A PROOF OF THE SUBADDITIVE ERGODIC THEOREM

ANDERS KARLSSON

For Wolfgang Woess on the occasion of his sixtieth birthday

ABSTRACT. This is a presentation of the subadditive ergodic theorem. A proof is given that is an extension of F. Riesz' approach to the Birkhoff ergodic theorem.

1. Introduction

Between the years 1938 and 1948 F. Riesz published works in ergodic theory [R] and found among other things several extensions of the mean ergodic theorem by elegantly simple and powerful arguments. In October 1931, von Neumann had proved the basic version of the mean ergodic theorem [vN32] inspired by Koopman's observations published in a note in June that same year [Ko31]. In his 1944 lecture notes [R44]¹, Riesz gave an insightful exposition of these developments. In particular, he presented there a new proof of the maximal ergodic lemma needed for Birkhoff's a.e. ergodic theorem, which is a deeper result than the mean ergodic theorem and dates November 1931 [B31]. This became one of the standard proofs of this theorem and was based on a lemma which is harder to formulate than to prove (see Lemma 3.2 below). He also remarked that his work on the mean ergodic theorem was suggested by the method of Carleman [C32]. In fact, Carleman announced in May 1931 results similar to Koopman's as well as a proof of the mean ergodic theorem. This announcement was published in June 1931 [C31] and the details can be found in $[C32]^2$.

The main purpose of this note is to show how Riesz's method extends to give a proof of the subadditive ergodic theorem of Kingman from 1968 [K68]. This was a by-product of the work [KM99] and all the

Date: December, 2014.

¹He was supposed to deliver these lectures in Geneva in the spring of 1944, but he was prevented from coming.

²Several early publications in the subject give due credit to Carleman, for example E. Hopf's important monograph *Ergodentheorie* from 1938. In most modern works however, including the standard reference [Kr85], Carleman is not mentioned.

details of the proof are included. Note that this proof at the same time gives a proof of Birkhoff's theorem.

Let throughout this paper (X, μ) be a measure space with $\mu(X) = 1$ and $T: X \to X$ a measure preserving transformation. Recall the following result:

Theorem (Birkhoff, 1931). Let $f \in L^1(X)$, then there is an integrable, a.e. T-invariant function \bar{f} such that

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

for a.e. x (the convergence also takes place in L^1). In fact $\|\bar{f}\|_1 \leq \|f\|_1$ and for any T-invariant set A, $\int_A f = \int_A \bar{f}$, in particular $\int_X f = \int_X \bar{f}$.

Note that if

(1.1)
$$c(n,x) = \sum_{k=0}^{n-1} f(T^k x)$$

then

$$c(n+m,x) = c(n,x) + c(m,T^nx),$$

and we then call c an additive cocycle. These are all of the form (1.1), for f(x) = c(1, x).

If for a sequence of functions $a(n, .) \in L^1(X)$, with integer n > 0 and a(0, x) = 0, we instead require

$$a(n+m,x) \le a(n,x) + a(m,T^n x),$$

then a is here called a *subadditive cocycle*. Assume that

$$\inf \frac{1}{n} \int_{Y} a(n,x) d\mu(x) > -\infty.$$

Then the following generalization of the Birkhoff ergodic theorem holds.

Theorem (Kingman, 1968). Under the above conditions, there is an integrable, a.e. T-invariant function \bar{a} such that

$$\lim_{n \to \infty} \frac{1}{n} a(n, x) = \bar{a}(x)$$

for a.e. x (the convergence also takes place in L^1). Moreover

$$\lim_{n \to \infty} \frac{1}{n} \int_A a(n, x) d\mu(x) = \int_A \bar{a}(x) d\mu(x)$$

for all T-invariant measurable sets A.

A draft of this paper was written in 1998 when I was a graduate student at Yale University. Some people have since told me that my text has been useful to them, so I thought that it might perhaps be worthwhile to publish a revised version. In addition, I thank the referee for useful comments leading to an improved text. I dedicate it to Wolfgang Woess, whom I have had the pleasure of knowing since more than a decade. I have at various times benefitted from his many insights within, as well as outside of, mathematics.

