Diploma Thesis

Ergodic Theorems on Amenable Groups

Felix Pogorzelski

Tübingen, September 7, 2010

supervised by

Prof. Dr. Rainer Nagel
Eberhard-Karls-Universität Tübingen

and

Prof. Dr. Daniel Lenz
Friedrich-Schiller-Universität Jena

Table of Contents

Danksagung

Zuerst danke ich meinen Eltern für ihre Unterstützung während meines ganzen Lebens. Nur durch ihr Vertrauen in mich sowie durch die mir überlassene Freiheit konnten meine akademischen Träume wahr werden.

Mein ganz besonderer Dank gilt meinem Betreuer Prof. Dr. Rainer Nagel, den ich nicht nur als großen Mathematiker, sondern auch als eine wunderbare Persönlichkeit schätzen gelernt habe. Seine Menschlichkeit, Geduld und Begeisterung habe ich stets als Quelle von Motivation und Energie für meine akademische Arbeit begriffen. Ich danke Prof. Nagel für die stetige Unterstützung meiner Studienpläne vom ersten Semester bis hin zu dieser Diplomarbeit.

Des Weiteren möchte ich mich herzlich bei Prof. Dr. Daniel Lenz bedanken, der mir dieses Thema vorgeschlagen und mich über mehrere Monate an der Universität Jena betreut hat. Dabei danke ich ihm vor allem für seine geduldige, motivierende und vor allem freundliche Art. In konstruktiven Diskussionen sowie durch die Gelegenheit, regelmäßig in der Arbeitsgruppe vorzutragen, habe ich durch Prof. Lenz sehr viel gelernt.

Für Korrekturen und für viele gute Vorschläge zu dieser Arbeit bedanke ich mich neben meinen Betreuern besonders bei Christina Botros, Dr. Tanja Eisner, Dr. Markus Haase, Matthias Keller, Marco Schreiber und Pavel Zorin-Kranich. Ohne ihre konstruktive Kritik wäre der Leser wohl an so mancher Stelle verzweifelt.

Allen Mitgliedern der Arbeitsgruppe Funktionalanalysis (AGFA) in Tübingen danke ich für viele mathematische Vorträge und Diskussionen, für gemeinsame Unternehmungen und für das gesellige Miteinander. Ich habe mich dort stets sehr wohl gefühlt.

In diesem Zusammenhang danke ich ebenso der Arbeitsgruppe Analysis an der Universität Jena, die mich sehr freundlich aufgenommen und mir in entspannter Atmosphäre viele neue Einblicke in die Welt der Graphen, sowie der mathematischen Physik gewährt hat.

Ein wichtiger Dank geht an alle Freunde, die mich im Mathematikstudium begleitet haben. Ob Kaffee trinken, Mittag essen, diskutieren, Vorlesungen nacharbeiten oder Übungsblätter bearbeiten - in der Gemeinschaft mit Chris, Dhia, Kaspar, Marco, Matthias, Matze, Pirmin, Sebastian, Stefan und Steffi hat es doppelt Spaß gemacht. Besonders möchte ich mich bei Markus Klein bedanken, der seine getexten Vorlesungsmitschriebe der Allgemeinheit zur Verfügung stellte und mir bei Fragen zu LaTeX immer zur Seite stand.

Für ihre Freundschaft in den letzten Jahren danke ich abschließend von ganzem Herzen Christoph, Julia, Matthias und Yvonne + Matthias. Sie haben mich an verschiedenen Stellen auf meinem Weg begleitet und mir dabei geholfen, mehr ich selbst zu werden.

Erklärung

Hiermit erkläre ich, die vorliegende Arbeit selbstständig verfasst zu haben und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt zu haben.

Tübingen, den 9. September 2010

Ort, Datum

Felix Pogorzelski

Zusammenfassung in deutscher Sprache

Die Ergodentheorie hat ihren Ursprung in der *Statistischen Mechanik*, welche das thermodynamische Verhalten physikalischer Systeme auf mikroskopischer Ebene beschreibt. Ausgehend von einschneidenden mathematischen Durchbrüchen in den 30er Jahren hat sich daraus im Laufe des 20. Jahrhunderts eine eigenständige mathematische Disziplin gebildet. Diese moderne Theorie verbindet Funktional- und Fourieranalysis, Wahrscheinlichkeitstheorie sowie Quantenmechanik und liefert mächtige Werkzeuge zur Lösung schwieriger Fragestellungen in vielen mathematischen Bereichen, darunter z.B. Zahlentheorie und Harmonische Analysis.

In der klassischen Ergodentheorie untersucht man sog. dynamische Systeme. In dem betrachteten Kontext verstehen wir darunter ein Tupel (X, φ), wobei $(X, \mathcal{B}, \mu)$ einen Wahrscheinlichkeitsraum darstellt und $\varphi : X \to X$ eine messbare Transformation beschreibt, welche das Maß μ erhält, d.h. $\mu(\varphi^{-1}A) = \mu(A)$ für alle $A \in \mathcal{B}$.
Für festes $1 \leq p < \infty$ induziert φ durch die Vorschrift $Tf := f \circ \varphi$ eine positive Kontraktion auf $L^p(X)$. Der diskrete zeitliche Durchschnitt eines Elementes $f \in L^p(X)$ bis zum Zeitpunkt $t = (N-1)$, $N \in \mathbb{N}$ ist dann durch das Cesàro Mittel $A_N := N^{-1} \sum_{n=0}^{N-1} T^n f$ gegeben. Eine natürliche Frage ist die nach dem Langzeitverhalten dieser Mittel. Eine erste Antwort geht dabei auf den Ungarn John von Neumann zurück, der für $f \in L^2(X)$ zeigte, dass die Ausdrücke $A_N f$ in der Norm konvergieren *(Mittelergodentheorem)*. Durch dieses Resultat inspiriert, gelang dem Amerikaner George David Birkhoff der Nachweis der punktweisen fast sicheren Konvergenz auf $L^1(X)$ *(Individuelles Ergodentheorem)*. In dieser Arbeit betrachten wir anstelle von φ maßerhaltende Wirkungen von amenablen (mittelbaren) Gruppen auf σ-endlichen Maßräumen.
In Kapitel 2 geben wir eine kurze Einführung in die Theorie mittelbarer Gruppen und stellen unter anderem fest, dass alle kompakten, alle abelschen und alle auflösbaren Gruppen zu dieser Klasse gehören. Falls die Gruppe G zusätzlich σ-kompakt ist, so wird ihre Mittelbarkeit charakterisiert durch die Existenz einer Folge $\{F_n\}$ kompakter Teilmengen aus G, welche sich durch spezielle Invarianzeigenschaften auszeichnen *(Følnerfolge)*. Details und Beispiele sind in Kapitel 3 zu finden. Mittels Følnerfolgen lassen sich abstrakte ergodische Mittel $A_N f$ durch

$$A_N f(x) := m_L(F_N)^{-1} \int_{F_N} f(g \cdot x)\, dm_L(g), \quad f \in L^p(X),$$

definieren, wobei $m_L(\cdot)$ das Linkshaarmaß auf G und $g \cdot x$ die Wirkung des Elements $g \in G$ auf $x \in X$ bezeichnet.

Im Folgenden werden die Konvergenzeigenschaften der $A_N f$ ausführlich untersucht. Als Höhepunkt der Arbeit beweisen wir mittels eines Transferenzprinzips eine Erweiterung von Elon Lindenstrauss' (Fields Medaillen Preisträger 2010 für Arbeiten über Maßrigidität in der Ergodentheorie und deren Anwendungen in der Zahlentheorie) gefeiertem Individuellen Ergodentheorem für fast alle σ-kompakten, mittelbaren Gruppen. Während Lindenstrauss nur den L^1-Fall auf Wahrscheinlichkeitsräumen betrachtet, weiten wir dieses Resultat auf alle L^p-Räume ($1 \leq p < \infty$) über σ-endlichen Maßräumen aus (siehe Korollar 8.8). Der Fall $p = \infty$ für unendliche Maßräume kann im Rahmen dieser Arbeit nicht behandelt werden, da hier zentrale Argumente wie Kompaktheitskriterien (siehe Kapitel 4) und bestimmte Ungleichungen (siehe z.B. Kapitel 6 und 7) nicht mehr gelten müssen.

Wir gehen in vier Schritten vor.

- In Kapitel 4 weisen wir die Normkonvergenz der abstrakten ergodischen Mittel durch einen allgemeinen Mittelergodensatz nach. Wir erhalten dabei eine Zerlegung des betrachteten Banachraumes in die direkte Summe zweier abgeschlossener Teilräume. Für den Beweis der punktweisen Konvergenz werden diese im Folgenden separat betrachtet.
- Wir erklären das Konzept einer L^p-Maximalungleichung in Kapitel 5. Mit der mittelergodischen Zerlegung beweisen wir das punktweise Ergodentheorem mit elementaren Rechnungen auf einem dichten Teilraum von $L^p(X)$. Wir zeigen ferner, dass wir die Maximalungleichung nutzen können, um das Konvergenzresultat auf den Abschluss (also auf den ganzen Raum) auszuweiten.
- Im folgenden Kapitel 6 wird das bereits angesprochene Transferenzprinzip erläutert. Wir zeigen, dass sog. Transferenzungleichungen hinreichend für L^p-Maximalungleichungen sind.
- Schließlich beweisen wir unter Ausnutzung der intrinsischen Gruppenstruktur die Gültigkeit der Transferenzungleichungen für alle $1 \leq p < \infty$. Dabei stellen wir zunächst die deterministischen, kombinatorischen Argumente von Weiss für abzählbare mittelbare Grupen vor (Kapitel 7). Danach werden Lindenstrauss' stochastische Methoden erläutert, welche letztlich die Transferenzungleichungen für die betrachteten σ-kompakten amenablen Gruppen liefern (Kapitel 8).

Abschließend erwähnen wir in Kapitel 9 mögliche Anwendungen des Individuellen Ergodentheorems.

1 Introduction

Ergodic theory has its origin in *statistical thermodynamics* which describes the thermodynamic behaviour of physical systems involving a large number of small particles. By the end of the 19^{th} century, one of the originators of *statistical mechanics*, the Austrian physicist LUDWIG BOLTZMANN (1844-1906) formulated the so-called *ergodic hypothesis*. It states that given a 'typical' transformation φ on a physical state space X (e.g. the motion of particles in a box filled with an ideal gas described by the change of their space and velocity coordinates), the trajectory (orbit) of each initial state hits every possible state of the system as time progresses. These systems were called *ergodic* by BOLTZMANN. However, it turned out that such 'ergodic systems' do not exist at all (cf. [35]). Consequently, the ergodic hypothesis was weakened and one conjectured that the orbit of each initial state gets 'close' to every element in the state space (*quasi ergodic hypothesis*). But what does 'close' mean in this setting? By the turn of the century, strong efforts were put into the development of a precise theory. At the beginning of the 1930s, fundamental contributions were made by mathematicians such as JOHN VON NEUMANN (1903-1957), NORBERT WIENER (1894-1964) and GEORGE DAVID BIRKHOFF (1884-1944). Linking probability theory, functional and Fourier analysis as well as quantum mechanics, the ergodic theory was born as a new mathematical discipline. Today, it provides powerful tools to approach difficult problems in various areas such as Lie theory, number theory or harmonic analysis.

