

FLUCTUATION BOUNDS FOR ERGODIC AVERAGES OF AMENABLE GROUPS

ANDREW WARREN

ABSTRACT. We study fluctuations of ergodic averages generated by actions of amenable groups. In the setting of an abstract ergodic theorem for locally compact second countable amenable groups acting on uniformly convex Banach spaces, we deduce a highly uniform bound on the number of fluctuations of the ergodic average for a class of Følner sequences satisfying an analogue of Lindenstrauss's temperedness condition. Equivalently, we deduce a uniform bound on the number of fluctuations over long distances for arbitrary Følner sequences. As a corollary, these results imply associated bounds for a continuous action of an amenable group on a σ -finite L^p space with $p \in (1, \infty)$.

In this article, we consider a problem at the interface of *effective ergodic theory* and the *ergodic theory of group actions*.

By *effective ergodic theory*, we mean the following programme: insofar as ergodic theory provides theorems which tell us about the long-term behaviour of dynamical systems, we may ask for *more quantitative, or computationally explicit*, analogues of those theorems. For instance, the mean ergodic theorem of von Neumann asserts that, whenever (S, μ) is a σ -finite measure space, $T : S \rightarrow S$ is a measure-preserving transformation, and $f \in L^2(S, \mu)$, the sequence of ergodic averages $A_n f := \frac{1}{n} \sum_{i=1}^{n-1} f \circ T^{-i}$ converges in L^2 norm; but *how fast* does this convergence occur?

Effective ergodic theory is complicated by some rather general negative results. Indeed, it is known [19] that in the aforementioned setting of the mean ergodic theorem, there is *no* uniform rate of convergence of the averages $A_n f$ in norm if one considers arbitrary functions f in $L^2(S, \mu)$. Consequently, if we are to find a quantitative analogue of the mean ergodic theorem, viz. one which is more explicit about the nature of the convergence of $A_n f$, it is necessary to look for more subtle convergence data than a uniform rate of convergence. One such form of convergence data is the *number of ε -fluctuations* of a sequence for each $\varepsilon > 0$, or possibly ε -fluctuations satisfying some side condition, such as the fluctuations being “over long distances”. This is exactly the type of convergence data that we investigate here; a precise explanation of these terms is given in Definition 1.

On the other hand, the *ergodic theory of group actions* seeks to modify the machinery of classical ergodic theory, by replacing the action of a single measure-preserving transformation with a measure-preserving action by a group. In fact, since the action by a single invertible measure-preserving transformation can be identified with a measure-preserving action of $\mathbb{Z}$, many theorems in the ergodic theory of group actions contain results from classical ergodic theory as special cases. In particular, this is typically the case where one is interested in the ergodic theory of a class of group actions where the groups involved belong to a class of groups containing $\mathbb{Z}$, such as abelian groups, nilpotent groups, or, in the case of the present article, amenable groups (see Definition 12 below).

Consider, therefore, the following version of the mean ergodic theorem for actions of amenable groups:

Theorem. *Let $L^p(S, \mu)$ be such that either S is σ -finite and $1 < p < \infty$ or $\mu(S) < \infty$ and $p = 1$, and let $x \in L^p(S, \mu)$. Let G be a locally compact second countable amenable group with Haar measure dg , let G act continuously on (S, μ) by measure-preserving transformations, and let (F_n) be a Følner sequence of compact subsets of G . Then $A_n x := \frac{1}{|F_n|} \int_{F_n} \pi(g^{-1})x dg$ converges in L^p .*

This result is originally due to Greenleaf [11], whose proof goes by way of an abstract Banach space analogue of the mean ergodic theorem which is simultaneously general enough to deduce the mean ergodic theorem for an amenable group acting on any reflexive Banach space or any

$L^1(\mu)$ with μ a finite measure. Central to Greenleaf's proof is a fixed point argument which, in particular, does not give any effective convergence information about the averages $A_n x$. (If one specialises to the case where $p = 2$, a simpler proof is available [7, Thm. 8.13], but this proof *also* does not give any effective convergence information.)

Here our aim is to give an effective analogue of Greenleaf's theorem. At the cost of some generality — here, we only consider actions of amenable groups on *uniformly convex* Banach spaces — we obtain an explicit *uniform fluctuation bound* for $(A_n x)$. In other words, we deduce a result of the following form:

Theorem. (“Main theorem”) *Let $L^p(S, \mu)$ be such that S is σ -finite and $1 < p < \infty$, and let $x \in L^p(S, \mu)$. Let G be a locally compact second countable amenable group with Haar measure dg , let G act continuously on (S, μ) by measure-preserving transformations, and let (F_n) be a Følner sequence of compact subsets of G . Then $A_n x := \frac{1}{|F_n|} \int_{F_n} \pi(g^{-1}) x dg$ converges in L^p with a uniform bound on the number of ε -fluctuations over long distances for each $\varepsilon > 0$; where the uniform bound (both the number of fluctuations and the “long distances”) depends exclusively, and explicitly, on: the choice of $p \in (1, \infty)$, the quantity $\|x\|_{L^p(S, \mu)}$, and the Følner sequence (F_n) .*

In fact, the dependence of the number of ε -fluctuations over long distances of $A_n x$ on the choice of Følner sequence (F_n) is only via a specific type of data from the Følner sequence, which expresses the *fact* that (F_n) is a Følner sequence in a slightly more quantitatively explicit fashion. We call this data the *Følner convergence modulus*; see Definition 15. It also turns out to be possible to delete the “over long distances” clause from the main theorem if, instead, one adds a side condition on the Følner sequence used, namely the condition that the Følner sequence be “fast”, as detailed in Definition 17.

We deduce this theorem as a special case of an “explicit fluctuation bound” analogue of an abstract mean ergodic theorem for a certain class of amenable group actions on uniformly convex Banach spaces, which we also call the “main theorem” of the article; this result is stated as Theorem 21 below, which also gives the explicit form of the uniform bound. The modified version of this theorem where the fluctuations are not “over long distances” but the Følner sequence is assumed to be “fast” is given as Corollary 22.

The plan of the article is as follows. In Section 1, we establish a number of background facts from functional analysis and the theory of amenable groups which are needed to state and prove the main theorem. In Section 2, we first provide a specialised proof of an abstract mean ergodic theorem for lcsc amenable groups acting on uniformly convex Banach spaces, and then modify this proof so as to be sufficiently quantitatively explicit that we are able to deduce the main theorem. Lastly, in Section 3, a discussion of some related literature and open problems is provided.

1. PRELIMINARIES

Definition 1. Fix an $\varepsilon > 0$. Given a sequence (x_n) in some metric space, we say that (x_n) has at most N ε -fluctuations if for every finite sequence $n_1 < n_2 < \dots < n_k$ such that for each $1 \leq i < k$, $d(x_{n_i}, x_{n_{i+1}}) \geq \varepsilon$, it always holds that $k < N$. A weaker notion is “ ε -fluctuations at distance β ”: given some function $\beta : \mathbb{N} \rightarrow \mathbb{N}$ with $\beta(n) > n$ for every n , we say that (x_n) has at most N ε -fluctuations at distance β if for every finite sequence $n_1 < n_2 < \dots < n_k$ with the property that $n_{i+1} \geq \beta(n_i)$ for every $1 \leq i < k$ such that for each $1 \leq i < k$, $d(x_{n_i}, x_{n_{i+1}}) \geq \varepsilon$, it always holds that $k < N$.

Remark. It holds that a sequence (x_n) is Cauchy (viz. that for every $\varepsilon > 0$ there exists an N such that for $m, n \geq N$, $d(x_m, x_n) < \varepsilon$) iff for every $\varepsilon > 0$ there exists some N' such that (x_n) has at most N' ε -fluctuations, iff for any $\varepsilon > 0$ and β with $\beta(n) > n$, there exists some N'' such that (x_n) has at most N'' ε -fluctuations at distance β . (More precisely: if (x_n) is Cauchy then for *any* $\beta : \mathbb{N} \rightarrow \mathbb{N}$ with $\beta(n) > n$, (x_n) has only finitely many ε -fluctuations at distance β for each $\varepsilon > 0$; whereas conversely, if there is *any* such β so that (x_n) has only finitely many ε -fluctuations at distance β for each $\varepsilon > 0$, then it follows that (x_n) is Cauchy.) Likewise, it is obvious that if for a specific sequence (x_n) we happen to know an explicit N witnessing the Cauchy property, then this N also serves as an explicit upper bound for N' , and likewise any explicit N' serves as an explicit upper bound on N'' (for any β). However the converses are all false in a strong sense: there exist examples of sequences where N' is a computable function of ε but N is not computable [1], and likewise with N'' (for $\beta \neq n + 1$) and N' respectively [18].

These phenomena are certainly present in ergodic theory. As mentioned in the introduction, it has long been known that a single measure-preserving transformation acting on (S, μ) , then

when $1 \leq p < \infty$, there exist functions $f \in L^p(S, \mu)$ for which the convergence indicated by the mean (and pointwise) ergodic theorem occurs arbitrarily slowly [19]. In other words $(A_n x)$ does not exhibit a uniform rate of convergence. However, it was shown by Avigad and Rute [1] that when $p \in (1, \infty)$ in this setting — in fact, more generally, if the acted-upon space $L^p(S, \mu)$ is replaced with any uniformly convex Banach space $\mathcal{B}$ with modulus of uniform convexity $u(\varepsilon)$ (see definition below) — then there exists a uniform bound on the number of ε -fluctuations in the sequence $(A_n x)$ which depends only on $u(\varepsilon)$ and $\|x\|/\varepsilon$.