2. A FEW EXAMPLES OF SUBADDITIVE COCYCLES

We give a few examples of subadditive cocycles in order to illustrate that Kingman's theorem is a significant extension of Birkhoff's theorem with many applications. In fact the origin of Kingman's theorem comes from probability theory (theory of percolation) in the works by Hammersley and Welsh.

2.1. Random products in a Banach algebra. Let $A: X \to B$ be a measurable map into a Banach algebra. Let

$$u(n, x) = A(T^{n-1}x)A(T^{n-2}x)...A(x),$$

then

$$a(n,x) = \log \|u(n,x)\|$$

is a subadditive cocycle, because $||AB|| \le ||A|| ||B||$. The corresponding convergence was first proved by Furstenberg and Kesten in 1960 for random products of matrices, of course without the use of the subadditive ergodic theorem. This application is used in some proofs of Oseledets' multiplicative ergodic theorem.

2.2. **Random walks.** Let G be a topological group and $h: X \to G$ a Borel measurable map. Let $v(n,x) = h(T^{n-1}x)...h(Tx)h(x)$, if $\{h \circ T^k\}$ are independent, then this is usually called a random walk. The range, that is how many points visited in G,

$$a(n, x) := \text{Card}\{v(i, x) : 1 \le i \le n\}$$

is a subadditive cocycle.

Assume that d is a left invariant metric on G, (e.g. a word metric in the case G is finitely generated) then the drift

$$b(n,x) := d(e,v(n,x))$$

is a subadditive cocycle, by the triangle inequality and the invariance of d. A third source of subadditivity is the entropy of the measures of the distribution of random walk after n steps, see [D80].

2.3. Metric theory of continued fractions. See [Ba97]. Let X = [0,1) and A be the Borel σ -algebra. For any x write it as

$$x = \frac{1}{\frac{1}{x}} = \frac{1}{\left[\frac{1}{x}\right] + \left\{\frac{1}{x}\right\}}.$$

Continuing this scheme gives the continued fraction expansion of x, $a(x) := \begin{bmatrix} \frac{1}{x} \end{bmatrix}$, etc. The relevant transformation of X is

$$Tx = \{\frac{1}{x}\}$$

unless x = 0 in which case Tx = 0. So $a_n(x) = a(T^{n-1}x)$.

The corresponding approximants $p_n(x)/q_n(x)$ are defined for all n if x is irrational, they are given by recusion formulas, and by

$$\frac{p_n(x)}{q_n(x)} = \frac{1}{a_1(x) + \frac{1}{a_2(x) + \frac{1}{a_n(x)}}}.$$

There is a unique T-invariant probability measure absolutely continuous with respect to Lebesgue measure, namely

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

called the Gauss measure.

By use of the subadditive ergodic theorem, Barbolosi proves that for all constants C > 0,

$$\lim_{n \to \infty} \frac{1}{n} \operatorname{Card} \{ 1 \le i \le n : |xq_i(x) - p_i(x)| \le \frac{C}{q_i(x)} \}$$

exists for a.e. x. The functions are not simply a subadditive cocycle, it is necessary to divide into odd and even index, due to the fact that approximants lie on alternating sides of x.

- 2.4. Further examples. There are too many applications of the sub-additive ergodic theorem to list here, let us just refer to [H72] and [D80].
 - 3. A PROOF OF KINGMAN'S THEOREM
- 3.1. Three elementary observations.

Proposition 3.1. Let $(v_n)_{n\geq 1}$ be a subadditive sequence of real numbers, that is $v_{n+m} \leq v_n + v_m$. Then the following limit exists

$$\lim_{n \to \infty} \frac{1}{n} v_n = \inf_{m > 0} \frac{1}{m} v_m \in \mathbb{R} \cup \{-\infty\}.$$

Proof. Given $\varepsilon > 0$, pick M such that $v_M/M \le \inf v_n/n + \varepsilon$. Decompose $n = k_n M + r_n$, where $0 \le r_n < M$. Hence $k_n/n \to 1/M$. Using the subadditivity and considering n big enough $(n > N(\varepsilon))$

$$\inf \frac{1}{m} v_m \leq \frac{1}{n} v_n = \frac{1}{n} v_{k_n M + r_n} \leq \frac{1}{n} (k_n v_M + v_{r_n})$$

$$\leq \frac{1}{M} v_M + \varepsilon \leq \inf \frac{1}{m} v_m + 2\varepsilon.$$

Since ε is at our disposal, the lemma is proved.