In classical ergodic theory one studies the state space and the transformation in a *dynamical system* (X, φ) which is supposed to run for a long time. In topological cases one often assumes that X is a compact Hausdorff space and that the map $\varphi : X \to X$ is a homeomorphism. For the purpose of this thesis, we restrict ourselves to the class of so-called *measure dynamical systems* given by a probability space $(X, \mathcal{B}, \mu)$ and by some transformation $\varphi : X \to X$ preserving the measure, i.e. $\mu(\varphi^{-1}A) = \mu(A)$ for all $A \in \mathcal{B}$. A transformation φ is then said to be *ergodic* if all φ-invariant measurable sets ($\varphi^{-1}A = A$) have either zero or full measure.

Fixing some $1 \leq p < \infty$, the transformation φ induces a positive contraction on $L^p(X)$ via $Tf := f \circ \varphi$ with $T\mathbb{1} = \mathbb{1}$, where $\mathbb{1}$ denotes the constant one-function on X. In this context one can interpret some $f \in L^p(X)$ as a map which assigns some physically observable number to almost every state of the space X (returning to the gas box system one might e.g. think of the kinetic energy). Hence, the measurement of the key number at time $n \in \mathbb{N}$ gives the value $T^n f(x) = f \circ \varphi^n(x)$ for the initial state $x \in X$. Thus, measuring at each point of time from $t = 0$ up to $t = N - 1$, $N \in \mathbb{N}$, the (unweighted) time average of the measurements will be denoted by the Cesàro means $A_N f := (1/N) \sum_{n=0}^{N-1} T^n f$. So the natural follow-up question is about the long term behaviour of the system, i.e. to understand what happens to the terms A_N as N tends to infinity. In 1931, VON NEUMANN gave a first answer by proving his *mean ergodic theorem* in the L^2-case. It states that the Cesàro means $A_N f$ converge to some almost surely φ-invariant limit f^* in $L^2(X)$-norm. Inspired by this result, Birkhoff managed by the end of the same year to prove the *pointwise ergodic theorem*, likewise called *individual ergodic theorem* by showing that for each $f \in L^1(X)$, the $A_N f$ converge to f^* pointwise almost surely. In particular, if the transformation φ is ergodic, then $f^*(x) = \int_X f \, d\mu$ for almost every $x \in X$ by the invariance of f^*. It follows that the trajectory $\{\varphi^n(x)\}_n$ returns to *every* set of positive measure infinitely often for almost every initial state x and therefore, a precise version of the quasi ergodic hypothesis was confirmed. While von Neumann's mean ergodic theorem (which also proves the quasi ergodic hypothesis) came out in January 1932 (see [42]), Birkhoff managed to publish his individual ergodic theorem already in December 1931 (see [2]). Historic facts even indicate that Birkhoff used his strong academic influence in order to delay the paper of the young von Neumann (cf. [30]). Although the latter kept on pretending not to care too much, the relation between these outstanding mathematicians was always affected by tensions.

During the 20$^{\text{th}}$ century, many generalizations of both ergodic theorems have been found. As shown in [19] by Hopf, the induced operator can e.g. be replaced by a positive isometry on $L^1(X)$ with the additional property that $T\mathbb{1} = \mathbb{1}$. In 1965, Garsia gave a far simpler proof of this fact in [15] which also lays the basis for many modern pointwise ergodic results. In light of that, Akcoglu significantly extended the work of Ionescu-Tulcea (cf. [22]) by showing that the pointwise ergodic theorem also holds in the case of a σ-finite measure space and for each positive contraction T on the corresponding reflexive L^p-spaces (cf. [1]). Moreover, convergence theorems for modified ergodic averages have been proven, see e.g. [20].

In this thesis, we devote ourselves to the case when the dynamics on the measure space is induced by actions of so-called *amenable groups* allowing for the definition of abstract ergodic averages. This class is large enough to comprise e.g. all compact as well as all abelian groups. As a main issue we examine the convergence properties of these averages. This has been done before by calderon for a rather narrow class of groups in [4] and by Emerson for a large class of σ-compact amenable groups in [11]. As a highlight, we give an extension of Lindenstrauss' (Fields Medalist 2010 for his work on measure rigidity in ergodic theory and its applications to number theory) celebrated general pointwise ergodic theorem for second countable amenable groups. While this result was originally stated for integrable functions on a probability space (see [29]), we use the *transfer principle* to show the validity of the theorem for all p-integrable functions on a σ-finite measure space (see Chapter 8). We will not be able to draw conclusions if $p = \infty$ on an infinite measure space because essential arguments such as compactness criteria (cf. Chapter 4) and certain inequalities (cf. Chapter 6 and Chapter 7) will fail in this case.

We proceed as follows. In Chapter 2, we explain in detail the notion of an amenable group and give significant examples. Next, we draw our attention to σ-compact groups, which are amenable if and only if there is a so-called *Følner sequence* of compact subsets of the group with certain invariance properties. A Følner sequence is the crucial ingredient to define abstract ergodic averages.
In Chapter 3, we give examples for such sequences and introduce growth conditions which will become important for pointwise ergodic theorems.

The Chapters 4 to 8 are devoted to the presentation of a standard technique for proving pointwise ergodic theorems. This procedure is divided into four parts.

(1) **Mean ergodicity**. One starts by proving a mean ergodic theorem. In Chapter 4, we amend an abstract version which is due to GREENLEAF in 1973 (cf. [18]). One major result will be that mean ergodicity (defined as the strong convergence of the ergodic averages) is equivalent to the direct decomposition of the Banach space into the direct sum of two closed subspaces with specific properties. For the proof of the individual ergodic theorem, it is convenient to treat these spaces separately.

(2) **Maximal inequality**. We explain the concept of an L^p*-maximal inequality* in Chapter 5. It turns out that using the mean ergodic decomposition in the L^p-case, the pointwise almost everywhere convergence of the abstract ergodic averages is obtained by elementary calculations on a dense subspace of $L^p(X)$. The maximal inequality will provide the tool needed to extend the convergence result to the closure.

(3) **Transfer principle**. In Chapter 6, we describe the *transfer principle* which gives a sufficient condition for the validity of an L^p-maximal inequality in form of a *transfer inequality*. The main conclusion here will be the fact that one can turn away from the action of the group on the measure space and consider instead the canonical action of the group on itself.

(4) **Transfer inequality**. Finally, we use combinatorial arguments and the intrinsic structure of the group to prove the transfer inequalities for all $1 \leq p < \infty$. We present a method of WEISS for countable amenable groups in Chapter 7. As already mentioned, using LINDENSTRAUSS' ideas, we treat the case of σ-compact amenable groups in Chapter 8 and prove the pointwise ergodic theorem in a rather general setting.

In a short outlook (Chapter 9), we give a brief outline of possible applications of the pointwise ergodic theorem as well as of the decompositions presented in Chapter 8.

2 Amenability

The following discussion of amenable groups is mainly based on the first two chapters of [17]. Deviations are labeled separately.
In this chapter, we introduce the notion of amenability for locally compact groups. In light of that, we verify the existence of so-called *left-invariant means* on (left-)invariant function spaces (cf. Definition 2.2) over the group. We will see that the property of amenability as specified in Definition 2.6 allows some flexibility concerning the invariance properties (cf. Lemma 2.3) as well as the domain of definition (cf. Theorem 2.5) of the mean. We present examples of amenable and non-amenable groups in the example Sections 2.4 and 2.7. Further, we examine specific nets of the group with weak and strong invariance properties as determined in Definition 2.9 and show that their existence is equivalent to amenability of the group (cf. Theorems 2.11 and 2.10, Corollary 2.12).

Definition 2.1
Let G be an arbitrary set and let Y be a closed subspace of $B(G)$, the space of all bounded, complex-valued functions on G, equipped with the sup-norm $\|\cdot\|_\infty$. Assume further that Y contains the constant functions and is closed under complex conjugation $z \mapsto \overline{z}$.
Then a positive linear functional m on Y is said to be a **mean** *if*

(i) $m(\overline{f}) = \overline{m(f)}$ *for all* $f \in Y$ *and*

(ii) $m(\mathbb{1}) = 1$,

where $\mathbb{1} := \mathbb{1}_G$ *denotes the constant one-function on* G.

It follows that $\|m\| := \sup_{f\in B(G),\,\|f\|_\infty=1} |m(f)| = 1$. By the Banach-Alaoglu Theorem, the closed set of means on Y forms a weak$*$-compact, convex set $\Sigma \in Y^*$, where we denote by Y^* the dual space of Y. If G is a group, one can define (left-)invariant function spaces and (left-)invariant means.

Definition 2.2
Let G be a group and $Y \subseteq B(G)$ be a closed subspace of $B(G)$. Then Y is called **left-invariant** *if for all $g \in G$ and for all $f \in Y$, $L_g f := f(g^{-1}\cdot) \in Y$.*
We say that m is a **left-invariant mean** *on the left-invariant space Y if m is a mean on Y and $m(L_g f) = m(f)$ for all $g \in G$ and for all $f \in Y$.*
The operator L_g is called left translation by $g \in G$.

Right-invariant function spaces are defined in an analogous manner. Note that the right translation by some $g \in G$ is denoted by $R_g(f) := f(\cdot g)$. Consequently, a space Y is called invariant if it is both left- and right-invariant. An invariant mean m is a mean on an invariant function space Y which is invariant under L_g and R_g for all $g \in G$.

Lemma 2.3
Let G be a group and Y be a left-invariant subspace of $B(G)$. Then the following statements are equivalent:

(i) There is a left-invariant mean on Y.

(ii) There is a right-invariant mean on $\tilde{Y} := \{f(\cdot^{-1}) \mid f \in Y\}$.

PROOF
Given a left-invariant mean m on Y, define

$$\tilde{m} : \tilde{Y} \to \mathbb{C} : \tilde{m}(f) := m(I(f)),$$

where $I(f) := f(\cdot^{-1})$. We see that

$$R_g[I(f)] = f(\cdot^{-1}g) = f((g^{-1}\cdot)^{-1}) = I[L_g(f)]$$

and thus

$$\begin{aligned} \tilde{m}(R_g[I(f)]) &= \tilde{m}(I[L_g(f)]) \\ &\overset{\text{Def.}}{=} m(L_g f) = m(f) \overset{\text{Def.}}{=} \tilde{m}(I(f)) \end{aligned}$$

for each $g \in G$ and every $f \in Y$. □

The existence of (left-)invariant means on $Y \subseteq B(G)$ cannot be confirmed for all groups as the following example shows.