Definition 2. A normed vector space $(\mathcal{B}, \|\cdot\|)$ is said to be uniformly convex if there exists a nondecreasing function $u(\varepsilon)$ such that for all $x, y \in \mathcal{B}$ with $\|x\| \leq \|y\| \leq 1$ and $\|x - y\| \geq \varepsilon$, it follows that $\|\frac{1}{2}(x + y)\| \leq \|y\| - u(\varepsilon)$. Such a function $u(\varepsilon)$ is then referred to as a *modulus of uniform convexity* for $\mathcal{B}$. We say that $\mathcal{B}$ is *p-uniformly convex* if $K\varepsilon^{p+1}$ is a modulus of uniform convexity for $\mathcal{B}$, where K is some constant.

Remark. There are a number of equivalent ways to define uniform convexity. We have chosen the preceding definition because it is the most convenient for our argument, but it is worth mentioning another characterisation which is perhaps more standard: a space $(\mathcal{B}, \|\cdot\|)$ is uniformly convex provided there is a nondecreasing function $\delta(\varepsilon)$ (also called a modulus of uniform convexity) such that for all $x, y \in \mathcal{B}$ with $\|x\|, \|y\| \leq 1$ and $\|x - y\| \geq \varepsilon$, it follows that $\|\frac{1}{2}(x + y)\| \leq 1 - \delta(\varepsilon)$. It is not hard to show (cf. [17, Lemma 3.2]) that a function $\delta(\varepsilon)$ is a modulus of convexity in this sense iff $u(\varepsilon) = \frac{\varepsilon}{2}\delta(\varepsilon)$ is a modulus of uniform convexity in the sense of our definition. (This indicates the origin of the off-by-one issue in our definition of *p*-uniform convexity.)

It is well known [6] that the L^p and ℓ^p spaces are uniformly convex when $p \in (1, \infty)$. Hanner showed [12] that for $p \in [2, \infty)$, the *sharp* modulus $\delta(\varepsilon)$ for L^p has an especially nice form, namely $\delta(\varepsilon) = 1 - (1 - (\frac{\varepsilon}{2})^p)^{1/p}$. In particular, this implies that $u(\varepsilon) = \frac{\varepsilon}{2} - \frac{\varepsilon}{2}(1 - (\frac{\varepsilon}{2})^p)^{1/p}$ is a modulus of uniform convexity, in our sense, for L^p with $p \in [2, \infty)$. (The same work shows that, even though L^p is also uniformly convex for $p \in (1, 2)$, the sharp modulus $\delta(\varepsilon)$ does not have as nice of an explicit form.)

We recall some basic notions from the theory of vector-valued integration. We shall closely follow the recent textbook by Hytönen et al. [13]; for the convenience of the reader, we will sometimes refer directly to specific definitions, theorems, etc. therein.

Consider some measure space $(G, \mathcal{A}, \mu)$ with some function $f : G \rightarrow \mathcal{B}$, with $(\mathcal{B}, \|\cdot\|)$ a Banach space. We say that $f(g)$ is a *simple function* with respect to the σ -algebra $\mathcal{A}$ and space $\mathcal{B}$, if it is of the form $\sum_{i=1}^N 1_{A_i}(g)b_i$, with $1_{A_i}(g)$ an indicator function for $A_i \in \mathcal{A}$, and $b_i \in \mathcal{B}$. We then say that a function f is *strongly measurable* if it is a pointwise limit of simple functions, i.e. if there exists a sequence f_n of simple functions such that for every $g \in G$, $\|f(g) - f_n(g)\| \rightarrow 0$ [13, Def. 1.1.4]. By contrast, a subtly different notion (but more standard in the vector integration literature) is that of μ -strong measurability, which only asserts this limit for μ -almost all $g \in G$, but requires that the sets A_i in the definition of simple function have finite μ -measure (see [13, Def. 1.1.13 and Def. 1.1.14]). We do not directly use μ -strong measurability in the main theorem of this paper — in particular, it is too weak of a form of measurability for Propositions 8 and 9 below. The relationship between strong measurability and μ -strong measurability is the following (quoting from [13, Prop. 1.1.16]):

Fact 3. Consider a measure space $(G, \mathcal{A}, \mu)$, a Banach space $\mathcal{B}$, and a function $f : G \rightarrow \mathcal{B}$.

- (1) If f is strongly μ -measurable, then f is μ -almost everywhere equal to a strongly measurable function.
- (2) If μ is σ -finite and f is μ -almost everywhere equal to a strongly measurable function, then f is strongly μ -measurable.

As a particular case of Fact 3, we see that when μ is σ -finite, a strongly measurable function is also μ -strongly measurable.

For μ -strongly measurable functions, one can define a form of integration, namely the *Bochner integral*, in direct analogy with the Lebesgue integral. Bochner integration is denoted by $\int_G f(g)d\mu$. A function is *Bochner μ -integrable* iff it is both μ -strongly measurable and $\int_G \|f(g)\|d\mu < \infty$, that is, $\|f\| : G \rightarrow \mathbb{R}$ is integrable in the Lebesgue sense [13, Prop. 1.2.2].

We record some other basic facts about the Bochner integral. (Each of these will be ultimately used in the proof of Lemma 19.)

Fact 4. Let (G, μ) be a measure space and $f : G \rightarrow \mathcal{B}$ be μ -strongly measurable.

- (1) $\|\int_G f(g)d\mu\| \leq \int_G \|f(g)\|d\mu$. [13, Prop. 1.2.2]
 (2) If $T \in \mathcal{L}(\mathcal{B}, \mathcal{B})$, then $T(\int_G f(g)d\mu) = \int_G Tf(g)d\mu$. [13, Eqn. 1.2]
 (3) Fubini's theorem holds for the Bochner integral. [13, Prop. 1.2.7] In particular, if (H, ν) is another measure space, and μ and ν are σ -finite, and $F : G \times H \rightarrow \mathcal{B}$ is Bochner integrable, then

$$\int_{G \times H} F d\mu \times d\nu = \int_H \left(\int_G F d\mu \right) d\nu = \int_G \left(\int_H F d\nu \right) d\mu.$$

A more general notion than strong measurability is *weak measurability*: we say that $f : G \rightarrow \mathcal{B}$ is *weakly measurable* if for every $b^* \in \mathcal{B}^*$, the function $b^* \circ f : G \rightarrow \mathbb{R}$ is measurable (in the ordinary sense as a function from $(G, \mathcal{A}, \mu)$ to $\mathbb{R}$ with the Borel σ -algebra). The following classical result indicates when weak measurability implies strong measurability.

Proposition 5. (Pettis measurability criterion [13, Thm. 1.1.6]) *Let (G, μ) be a measure space and $\mathcal{B}$ a Banach space. For a function $f : G \rightarrow \mathcal{B}$ the following are equivalent:*

- (1) f is strongly measurable.
- (2) f is weakly measurable, and $f(G)$ is separable in $\mathcal{B}$.

An easy consequence of the Pettis measurability criterion is the following.

Proposition 6. *If the measure space (G, μ) is also a separable topological space and every Borel set in G is μ -measurable, and $f : G \rightarrow \mathcal{B}$ is continuous, then f is strongly measurable.*

Proof. Observe that for any $b^* \in \mathcal{B}^*$, $b^* \circ f$ is a composition of continuous functions, and is therefore continuous. Hence f is weakly measurable. Moreover, it holds that the continuous image of a separable space is separable. $\square$

Before the next proposition, it will be convenient to introduce the following terminology.

Definition 7. If $(G, \mathcal{A}, \mu)$ is a measure space which is also a topological space, and $\mathcal{A}$ extends the Borel σ -algebra on G , then we say that $f : G \rightarrow \mathcal{B}$ is *weakly Borel* if $b^* \circ f : G \rightarrow \mathbb{R}$ is Borel, for all $b^* \in \mathcal{B}^*$. Likewise we say that $f : G \rightarrow \mathcal{B}$ is *strongly Borel* provided it is weakly Borel and that $f(G)$ is separable in $\mathcal{B}$.

Of course, in the case where $\mathcal{A}$ is precisely the Borel σ -algebra on G , then the notions of weakly and strongly Borel functions $f : G \rightarrow \mathcal{B}$ coincide with weak and strong measurability of f .

Proposition 8. *If $(G, \mathcal{A}, \mu)$ and $(H, \mathcal{C}, \nu)$ are both topological spaces where every Borel set is measurable, $\phi : H \rightarrow G$ is Borel, and $f : G \rightarrow \mathcal{B}$ is strongly Borel, then $f \circ \phi$ is strongly Borel (in particular strongly measurable).*

Proof. By hypothesis, f is also weakly Borel, so for any $b^* \in \mathcal{B}^*$, we have that $b^* \circ f$ is Borel. Therefore $b^* \circ (f \circ \phi) = (b^* \circ f) \circ \phi$ is Borel, and thus $f \circ \phi$ is weakly Borel. Now, let $A = \phi(H)$ be the image of ϕ in G . Note that $(f \circ \phi)(H) = f(A)$. Since $f(G)$ is separable in $\mathcal{B}$, it follows that $f(A)$ is also separable in $\mathcal{B}$ since it is contained in $f(G)$. $\square$

We now fix some notation and terminology regarding group actions on Banach spaces and measure spaces.