Note that this proposition implies that for any T-invariant set A,

$$\lim_{n\to\infty}\frac{1}{n}\int_A a(n,x)d\mu(x)=\inf_n\frac{1}{n}\int_A a(n,x)d\mu(x),$$

when X = A denote this value by $\gamma(a)$.

F. Riesz noted in [R44] that the following simple lemma can be used to prove Birkhoff's theorem, the lemma is sometimes called Riesz's combinatorial lemma or the lemma about leaders.

Lemma 3.2. Call the term c_u a leader in the finite sequence $c_0, c_1, ..., c_{n-1}$ if one of the sums

$$c_u, c_u + c_{u+1}, ..., c_u + ... + c_{n-1}$$

is negative. Then the sum of the leaders is non-positive. (An empty sum is by convention 0).

Proof. Proof by induction. If n = 1, then either $c_0 \ge 0$, in which case the sum is empty, or $c_0 < 0$, in which case the sum equals $c_0 < 0$.

Assume that the statement is true for integers smaller than n. Consider the two cases, c_0 is or is not a leader. If c_0 is not a leader then all leaders are among $c_1, ..., c_{n-1}$ in which case the induction hypothesis applies.

If c_0 is a leader, then pick the smallest integer k such that $c_0+...+c_k < 0$, then each c_i , $i \le k$ is a leader. If not then $c_i + ... + c_k \ge 0$, but by minimality $c_0 + ... + c_{i-1} \ge 0$, which is a contradiction.

Hence $c_0, ..., c_k$ are all leaders and $c_0 + ... + c_k < 0$, the remaining leaders (if any) are leaders of $c_k, ..., c_n$, for which the induction hypothesis applies.

Proposition 3.3. Let a(n,x) be a subadditive cocycle as in the introduction. Then the functions

$$f(x) = \limsup_{n \to \infty} \frac{1}{n} a(n, x)$$

and

$$g(x) = \liminf_{n \to \infty} \frac{1}{n} a(n, x)$$

are a.e T-invariant.

Proof. Note that $f(Tx) \ge f(x)$ and $g(Tx) \ge g(x)$ because of the subadditivity

$$a(n,Tx) \ge a(n+1,x) - a(1,x)$$

and in the case of limsup (same for liminf)

$$\lim \sup (\frac{1}{n}a(n+1,x) - \frac{1}{n}a(1,x)) = f(x).$$

Now integrate

$$\int_{X} \left[f(Tx) - f(x) \right] d\mu(x) = 0$$

by the *T*-invariance, but the integrand is non-negative, hence f(Tx) - f(x) = 0 a.e.

3.2. The maximal ergodic inequality. The following key lemma will be proved by an extension of the argument of F. Riesz. It thus avoids use of the usual maximal ergodic inequality and it is not more difficult than Derriennic's proof.

Lemma 3.4 (Derriennic, 1975). Let a(n,x) be a subadditive ergodic cocycle as in the introduction. Let

$$B = \{x : \liminf_{n \to \infty} \frac{1}{n} a(n, x) < 0\}$$

then

$$\lim_{n \to \infty} \frac{1}{n} \int_B a(n, x) d\mu \le 0.$$

Proof. For each n, let

$$\Psi_n = \{x : \inf_{1 \le k \le n} a(n, x) - a(n - k, T^k x) < 0\}.$$

Note that

$$A_n := \{x : \inf_{1 \le k \le n} a(k, x) < 0\} \subset \Psi_n$$

by subadditivity. Note also that $A_n \subset A_{n+1}$ and $B \subset \bigcup A_n$.