Example 2.4
Let $G = \mathbb{F}_2$ be the free group on two generators a, b and assume that m is a left-invariant mean on $B(G)$. Then one can define a left-invariant, finitely additive function μ on the powerset of $\mathbb{F}_2$ by $\mu(A) = m(\mathbb{1}_A)$ for $A \subseteq \mathbb{F}_2$.
By definition, we can decompose $\mathbb{F}_2$ into disjoint subsets $\{H_i \mid i \in \mathbb{Z}\}$ with H_i containing exactly the elements expressed as a reduced word by

$$g = a^i b^{i_1} a^{i_2} \ldots, \; i_1 \neq 0 \text{ if } g \neq a^i.$$

Then the mappings $\lambda_a : H_i \to H_{i+1} : g \mapsto ag$, $(i \in \mathbb{Z})$ and $\lambda_b : H_i \to H_0 : g \mapsto bg$, $(i \in \mathbb{Z} \setminus \{0\})$ are well defined. Since μ is left-invariant and λ_a is bijective, we must have $\mu(H_i) = \mu(H_{i+1})$ for all integers. By the fact that m is a mean, $\mu(G) = 1$ and by disjointness of the H_i we conclude that $\mu(H_i) = 0$ for all $i \in \mathbb{Z}$. Now λ_b maps every set H_i with $i \neq 0$ into H_0, so that $\mu(H_0) \geq \mu(\cup_{i\neq 0} H_i)$. But since μ is finitely additive, one obtains $\mu(H_0) + \mu(\cup_{i\neq 0} H_i) = \mu(G) = 1$ and by the preceding inequality this implies $\mu(H_0) \geq 1/2$. This is a contradiction to our previous result stating that $\mu(H_i) = 0$ for all $i \in \mathbb{Z}$. Hence, such a mean m cannot exist.

So far, we have restricted our attention to means on $B(G)$. In the following we take G to be a locally compact group, i.e. a topological Hausdorff group with the property that each point in G has a compact neighborhood.

Then the space $B(G)$ does not reflect the topological structure of G. We now consider subspaces Y which take into account the topology of the group. A good choice for an appropriate space Y is $CB(G)$, the space of all bounded, *continuous* functions on G.
A function $f : G \to \mathbb{C}$ is called right uniformly continuous if for some given $\varepsilon > 0$, there is a neighborhood $U(\varepsilon)$ of the unit in G such that $|L_h f(g) - f(g)| < \varepsilon$ for all $g \in G$ and all $h \in U(\varepsilon)$. So one can examine (left-)invariant means on $Y = UCB_r(G)$, the ($\| \cdot \|_\infty$-closed) space of all right uniformly continuous functions.
Another reasonable space is $L^\infty(G, m_L)$, the collection of all m_L-essentially bounded functions (equivalence classes) on G with respect to m_L, where m_L is the left Haar measure on G (see [9], Section 1.3).
In fact, the existence of some left-invariant mean on one of these three spaces implies the existence of a left-invariant mean on all of the others.

Theorem 2.5
The following statements are equivalent:

(i) There is a left-invariant mean on $CB(G)$.

(ii) There is a left-invariant mean on $UCB_r(G)$.

(iii) There is a left-invariant mean on $L^\infty(G, m_L)$.

PROOF
See [17], Theorem 2.2.1. □

In the following, we simply write $L^\infty(G)$ instead of $L^\infty(G, m_L)$. Let us introduce the notion of amenability for locally compact groups.

Definition 2.6
A locally compact group G is said to be **amenable** *if there exists a left-invariant mean m on $L^\infty(G)$, i.e. m is a mean and for the operator L_g from above, considered on $L^\infty(G)$, we have*

$$m(L_g f) = m(f)$$

for all $g \in G$ and all $f \in L^\infty(G)$.

We give some examples and refer to certain stability conditions for amenability such as inheritance on closed subgroups (cf. [37], Theorem 1.2.7) and division by closed normal subgroups (cf. [37], Theorem 1.2.10).

Examples 2.7 (cf. [37], Examples 1.1.5 and 1.2.11)

(1) Each compact group is amenable. Since the Haar measure m_L is a Radon measure, we have $m_L(G) < \infty$ and thus $L^\infty(G) \subseteq L^1(G)$. By the left-invariance of the Haar measure,

$$m : L^\infty(G) \to \mathbb{C} : m(f) = \frac{1}{|G|} \int_G f(g) dm_L(g)$$

is a left-invariant positive functional with $m(\mathbb{1}) = 1$, hence a left-invariant mean.

(2) Each locally compact, abelian group is amenable. As mentioned above, the set Σ of all means on $L^\infty(G)$ is weak$*$-compact and convex. For each $g \in G$, define a linear map

$$T_g : L^\infty(G)^* \to L^\infty(G)^* : m \mapsto T_g m$$

through

$$\langle f, T_g m\rangle_{L^\infty,(L^\infty)^*} = \langle L_g f, m\rangle_{L^\infty,(L^\infty)^*}.$$

If $m \in \Sigma$, then $T_g m$ is evidently a mean on $L^\infty(G)$, i.e. $T_g(\Sigma) \subseteq \Sigma$ for all $g \in G$. Further, the maps T_g are weak$*$-continuous. Since G is abelian, we can apply the Markov-Kakutani fixed-point theorem ([10], Theorem V.10.6). Hence there must be some $m^* \in \Sigma$ such that $T_g m^* = m^*$ for all $g \in G$ and we conclude that m^* is an invariant mean on $L^\infty(G)$.

(3) A group G is called solvable if there are normal subgroups $N_0, N_1, \ldots, N_n$ of G with

$$\{1\} = N_0 \subseteq N_1 \subseteq \cdots \subseteq N_n = G$$

and N_j/N_{j-1} is abelian for $1 \leq j \leq n$. Solvable groups are amenable if we endow such a group G with the discrete topology. To see this, note that the property of being solvable does not depend on the topology and use the fact that amenability is compatible with division by closed normal subgroups, i.e. if N is such a subgroup and both N and G/N are amenable, so is G (cf. [37], Theorem 1.2.10). Combining this result with example (2), we observe that discrete solvable groups are amenable. Hence, there is some left-invariant mean m on $B(G)$. If we examine other topologies on G, we have that $CB(G) \subseteq B(G)$ and we conclude that the restriction of m is a left-invariant mean on $CB(G)$. Therefore, all solvable locally compact groups are amenable.
We present two concrete examples.

(a) We denote by H the *Heisenberg* group, defined as

$$H := \left\{ \begin{pmatrix} 1 & x & y \\ 0 & 1 & z \\ 0 & 0 & 1 \end{pmatrix} \middle| x, y, z \in \mathbb{R} \right\}$$

with the matrix multiplication as group operation. Put $N_0 = \{I_3\}, N_1 = [H, H]$ and $N_2 = H$, where I_3 is the unit matrix in $\mathbb{R}^{3\times 3}$ and $[H, H]$ is the commutator subgroup of H. It is easily verified that these normal subgroups of H satisfy the requirements for solvability of H. Therefore, the Heisenberg group is amenable.

(b) The *Lamplighter* group is denoted by

$$G := \mathbb{Z} \wr \mathbb{Z}_2 := \{(m, a) \,|\, m \in \mathbb{Z},\, a \in \bigoplus_{i\in\mathbb{Z}} \mathbb{Z}_2\},$$

where $\mathbb{Z}_2$ is the cyclic group of order two. Thus, we have $G = \{(m, a) \,|\, m \in \mathbb{Z}, a \in \mathbb{Z}_2^{\mathbb{Z}}\}$ as a set. With σ as the left shift on the space of all 0-1-sequences over $\mathbb{Z}$ (i.e. $\sigma((x_n)_n) = (x_{n+1})_n$), the group operation is

$$(m, a) \cdot (n, b) := (m + n, \sigma^n a + b).$$

This group is solvable as the wreath product of two solvable groups and hence amenable. For a more detailed discussion of wreath products and its properties, see [36], p. 172 ff.

(4) As we have seen before, the free group $\mathbb{F}_2$ on two generators is not amenable.

(5) Using the theory of linear fractional transformations (see [6], III.3.) one can show that the groups $SL(n, \mathbb{C})$ (all complex $(n \times n)$-matrices with determinant equal to one) and $GL(n, \mathbb{C})$ (all complex, invertible $(n \times n)$-matrices) with their natural locally compact topologies are not amenable for $n \geq 2$.
To see this, note first that for every matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL(2, \mathbb{C}),$$

there is a map

$$h_A : \mathbb{C} \cup \{\infty\} \to \mathbb{C} \cup \{\infty\} : z \mapsto \frac{az + b}{cz + d},$$

where we put $h_A(\infty) := a/c$ and $h_A(-d/c) := \infty$. These maps h_A form a group $\mathcal{G}$ with composition as group multiplication. The non-amenability of the group $SL(n, \mathbb{C})$ can now be proven as follows (cf. [37], Exercise 1.2.6).
Observing that the assignment

$$\mathcal{M} : GL(2, \mathbb{C}) \to \mathcal{G} : A \mapsto h_A$$

is a group homomorphism, one defines the group $PSL(2, \mathbb{R})$ as the image of $SL(2, \mathbb{R})$ under $\mathcal{M}$. Further, one can show $PSL(2, \mathbb{R}) \cong SL(2, \mathbb{R})/\approx$, where $A \approx B$ if and only if $A = -B$ or $A = B$ for $A, B \in SL(n, \mathbb{R})$. With the quotient topology (the finest topology making the canonical map $SL(2, \mathbb{R}) \to PSL(2, \mathbb{R})$ continuous), $PSL(2, \mathbb{R})$ becomes a topological group.
Using the theory of fractional linear transformations, we can find elements $h_1, h_2 \in PSL(2, \mathbb{R})$ generating a subgroup which is isomorphic to $\mathbb{F}_2$. Hence, by Theorem 1.2.7 of [37], the group $PSL(2, \mathbb{R})$ cannot be amenable. It follows that none of the groups $SL(2, \mathbb{R})$, $SL(2, \mathbb{C})$, $GL(2, \mathbb{R})$ and $GL(2, \mathbb{C})$ is amenable. Interpreting these elements as subgroups of the corresponding higher-dimensional type, we note that the groups $SL(n, \mathbb{R})$, $GL(n, \mathbb{R})$, $SL(n, \mathbb{C})$ as well as $GL(2, \mathbb{C})$ are not amenable for all $n \geq 2$.

(6) The matrix groups $SL(2, \mathbb{R})$ and $GL(2, \mathbb{R})$, endowed with the discrete topology, contain a subgroup which is isomorphic to $\mathbb{F}_2$. As a consequence of Theorem 1.2.7 in [37], they are not amenable.