A locally compact group G will always come equipped with a Haar measure. In the countable discrete case this coincides with the counting measure. Regardless of whether the group is discrete or continuous, we will use the notations dg and $|\cdot|$ interchangeably to refer to the Haar measure. Conventions differ on the issue of whether the measure space (G, dg) is understood to come equipped with the Borel σ -algebra or its completion; in the latter case, the assertion that a function $f : G \rightarrow \mathcal{B}$ is strongly Borel is stronger than the assertion that f is strongly measurable. Therefore, in what follows, we assume only that the σ -algebra on G contains the Borel σ -algebra, and carefully note when a function $f : G \rightarrow \mathcal{B}$ is assumed to be strongly Borel rather than strongly measurable. We use the abbreviation *lcsc* for topological groups which are locally compact and second countable; importantly, for lcsc groups it always holds that the Haar measure is σ -finite.

In general, we say that a group G acts on a Banach space $(\mathcal{B}, \|\cdot\|)$ if there is a function $\pi(g)$ that returns an operator on $\mathcal{B}$ for every $g \in G$, $\pi(e)$ is the identity operator, and for all $g, h \in G$, $\pi(g)\pi(h) = \pi(gh)$. Together these imply that $\pi(g)^{-1} = \pi(g^{-1})$. We say that G *acts linearly* on $\mathcal{B}$ provided that in addition, π maps from G to the space $\mathcal{L}(\mathcal{B}, \mathcal{B})$ of linear operators on $\mathcal{B}$. Writing $\mathcal{L}_1(\mathcal{B}, \mathcal{B})$ to indicate the set of all linear operators from $\mathcal{B}$ to $\mathcal{B}$ with supremum norm 1, another

way to say that G acts both linearly and with unit norm on $\mathcal{B}$ is to say that G acts on $\mathcal{B}$ via $\pi : G \rightarrow \mathcal{L}_1(\mathcal{B}, \mathcal{B})$.¹ Likewise, we say that a topological group G *acts continuously* on $\mathcal{B}$ provided that for every $x \in \mathcal{B}$, if $g \rightarrow e$ then $\|\pi(g)x - x\| \rightarrow 0$. In other words $g \mapsto \pi(g)x$ is continuous from G to $\mathcal{B}$. In the case where G also acts linearly (resp. and with unit norm) on $\mathcal{B}$, this is equivalent to requiring that $\pi : G \rightarrow \mathcal{L}(\mathcal{B}, \mathcal{B})$ (resp. $\pi : G \rightarrow \mathcal{L}_1(\mathcal{B}, \mathcal{B})$) is continuous when $\mathcal{L}(\mathcal{B}, \mathcal{B})$ is equipped with the strong operator topology.

Relatedly, in the case where a group G acts on a measure space $(S, \mathcal{A}, \mu)$, we say that G *acts by measure-preserving transformations* if $\mu(g^{-1} \cdot A) = \mu(A)$ for every $g \in G$ and $A \in \mathcal{A}$. Likewise, we say that G *acts continuously* on $(S, \mathcal{A}, \mu)$ provided that for every $A \in \mathcal{A}$, if $g \rightarrow e$ then $\mu(A \Delta gA) \rightarrow 0$. This notion of continuous group action is related to our previous notion of continuous action on a Banach space by Proposition 10 below (which may be taken as a justification for the terminology).

Furthermore, we say that if G is understood as a measure space, then G acts *strongly* on a Banach space $\mathcal{B}$ provided that for every $x \in \mathcal{B}$, $g \mapsto \pi(g)x$ is strongly measurable from G to $\mathcal{B}$.² Likewise, we say that G *acts Borel strongly* on $\mathcal{B}$ provided that for every $x \in \mathcal{B}$, $g \mapsto \pi(g)x$ is strongly Borel from G to $\mathcal{B}$. Our argument will simultaneously require that G acts Borel strongly on $\mathcal{B}$, and linearly and with unit norm. To refer to this last condition, we will say that G acts Borel strongly on $\mathcal{B}$ via the representation $\pi : G \rightarrow \mathcal{L}_1(\mathcal{B}, \mathcal{B})$.

In the situation where G acts Borel strongly on $\mathcal{B}$ via the representation $\pi : G \rightarrow \mathcal{L}_1(\mathcal{B}, \mathcal{B})$, we can immediately deduce the following facts, which will be used in the proof of Lemma 19.

Proposition 9. *Suppose that G acts Borel strongly on $\mathcal{B}$ via the representation $\pi : G \rightarrow \mathcal{L}_1(\mathcal{B}, \mathcal{B})$. Then,*

- (1) *for each $A \subset G$ with $|A| < \infty$, we have that $1_A \pi(g^{-1})x$ is Bochner integrable for each $x \in \mathcal{B}$, i.e. $1_A \pi(g^{-1})x$ is μ -strongly measurable and $\int_A \|\pi(g^{-1})x\| dg < \infty$. Moreover,*
- (2) *for each $A, B \subset G$ with $|A|, |B| < \infty$ and for each $x \in \mathcal{B}$, we have that $1_{A \times B} \pi((hg)^{-1})x$ is Bochner integrable.*

Proof. (1) Since $\pi(\cdot)x$ is a strongly Borel function from G to $\mathcal{B}$, and $g \mapsto g^{-1}$ is a continuous and therefore Borel function from G to G , it follows from Proposition 8 that $g \mapsto \pi(g^{-1})x$ is strongly measurable (in fact strongly Borel). Since the Haar measure on G is σ -finite, this implies that $g \mapsto \pi(g^{-1})x$ is μ -strongly measurable, by Fact 3. For the second condition of Bochner integrability, it suffices to observe that since $\pi(g^{-1}) \in \mathcal{L}_1(\mathcal{B}, \mathcal{B})$ for every $g \in G$,

$$\int_A \|\pi(g^{-1})x\| dg \leq \int_A \|x\| dg < \infty.$$

(2) Since G is a topological group, it holds that group multiplication is continuous. Therefore $(g, h) \mapsto \pi((hg)^{-1})x$ is strongly measurable (in fact strongly Borel) thanks to Proposition 8, since it is a composition of the continuous (and therefore Borel) multiplication function and the strongly Borel function $\pi(\cdot)x$. Consequently, Fact 3 implies that $(g, h) \mapsto \pi((hg)^{-1})x$ is also μ -strongly measurable. We also have Bochner integrability because again,

$$\int_{A \times B} \|\pi((hg)^{-1})x\| dg \times dh \leq \int_{A \times B} \|x\| dg \times dh < \infty. \quad \square$$

Let us tie all this discussion of Bochner integration and group actions on Banach spaces back to our original setting of interest. In the “concrete version” of Greenleaf’s mean ergodic theorem, stated in the introduction, G acts continuously and by measure-preserving transformations on a σ -finite measure space (S, μ) , and $f \in L^p$ with $p \in (1, \infty)$. In the next proposition and corollary, we briefly check that in fact, this particular situation is actually covered by our abstract “group action on Banach space” framework.

¹We remark that any group that acts via a representation $\pi : G \rightarrow \mathcal{L}(\mathcal{B}, \mathcal{B})$ such that every $\pi(g)$ is *nonexpansive* actually does so via $\pi : G \rightarrow \mathcal{L}_1(\mathcal{B}, \mathcal{B})$, by the fact that $\pi(g^{-1}) = \pi(g)^{-1}$ and the general fact about linear operators that $\|T^{-1}\| \geq \|T\|^{-1}$. Nonexpansivity is required for the proof of our main result.

²By analogy with the previous situation where G acts continuously on $\mathcal{B}$: in the case where G also acts linearly (resp. and with unit norm) on $\mathcal{B}$, this is equivalent to requiring that $\pi : G \rightarrow \mathcal{L}(\mathcal{B}, \mathcal{B})$ (resp. $\pi : G \rightarrow \mathcal{L}_1(\mathcal{B}, \mathcal{B})$) satisfies a “strongly measurable with respect to the strong operator topology” like condition: for a precise statement, see [13, Cor. 1.4.7]. However we do not use this alternative, strong operator topology based, characterisation in our argument.

Proposition 10. *Suppose that G is a topological group which acts continuously and by measure-preserving transformations on a σ -finite measure space (S, μ) . Denote this action by*

$$G \times S \rightarrow S$$

$$(g, s) \mapsto g \cdot s.$$

Fix $p \in [1, \infty)$.

- (1) *This action of G on S induces an action of G on $L^p(S, \mu)$, given by*

$$G \times L^p(S, \mu) \rightarrow L^p(S, \mu)$$

$$(g, f) \mapsto f \circ g^{-1}$$

where for a fixed $g \in G$, the $L^p(S, \mu)$ function $f \circ g^{-1}$ is defined in the following manner: $(f \circ g^{-1})(s) := f(g^{-1} \cdot s)$. In other words, there is a representation π of G into the space of operators on $L^p(S, \mu)$, with $\pi(g)f = f \circ g^{-1}$.