For each n, let

$$b_n(x) = a(n, x) - a(n - 1, Tx).$$

Because of telescoping we have that

$$a(n,x) - a(n-k,T^kx) = b_n(x) + b_{n-1}(Tx) + \dots + b_{n-k+1}(T^{k-1}x)$$

and in particular, recall that a(0, x) = 0,

$$a(n,x) = \sum_{0 \le k \le n-1} b_{n-k}(T^k x).$$

By definition, $T^k x \in \Psi_{n-k}$ means that there is a $j, k \leq j \leq n-1$ such that

$$b_{n-k}(T^k x) + \dots + b_{n-j}(T^j x) < 0.$$

Hence by the lemma about leaders applied to $c_k = b_{n-k}(T^k x)$ we have

$$\sum_{0 \le k \le n-1, T^k x \in \Psi_{n-k}} b_{n-k}(T^k x) \le 0.$$

Therefore, using the T-invariance of μ and B,

$$0 \geq \int_{B} \sum_{0 \leq k \leq n-1, T^{k}x \in \Psi_{n-k}} b_{n-k}(T^{k}x) d\mu(x) =$$

$$= \sum_{0 \leq k \leq n-1} \int_{B \cap T^{-k}\Psi_{n-k}} b_{n-k}(T^{k}x) d\mu(x) =$$

$$= \sum_{0 \leq k \leq n-1} \int_{T^{k}B \cap \Psi_{n-k}} b_{n-k}(x) d\mu(x) =$$

$$= \sum_{i=1}^{n} \int_{B \cap \Psi_{i}} b_{i}(x) d\mu(x).$$

On the other hand, again by the T-invariance

$$\frac{1}{n} \int_{B} a(n,x) d\mu(x) = \frac{1}{n} \sum_{i=1}^{n} \int_{B} b_{i}(x) d\mu(x) =$$

$$= \frac{1}{n} \sum_{i=1}^{n} \int_{B \cap \Psi_{i}} b_{i}(x) d\mu(x) + \frac{1}{n} \sum_{i=1}^{n} \int_{B - (B \cap \psi_{i})} b_{i}(x) d\mu(x) \le$$

$$\leq 0 + \frac{1}{n} \sum_{i=1}^{n} \int_{B - (B \cap A_{i})} a^{+}(1,x) d\mu(x)$$

since $b_j(x) \leq a(1,x) \leq a^+(1,x) := \max\{0,a(1,n)\}$, which is positive and $A_i \subset \Psi_i$. Now since $a^+ \in L^1$, $A_i \subset A_{i+1}$ and $B \subset \bigcup_{i \geq 1} A_i$ it follows

when taking limsup, that

$$\limsup \frac{1}{n} \int_B a(n,x) d\mu \le 0.$$

Since B is invariant, the limsup actually is the limit, by Proposition 3.1.

Proposition 3.5. Let a(n, x) be a subadditive ergodic cocycle as in the introduction. Let

$$B = \{x : \liminf_{n \to \infty} \frac{1}{n} a(n, x) < \lambda\}$$

then

$$\lim_{n \to \infty} \frac{1}{n} \int_{B} a(n, x) d\mu \le \lambda \mu(B).$$

Proof. Apply the lemma to $a(n,x)-n\lambda$, which is a subadditive cocycle.

3.3. The proof of a.e. convergence. First we establish the result for additive cocycles $c_n = c(n, x)$. The point is that $-c_n$ is again additive, hence in particular the previous proposition applies to $-c_n$ as well.

Let

$$E_{\alpha,\beta} = \{x : \liminf_{n} \frac{1}{n} c_n < \alpha < \beta < \limsup_{n} \frac{1}{n} c_n \}$$

and by Proposition 3.3 this set is T-invariant. Hence we can apply Proposition 3.5 with $X = E_{\alpha,\beta}$. If we let $E := \{x : \liminf \frac{1}{n} c_n < \alpha\}$, then $E \cap E_{\alpha,\beta} = E_{\alpha,\beta}$, and this gives

$$\int_{E_{\alpha,\beta}} c_1 d\mu = \lim_{n \to \infty} \frac{1}{n} \int_{E_{\alpha,\beta}} c_n d\mu \le \alpha \mu(E_{\alpha,\beta}).$$

And similarly for $-c_n$,

$$-\int_{E_{\alpha,\beta}} c_1 d\mu = \lim_{n \to \infty} \frac{1}{n} \int_{E_{\alpha,\beta}} -c_n d\mu \le -\beta \mu(E_{\alpha,\beta}).$$