In the following, we look at the functions

$$P(G) := \{f \in L^1(G) \mid f \geq 0, \|f\|_{L^1} = 1\}.$$

One can consider the elements in $P(G)$ as objects in $(L^1(G)^*)^* = (L^\infty(G))^*$. Since $\|f\|_{L^1(G)} = 1$ for $f \in P(G)$, they can be identified with means on $L^\infty(G)$.

Proposition 2.8
The set $P(G)$ is weak$*$-dense in Σ, the collection of all means on $L^\infty(G)$.

Proof
Assume that $\overline{P(G)}^* \neq \Sigma$. As $\overline{P(G)}^*$ is a weak$*$-closed, convex set in Σ we can then find some $m_0 \in \Sigma \setminus \overline{P(G)}^*$. By the Hahn-Banach Separation Theorem (cf. [44], Theorem VIII 2.12), there is some $f_0 \in (L^\infty(G)^*_{\sigma((L^\infty)^*,L^\infty)})^*$ such that

$$\text{Re } m_0(f_0) > 1 \geq \text{Re } m(f_0)$$

for all $m \in \overline{P(G)}^*$. With the fact that $f_0 \in (L^\infty(G)^*_{\sigma((L^\infty)^*,L^\infty)})^* = L^\infty(G)$ (cf. [44], Corollary VIII 3.4), this contradicts the definition of a mean: the first inequality implies $\|f_0\|_{L^\infty} > 1$, i.e. there is some measurable set $M \subseteq G$ of positive finite (Haar)measure where f_0 takes values greater than one. Obviously, the function $|M|^{-1} \cdot \mathbb{1}_M$ is in $P(G)$ and by identification, defines a mean m_M on $L^\infty(G)$. But by choice of the set M, we have $m_M(f_0) > 1$, which is a contradiction to the second inequality. □

Definition 2.9
A net $\{\Phi_\alpha\} \subset P(G)$ is **weakly (strongly) convergent to left-invariance** *if $L_g\Phi_\alpha - \Phi_\alpha \to 0$ weak$*$ in $L^\infty(G)$ (in $\|\cdot\|_{L^1}$-norm) for each $g \in G$.*

In 1957, Day established in [7] the following link between the set $P(G)$ and amenability for discrete groups. This result was extended in 1966 to the general case by Hulanicki (cf. [21]).

Theorem 2.10
There is a net in $P(G)$ which is weakly convergent to left-invariance if and only if G is amenable.

Proof
If the net $\{\Phi_\alpha\}$ converges weakly to left-invariance, its elements, identified with means on $L^\infty(G)$, lie in the weak$*$-compact set Σ of all means. So one can extract a subnet which we also call $\{\Phi_\alpha\}$ that is weak$*$ convergent to some mean $m \in \Sigma$. For this mean, one computes for each $g \in G$ and each $f \in L^\infty(G)$ that

$$\begin{aligned} m(L_g f) - m(f) &= \lim_\alpha [\Phi_\alpha(L_g f) - \Phi_\alpha(f)] \\ &= \lim_\alpha \langle L_g f - f, \Phi_\alpha \rangle \\ &= \lim_\alpha \langle f, L_{g^{-1}}\Phi_\alpha - \Phi_\alpha \rangle = 0, \end{aligned}$$

where $\langle \cdot, \cdot \rangle$ stands for the natural pairing of $L^\infty(G)$ with $L^\infty(G)^*$.
Conversely, if $m \in \Sigma$ is left-invariant, the weak$*$-density of $P(G)$ in Σ (cf. Proposition 2.8) insures that by identification of L^1-functions with means we can find a weak$*$-convergent net $\{\Phi_\alpha\} \subseteq P(G)$ such that $\Phi_\alpha \to m$. We have to show that this net is weakly convergent to left invariance. Indeed, for each $f \in L^\infty(G)$ and every $g \in G$, we have $m(L_g f) = m(f)$,

which implies

$$\begin{aligned}\langle f, L_{g^{-1}}\Phi_\alpha - \Phi_\alpha\rangle &= \langle f, L_{g^{-1}}\Phi_\alpha\rangle - \langle f, \Phi_\alpha\rangle \\ &= \langle L_g f, \Phi_\alpha\rangle - \langle f, \Phi_\alpha\rangle \\ &= \langle L_g f, \Phi_\alpha\rangle - \langle L_g f, m\rangle + \langle f, m\rangle - \langle f, \Phi_\alpha\rangle \\ &\to 0 + 0 = 0,\end{aligned}$$

where again $\langle\cdot,\cdot\rangle$ denotes the natural pairing of $L^\infty(G)$ with $L^\infty(G)^*$. □

It follows from the first part of the proof that each weak∗-limit point of $\{\Phi_\alpha\}$ is a left-invariant mean. In general, these limit points are far from being unique (cf. [17], p. 4).

Surprisingly, the following is true.

Theorem 2.11
The following statements are equivalent:

(i) There is a net in $P(G)$ weakly convergent to left-invariance.

(ii) There is a net in $P(G)$ strongly convergent to left-invariance.

Proof
Let $\{\Phi_\alpha\} \subseteq P(G)$ be weakly convergent to left-invariance. For every $g \in G$, take a copy of $L^1(G)$ and construct the locally convex product space $E := \prod_{g\in G} L^1(G)$ carrying the canonical product (norm-)topology. Further, we define the linear map

$$T : L^1(G) \to E : (Tf)_g := L_g f - f.$$

Now the $\sigma(E, E^*)$-topology coincides with the product of the $\sigma(L^1(G), L^\infty(G))$-topologies (cf. [23], Chapter 5, Section 17.13 (iii)).
Since by assumption $L_g\Phi_\alpha - \Phi_\alpha \to 0$ weak∗ in $L^\infty(G)^*$ for each $g \in G$, we conclude that

$$\langle f, L_g\Phi_\alpha\rangle_{L^\infty,(L^\infty)^*} - \langle f, \Phi_\alpha\rangle_{L^\infty,(L^\infty)^*} \to 0,$$

whenever $f \in L^\infty(G)$. We can interpret this convergence as $L_g\Phi_\alpha - \Phi_\alpha \to 0$ weakly in $L^1(G)$ ($\sigma(L^1(G), L^\infty(G))$-topology) for every $g \in G$. In light of that, the zero element $(0)_E$ of E is contained in $\overline{T(P(G))} \subseteq E$, where the closure is taken with respect to the weak topology on E. Since E is locally convex and the set $T(P(G))$ is convex, its weak and strong closure must coincide (cf. e.g. [23], Chapter 5, Section 17.1). Hence there is some net $\{\Psi_\alpha\}$ in $P(G)$ such that $T(\Psi_\alpha) \to (0)_E$ strongly, which means by the definition of the topology on E that

$$\lim_\alpha \|L_g\Psi_\alpha - \Psi_\alpha\|_{L^1(G)} = 0$$

for all $g \in G$. This finishes the proof.
The converse direction is trivial. □

Corollary 2.12
A locally compact group G is amenable if and only if there is a net $\{\Phi_\alpha\} \subseteq P(G)$ that converges strongly to left invariance.

Proof
This is a direct consequence of the preceding Theorems 2.10 and 2.11. □

3 Følner conditions

In the following, we present an important characterization of amenability for σ-compact amenable groups (see Definition 3.3). We start with a simple example motivating the main theorem (3.5) of this chapter. It shows that for σ-compact groups, amenability is equivalent to the existence of a so-called *Følner sequence* which can be seen as a special case of a net converging strongly to left-invariance. Examples for Følner sequences are given in Section 3.7. Further, we draw our attention to sequences with specific growth restrictions and introduce the *Tempelman* condition as well as the *Shulman* condition (cf. Definition 3.8). As Lemma 3.10 shows, Følner sequences with Shulman condition always exist in amenable σ-compact groups. To enter the world of ergodic theory and measure dynamical systems, we explain the notion of a group action on a measure space by measure preserving transformations (see Definition 3.11 and Examples 3.12). Finally, we give a concrete example of a group action in Theorem 3.13 showing that it is not possible to prove a pointwise ergodic theorem of Birkhoff type along arbitrary Følner sequences. We conclude that additional restrictions on the sequence as introduced above are indeed necessary for pointwise almost everywhere convergence.

Example 3.1
Consider $G = (\mathbb{R}, +)$ with the natural topology, as well as the sets $F_n := [-n, n] \subseteq \mathbb{R}$. Then it is not hard to show that for an arbitrary $\varepsilon > 0$ and an arbitrary compact set $K \subseteq \mathbb{R}$, one can find some $n_0 := n_0(\varepsilon, K) \in \mathbb{N}$ such that

$$\frac{|F_n \triangle xF_n|}{|F_n|} < \varepsilon \tag{3.1}$$

for all $x \in K$, whenever $n \geq n_0$. Here and also throughout the remainder of our elaborations, the symbol $\triangle$ stands for the symmetric difference of sets. To justify the claim we recall that every compact $K \subset \mathbb{R}$ is bounded, i.e. there is some $c > 0$ such that $|x| \leq c$ for all $x \in K$. Note that $xF_n = [-n + x, n + x]$ and that $|F_n \triangle xF_n| \leq |F_n \triangle cF_n|$ which means that

$$|F_n \triangle xF_n| \leq 2c$$

for all $x \in K$. We now put $\Phi_n := |F_n|^{-1}\mathbb{1}_{F_n}$. Then for each $x \in K$

$$\begin{aligned} \lim_{n\to\infty} \|L_x\Phi_n - \Phi_n\|_{L^1} &= \lim_{n\to\infty} \frac{|F_n \triangle xF_n|}{|F_n|} \\ &\leq \lim_{n\to\infty} \frac{2c}{|F_n|} = 0. \end{aligned}$$

Since K was arbitrary, the sequence $\{\Phi_n\}$ is strongly convergent to left-invariance and hence, $(\mathbb{R}, +)$ is amenable by Corollary 2.12. This is of course not surprising, as we have already seen that abelian groups are amenable. However, this construction raises the question if there is a connection between amenability and condition (3.1). This is indeed the case, as the following theorem shows.

Theorem 3.2
A locally compact group G is amenable if and only if the Følner condition (FC for short) holds, i.e. if for each $\varepsilon > 0$ and every compact set $K \subseteq G$ there exists some non-empty compact set $F \subseteq G$ of positive measure such that

$$\frac{|F \triangle gF|}{|F|} < \varepsilon \quad \textit{for all } g \in K \qquad \text{(FC)}.$$

Remark
For discrete groups, this theorem is due to Erling Følner, a Danish mathematician (cf. [13]). It was extended to the general topological case by efforts of Namioka in [32] and Ryll-Nardzewski (cf. [17], Theorem 3.6.3).