- (2) *Moreover, G acts linearly and isometrically on $L^p(S, \mu)$, in the sense that $\|\pi(g)f\|_p = \|f\|_p$ for all $f \in L^p(S, \mu)$; so in particular, $\pi : G \rightarrow \mathcal{L}_1(L^p(S, \mu), L^p(S, \mu))$.*
 (3) *In addition, G acts continuously on $L^p(S, \mu)$, that is, for each $f \in L^p(S, \mu)$, $g \mapsto \pi(g)f$ is a continuous function from G to $L^p(S, \mu)$.*

Remark. This proposition is basically a modification of the rather classical *Koopman operator* formalism; except here, instead of a single measure-preserving transformation, we have a topological group of measure-preserving transformations. In the former case, the argument is quite standard; its generalisation to a group action is straightforward, but we explicitly go through the proof here for completeness. (The argument is nearly given in [7, Ch. 8], for instance, albeit using a more restrictive definition of continuous G -action on a measure space.)

Additionally, we note that the statement of the proposition includes the case $p = 1$, which is excluded from our main theorem; nor does the proposition require that G be lcsc or amenable.

Proof. (1) Here, we need only to check explicitly that $(g, f) \mapsto f \circ g^{-1}$ is indeed a group action of G on $L^p(S, \mu)$. To wit, we must show that $\pi(e)f = f$, and $\pi(gh)f = \pi(g)\pi(h)f$, for every $f \in L^p(S, \mu)$.

The first is obvious, since

$$\pi(e)f(s) = (f \circ e^{-1})(s) = f(e^{-1} \cdot s)$$

and $e^{-1} \cdot s = e \cdot s = s$ for every $s \in S$, since $(g, s) \mapsto g \cdot s$ is a group action on S . Likewise, the second condition also follows directly from the fact that $(g, s) \mapsto g \cdot s$ is a group action on S :

$$\pi(gh)f(s) = (f \circ (gh)^{-1})(s) = f((gh)^{-1} \cdot s) = f(h^{-1} \cdot (g^{-1} \cdot s)) = \pi(g)f(h^{-1} \cdot s) = \pi(g)\pi(h)f(s).$$

- (2) To see that $\pi(g)$ is a linear operator on $L^p(S, \mu)$ for each $g \in G$, simply note that

$$\pi(g^{-1})(cf_1 + f_2)(s) := (cf_1 + f_2)(g^{-1} \cdot s) = cf_1(g^{-1} \cdot s) + f_2(g^{-1} \cdot s) = c\pi(g^{-1})f_1 + \pi(g^{-1})f_2.$$

The isometry property follows immediately from the change of variable formula and the fact that the action of G on (S, μ) is measure preserving. Explicitly: fix $h \in G$, and compute that

$$\|\pi(h)f\|_p^p = \int_S |f(h^{-1} \cdot s)|^p d\mu(s) = \int_S |f(s)|^p d\mu(h \cdot s) = \int_S |f(s)|^p d\mu(s) = \|f\|_p^p.$$

- (3) We need to show that, for arbitrary $f \in L^p(S, \mu)$, if $g \rightarrow e$ then $\|\pi(g)f - f\|_p \rightarrow 0$. We first prove the claim for f an indicator function; then, for simple functions; then pass to general functions in $L^p(S, \mu)$ by a density argument.

Fix a measurable set $A \subseteq S$ with $\mu(A) < \infty$. By the assumption that the action of G on (S, μ) is continuous, it holds that: for every $\varepsilon > 0$, there exists a $U \ni e$ such that for all $g \in U$, $\mu(gA \Delta A) < \varepsilon$.

Let $\chi_A(s)$ denote the indicator function for A . Let U be a neighbourhood of the origin such that for every $g \in U$, $\mu(gA \Delta A) < \varepsilon^p$. Let $g \in U$. Observe that $\pi(g)\chi_A = \chi_A(g^{-1}(s)) = \chi_{gA}(s)$. Since $g \in U$, we have that $\mu(A \Delta gA) < \varepsilon^p$. But observe that

$$\|\chi_A - \chi_{gA}\|_p^p = \int_{A \Delta gA} 1 d\mu < \varepsilon^p.$$

Since our choice of ε was arbitrary, we conclude that $g \mapsto \pi(g)f$ is continuous when f is an indicator function of a set of finite measure.

Now let f be a simple function of the form $\sum_{i=1}^k c_i \chi_{A_i}$. Then, $\pi(g)f = \sum_{i=1}^k c_i \chi_{gA_i}$. In this case, it suffices to pick a neighbourhood $U \ni e$ which is small enough that for each i , we have that

$$\mu(A_i \Delta g_i A_i) < \left(\frac{\varepsilon}{k c_i} \right)^p.$$

Indeed, from the triangle inequality, and the previous result for indicator functions, we observe that

$$\|f - \pi(g)f\|_p \leq \sum_{i=1}^k \|c_i (\chi_{A_i} - \chi_{gA_i})\|_p < \sum_{i=1}^k c_i \left(\frac{\varepsilon}{k c_i} \right) = \varepsilon.$$

Lastly, take $f \in L^p(S, \mu)$ to be arbitrary. Suppose that f_0 is a simple function such that $\|f - f_0\|_p < \varepsilon/3$. By the previous case, we can pick a neighbourhood $U \ni e$ such that $\|f_0 - \pi(g)f_0\|_p < \varepsilon/3$ for all $g \in U$. Now, observe that

$$\|f - \pi(g)f\|_p \leq \|f - f_0\|_p + \|f_0 - \pi(g)f_0\|_p + \|\pi(g)f_0 - \pi(g)f\|_p.$$

We have already seen that the first two terms are each $< \varepsilon/3$. For the last term, observe that

$$\|\pi(g)f_0 - \pi(g)f\|_p = \|\pi(g)(f_0 - f)\|_p = \|f_0 - f\|_p$$

by part (2) of the proposition. Consequently, $\|\pi(g)f_0 - \pi(g)f\|_p < \varepsilon/3$ as well, so that $\|f - \pi(g)f\|_p < \varepsilon$ as desired. $\square$

Corollary 11. *Let G be a lcsc group acting continuously by measure-preserving transformations on a σ -finite measure space (S, μ) . Fix $p \in [1, \infty)$. Then the induced action of G on $L^p(S, \mu)$ is linear, strongly Borel, and has unit norm.*

Proof. That the induced action of G on $L^p(S, \mu)$ is linear with unit norm is immediate from Proposition 10 (2). From Proposition 10 (3), we know that $g \mapsto \pi(g)f$ is continuous from G to $L^p(S, \mu)$ for each f ; since G is separable, Proposition 6 indicates that $g \mapsto \pi(g)f$ is strongly measurable. Thus, we need only to check that $g \mapsto \pi(g)f$ is strongly Borel, as per Definition 7.

The continuity of $g \mapsto \pi(g)f$ and the separability of G imply that the image of G under this mapping is separable in $L^p(S, \mu)$. Likewise, since $g \mapsto \pi(g)f$ is continuous, post-composing the mapping with an element of the dual space of $L^p(S, \mu)$ results in a continuous map from G to $L^p(S, \mu)$, hence automatically Borel also. Therefore $g \mapsto \pi(g)f$ is both weakly Borel and has separable image, for any choice of f ; so we've shown that G acts Borel strongly on $L^p(S, \mu)$. $\square$

We now turn our attention to amenable groups. The following serves as our preferred characterisation of amenability.

Definition 12. (1) Let G be a countable discrete group. A sequence (F_n) of finite subsets of G is said to be a Følner sequence if for every $\varepsilon > 0$ and finite $K \subset G$, there exists an N such that for all $n \geq N$ and for all $k \in K$, $|F_n \Delta k F_n| < |F_n| \varepsilon$.

(2) Let G be a locally compact second countable (lcsc) group with Haar measure $|\cdot|$. A sequence (F_n) of compact subsets of G is said to be a Følner sequence if for every $\varepsilon > 0$ and compact $K \subset G$, there exists an N such that for all $n \geq N$, there exists a subset K' of K with $|K \setminus K'| < |K| \varepsilon$ such that for all $k \in K'$, $|F_n \Delta k F_n| < |F_n| \varepsilon$.

Remark. It has been observed, for instance, by Ornstein and Weiss [23] that (2) is one of several equivalent “correct” generalisations of (1) to the lcsc setting. Note however, that we do not assume (F_n) is nested ($F_i \subset F_{i+1}$ for all $i \in \mathbb{N}$) or exhausts G ($\bigcup_{n \in \mathbb{N}} F_n = G$), nor do we assume, in the lcsc case, that G is unimodular. (Each of these is a common additional technical assumption when working with amenable groups.) Conversely, some authors use a version of (2) where the sets in (F_n) are merely assumed to have finite volume, rather than compact; thanks to the regularity of the Haar measure, our definition results in no loss of generality.

Definition 13. If G is either a countable discrete or lcsc amenable group, and has some distinguished Følner sequence (F_n) , and acts Borel strongly on $\mathcal{B}$ via a representation $\pi : G \rightarrow \mathcal{L}_1(\mathcal{B}, \mathcal{B})$, then we define the n th ergodic average operator as follows: $A_n x := \int_{F_n} \pi(g^{-1}) x dg$. Here $\int$ denotes the normalised integral, that is, $\int_A f(g) dg := \frac{1}{|A|} \int_A f(g) dg$.