This yields a contradiction unless $\mu(E_{\alpha,\beta}) = 0$, because

$$\beta\mu(E_{\alpha,\beta}) \le \int_{E_{\alpha,\beta}} c_1 d\mu \le \alpha\mu(E_{\alpha,\beta})$$

but $\beta > \alpha$

Now let $a_n(x) = a(n, x)$, be a subadditive cocycle satisfying the integrability assumptions in the introduction. Consider

$$v_n(x) = a_n(x) - \sum_{i=0}^{n-1} a_1(T^i x).$$

Note that v_n is a subadditive cocycle and $v_n \leq 0$. Let $g(x) = \liminf_n \frac{1}{n} v_n(x)$ and $f(x) = \limsup_n \frac{1}{n} v_n(x)$. For an arbitrary $\alpha > 0$, we want to show that $B := \{x : f(x) - g(x) > \alpha\}$ has measure zero. By the additive cocycle case above we know that $\frac{1}{n} \sum_{i=0}^{n-1} a_1(T^i x)$ converges a.e., so taken together this would show that $\frac{1}{n} a_n(x)$ converges a.e. as desired.

Fix $\varepsilon > 0$. Pick M large enough so that for all m > M,

$$\frac{1}{m} \int_X v_m \le \gamma(v) + \varepsilon.$$

Let

$$g^M(x) := \liminf_{n \to \infty} \frac{1}{nM} v_{nM}(x)$$

and

$$f^{M}(x) := \limsup_{n \to \infty} \frac{1}{nM} v_{nM}(x).$$

On the one hand, since nM is a subsequence of n, we have that $g^M(x) \ge g(x)$ and $f^M(x) \le f(x)$. On the other hand, by the subadiditvity and non-positivity of v_n , for all $0 \le k < M$,

$$v_{(n+1)M}(x) \le v_{nM+k}(x) + v_{M-k}(T^{nM+k}x) \le v_{nM+k}(x)$$

and

$$v_{nM+k}(x) \le v_{nM}(x) + v_k(T^{nM}x) \le v_{nM}(x).$$

Therefore in view of that $(n+1)nM/(nM+k) \to 1$ and $nM/(nM+k) \to 1$, we can conclude that $g^M(x) = g(x)$ and $f^M(x) = f(x)$. Let

$$v_n^M(x) := v_{nM}(x) - \sum_{i=0}^{n-1} v_M(T^{iM}x),$$

which again is subadditive and non-positive. We then have that:

$$f - g = f^M - g^M \le -\liminf_{n \to \infty} \frac{1}{nM} v_n^M,$$

because $v_n^M \leq 0$ and the established convergence for additive cocycles. This means that

$$B := \{x : f - g > \alpha\} \subset \{x : \liminf_{n \to \infty} \frac{1}{n} v_n^M(x) < -M\alpha\} =: E.$$

Note that

$$0 \ge \gamma(v^M) = M\gamma(v) - \int_X v_M \ge -M\varepsilon.$$

Now applying Proposition 3.5 we get

$$-M\alpha\mu(E) \ge \lim \frac{1}{n} \int_E v_n^M \ge \lim \frac{1}{n} \int_X v_n^M \ge -M\varepsilon.$$

Hence

$$\mu(E) \leq \frac{\varepsilon}{\alpha}$$

and letting $\varepsilon \to 0$ we conclude that $\mu(B) = 0$ for any $\alpha > 0$ as required.

The limit is almost everywhere T-invariant, by Proposition 3.3 and integrable by Fatou's lemma

$$\int \liminf -\frac{1}{n}v_n \le \liminf \int -\frac{1}{n}v_n < \infty.$$

In general, $\int |a(n,x)| \le n \int |a(1,x)|$.

4. Appendix: Garsia's proof of the maximal ergodic Lemma

This appendix, which is not used above, notes that Garsia's celebrated argument for Birkhoff's ergodic theorem [G70] has a minor subadditive extension, although not strong enough to yield Kingman's theorem.

Lemma 4.1. Let a(n,x) be a subadditive ergodic cocycle as in the introduction. For $1 \le n \le \infty$ define

$$E_n = \{x : \sup_{1 \le k \le n} a(k, x) \ge 0\}.$$

Then

$$\int_{E_n} a(1,x) d\mu \ge 0 \ and \ \int_{E_\infty} a(1,x) d\mu \ge 0.$$

Proof. Let

$$h_n(x) = \sup_{1 \le k \le n} a(k, x),$$

SO

$$h_n^+(x) - a(k, x) \ge 0$$

for $k \leq n$. By positivity and linearity of T, viewed as (Koopman) operator on functions,

$$Th_n^+ \ge Ta_k$$

for $k \le n$. $(a_k(x) = a(k, x))$.