Proof (of '$\Longleftarrow$')
Using Corollary 2.12, it is not hard to see that (FC) implies amenability. Namely, take the net I consisting of all pairs (ε, K), where ε is a positive number and $K \subseteq G$ is a compact set. This collection is partially ordered; for $\alpha = (\varepsilon_\alpha, K_\alpha)$ and $\beta = (\varepsilon_\beta, K_\beta)$ in I we say $\alpha \geq \beta$ if $\varepsilon_\alpha \leq \varepsilon_\beta$ and $K_\alpha \supseteq K_\beta$.
By (FC), we find for each $\gamma = (\varepsilon_\gamma, K_\gamma) \in I$ some compact set F_γ such that

$$\frac{|F_\gamma \triangle gF_\gamma|}{|F_\gamma|} < \varepsilon_\gamma$$

for all $g \in K_\gamma$. Define $\Phi_\gamma := |F_\gamma|^{-1} \mathbb{1}_{F_\gamma}$. Then by the fact that each point $g \in G$ has a compact neighborhood we have that

$$\lim_\gamma \|L_g \Phi_\gamma - \Phi_\gamma\|_{L^1} = \lim_\gamma |F_\gamma|^{-1} \cdot |F_\gamma \triangle gF_\gamma| \overset{\text{(FC)}}{=} 0$$

and the net $\{\Phi_\gamma\} \subseteq P(G)$ converges (strongly) to left-invariance. Hence, G is amenable.
A direct proof of the converse direction is rather technical and shall be omitted. The interested reader may consult ([41], Theorem 4.7). □

Remark
Theorem 3.2 shows that amenability of a locally compact group is equivalent to the existence of some net $\{F_\gamma\}$ consisting of compact sets which are asymptotically relatively invariant under translations by compact sets (with respect to the Haar measure). This makes sure that the sets F_γ have to grow in a uniform manner (see picture 3.1 for $G = (\mathbb{R}^2, +)$).

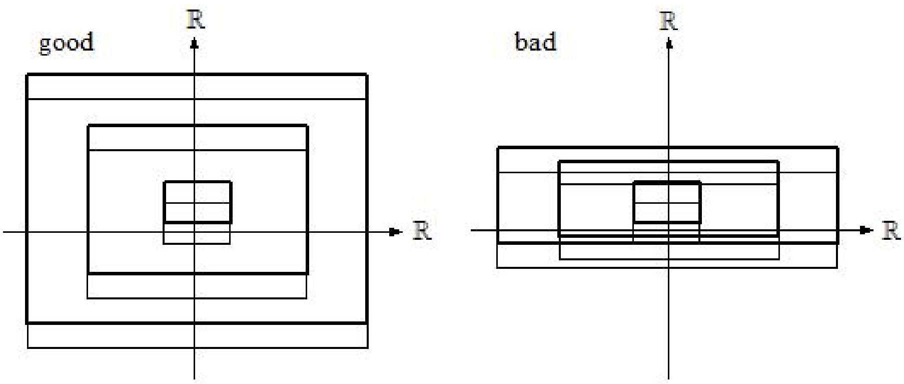

Figure 3.1: Example for $G = (\mathbb{R}^2, +)$; in bold face: translations by singletons $\{(0, c)\}, c > 0$

Note that there are many characterizations of amenability for locally compact groups. For our purposes, the Følner condition is the most adequate one. Also, we will draw our attention to σ-compact groups defined as follows.

Definition 3.3
A group G is called σ-compact if it can be represented as a countable union of compact sets.

Clearly, the groups $\mathbb{R}^d$ and $\mathbb{Z}^d$ are σ-compact. The same is true for the Lamplighter group which can be written as a countable union of finite sets. As an abstract example, consider the following proposition:

Proposition 3.4
Any locally compact, second countable group is σ-compact.

Proof
It follows from elementary topology that by the Hausdorff property of the group, we can find for each $g \in G$ some open neighborhood U_g with compact closure. So if $\{V_n\}$ is a countable base of the topology of G, then one can choose some $n(g) \in \mathbb{N}$ with $g \in V_{n(g)} \subseteq U_g$. Taking closures, we see that $\overline{V_{n(g)}} \subseteq \overline{U_g}$. The sets $\overline{U_g}$ are compact by the choice of the U_g and since G is a Hausdorff group, $\overline{V_{n(g)}}$ must be compact as a closed subset of a compact set. We conclude $G = \cup_{g \in G}\, g \subseteq \cup_{g \in G} V_{n(g)} \subseteq \cup_{n \in N} \overline{V_n} \subseteq G$ and the proposition is proven. □

Emerson and Greenleaf showed in [12] that for σ-compact groups, amenability is equivalent to the existence of a *sequence* of compact sets such that the corresponding normalized characteristic functions converge strongly to left-invariance in a uniform manner.

Theorem 3.5
A σ-compact group G is amenable if and only if there exists a sequence $\{F_n\}$ of non-empty compact sets in G of positive measure such that for every compact set $\emptyset \neq K \subseteq G$ we have

$$\lim_{n\to\infty} \frac{|F_n \triangle K F_n|}{|F_n|} = 0.$$

Proof
See [12], Theorem 3.2.1. □

Definition 3.6
A sequence $\{F_n\}$ *as in Theorem 3.5 is called* **Følner sequence**.

As we have seen in the Theorems 2.10 and 2.11, Følner sequences provide useful tools to define left-invariant means on $L^\infty(G)$. We give some examples of Følner sequences for various amenable groups.

Examples 3.7

(1) For any compact group (e.g. $G = \mathbb{T}^d$), the sequence $\{F_n\}_n = \{G\}_n$ is a Følner sequence.

(2) For $G = (\mathbb{Z}^d, +), (\mathbb{R}^d, +)$, the canonical cubes $\{F_n\}$ with $F_n = \{-n, -(n-1), \ldots, (n-1), n\}^d$ resp. $F_n = [-n, n]^d$ are Følner sequences. For $G = (\mathbb{Z}, +)$, the sequence defined by $G_n = \{n^2, n^2+1, \ldots, n^2+n\}$ is a Følner sequence.

(3) Consider the Lamplighter group and set

$$F_n := \Big\{(m,a) \in G \,\Big|\, |m| \leq n,\, a = \sum_{k=-2n}^{2n} \alpha_k e_k,\, \alpha_k \in \{0,1\}\Big\},$$

with $e_k = (\delta_{jk})_{j\in\mathbb{Z}}$ and δ_{jk} the Kronecker symbol. We claim that $\{F_n\}$ is a Følner sequence.
Let $g \in G$, i.e. $g = (l, b)$ with $l \in \mathbb{Z}$ and $b = \sum_{k=-r}^{r} \beta_k e_k$ for some $r \in \mathbb{Z}$, $\beta_k \in \{0,1\}$. Let further $f = (m, a) \in F_n$. We compute

$$\begin{aligned} g \cdot f &= (l,b)\cdot(m,a) = (l+m, \sigma^m b + a) \\ &= \Big(l+m, \sum_{k=-r-m}^{r-m} \beta_k e_k + \sum_{k=-2n}^{2n} \alpha_k e_k\Big). \end{aligned}$$

Note that for $n > r$, one obtains with $|m| \leq n$ (def. of F_n) that

$$-2n < -r - m < r - m < 2n.$$

Hence,

$$\sum_{k=-r-m}^{r-m} \beta_k e_k + \sum_{k=-2n}^{2n} \alpha_k e_k = \sum_{k=-2n}^{2n} \gamma_k e_k$$

with $\gamma_k \in \{0,1\}$ for $-2n \leq k \leq 2n$.
It follows from the definition of the F_n that $gf \in F_n$ as long as $|l+m| \leq n$. The condition that for fixed $g = (l,b) \in G$, we have $f = (m,a) \in F_n$ with $|l+m| > n$ (and thus $gf \notin F_n$) implies that for any choice of l, we have exactly $|l|$ choices for m. Therefore,

$$|F_n \triangle gF_n| \leq |l| \cdot 2^{4n+1},$$

which with $|F_n| = (2n+1) \cdot 2^{4n+1}$ implies that

$$\frac{|F_n \triangle g F_n|}{|F_n|} \leq \frac{|l|}{2n+1} \to 0, \ n \to \infty.$$

Hence, $\{F_n\}_n$ is a Følner sequence.

We will see in a moment that it is not possible to prove pointwise convergence results along arbitrary Følner sequences of amenable groups acting on a Lebesgue space. In fact, some growth restrictions on the sequence are required. We present the following two conditions.

Definition 3.8
A Følner sequence $\{F_n\}_n$ satisfies the **Tempelman condition** *if there exists some $C > 0$ such that for all $n \in \mathbb{N}$*

$$\left| \bigcup_{k \leq n} F_k^{-1} F_n \right| \leq C |F_n|. \tag{3.2}$$

It satisfies the **Shulman condition** *if there exists some $C > 0$ such that for all $n \in \mathbb{N}$*

$$\left| \bigcup_{k < n} F_k^{-1} F_n \right| \leq C |F_n|. \tag{3.3}$$

Følner sequences satisfying the Shulman condition are called **tempered**.

Evidently, the Tempelman condition implies the Shulman condition. The converse is not true in general. Let us consider some examples.

Examples 3.9

(1) The canonical cubes, defined above, satisfy the Shulman as well as the Tempelman condition.

(2) We have already mentioned that the sequence $F_n := \{n^2, n^2+1, \ldots, n^2+n\}$ is a Følner sequence in $G = \mathbb{Z}$. It neither satisfies the Tempelman, nor the Shulman condition. To see this, we remark that

$$F_k^{-1} F_n = \{-k^2 - k + n^2, \ldots, -k^2 + n + n^2\}.$$

It is easily verified that with $k < n$, $-(k+1)^2 + n + n^2 \geq -k^2 - k + n^2$ and we conclude that $F_{k-1}^{-1} F_n \cap F_k^{-1} F_n \neq \emptyset$ for $2 \leq k \leq (n-1)$. In light of that, we have

$$\begin{aligned} \bigcup_{1 \leq k < n} F_k^{-1} F_n &= \{-(n-1)^2 - (n-1) + n^2, \ldots, -1^2 + n + n^2\} \\ &= \{n, \ldots, n^2 + n - 1\}, \end{aligned}$$

so that

$$\liminf_{n \to \infty} \frac{|\cup_{k<n} F_k^{-1} F_n|}{|F_n|} = \liminf_{n \to \infty} \frac{n^2}{n+1} = \infty.$$

Hence, $\{F_n\}$ is not tempered and in particular, it does not satisfy the Tempelman condition.