Proposition 14. *With the notation above, $\|A_n\|_{\mathcal{L}(\mathcal{B}, \mathcal{B})} \leq 1$.*

Proof. Evidently A_n is a linear operator from $\mathcal{B}$ to $\mathcal{B}$. From Proposition 9 we know that $1_{F_n}\pi(g^{-1})x$ is Bochner integrable, and in particular

$$\|A_n x\| := \left\| \int_{F_n} \pi(g^{-1})x dg \right\| \leq \int_{F_n} \|\pi(g^{-1})x\| dg \leq \int_{F_n} \|x\| dg = \|x\|. \quad \square$$

A key piece of quantitative information for us will be how large N has to be if g is chosen to be an element of (F_n) , in order for $|F_N \Delta g F_N|/|F_N|$ to be small. This information is encoded by the following modulus:

Definition 15. Let G be an amenable group, either countable discrete or lcsc, with Følner sequence (F_n) . A *Følner convergence modulus* $\beta(n, \varepsilon)$ for (F_n) returns an integer N such that:

(1) If G is countable discrete,

$$(\forall m \geq N)(\forall g \in F_n) [|F_m \Delta g F_m| < |F_m| \varepsilon].$$

(2) If G is lcsc,

$$(\forall m \geq N)(\exists F'_n \subset F_n)(\forall g \in F'_n) [|F_n \setminus F'_n| < |F_n| \varepsilon \wedge |F_m \Delta g F_m| < |F_m| \varepsilon].$$

We remark that if (F_n) is an *increasing* Følner sequence (that is, $F_n \subset F_m$ for all $n \leq m$) then it follows trivially that $\beta(n, \varepsilon)$ is a nondecreasing function for any fixed ε . However, in what follows we do not always assume that (F_n) is increasing. In some instances it is technically convenient to assume that $\beta(n, \varepsilon)$ is nondecreasing; in this case, we can upper bound $\beta(n, \varepsilon)$ using an “envelope” of the form $\tilde{\beta}(n, \varepsilon) = \max_{1 \leq i \leq n} \beta(i, \varepsilon)$. Hence, in any case we are free to assume that $\beta(n, \varepsilon)$ is nondecreasing in n if necessary.

It should be clear that if we have a “sufficiently explicit” amenable group G with a “sufficiently explicit” Følner sequence (F_n) , then we can explicitly write down a Følner convergence modulus $\beta(n, \varepsilon)$ for (F_n) . What is meant by this? Simply, the following: if we know “explicitly” that (F_n) is a Følner sequence, this amounts to knowing “explicitly”, for each $g \in G$, that $|F_n \Delta g F_n|/|F_n| \rightarrow 0$ (resp. the analogous statement for when G is lcsc), which in turn amounts to being able to write down how exactly this convergence occurs in an explicit way, i.e. that one can explicitly write down a rate of convergence $N(g, \varepsilon)$ for $|F_n \Delta g F_n|/|F_n|$. But a Følner convergence modulus $\beta(n, \varepsilon)$ is just a function which dominates $N(g, \varepsilon)$ for every $g \in F_n$ — in particular we can take $\beta(n, \varepsilon) = \max_{g \in F_n} N(g, \varepsilon)$.

Likewise (but less heuristically), it is easy to see that we can select a Følner sequence in such a way that $\beta(n, \varepsilon)$ can be chosen to be a computable function (for an appropriate restriction on the domain of the second variable). The following argument has essentially already been observed by previous authors [4, 5, 22] working with slightly different objects, but we include it for completeness. (In the interest of simplicity, we restrict our attention to countable discrete groups; one can say something similar for lcsc groups with computable topology, but we do not pursue this here.)

Proposition 16. *Let G be a countable discrete finitely generated amenable group with the solvable word property. Fix $k \in \mathbb{N}$. Then G has a Følner sequence (F_n) such that $\beta(n, k^{-1}) = \max\{n + 1, 3k\}$ is a Følner convergence modulus for (F_n) . Moreover (F_n) can be chosen in a computable fashion.*

Proof. Fix a computable enumeration of the elements of G , as well as a computable enumeration of the finite subsets of G . The solvable word property ensures that we can do this, and also that the cardinality of $F \Delta g F$ can always be computed for any $g \in G$ and finite set F . So, take F_1 to be an arbitrary finite subset of G containing g_1 , the first element of G . Given F_{n-1} , take $\tilde{F}_n$ to be the least (with respect to the enumeration) finite subset of G containing F_{n-1} , such that for all $g \in F_{n-1}$, $|\tilde{F}_n \Delta g \tilde{F}_n| < |\tilde{F}_n|/n$. Such an $\tilde{F}_n$ exists since G is amenable. Then, put $F_n = \tilde{F}_n \cup \{g_n\}$ where g_n is the n th element of G . This is indeed a Følner sequence: for a fixed g , we see that $|F_n \Delta g F_n|/|F_n| < 3/n$ for all n greater than the first N such that $g \in F_N$, because

$$\frac{|F_n \Delta g F_n|}{|F_n|} \leq \frac{2 + |\tilde{F}_n \Delta g \tilde{F}_n|}{|F_n|} \leq \frac{2}{|F_n|} + \frac{|\tilde{F}_n \Delta g \tilde{F}_n|}{|\tilde{F}_n|} < \frac{3}{n}.$$

Hence $|F_n \Delta g F_n|/|F_n| \rightarrow 0$. Moreover, we see that if $m \geq \max\{n + 1, 3k\}$, then

$$(\forall g \in F_n) \quad |F_m \Delta g F_m| < 3|F_m|/m \leq |F_m|/k$$

and so $\beta(n, k^{-1}) := \max\{n + 1, 3k\}$ is indeed a Følner convergence modulus for (F_n) . $\square$

Remark. The previous proposition is not sharp. It has been shown that there are groups *without* the solvable word property which nonetheless have computable Følner sequences with computable convergence behaviour [5]. (The cited paper uses a different explicit modulus of convergence for Følner sequences than the present paper, although the argument carries over to our setting without modification.)

Finally, it will be convenient to introduce the notion of a Følner sequence where the gap between n and $\beta(n, \varepsilon)$ is controlled.

Definition 17. Let G be a countable discrete or lcsc amenable group and (F_n) a Følner sequence. Let $\lambda \in \mathbb{N}$ and $\varepsilon > 0$. We say that (F_n) is a (λ, ε) -fast Følner sequence if

- (1) For G countable and discrete, it holds that for all $n \in \mathbb{N}$ that for all $m \geq n + \lambda$, for all $g \in F_n$, $|F_m \Delta g F_m|/|F_m| < \varepsilon$.
- (2) For G lcsc, it holds that for all $n \in \mathbb{N}$ that for all $m \geq n + \lambda$, there exists a set $F'_n \subset F_n$ such that $|F_n \setminus F'_n| < |F_n| \varepsilon$, so that for all $g \in F'_n$, $|F_m \Delta g F_m|/|F_m| < \varepsilon$.

In other words, a Følner sequence (F_n) is (λ, ε) -fast provided that $\beta(n, \varepsilon) = n + \lambda$ is a Følner convergence modulus for (F_n) .

It is clear that any Følner sequence can be refined into a (λ, ε) -fast Følner sequence.

Proposition 18. Given $\lambda \in \mathbb{N}$ and $\varepsilon > 0$, any Følner sequence can be refined into a (λ, ε) -fast Følner sequence.

Proof. It suffices to produce a $(1, \varepsilon)$ -fast refinement. For simplicity, we only state the argument for the case where G is countable and discrete.

Suppose we have already selected the first j Følner sets in our refinement $F_{n_1}, \dots, F_{n_j}$. Then, take $F_{n_{j+1}}$ to be the next element of the sequence (F_n) after n_j such that, for all $g \in \bigcup_{i=1}^j F_{n_i}$, $|F_{n_{j+1}} \Delta g F_{n_{j+1}}|/|F_{n_{j+1}}| < \varepsilon$. Such a term exists since (F_n) is a Følner sequence. $\square$

Less obvious is the relationship between a Følner sequence being fast and the property of being *tempered* which is used in Lindenstrauss's pointwise ergodic theorem [20], although they are somewhat similar in spirit. Nonetheless, we quickly observe that the previous proposition indicates that any tempered Følner sequence can be refined into a Følner sequence which is both tempered and (λ, ε) -fast, simply by the fact that any subsequence of a tempered Følner sequence is again tempered.

2. THE MAIN THEOREM

Frequently in ergodic theory, one argues that if $K \gg N$, then $A_K A_N x \approx A_K x$. The following lemma makes this precise in terms of the modulus β .

Lemma 19. Let $(\mathcal{B}, \|\cdot\|)$ be a normed vector space. Let G be a lcsc amenable group with Følner sequence (F_n) , and let G act Borel strongly on $\mathcal{B}$ via the representation $\pi : G \rightarrow \mathcal{L}_1(\mathcal{B}, \mathcal{B})$. Fix $N \in \mathbb{N}$ and $\eta > 0$. Let β be the Følner convergence modulus and suppose $K \geq \beta(N, \eta)$. Then for any $x \in \mathcal{B}$, $\|A_K x - A_K A_N x\| < 3\eta\|x\|$. (If G is countable discrete, the “strongly Borel” part is trivially satisfied, and we have the sharper estimate $\|A_K x - A_K A_N x\| < \eta\|x\|$.)

