_t) and for the functions of the form $G \circ T_s - G$ with G bounded ;
- square integrable invariant functions and functions of the form $G \circ T_s - G$ (with G bounded) generate a dense subspace of $L^2(m)$, hence of $L^1(m)$. (cf second step in the proof of von Neumann ergodic theorem.)

So, let us prove the claim. Consider an integrable function F which can be approximated in $L^1(m)$ by functions satisfying (2). Fix $\varepsilon > 0$ and $G \in L^1(m)$ such that

$$\int_X |F - G| dm \leq \varepsilon \quad \text{and} \quad \lim_{S \rightarrow +\infty} \frac{1}{S} \int_0^S G(T_t x) dt \text{ exists almost everywhere.}$$

Fix $\delta > 0$. Using triangular inequality, we see that the set

$$\left\{ x \in X \mid \limsup_{S, R \rightarrow +\infty} \left| \frac{1}{S} \int_0^S F(T_t x) dt - \frac{1}{R} \int_0^R F(T_t x) dt \right| > 3\delta \right\}$$

is included in the union of the three following ones

$$E_1 := \left\{ x \in X \mid \limsup_{S \rightarrow +\infty} \left| \frac{1}{S} \int_0^S F(T_t x) dt - \frac{1}{S} \int_0^S G(T_t x) dt \right| > \delta \right\}$$

$$E_2 := \left\{ x \in X \mid \limsup_{S, R \rightarrow +\infty} \left| \frac{1}{S} \int_0^S G(T_t x) dt - \frac{1}{R} \int_0^R G(T_t x) dt \right| > \delta \right\}$$

$$E_3 := \left\{ x \in X \mid \limsup_{R \rightarrow +\infty} \left| \frac{1}{R} \int_0^R F(T_t x) dt - \frac{1}{R} \int_0^R G(T_t x) dt \right| > \delta \right\}$$

By hypothesis, the set E_2 has zero m -measure. The sets E_1 and E_3 are the same and they are included in the set

$$E := \left\{ x \in X \mid \sup_S \frac{1}{S} \int_0^S |F - G|(T_t x) dt > \delta \right\}$$

We can apply Lemma 2 and using the fact that $\|F - G\|_1 \leq \varepsilon$ we obtain

$$m(E) \leq \frac{6}{\delta} \varepsilon.$$

We conclude that

$$m \left(\left\{ x \in X \mid \limsup_{S, R \rightarrow +\infty} \left| \frac{1}{S} \int_0^S F(T_t x) dt - \frac{1}{R} \int_0^R F(T_t x) dt \right| > 3\delta \right\} \right) \leq \frac{6}{\delta} \varepsilon.$$

But ε is arbitrarily small hence the measure on the left hand side is zero. Taking the union of these sets along a sequence of δ 's going to zero, we obtain that, for m -almost all x

$$\limsup_{S, R \rightarrow +\infty} \left| \frac{1}{S} \int_0^S F(T_t x) dt - \frac{1}{R} \int_0^R F(T_t x) dt \right| = 0,$$

which gives the announced result, thanks to Cauchy's criterion. \square

Two auxiliary lemmas will be useful in order to prove Lemma 2.

Lemma (Vitali's covering lemma). *Let I_1, I_2, \dots, I_n be a finite family of finite length open intervals on the real line. There exists a subfamily $I_{j_1}, I_{j_2}, \dots, I_{j_m}$ of two by two disjoint intervals such that*

$$\bigcup_{k=1}^n I_k \subset \bigcup_{\ell=1}^m 3I_{j_\ell}.$$

(Notation : if $I = (a - r, a + r)$ then $3I := (a - 3r, a + 3r)$)

This is a very useful result, but not difficult to prove : exercise or <https://en.wikipedia.org/wiki/Vitali-covering-lemma> .

Lemma (Hardy-Littlewood maximal inequality). *We denote by λ the Lebesgue measure on \mathbb{R} . For any positive measurable function f on \mathbb{R} and any $\alpha > 0$, we have*

$$\lambda \left(\left\{ t \in \mathbb{R} \mid \sup_{r>0} \frac{1}{r} \int_{-r}^r f(t+s) ds > \alpha \right\} \right) \leq \frac{6}{\alpha} \int_{\mathbb{R}} f(s) ds.$$

Proof of the HL-maximal inequality. Let K be a compact subset of

$$(3) \quad \left\{ t \in \mathbb{R} \mid \sup_{r>0} \frac{1}{r} \int_{-r}^r f(t+s) ds > \alpha \right\}.$$

For any $t \in K$, there exists $r(t) > 0$ such that

$$\frac{1}{r(t)} \int_{-r(t)}^{r(t)} f(t+s) \, ds > \alpha.$$

From the covering of K by open intervals $(t-r(t), t+r(t))$, with $t \in K$, we can extract a finite covering $(t_k-r(t_k), t_k+r(t_k))$, $1 \leq k \leq n$. Vitali covering lemma tells us now that we can select a subfamily $(t_{j_\ell}-r(t_{j_\ell}), t_{j_\ell}+r(t_{j_\ell}))$, $1 \leq \ell \leq m$ of two by two disjoint intervals such that

$$K \subset \bigcup_{\ell=1}^m (t_{j_\ell} - 3r(t_{j_\ell}), t_{j_\ell} + 3r(t_{j_\ell})).$$

Remind that, for any $t \in K$

$$r(t) < \frac{1}{\alpha} \int_{t-r(t)}^{t+r(t)} f(s) \, ds.$$

So we have

$$\lambda(K) \leq \sum_{\ell=1}^m 6r(t_{j_\ell}) < 6 \sum_{\ell=1}^m \frac{1}{\alpha} \int_{t_{j_\ell}-r(t_{j_\ell})}^{t_{j_\ell}+r(t_{j_\ell})} f(s) \, ds \leq \frac{6}{\alpha} \int_{\mathbb{R}} f(s) \, ds.$$

(The last inequality uses the fact that the intervals are disjoint.)