(3) Again, consider the Lamplighter group G with the Følner sequence $\{F_n\}_n$ from the previous example. We claim that this sequence does not satisfy the Tempelman condition, but contains a tempered subsequence (cf. [3], example 1.50).
Note that the inverse element of some $g = (l, b) \in G$ is $g^{-1} = (l, b)^{-1} = (-l, \sigma^{-l}b)$. In light of that we write the set $F_m^{-1}F_n$ as

$$\begin{aligned} F_m^{-1}F_n &= \left\{ \Big(k-l, \sigma^{k-l}b + c\Big) \,\Big|\, |l| \le m,\ |k| \le n\,, b = \sum_{s=-2m}^{2m} \beta_s e_s,\ c = \sum_{t=-2n}^{2n} \gamma_t e_t \right\} \\ &= \left\{ \left(k-l, \sum_{s=-2m}^{2m} \beta_s e_{s+(k-l)} + \sum_{t=-2n}^{2n} \gamma_t e_t\right) \,\Big|\, |l| \le m,\ |k| \le n \right\} \\ &= \left\{ \left(d, \sum_{s=-2m+d}^{2m+d} \beta_{(s-d)} e_s + \sum_{t=-2n}^{2n} \gamma_t e_t\right) \,\Big|\, |d| \le n+m \right\}, \end{aligned} \tag{3.4}$$

where of course, we take all combinations for $\beta_s, \gamma_t \in \{0, 1\}$. Thus, for $m = n$, we obtain

$$F_n^{-1}F_n = \left\{ \left(d, \sum_{s=\min\{-2n,-2n+d\}}^{\max\{2n,2n+d\}} \delta_s e_s\right) \,\Big|\, |d| \le 2n,\ \delta_s \in \{0, 1\} \right\}.$$

Therefore,

$$|F_n^{-1}F_n| = \sum_{d=-2n}^{2n} 2^{4n+d+1} = 2^{2n+1}(2^{4n+1} - 1)$$

and it follows that

$$\liminf_{n\to\infty} \frac{|F_n^{-1}F_n|}{|F_n|} = \liminf_{n\to\infty} \frac{2^{2n+1}(2^{4n+1} - 1)}{(2n+1)2^{4n+1}} = \infty,$$

so that the sequence $\{F_n\}_n$ cannot satisfy (3.2).
However, by extracting a subsequence $\{G_k\}$, we can construct a tempered Følner sequence in G. Note first that with $d \le n + m$, it is true that $2m + d \le 3m + n$ and $-2m + d \ge -3m - n$. Thus, if we choose $n \ge 3m$, equality (3.4) transforms to

$$F_m^{-1}F_n = \left\{ \left(d, \sum_{s=-2n}^{2n} \delta_s e_s\right) \,\Big|\, |d| \le m+n,\ \delta_s \in \{0, 1\} \right\}.$$

Consequently, one obtains in this situation that

$$|F_m^{-1}F_n| = (2n + 2m + 1)2^{4n+1}$$

and hence

$$\frac{|F_m^{-1}F_n|}{|F_n|} = \frac{2n+2m+1}{2n+1} = 1 + \frac{2m}{2n+1}. \tag{3.5}$$

Now define $G_k := F_{n_k}$, where $n_k := 3^k$. Since then $n_k = 3n_{k-1}$, we can use (3.5) to conclude

$$\frac{|\cup_{l<k} G_l^{-1}G_k|}{|G_k|} = \frac{|G_{k-1}^{-1}G_k|}{|G_k|} = \frac{|F_{n_{k-1}}F_{n_k}|}{|F_{n_k}|} = 1 + \frac{2 \cdot 3^{k-1}}{2 \cdot 3^k + 1} < 2$$

for all $k \in \mathbb{N}$. Recalling inequality (3.3), we see that $\{G_k\}_k$ is a tempered Følner sequence.

The following lemma shows that the construction in the latter example is 'typical' for amenable groups. Indeed, from *every* Følner sequence, one can extract a tempered subsequence.

Lemma 3.10 (Lindenstrauss, cf. [29])
Let G be an amenable group G with a Følner sequence $\{F_n\}_n$. Then there is some subsequence $\{F_{n_k}\}_k$ satisfying the Shulman condition (3.3).

Proof
We proceed inductively. Set $n_1 := 1$. If $n_1, \dots n_{j-1}\,(j \geq 2)$ are chosen, define $\tilde{F}_j := \cup_{i=1}^{j-1} F_{n_i}^{-1}$. Clearly, this set is compact. It follows from the Følner property (Theorem 3.5) with $\varepsilon = 1$ that there is some n_j large enough such that

$$\begin{aligned} |\cup_{i<j} F_{n_i}^{-1}F_{n_j}| = |\tilde{F}_j F_{n_j}| &\leq |F_{n_j}| + |F_{n_j} \triangle \tilde{F}_j F_{n_j}| \\ &\overset{(\mathrm{FC})}{\leq} |F_{n_j}| + \varepsilon |F_{n_j}| = 2|F_{n_j}|. \end{aligned}$$

Hence, $\{F_{n_j}\}_j$ is tempered with constant $C = 2$. □

It turns out that a statement of this kind cannot be proven for the Tempelman condition. Lindenstrauss showed in [29] (Corollary 5.5) that the Lamplighter group does not contain a Følner sequence satisfying condition (3.2). Therefore, the condition on such a sequence $\{F_n\}_n$ to be tempered is not only milder than the Tempelman restriction but is also natural in the sense that it always exists in an amenable σ-compact group.

In order to link the theory of σ-compact amenable groups with ergodic theory, we need the notion of a group action on a measure space by measure preserving transformations.

Definition 3.11 (measure preserving action of a group)
Let G be a σ-compact group and let $(X, \mathcal{B}, \mu)$ be some σ-finite measure space. We say that G acts on X by measure preserving transformations if there is a map

$$\pi : G \times X \to X$$

with the following properties:

(i) π is $(\mathcal{J} \times \mathcal{B})$-$\mathcal{B}$-measurable, where $\mathcal{J}$ is the Borel σ-algebra on G,

(ii) $\pi(e, x) = x$ for each $x \in X$, where e is the unit element in G,

(iii) $\pi(g_1, \pi(g_2, x)) = \pi(g_1 g_2, x)$ *for each* $x \in X$ *and all* $g_1, g_2 \in G$,

(iv) $\mu(\pi(g, A)) = \mu(A)$ *for each* $g \in G$ *and each* $A \in \mathcal{B}$.

For short, we introduce the notation

$$g \cdot x := gx := \pi(g, x),\ g \in G,\ x \in X.$$

Examples 3.12

(1) We start with a standard example, the rotation on the circle. Put $G := (\mathbb{Z}, +)$ and fix $\alpha \in [0, 1)$. Further, consider the measure space $(\mathbb{T}, \mathcal{B}(\mathbb{T}), \mathcal{L}(\mathbb{T}))$, where $\mathbb{T}$ is the one-dimensional torus $\mathbb{T} := \mathbb{R}/\mathbb{Z}$, endowed with the Borel σ-algebra $\mathcal{B}(\mathbb{T})$ and the Lebesgue measure $\mathcal{L}(\mathbb{T})$. Then G acts on $\mathbb{T}$ by measure preserving transformations via

$$\pi_\alpha : G \times \mathbb{T} \to \mathbb{T} : (n, z) \mapsto n \cdot \alpha + z \ (\mathrm{mod}\ 1).$$

For $z \in \mathbb{T}$, we call the set $O_z^{(\alpha)} := \{\pi_\alpha(n, z) \,|\, n \in \mathbb{Z}\}$ the orbit of the element z. It is a well-known fact that for $\alpha \in \mathbb{R} \setminus \mathbb{Q}$, the orbit $O_z^{(\alpha)}$ is dense in $\mathbb{T}$ for each $z \in \mathbb{T}$ (see e.g. [14], Example 2.19).

(2) Every σ-compact group G with Borel σ-algebra $\mathcal{J}$ and left Haar measure $m_L(\cdot)$ acts on itself via group multiplication. By the left-invariance of m_L, we have $m_L(g_0 B) = m_L(B)$ for all $g_0 \in G$ and every $B \in \mathcal{J}$ and the action is measure preserving.

Continuing the first example, we fix some irrational value $\alpha \in [0, 1)$ and consider the function $f(x) := x^{-1/2} \in L^1(\mathbb{T})$. By the Birkhoff Ergodic Theorem (see e.g. [14], Theorem 10.1 and below), the limit

$$\lim_{n \to \infty} \frac{1}{n} \sum_{j=0}^{n-1} f(\pi_\alpha(j, z))$$

exists for $\mathcal{L}(\mathbb{T})$-almost every $z \in \mathbb{T}$. The following theorem shows that this is not the case for all Følner sequences in $\mathbb{Z}$.

Theorem 3.13 (cf. [11], Theorem 1)

Consider the action of the integers $\mathbb{Z}$ *on the torus* $\mathbb{T}$ *as defined in Example 3.12 (1), where* α *is any fixed irrational number in* $[0, 1)$*. Moreover let* $f(z) = z^{-1/2} \in L^1(\mathbb{T})$*. Then there is a Følner sequence* $\{F_n\}$ *in* $\mathbb{Z}$ *such that*

$$\limsup_{n \to \infty} |F_n|^{-1} \sum_{m \in F_n} f(\pi_\alpha(m, z)) = \infty$$

for all $z \in \mathbb{T}$.

Proof

Set $E_n := \{0, \pm 1, \pm 2, \pm 3, \dots, \pm n!\}$. Beginning at the point $p_1 := 0$ and in counter clockwise direction, we lay off adjacent closed 'intervals' I_n of length $1/n$ ($n \in \mathbb{N}$) on $\mathbb{T}$. We denote the left-hand endpoint of the interval I_n by p_n. The divergence of the harmonic series insures that each point of $\mathbb{T}$ is contained in infinitely many of the I_n. Moreover, we have seen in

Example 3.12 (1) that for all $z \in \mathbb{T}$, the transformation $z \mapsto z + \alpha \pmod 1$ has dense orbit in $\mathbb{T}$ such that for each endpoint p_n, there exist infinitely many integers $m \in \mathbb{Z}$ such that $\pi_\alpha(m, p_n) \in [0, 1/n]$. We construct the Følner sequence $\{F_n\}$ in a recursive manner.
Put $F_1 := E_1$ and assume that for $n > 1$ and $k < n$, all F_k have been constructed. Denote by D_n an arbitrary collection of exactly $\operatorname{int}(n!n^{-1/4})$ distinct integers $m \in \mathbb{Z}$ for which $\pi_\alpha(m, p_n)$ is contained in $[0, 1/n]$, where

$$\operatorname{int}(r) := \max\{m \in \mathbb{N} \mid m \leq r\}$$

is the integer function defined on $\mathbb{R}_0^+$. Further, define $F_n := E_n \cup F_{n-1} \cup D_n$. One verifies by induction that $|F_k| \leq 3k!$ for all $k \geq 1$. To see this, note that $|F_1| = 3$ and recursively, we have

$$\begin{aligned} |F_k| &\leq 2(k! - (k-1)!) + |F_{k-1}| + k!k^{-1/4} \\ &= 2k! + k! \left(\frac{|F_{k-1}|}{k!} + \frac{1}{k^{1/4}} - \frac{2}{k} \right) \end{aligned} \tag{3.6}$$

for all $k \geq 2$. For the first elements of the sequence, we compute

$$\begin{aligned} |F_1| &= 3 = 3 \cdot 1! \\ |F_2| &= 2 + 3 + \operatorname{int}(2 \cdot 2^{-1/4}) = 6 = 3 \cdot 2! \\ |F_3| &= 8 + (5 + 2 \cdot 2^{-1/4}) + 3 \cdot 3^{-1/4} \leq 18 = 3 \cdot 3!. \end{aligned}$$