Proof. From the definition of Følner convergence modulus, we know that there exists an $F'_N \subset F_N$ such that $|F_N \setminus F'_N| < |F_N| \eta$ and such that for all $h \in F'_N$, $|F_K \Delta h F_K| < |F_K| \eta$. Now perform the following computation (justification for each step addressed below):

$$\begin{aligned}
\|A_K x - A_K A_N x\| &:= \left\| \int_{F_K} \pi(g^{-1}) x dg - \int_{F_K} \pi(g^{-1}) \left(\int_{F_N} \pi(h^{-1}) x dh \right) dg \right\| \\
&= \left\| \int_{F_K} \pi(g^{-1}) x dg - \int_{F_K} \left(\int_{F_N} \pi(g^{-1}) (\pi(h^{-1}) x) dh \right) dg \right\| \\
&= \left\| \int_{F_K} \pi(g^{-1}) x dg - \int_{F_K} \left(\int_{F_N} \pi((hg)^{-1}) x dh \right) dg \right\| \\
&= \left\| \int_{F_N} \left(\int_{F_K} \pi(g^{-1}) x dg \right) dh - \int_{F_N} \left(\int_{F_K} \pi((hg)^{-1}) x dg \right) dh \right\|
\end{aligned}$$

$$\begin{aligned}
&\leq \int_{F_N} \left\| \int_{F_K} \pi(g^{-1})xdg - \int_{F_K} \pi((hg)^{-1})xdg \right\| dh \\
&= \int_{F_N} \left\| \int_{F_K} \pi(g^{-1})xdg - \int_{hF_K} \pi(g^{-1})xdg \right\| dh \\
&\leq \int_{F_N} \left(\frac{1}{|F_K|} \int_{F_K \Delta hF_K} \|\pi(g^{-1})x\| dg \right) dh \\
&\leq \int_{F_N} \left(\frac{1}{|F_K|} \int_{F_K \Delta hF_K} \|x\| dg \right) dh \\
&= \int_{F_N} \frac{1}{|F_K|} (|F_K \Delta hF_K| \|x\|) dh \\
&< \frac{1}{|F_N|} \left[\int_{F'_N} \eta \|x\| dh + \int_{F_N \setminus F'_N} \left(\frac{1}{|F_K|} |F_K \Delta hF_K| \|x\| \right) dh \right] \\
&\leq \eta \|x\| + \frac{1}{|F_N|} \int_{F_N \setminus F'_N} (2\|x\|) dh \leq 3\eta \|x\|.
\end{aligned}$$

If G is countable discrete, we instead assume that for all $h \in F_N$ (rather than F'_N), $|F_K \Delta hF_K| < |F_K|\eta$. Therefore, the penultimate line reduces to $\frac{1}{|F_N|} \int_{F_N} \eta \|x\| dh$, and the last line reduces to $\eta \|x\|$.

Finally let's discuss which properties of the Bochner integral we had to use. Thanks to Proposition 9, we have that $g \mapsto 1_{F_K} \pi(g^{-1})x$ is Bochner integrable. Given that, for each g , $\pi(g)$ is a bounded linear operator, then indeed it follows that $\pi(g^{-1}) \left(\int \pi(h^{-1})xdh \right) = \int (\pi(g^{-1})\pi(h^{-1})x)dh$. From Fubini's theorem and the fact (also from Proposition 9) that $(g, h) \mapsto 1_{F_N \times F_K} \pi((hg)^{-1})x$ is Bochner integrable, we see that $\int_{F_K} \int_{F_N} \pi((hg)^{-1})xdh dg = \int_{F_N} \int_{F_K} \pi((hg)^{-1})xdg dh$. Lastly, we repeatedly invoked the fact that $\|\int_A f(g)dg\| \leq \int_A \|f(g)\|dg$. It's worth noting that in the case where G is countable discrete, only the first fact (that G acts by bounded linear operators with unit norm) is needed as an assumption, as the latter two properties hold trivially for finite averages. $\square$

Remark. It is possible to generalise this argument to the case where the action of G is “power bounded” in the sense that there is some uniform constant C such that for (dg -almost) all $g \in G$, $\|\pi(g)\| \leq C$. However the argument for our main theorem necessitates setting $C = 1$.

The following argument is a generalisation of a proof of Garrett Birkhoff [2] to the amenable setting. The statement of the theorem is weaker than results which are already in Greenleaf's article [11], but we include the argument for several reasons. One is that it is very short; another is that we will ultimately derive a bound on ε -fluctuations via a modification of this proof; and finally, the proof indicates additional information about the limiting behaviour of the norm of $A_n x$, namely that $\lim_n \|A_n x\| = \inf_n \|A_n x\|$.

Theorem 20. *Let G be a lcsc amenable group with compact Følner sequence (F_n) , and let $\mathcal{B}$ be a uniformly convex Banach space such that G acts Borel strongly on $\mathcal{B}$ via the representation $\pi : G \rightarrow \mathcal{L}_1(\mathcal{B}, \mathcal{B})$. Then for every $x \in \mathcal{B}$, the sequence of averages $(A_n x)$ converges in norm $\|\cdot\|_{\mathcal{B}}$.*

Proof. Without loss of generality, we assume $\|x\| \leq 1$ (otherwise, simply replace x with $x/\|x\|$). Define $L := \inf_n \|A_n x\|$. Fix an arbitrary $\varepsilon_0 > 0$, and let N be some index such that $\|A_N x\| < L + \varepsilon_0$. Let u denote the modulus of uniform convexity. Fix a Følner convergence modulus $\beta(n, \varepsilon)$ for (F_n) , and let $\eta > 0$ be arbitrary. Suppose $M \geq \beta(N, \eta/(3\|x\|))$ is an index such that $\|A_N x - A_M x\| > \delta$. (If no such δ exists then this means that after $\beta(N, \eta/(3\|x\|))$, the sequence has converged to within δ .) Then, from uniform convexity, we know that

$$\left\| \frac{1}{2}(A_N x + A_M x) \right\| \leq \max\{\|A_N x\|, \|A_M x\|\} - u(\delta).$$

Additionally, it follows from Lemma 19 that $\|A_M x - A_M A_N x\| < \eta$, and thus

$$\left\| \frac{1}{2}(A_N x + A_M x) \right\| < \max\{\|A_N x\|, \|A_M A_N x\| + \eta\} - u(\delta).$$

But $\|A_M A_N x\| \leq \|A_N x\|$, so this implies

$$\left\| \frac{1}{2}(A_N x + A_M x) \right\| < \|A_N x\| + \eta - u(\delta).$$

In turn, we know that $\|A_N x\| < L + \varepsilon_0$, so

$$\left\| \frac{1}{2}(A_N x + A_M x) \right\| < L + \varepsilon_0 + \eta - u(\delta).$$

In fact, it follows that $\|\frac{1}{2}A_K(A_N x + A_M x)\| < L + \varepsilon_0 + \eta - u(\delta)$ also, for any index K , since $\|A_K\| \leq 1$. Now, choosing $K \geq \max\{\beta(N, \eta/(3\|x\|)), \beta(M, \eta/(3\|x\|))\}$, we have that both $\|A_K x - A_K A_N x\| < \eta$ and $\|A_K x - A_K A_M x\| < \eta$. Thus,

$$\begin{aligned} \|A_K x\| &= \left\| \frac{1}{2}(A_K x - A_K A_N x) + \frac{1}{2}(A_K x - A_K A_M x) + \frac{1}{2}(A_K A_N + A_K A_M) \right\| \\ &\leq \eta + \left\| \frac{1}{2}A_K(A_N x + A_M x) \right\| \\ &< 2\eta + L + \varepsilon_0 - u(\delta). \end{aligned}$$

Since η can be chosen to be arbitrarily small (this merely implies that K and M are very large), we see that $\limsup_K \|A_K x\| \leq L + \varepsilon_0 - u(\delta)$. But since our choice of ε_0 was arbitrary, it follows that in fact $\limsup_K \|A_K x\| \leq L$.

Moreover this implies that $(A_n x)$ converges in norm. For if this were not the case, then we could find some $\delta_0 > 0$ such that $\|A_n x - A_m x\| > \delta_0$ infinitely often; in fact, there must be some $\delta_0 > 0$ such that $\|A_n x - A_m x\| > \delta_0$ infinitely often, *with the further restriction*³ that $m \geq \beta(n, \eta/(3\|x\|))$. Now, pick η and ε_0 small enough that $2\eta + \varepsilon_0 < u(\delta_0)$, and pick both n and m such that: m and n are sufficiently large that $\|A_n x\|, \|A_m x\| < L + \varepsilon_0$ (which we can do since $\limsup_K \|A_K x\| \leq L$); and, moreover, pick m and n in such a way that $\|A_n x - A_m x\| > \delta_0$ and $m \geq \beta(n, \eta/(3\|x\|))$. The computation above shows that for K larger than $\beta(m, \eta/(3\|x\|))$ and $\beta(n, \eta/(3\|x\|))$ we have that $\|A_K x\| < 2\eta + L + \varepsilon_0 - u(\delta_0)$; but since we chose $2\eta + \varepsilon_0 < u(\delta_0)$, this implies that $\|A_K x\| < L$, which contradicts the definition of L . $\square$

We now proceed to deriving a quantitative analogue of this result. We first do so “at distance β ”, and then recover a global bound in the case where the Følner sequence is fast (in the sense of Definition 17). The only really non-explicit part of the proof of the preceding theorem was the step where we used the fact that an infimum of a real sequence exists. Therefore the principal innovation in what follows is the avoidance of the direct invocation of this fact.