Thanks to inner regularity of the Lebesgue measure, this upper bound uniform on all compact sets included in the set (3) gives the announced result. \square

Proof of the maximal inequality. Hardy-Littlewood maximal inequality concerns the action of \mathbb{R} on itself by translation. In order to obtain the maximal inequality for the action of \mathbb{R} on the measure space (X, m) we shall apply Hardy-Littlewood on each trajectory of the flow and integrate the result on X . This technic is known as Calderon transfer principle.

The function F is positive and measurable. We fix $\alpha > 0$, $S > 0$ and $R > 0$, we apply Lemma 2 to the function $t \mapsto F(T_t x) \mathbf{1}_{[0, S+R]}(t)$ and obtain, for a given x ,

$$\lambda \left(\left\{ t \in \mathbb{R} \mid \sup_{r>0} \frac{1}{r} \int_{-r}^r F(T_{t+s} x) \mathbf{1}_{[0, S+R]}(t+s) \, ds > \alpha \right\} \right) \leq \frac{6}{\alpha} \int_0^{S+R} F(T_t x) \, dt.$$

A fortiori, we have

$$\lambda \left(\left\{ t \in [0, S] \mid \sup_{0 < r \leq R} \frac{1}{r} \int_0^r F(T_{t+s} x) \mathbf{1}_{[0, S+R]}(t+s) \, ds > \alpha \right\} \right) \leq \frac{6}{\alpha} \int_0^{S+R} F(T_t x) \, dt,$$

but if $t \leq S$ and $s \leq r \leq R$ then $t+s \leq S+R$, so that we have in fact

$$\lambda \left(\left\{ t \in [0, S] \mid \sup_{0 < r \leq R} \frac{1}{r} \int_0^r F(T_{t+s} x) \, ds > \alpha \right\} \right) \leq \frac{6}{\alpha} \int_0^{S+R} F(T_t x) \, dt.$$

If we integrate on X each term of this inequality, use Fubini theorem and the invariance of the measure m by the flow (T_t) , we obtain

$$S \cdot m \left(\left\{ x \in X \mid \sup_{0 < r \leq R} \frac{1}{r} \int_0^r F(T_s x) \, ds > \alpha \right\} \right) \leq (S + R) \cdot \frac{6}{\alpha} \int_X F(x) \, dm(x).$$

Now, we let S go to infinity and we have

$$m \left(\left\{ x \in X \mid \sup_{0 < r \leq R} \frac{1}{r} \int_0^r F(T_s x) \, ds > \alpha \right\} \right) \leq \frac{6}{\alpha} \int_X F(x) \, dm(x).$$

When R increases along a sequence going to infinity, we have a uniform upper bound for the measures of an increasing sequences of set. This upper bound applies to the measure of the union of these sets. This gives the announced result

$$m \left(\left\{ x \in X \mid \sup_{0 < r} \frac{1}{r} \int_0^r F(T_s x) \, ds > \alpha \right\} \right) \leq \frac{6}{\alpha} \int_X F(x) \, dm(x).$$

□

3. A FEW REFERENCES

Historical references are

Birkhoff, G.D. Proof of the ergodic theorem. *Proc. Nat. Acad. Sci. U.S.A.* 17 (1931), 656-660.

von Neumann, J. Proof of the quasi-ergodic hypothesis. *Proc. Nat. Acad. Sci. U.S.A.* 18 (1932), 70-82.

Khinchine, A. Zu Birkhoffs Lösung des Ergodenproblems. *Math. Ann.* 107 (1933), 485-488.

Yosida, K. and S. Kakutani. Birkhoff's ergodic theorem and the maximal ergodic theorem. *Proc. Imp. Acad. Tokyo* 15 (1939), 165-168.

Wiener, N. The ergodic theorem. *Duke Math. J.* 5 (1939), 1-18.

Many treatises give the basics of ergodic theory, including the ergodic theorems. For example

Cornfeld, I.P., S.V. Fomin and Ya.G. Sinai. *Ergodic Theory*. Grundlehren der mathematischen Wissenschaften 245. Springer-Verlag, New York. 1982.

Petersen K. *Ergodic Theory*. Cambridge Studies in Advanced Mathematics, Cambridge University Press, 1983.

de la Rue, T. *Introduction à la théorie ergodique*. <http://delarue.perso.math.cnrs.fr/te.html>

Coudène Y. *Théorie ergodique et systèmes dynamiques*. Savoirs Actuels, CNRS Éditions, EDP Sciences, 2013.

Finally, let us mention three books devoted to extensions of ergodic theorems to various situations, in particular in operator theory, in infinite measure and for actions of general groups.

Krengel U. *Ergodic Theorems*. De Gruyter Studies in Mathematics 6. Walter de Gruyter and Co. Berlin, 1985.

Aaronson J. *An Introduction to Infinite Ergodic Theory*. Mathematical Surveys and Monographs 50. Amer. Math. Soc., Providence, RI. 1997.

Anantharaman C., J.-P. Anker, M. Babillot, A. Bonami, B. Demange, S. Grellier, F. Havard, P. Jaming, E. Lesigne, P. Maheux, J.-P. Otal, B. Schapira and J.-P. Schreiber *Théorèmes ergodiques pour des actions de groupes*. Monographie 41 de l'Enseignement Mathématique. Genève, 2010.

emmanuel.lesigne@lmpt.univ-tours.fr

Institut Denis Poisson

Université de Tours et CNRS

France