Further, using equality (3.6), we see inductively that

$$\begin{aligned} \left(\frac{|F_{k-1}|}{k!} + \frac{1}{k^{1/4}} - \frac{2}{k} \right) &\leq \left(\frac{3(k-1)!}{k!} - \frac{2}{k} + \frac{1}{k^{1/4}} \right) \\ &= k^{-1} + k^{-1/4} < 1 \end{aligned}$$

and thus $|F_k| \leq 3k!$ for all $k \geq 4$ as claimed.
It follows that $|F_{n-1} \cup D_n| \leq 3(n-1)! + n!n^{-1/4}$ for $n \geq 2$ and noting that $|E_n| = 2n! + 1$, one computes that

$$\begin{aligned} \frac{|F_n \triangle z F_n|}{|F_n|} &\leq \frac{|E_n \cup (F_{n-1} \cup D_n) \triangle z(E_n \cup (F_{n-1} \cup D_n))|}{|E_n|} \\ &\leq \frac{|E_n \triangle z E_n|}{|E_n|} + \frac{|(F_{n-1} \cup D_n) \triangle z(F_{n-1} \cup D_n)|}{|E_n|} \\ &\leq \frac{|E_n \triangle z E_n|}{|E_n|} + 2 \frac{|F_{n-1} \cup D_n|}{|E_n|} \\ &\leq \frac{|E_n \triangle z E_n|}{|E_n|} + \frac{3}{n} + n^{-1/4} \end{aligned} \tag{3.7}$$

for each $z \in \mathbb{T}$, where we used the general inclusion

$$(A \cup B) \triangle (C \cup D) \subseteq (A \triangle C) \cup (B \triangle D)$$

for sets A, B, C and D. As $\{E_n\}$ is a Følner sequence, the expression in (3.7) converges to zero as $n \to \infty$ and we conclude that $\{F_n\}$ is indeed a Følner sequence.

Now let $z \in \mathbb{T}$ be fixed and take one of the (infinitely many) 'intervals' I_n containing z. Let us estimate the ergodic average over the set F_n evaluated at z. By construction, we have $\pi_\alpha(m, p_n) \in [0, 1/n]$ for every $m \in D_n$. Since p_n is the left-hand endpoint of the 'interval' I_n of length $1/n$, one obtains for each $m \in D_n$ that $p + m\alpha \pmod 1 \in [0, 2/n]$ for all $p \in I_n$. Therefore, as $z \in I_n$, we have

$$f(z + m\alpha) \geq (2/n)^{-1/2} = (n/2)^{1/2}$$

whenever $m \in D_n$. Recalling the fact that $|F_n| \leq 3n!$, this implies

$$\begin{aligned} |F_n|^{-1} \sum_{m \in F_n} f(z + m\alpha) &\geq \frac{1}{3n!} \sum_{m \in D_n} f(z + m\alpha) \\ &\geq \frac{|D_n|(n/2)^{1/2}}{3n!} \\ &\geq \frac{n! n^{-1/4} n^{1/2}}{6n!} = \frac{n^{1/4}}{6}. \end{aligned}$$

As mentioned above, there are infinitely many such $n \in \mathbb{N}$ and thus, the theorem is proven. □

Theorem 3.13 shows that additional growth conditions on the Følner sequence are indeed needed for pointwise convergence. Before turning to pointwise ergodic theorems, we will prove a mean ergodic theorem for σ-compact, amenable groups in the next chapter. No restrictions on $\{F_n\}$ are required here. We will exploit the Shulman condition in Chapters 7 and 8, where we prove the transfer inequality (6.1).

4 A mean ergodic theorem

This chapter is devoted to general mean ergodic theorems for amenable σ-compact groups. We start in a rather abstract setting which requires the notion of a weakly measurable action of a group on a Banach space Y (cf. Definition 4.1). Using Følner sequences, we define abstract ergodic averages which exist in the weak sense only, but have a strong representation. Moreover, we show in Theorem 4.2 that if the closed convex hull of an orbit (under the group action) of some element in Y is weakly compact, then it must contain some G-invariant element. We will see in the Abstract Mean Ergodic Theorem 4.3 that the existence of such a fixed point is equivalent to the convergence of (the strong representation of) the corresponding ergodic average to this fixed point in the Banach space norm. In particular, if each orbit contains a G-invariant element, then the representations of the ergodic averages converge in the strong operator topology to some projection on the fixed space of the group.
After this and also for the remainder of this thesis, we turn to a special case of the situation described above. In light of that, the Banach space Y will be some L^p-space $(1 \leq p < \infty)$ over some σ-finite measure space X. Further, the weakly measurable action on $L^p(X)$ is induced by an action of the amenable group on the measure space X by measure preserving transformations (see Proposition 4.4). It is proven in Lemma 4.5 that the corresponding orbits satisfy the weak compactness condition if in the case $p = 1$, we assume additionally that the measure space is finite. Exploiting the Abstract Mean Ergodic Theorem 4.3, we finally derive a general mean ergodic theorem for L^p-spaces in Corollary 4.6. We also give a counter example in the case of an infinite measure space for $p = 1$ (cf. Example 4.7).

Definition 4.1 (weakly measurable action of a group on a Banach space)
Let G be a σ-compact group and let $(Y, \|\cdot\|_Y)$ be a Banach space. Then the map

$$T : G \times Y \to Y : (g, f) \mapsto T_g f$$

is called a weakly measurable action of G as a bounded family $\{T_g\}$ of operators on Y if

(i) $T_g : Y \to Y$ *is a linear operator for each* $g \in G$ *and* $\sup_{g \in G} \|T_g\| < \infty$,

(ii) $T_e f = f$ *for each* $f \in Y$, *where e is the unit element in G,*

(iii) $T_{g_1}(T_{g_2} f) = T_{g_1 g_2} f$ *for each* $f \in Y$ *and all* $g_1, g_2 \in G$,

(iv) the map

$$\Phi_{f,h} : G \to \mathbb{C} : g \mapsto \langle T_g f, h \rangle_{Y,Y^*}$$

is $\mathcal{J}$-$\mathcal{B}(\mathbb{C})$-measurable for each $f \in Y$ and each $h \in Y^*$, where $\mathcal{J}$ and $\mathcal{B}(\mathbb{C})$ are the Borel σ-algebras of G and $\mathbb{C}$ respectively.

Further, we denote by

$$\mathrm{Fix}(T_G) := \{f \in Y \mid T_g f = f \text{ for all } g \in G\}$$

the fixed space of G. Analogously, if T_g^* is the adjoint operator of T_g, one defines

$$\mathrm{Fix}(T_G^*) := \{h \in Y^* \mid T_g^* h = h \text{ for all } g \in G\}.$$

We prove the following essential result.

Theorem 4.2 (Existence of a fixed point, Greenleaf 1973)
We assume that a σ-compact, amenable group G acts weakly measurably on the Banach space $(Y, \|\cdot\|_Y)$ by a family $\{T_g\}$ of bounded operators on Y with $A := \sup_{g\in G} \|T_g\| < \infty$. Let m be a left invariant mean on $L^\infty(G)$ and assume further that $f \in Y$ is an element such that the closed, convex hull of the G-orbit of f, denoted as $C_f := \overline{\mathrm{co}}\{T_g f \mid g \in G\}$ is weakly compact. If we put $\Phi_{f,h}(g) := \langle T_g f, h\rangle_{Y,Y^*}$, then $T_m(f)$, defined by the equation

$$\langle h, T_m(f)\rangle_{Y^*,Y^{**}} = m(\Phi_{f,h}) \text{ for all } h \in Y^*,$$

determines an element f^* in Y which belongs to $\mathrm{Fix}(T_G) \cap C_f$.
In particular, $\mathrm{Fix}(T_G) \cap C_f \neq \emptyset$.

Remark
Note that for all $g \in G$, we have $|\Phi_{f,h}(g)| \leq A\|f\|_Y\|h\|_{Y^*}$ so that $\Phi_{f,h} \in L^\infty(G)$ and $T_m(f)$ is well-defined.

PROOF
If $h \in Y^*$, then

$$\begin{aligned} |\langle h, T_m(f)\rangle_{Y^*,Y^{**}}| = |m(\langle T_\cdot f, h\rangle_{Y,Y^*})| &\leq \|m\| \cdot A\|f\|_Y\|h\|_{Y^*} \\ &= A\|f\|_Y\|h\|_{Y^*} \end{aligned}$$

and therefore, $T_m(f) \in Y^{**}$. We consider Y^{**} with the weak$*$-topology $\sigma(Y^{**}, Y^*)$ and let $j : Y \to Y^{**}$ be the canonical injection. Now the set C_f is $\sigma(Y, Y^*)$-compact and thus, by the weak-(weak$*$)-continuity of j, we obtain that $j(C_f)$ is a convex, weak$*$-compact set. By [44], Corollary VIII 3.4, we have $(Y^{**}, \sigma(Y^{**}, Y^*))^* = Y^*$. So if $T_m(f) \notin j(C_f)$, by the Hahn-Banach separation theorem (cf. [44], VIII 2.12), there must be some $h \in Y^*$ and some $\varepsilon_0 > 0$ such that

$$\mathrm{Re}\langle h, T_m f\rangle_{Y^*,Y^{**}} \geq \varepsilon_0 + \sup\{\mathrm{Re}\langle z, h\rangle_{Y,Y^*} \mid z \in C_f\}. \tag{4.1}$$