Fix a nondecreasing Følner convergence modulus β for (F_n) . Initially let us consider the case where $\|x\| \leq 1$. Suppose that $\|A_{n_0} x - A_{n_1} x\| \geq \varepsilon$. Moreover, we pick some $\eta < u(\varepsilon)/2$, and suppose that $n_1 \geq \beta(n_0, \eta/3\|x\|)$. Then the computation from the previous proof shows that

$$\left\| \frac{1}{2}(A_{n_0} x + A_{n_1} x) \right\| < \|A_{n_0} x\| + \eta - u(\varepsilon).$$

More generally, if $\|A_{n_i} x - A_{n_{i+1}} x\| \geq \varepsilon$ with $n_{i+1} \geq \beta(n_i, \eta/3\|x\|)$, it follows that

$$\left\| \frac{1}{2}(A_{n_i} x + A_{n_{i+1}} x) \right\| < \|A_{n_i} x\| + \eta - u(\varepsilon).$$

Now, choosing $k \geq \max\{\beta(n_{i+1}, \eta/(3\|x\|)), \beta(n_i, \eta/(3\|x\|))\}$, we have that both $\|A_k x - A_k A_{n_i} x\| < \eta$ and $\|A_k x - A_k A_{n_{i+1}} x\| < \eta$. Thus,

$$\begin{aligned} \|A_k x\| &= \left\| \frac{1}{2}(A_k x - A_k A_{n_i} x) + \frac{1}{2}(A_k x - A_k A_{n_{i+1}} x) + \frac{1}{2}(A_k A_{n_i} x + A_k A_{n_{i+1}} x) \right\| \\ &\leq \eta + \left\| \frac{1}{2}A_k(A_{n_i} x + A_{n_{i+1}} x) \right\| \\ &< 2\eta + \|A_{n_i} x\| - u(\varepsilon). \end{aligned}$$

Therefore let n_{i+2} equal the least index greater than $\max\{\beta(n_{i+1}, \eta/(3\|x\|)), \beta(n_i, \eta/(3\|x\|))\}$ (equivalently, greater than $\beta(n_{i+1}, \eta/(3\|x\|))$), since β is nondecreasing in n such that $\|A_{n_{i+1}} x -$

³Note, here, that we are invoking the fact that a sequence in a metric space converges iff, for any $\beta : \mathbb{N} \rightarrow \mathbb{N}$ with $\beta(n) > n$, the sequence contains only finitely many δ -fluctuations at distance β , for each $\delta > 0$; cf. discussion of this equivalence in the remark immediately following Definition 1.

$\|A_{n_{i+2}}x\| \geq \varepsilon$. The previous calculation shows that $\|A_{n_{i+2}}x\| < \|A_{n_i}x\| + 2\eta - u(\varepsilon)$. More generally, we have that

$$\|A_{n_i}x\| < \|A_{n_0}x\| - \frac{i}{2}(u(\varepsilon) - 2\eta) \quad i \text{ even}$$

$$\|A_{n_i}x\| < \|A_{n_1}x\| - \frac{i-1}{2}(u(\varepsilon) - 2\eta) \quad i \text{ odd}$$

So simply from the fact that $\|A_{n_i}x\| \geq 0$, these expressions derive a contradiction on the least i such that

$$\max\{\|A_{n_0}x\|, \|A_{n_1}x\|\} < \frac{i-1}{2}(u(\varepsilon) - 2\eta)$$

since this would imply that $\|A_{n_i}(x)\| < 0$. That is, the contradiction implies that the n_i th epsilon fluctuation could not have occurred. We have no *a priori* information on the values of $\|A_{n_0}x\|$ and $\|A_{n_1}x\|$, except that both are at most $\|x\|$. Therefore, we have the following uniform bound:

$$i \leq \left\lfloor \frac{\|x\|}{\frac{1}{2}u(\varepsilon) - \eta} + 1 \right\rfloor$$

where i tracks the indices of the subsequence along which ε -fluctuations occur. This is actually *one more* than the number of ε -fluctuations, so instead we have that the number of ε -fluctuations is bounded by $\left\lfloor \frac{\|x\|}{\frac{1}{2}u(\varepsilon) - \eta} \right\rfloor$.

If we happen to have any lower bound on the infimum of $\|A_n x\|$, we can sharpen the previous calculation. Instead of using the fact that for all n , $\|A_n x\| \geq 0$, we use the fact that $\|A_n x\| \geq L$ for some L . To wit, if i is large enough that

$$\|x\| < \frac{i-1}{2}(u(\varepsilon) - 2\eta) + L$$

then this would imply that $\|A_{n_i}x\| < L$, a contradiction. Therefore we have the bound

$$i \leq \left\lfloor \frac{\|x\| - L}{\frac{1}{2}u(\varepsilon) - \eta} + 1 \right\rfloor$$

and so the number of ε -fluctuations is bounded by $\left\lfloor \frac{\|x\| - L}{\frac{1}{2}u(\varepsilon) - \eta} \right\rfloor$.

In the case where $\|x\| > 1$, a small modification must be made. We can make the substitution $x' = x/\|x\|$, so that $\|A_{n_i}x' - A_{n_{i+1}}x'\| \geq \varepsilon/\|x\|$, and conclude (provided that $\eta < \frac{1}{2}u(\varepsilon/\|x\|)$) that $(A_n x')$ has at most $\left\lfloor \frac{\|x'\|}{\frac{1}{2}u(\varepsilon/\|x\|) - \eta} \right\rfloor \varepsilon/\|x\|$ -fluctuations at distance $\beta(n, \eta/3\|x'\|)$, or rather at most $\left\lfloor \frac{1}{\frac{1}{2}u(\varepsilon/\|x\|) - \eta} \right\rfloor \varepsilon/\|x\|$ -fluctuations at distance $\beta(n, \eta/3)$ since $\|x'\| = 1$. But since $\|A_{n_i}x' - A_{n_{i+1}}x'\| \geq \varepsilon/\|x\|$ iff $\|A_{n_i}x - A_{n_{i+1}}x\| \geq \varepsilon$, this actually tells us that $(A_n x)$ has at most $\left\lfloor \frac{1}{\frac{1}{2}u(\varepsilon/\|x\|) - \eta} \right\rfloor \varepsilon$ -fluctuations at distance $\beta(n, \eta/3)$. Likewise, if we happen to have a lower bound L on the infimum of $\|A_n x\|$, since $\inf_n \|A_n x\| \geq L$ iff $\inf_n \|A_n x'\| \geq L/\|x\|$ we have that $(A_n x)$ has at most $\left\lfloor \frac{1 - L/\|x\|}{\frac{1}{2}u(\varepsilon/\|x\|) - \eta} \right\rfloor \varepsilon$ -fluctuations at distance $\beta(n, \eta/3)$.

To summarise, we have shown that:

Theorem 21. (“Main theorem”) *Let $\mathcal{B}$ be a uniformly convex Banach space with modulus u . Fix $\varepsilon > 0$ and $x \in \mathcal{B}$. Pick some $\eta > 0$ such that $\eta < \frac{1}{2}u(\varepsilon)$ if $\|x\| \leq 1$, or $\eta < \frac{1}{2}u(\varepsilon/\|x\|)$ if $\|x\| > 1$. Then if G is a lcsc amenable group that acts Borel strongly on $\mathcal{B}$ via the representation $\pi : G \rightarrow \mathcal{L}_1(\mathcal{B}, \mathcal{B})$, with Følner sequence (F_n) , the sequence $(A_n x)$ has at most $\left\lfloor \frac{\|x\|}{\frac{1}{2}u(\varepsilon) - \eta} \right\rfloor \varepsilon$ -fluctuations at distance $\beta(n, \eta/3\|x\|)$ if $\|x\| \leq 1$, or at most $\left\lfloor \frac{1}{\frac{1}{2}u(\varepsilon/\|x\|) - \eta} \right\rfloor \varepsilon$ -fluctuations at distance $\beta(n, \eta/3)$ if $\|x\| > 1$. If we know that $\inf \|A_n x\| \geq L$, then we can sharpen the bound to $\left\lfloor \frac{\|x\| - L}{\frac{1}{2}u(\varepsilon) - \eta} \right\rfloor$ in the case where $\|x\| \leq 1$, or $\left\lfloor \frac{1 - L/\|x\|}{\frac{1}{2}u(\varepsilon/\|x\|) - \eta} \right\rfloor$ in the case where $\|x\| > 1$.*

Corollary 22. *In the above setting, suppose that (F_n) is $(\lambda, \eta/3\|x\|)$ -fast if $\|x\| \leq 1$, or $(\lambda, \eta/3)$ -fast if $\|x\| > 1$ respectively. Then the sequence $(A_n x)$ has at most $\lambda \cdot \left\lfloor \frac{\|x\|}{\frac{1}{2}u(\varepsilon) - \eta} \right\rfloor + \lambda \varepsilon$ -fluctuations (resp. $\lambda \cdot \left\lfloor \frac{1}{\frac{1}{2}u(\varepsilon/\|x\|) - \eta} \right\rfloor + \lambda \varepsilon$ -fluctuations).*

Proof. We know from the theorem that, if $\|x\| \leq 1$, there are at most $\left\lfloor \frac{\|x\|}{\frac{1}{2}u(\varepsilon) - \eta} \right\rfloor$ ε -fluctuations at distance λ . This leaves the possibility that there are some ε -fluctuations in the $\left\lfloor \frac{\|x\|}{\frac{1}{2}u(\varepsilon) - \eta} \right\rfloor$ many gaps of width λ , and also that there are some ε -fluctuations in between the last possible index n_i given by the previous theorem, and the index n_{i+1} at which contradiction is achieved. This last interval at the end is at most λ wide as well. The reasoning for the $\|x\| > 1$ case is identical. $\square$

Example 23. We mention two special cases of interest in the above setting. For brevity we restrict ourselves to the case where $\|x\| \leq 1$ and $L = 0$, but the reader can make the obvious substitutions.