Hence

$$a_1 + Th_n^+ \ge a_1 + Ta_k \ge a_{k+1}$$

by subadditivity. So

$$a_1 \ge a_{k+1} - Th_n^+$$

for all $k, 1 \le k \le n$ and trivially for k = 0. Therefore

$$a_1 \ge \sup_{0 \le k \le n-1} a_{k+1} - Th_n^+ = h_n - Th_n^+.$$

Now integrate

$$\int_{E_n} a_1 \ge \int_{E_n} h_n - \int_{E_n} T h_n^+ = \int_X h_n^+ - \int_{E_n} T h_n^+$$

$$\ge \int_X h_n^+ - \int_X T h_n^+ \ge 0,$$

because T is measure preserving (or contractive as in [G70] would be enough) and $h_n^+ \geq 0$.

The statement about $E_{\infty} = \bigcup E_n$ follows from passing to the limit.

REFERENCES

- [Ba97] Barbolosi, D. Une application du théorème ergodique sous-additif à la théorie métrique des fractions continues. J. Number Theory 66 (1997), no. 1, 172–182.
- [B31] Birkhoff, G.D. Proof of the ergodic theorem. *Proc. Nat. Acad. Sci. USA* 17 (1931) 656–660.
- [C31] Carleman, Torsten, Application de la théorie des équations intégrales singulières aux équations différentielles de la dynamique, Ark. Mat. Astr. Fys. 22, 1931.
- [C32] Carleman, Torsten, Application de la théorie des équations intégrales linéaires aux systèmes d'équations différentielles non linéaires. Acta Math. 59 (1932), no. 1, 63–87.
- [D75] Derriennic, Yves, Sur le théorème ergodique sous-additif. C. R. Acad. Sci. Paris Sér. A-B 281 (1975), no. 22, Aii, A985–A988.
- [D80] Derriennic, Yves, Quelques applications du théorème ergodique sous-additif. In: Conference on Random Walks (Kleebach, 1979) pp. 183–201, 4, Astérisque, 74, Soc. Math. France, Paris, 1980.
- [G70] Garsia, Adriano M. Topics in almost everywhere convergence. Lectures in Advanced Mathematics, 4 Markham Publishing Co., Chicago, Ill. 1970 x+154 pp.
- [H72] Hammersley, J. M. A few seedlings of research. In: *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* (Univ. California, Berkeley, Calif., 1970/1971), Vol. I: Theory of statistics, pp. 345–394. Univ. California Press, Berkeley, Calif., 1972.
- [KM99] Karlsson, Anders; Margulis, Gregory A., A multiplicative ergodic theorem and nonpositively curved spaces. Comm. Math. Phys. 208 (1999), no. 1, 107–123.
- [K68] Kingman, J. F. C. The ergodic theory of subadditive stochastic processes. *J. Roy. Statist. Soc.* Ser. B 30 1968 499–510.
- [Ko31] Koopman, B.O. Hamiltionian systems and transformations in Hilbert space. Proc. Nat. Acad. Sci. USA, 17 (1931) pp. 315–318.
- [Kr85] Krengel, Ulrich, *Ergodic theorems*. With a supplement by Antoine Brunel. de Gruyter Studies in Mathematics, 6. Walter de Gruyter & Co., Berlin, 1985. viii+357 pp.

- [vN32] von Neumann, J. Proof of the quasi-ergodic hypothesis. *Proc. Nat. Acad. Sci. USA* , 18 (1932) pp. 70–82.
- [R] Riesz, Frigyes, Oeuvres complètes. Publiées sur l'ordre de l'Académie des Sciences de Hongrie par Ákos Császár. 2 Vols. Akadémiai Kiadó, Budapest 1960 1601 pp. (4 plates).
- $[\mathrm{R}44]$ F. Riesz, Sur la théorie ergodique, Comment. Math. Helv. 17 (1944-1945), 221-239

Université de Genève and Uppsala universitet