- (1) In the case where $\mathcal{B}$ is p -uniformly convex with constant K , then in this setting $(A_n x)$ has at most $\lambda \cdot \left\lfloor \frac{\|x\|}{\frac{K}{2}\varepsilon^{p+1} - \eta} \right\rfloor + \lambda \varepsilon$ -fluctuations.
- (2) In the case where $\mathcal{B} = L^p(\mu)$ for $p \in [2, \infty)$, then in this setting $(A_n x)$ has at most $\lambda \cdot \left\lfloor \frac{\|x\|}{\frac{\varepsilon}{4} - \frac{\varepsilon}{4}(1 - (\frac{\varepsilon}{2})^{1/p}) - \eta} \right\rfloor + \lambda \varepsilon$ -fluctuations.

3. DISCUSSION

The prospect of obtaining quantitative convergence information for ergodic averages via a modification of Garrett Birkhoff's argument [2] has been previously explored, in the $\mathbb{Z}$ -action setting, by Kohlenbach and Leuştean [17] and subsequently by Avigad and Rute [1].

Our proof was carried out in the setting where the acted upon space was assumed to be uniformly convex, and indeed our bound on the number of fluctuations explicitly depends on the modulus of uniform convexity. Nonetheless, it is natural to ask whether an analogous result might be obtained for a more general class of acted upon spaces.

However, it has already been observed, in the case where $G = \mathbb{Z}$, that there exists a separable, reflexive, and strictly convex Banach space $\mathcal{B}$ such that for every N and $\varepsilon > 0$, there exists an $x \in \mathcal{B}$ such that $(A_n x)$ has at least N ε -fluctuations [1]. This counterexample applies equally to bounds on the rate of metastability (for more on metastability and its relationship with fluctuation bounds, see for instance [1, Section 5]). However, this counterexample does not directly eliminate the possibility of a fluctuation bound for $\mathcal{B} = L^1(X, \mu)$, so the question of a “quantitative L^1 mean ergodic theorem for amenable groups” remains unresolved.

What about the choice of acting group? Our assumptions on G (amenable and countable discrete or lsc) were selected because this is the most general class of groups which have Følner sequences. Our argument depends essentially on Følner sequences; indeed, proofs of ergodic theorems for actions of non-amenable groups have a qualitatively different structure. Remarkably, there are certain classes of non-amenable groups whose associated ergodic theorems have much stronger convergence behaviour than the classical ($G = \mathbb{Z}$) setting; for recent progress on quantitative ergodic theorems in the non-amenable setting, we refer the reader to the book and survey article of Gorodnik and Nevo [9, 10].

We should also mention quantitative bounds for *pointwise* ergodic theorems. For $G = \mathbb{Z}$ such results go as far back as Bishop's upcrossing inequality [3]. Inequalities of this type have also been found for $\mathbb{Z}^d$ by Kalikow and Weiss (for $F_n = [-n, n]^d$) [16]; more recently Moriakov has modified the Kalikow and Weiss argument to give an upcrossing inequality for symmetric ball averages in groups of polynomial growth [21], and (simultaneously with the preparation of this article) Gabor [8] has extended this strategy to give an upcrossing inequality for countable discrete amenable groups with Følner sequences satisfying a strengthening of Lindenstrauss's temperedness condition.

For both norm and pointwise convergence of ergodic averages, it is sometimes possible to deduce convergence behaviour which is stronger than ε -fluctuations and/or upcrossings but weaker than an explicit rate of convergence, namely that there exists an appropriate *variational inequality*. Jones et al. have succeeded in proving numerous variational inequalities, both for norm and

pointwise convergence, for a large class of Følner sequences in $\mathbb{Z}$ and $\mathbb{Z}^d$ [14, 15]. (In particular, the result proved in the present article is known not to be sharp when specialised to the case where $G = \mathbb{Z}^d$ and $\mathcal{B} = L^p$, due to these existing results.) However, their methods, which rely on a martingale comparison estimate and a Calderón-Zygmund decomposition argument, exploit numerous incidental geometric properties of $\mathbb{Z}^d$ which do not hold for many other groups. It would be interesting to determine which other groups enjoy similar variational inequalities.

ACKNOWLEDGEMENTS

The author thanks his advisor Jeremy Avigad, under whom this work was completed, for his guidance and support. The author also thanks Clinton Conley, Yves Cornuier, and Henry Towsner for helpful discussions. In addition, the author thanks the anonymous referee, whose comments led to a number of improvements to the exposition of the article.

REFERENCES

1. Jeremy Avigad and Jason Rute, *Oscillation and the mean ergodic theorem for uniformly convex Banach spaces*, Ergodic Theory and Dynamical Systems **35** (2015), no. 4, 1009–1027.
2. Garrett Birkhoff, *The mean ergodic theorem*, Duke Mathematical Journal **5** (1939), no. 1, 19, 20.
3. Errett Bishop, *A constructive ergodic theorem*, Journal of Mathematics and Mechanics **17** (1968), no. 7, 631–639.
4. Matteo Cavaleri, *Computability of Følner sets*, International Journal of Algebra and Computation **27** (2017), no. 07, 819–830.
5. ———, *Følner functions and the generic word problem for finitely generated amenable groups*, Journal of Algebra **511** (2018), 388–404.
6. James A. Clarkson, *Uniformly convex spaces*, Transactions of the American Mathematical Society **40** (1936), no. 3, 396–414.
7. Manfred Einsiedler and Thomas Ward, *Ergodic theory: with a view towards number theory*, vol. 259, Springer, 2010.
8. Uri Gabor, *Fluctuations of ergodic averages for amenable group actions*, arXiv preprint arXiv:1902.07912 (2019).
9. Alexander Gorodnik and Amos Nevo, *The ergodic theory of lattice subgroups*, Princeton University Press, 2009.
10. ———, *Quantitative ergodic theorems and their number-theoretic applications*, Bulletin of the American Mathematical Society **52** (2015), no. 1, 65–113.
11. Frederick P. Greenleaf, *Ergodic theorems and the construction of summing sequences in amenable locally compact groups*, Communications on Pure and Applied Mathematics **26** (1973), no. 1, 29–46.
12. Olof Hanner, *On the uniform convexity of L^p and l^p* , Arkiv för Matematik **3** (1956), no. 3, 239–244.
13. Tuomas Hytönen, Jan van Neerven, Mark Veraar, and Lutz Weis, *Analysis in Banach spaces: Volume I: Martingales and Littlewood-Paley theory*, vol. 63, Springer, 2016.
14. Roger L. Jones, Robert Kaufman, Joseph M. Rosenblatt, and Máté Wierdl, *Oscillation in ergodic theory*, Ergodic Theory and Dynamical Systems **18** (1998), no. 4, 889–935.
15. Roger L. Jones, Joseph M. Rosenblatt, and Máté Wierdl, *Oscillation in ergodic theory: higher dimensional results*, Israel Journal of Mathematics **135** (2003), no. 1, 1–27.
16. Steven Kalikow and Benjamin Weiss, *Fluctuations of ergodic averages*, Illinois Journal of Mathematics **43** (1999), no. 3, 480–488.
17. Ulrich Kohlenbach and Laurențiu Leuştean, *A quantitative mean ergodic theorem for uniformly convex Banach spaces*, Ergodic Theory and Dynamical Systems **29** (2009), no. 6, 1907–1915.
18. Ulrich Kohlenbach and Pavol Safarik, *Fluctuations, effective learnability and metastability in analysis*, Annals of Pure and Applied Logic **165** (2014), no. 1, 266–304.
19. Ulrich Krengel, *On the speed of convergence in the ergodic theorem*, Monatshefte für Mathematik **86** (1978), no. 1, 3–6.
20. Elon Lindenstrauss, *Pointwise theorems for amenable groups*, Inventiones Mathematicae **146** (2001), no. 2, 259–295.
21. Nikita Moriakov, *Fluctuations of ergodic averages for actions of groups of polynomial growth*, Studia Mathematica **240** (2018), no. 3, 255–273.
22. Nikita Moriakov, *On effective Birkhoff’s ergodic theorem for computable actions of amenable groups*, Theory of Computing Systems **62** (2018), no. 5, 1269–1287.
23. Donald S. Ornstein and Benjamin Weiss, *Entropy and isomorphism theorems for actions of amenable groups*, Journal d’Analyse Mathématique **48** (1987), no. 1, 1–141.

CARNEGIE MELLON UNIVERSITY, 5000 FORBES AVENUE, PITTSBURGH, PA 15213, USA
 Email address: awaren1@andrew.cmu.edu