However, it is clear that $\mathrm{Re}\langle T_g f, h\rangle_{Y,Y^*} \leq \theta := \sup\{\mathrm{Re}\langle z, h\rangle_{Y,Y^*} \mid z \in C_f\}$ for all $g \in G$ and hence, since m is a mean,

$$\begin{aligned} \mathrm{Re}\langle h, T_m(f)\rangle_{Y^*,Y^{**}} &= \mathrm{Re}\left[m(\langle T_\cdot f, h\rangle_{Y,Y^*})\right] \\ &= m(\mathrm{Re}\,\langle T_\cdot f, h\rangle_{Y,Y^*}) \leq \theta \cdot m(\mathbb{1}) = \theta, \end{aligned}$$

which contradicts inequality (4.1). Therefore, $T_m(f) \in j(C_f)$.
We still have to show that $f^* := j^{-1}(T_m(f))$ is a fixed point in C_f. Since m is left-invariant

by assumption and by the fact that j is an isometry, we have for all $g \in G$ and all $h \in Y^*$ that

$$\begin{aligned}\langle T_g f^*, h\rangle_{Y,Y^*} = \langle f^*, T_g^* h\rangle_{Y,Y^*} &= \langle T_g^* h, T_m(f)\rangle_{Y^*,Y^{**}} \\ &\overset{\text{Def.}}{=} m(\langle T_\cdot f, T_g^* h\rangle_{Y,Y^*}) \\ &= m(\langle T_{g\cdot} f, h\rangle_{Y,Y^*}) \\ &= m(L_{g^{-1}}\Phi_{f,h}) \\ &\overset{\text{left-inv.}}{=} m(\Phi_{f,h}) \\ &\overset{\text{Def.}}{=} \langle h, T_m(f)\rangle_{Y^*,Y^{**}} \\ &= \langle f^*, h\rangle_{Y,Y^*},\end{aligned}$$

where Φ is defined as above and T_g^* denotes the adjoint operator of T_g. In light of that, f^* is a fixed point in C_f. □

Remark
The techniques of the proof have been used before; in 1961, Day considered a compact subset K of a locally convex linear topological space Y and a semigroup $\mathcal{H}$ of continuous affine transformations of K into itself. If this semigroup contains a left-invariant mean, then there is some $\mathcal{H}$-fixed point K (cf. [8], Theorem 1). With this result, abstract fixed-point theorems for amenable semigroups of uniformly bounded linear operators on a Banach space Y (amenable means here that there is some projection $P \in \overline{\text{co}}\,\mathcal{H}$ (strong operator topology) such that $TP = PT = P$ for all $T \in \overline{\text{co}}\,\mathcal{H}$) can be proven. For a more detailed discussion, see for example [31], Lemma 1.6 or [25], Section 6.4.1.

The following theorem shows that the existence of such a fixed point is equivalent to a mean ergodic theorem for amenable (σ-compact) groups.

Theorem 4.3 (Abstract mean ergodicity)
Let G be a σ-compact amenable group and let $(Y, \|\cdot\|_Y)$ be some Banach space. Assume further that G acts weakly measurably as a uniformly bounded family $\{T_g\}$ of continuous, linear operators on Y with $A := \sup_{g\in G}\|T_g\| < \infty$. For a Følner sequence $\{F_n\}$ in G, we define the ergodic averages $A_n f$ on Y as

$$A_n f := |F_n|^{-1}\int_{F_n} T_{g^{-1}} f \, dm_L(g).$$

Then the following statements are equivalent:

(i) *For all $f \in Y$ there is some $f^* \in Y$ such that $\text{Fix}(T_G) \cap C_f = \{f^*\}$, where $C_f := \overline{\text{co}}\{T_g f \mid g \in G\}$.*

(ii) *For all $f \in Y$ there is some $f^* \in \text{Fix}(T_G) \cap C_f$.*

(iii) *$A_n \to P$ in the strong operator topology and $\text{ran}(P) \subseteq \text{Fix}(T_G)$.*

(iv) *There is a bounded projection P on Y with $\text{ran}(P) = \text{Fix}(T_G)$ and $\text{ran}(I - P) = \ker(P) = L_0$, where $L_0 := \overline{\text{lin}}\{f - T_g f \mid f \in Y, g \in G\}$.*

(v) $\mathrm{Fix}(T_G)$ *separates* $\mathrm{Fix}(T_G^*)$, *i.e. for every* $0 \neq h \in \mathrm{Fix}(T_G^*)$ *there exists* $f^* \in \mathrm{Fix}(T_G)$ *such that* $\langle f^*, h\rangle_{Y,Y^*} \neq 0$.

(vi) $Y = \mathrm{Fix}(T_G) \oplus L_0$.

Remark
Note that the ergodic averages are in the first instance only defined in the weak sense, that is $A_n f \in Y^{**}$ for all $n \in \mathbb{N}$ and every $f \in Y$. However, we can identify them with elements in Y which allows us to use the common calculating rules. Hence, without loss of generality, we can assume that $A_n f \in Y$.
To see this, recall that the Følner sequence $\{F_n\}$ can be interpreted as a net $\{q_n\}$ in $P(G)$ converging strongly to left-invariance (see Corollary 2.12). So let $q := |F_n|^{-1}\mathbb{1}_{F_n}$ for some element of the given Følner sequence. By identification, q determines a mean on $L^\infty(G)$ and with the definition of $T_m(f)$ as in Theorem 4.2, we observe that

$$\begin{aligned}\langle h, T_m(f)\rangle_{Y^*,Y^{**}} &= m(\Phi_{f,h}) = \langle \Phi_{f,h}, q\rangle_{L^\infty(G),L^\infty(G)^*} \\ &= \langle h, |F_n|^{-1}\int_G \mathbb{1}_{F_n}(\cdot)\, T_\cdot f\, dm_L(\cdot)\ \rangle_{Y^*,Y^{**}} \end{aligned} \tag{4.2}$$

for all $h \in Y^*$ and hence, we have $D_n f := |F_n|^{-1}\int_{F_n} T_g f\, dm_L(g) = T_m(f)$ weakly. Again by Theorem 4.2, we obtain that $\tilde{D}_n f := j^{-1}(T_m(f)) \in Y$ and that

$$\langle h, T_m(f)\rangle_{Y^*,Y^{**}} = \langle \tilde{D}_n f, h\rangle_{Y,Y^*}$$

for each $h \in Y^*$, where j is the canonical isometry between Y and Y^{**}. Thus, we conclude with equality (4.2) that for each $n \in \mathbb{N}$, the expression $\tilde{D}_n f$ is a representation in Y of the weakly defined $D_n f$. Finally, note that if $\{T_g\}$ acts weakly measurably as a family of uniformly bounded operators on Y, then so does the family $\{T_{g^{-1}}\}$. Therefore, we can assume with no loss of generality that the ergodic averages $A_n f$ are defined in the strong sense for all $n \in \mathbb{N}$ and all $f \in Y$.

Proof (of Theorem 4.3)
(i) $\Rightarrow$ (ii) is trivial.

(ii) $\Rightarrow$ (iii) Let $f \in Y$ and take $f^* \in \mathrm{Fix}(T_G) \cap C_f$, which exists by assumption. By definition of C_f, for an arbitrary $\varepsilon > 0$ we can find a finite convex combination $f' = \sum_i \alpha_i T_{g_i} f$ with $g_i \in G$, $\sum_i \alpha_i = 1$ and $\|f^* - f'\|_Y < \varepsilon$. Then

$$\begin{aligned} A_n f' &= |F_n|^{-1}\int_{F_n} T_{h^{-1}}\Big(\sum_i \alpha_i T_{g_i} f\Big)\, dm_L(h) \\ &= |F_n|^{-1}\sum_i \alpha_i \int_{F_n} T_{h^{-1}g_i} f\, dm_L(h) \\ &= \sum_i \alpha_i |F_n|^{-1}\int_{g_i^{-1}F_n} T_{h'^{-1}} f\, dm_L(h') \end{aligned}$$

for every $n \in \mathbb{N}$. By the Følner property, Theorem 3.5, we can choose n_0 large enough such that

$$\frac{|F_n \triangle g_i^{-1} F_n|}{|F_n|} \leq \varepsilon$$

for all (finitely many) i, whenever $n \geq n_0$. Thus, for $n \geq n_0$,

$$\begin{aligned} \|A_n f - A_n f^{'}\|_Y &= \left\| \sum_i \alpha_i |F_n|^{-1} \left(\int_{F_n} T_{h^{-1}} f \, dm_L(h) - \int_{g_i^{-1} F_n} T_{h'^{-1}} f \, dm_L(h') \right) \right\|_Y \\ &\leq \sum_i \alpha_i \frac{|F_n \triangle g_i^{-1} F_n|}{|F_n|} A \|f\|_Y \leq \varepsilon A \|f\|_Y . \end{aligned} \tag{4.3}$$

Moreover, since f^* is a fixed point,

$$A_n f^* = |F_n|^{-1} \int_{F_n} T_{g^{-1}} f^* \, dm_L(g) = |F_n|^{-1} \int_{F_n} f^* \, dm_L(g) = f^* \tag{4.4}$$

for each $n \in \mathbb{N}$. Hence, using (4.4), as well as the uniform boundedness of $\{T_g\}$, we compute

$$\begin{aligned} \|A_n f^{'} - f^*\|_Y &= \|A_n f^{'} - A_n f^*\|_Y \\ &= \left\| |F_n|^{-1} \int_{F_n} T_{h^{-1}} (f^{'} - f^*) \, dm_L(h) \right\|_Y \\ &\leq |F_n|^{-1} \int_{F_n} \|T_{h^{-1}} (f^{'} - f^*)\|_Y \, dm_L(h) \leq \varepsilon A \end{aligned} \tag{4.5}$$

for every $n \in \mathbb{N}$. Combining the inequalities (4.3) and (4.5), we conclude that $\|A_n f - f^*\|_Y \leq \varepsilon A(\|f\|_Y + 1)$ for $n \geq n_0$. Since ε was arbitrary, the ergodic averages $A_n f$ converge to f^* in norm. Note that this also shows that the fixed point in C_f is unique. Further, defining $Pf := f^*$ we observe that P is a linear operator on Y and since $\|P\| \leq \sup_{n \in \mathbb{N}} \|A_n\| \leq A$ we see that $A_n \overset{n \to \infty}{\to} P$ in the strong operator topology. Since Pf is a fixed point for each $f \in Y$, we conclude that $\operatorname{ran}(P) \subseteq \operatorname{Fix}(T_G)$.

(iii) $\Rightarrow$ (iv) If P is the strong limit of the sequence A_n, then $\|P\| \leq \sup_{n \in \mathbb{N}} \|A_n\| \leq A$ which implies that P is bounded. Let $f \in Y$ be arbitrary. Since Pf is a fixed point we observe by repeating calculation (4.4), that $A_n(Pf) = Pf$ for all $n \in \mathbb{N}$. By the strong convergence $A_n \to P$, we conclude $P^2 = P$. Hence P is a bounded projection with $\operatorname{ran}(P) \subseteq \operatorname{Fix}(T_G)$. Further, if $f^* \in \operatorname{Fix}(T_G)$, then $A_n f^* = f^*$ for all $n \in \mathbb{N}$ by the same calculation as in (4.4). By convergence in the strong operator topology, $Pf^* = f^*$ and thus, $\operatorname{Fix}(T_G) \subseteq \operatorname{ran}(P)$. It follows that $\operatorname{ran}(P) = \operatorname{Fix}(T_G)$.
Moreover, with $A_n \to P$ (strong operator topology), we obtain for all $g \in G$ and all $f \in Y$

$$\begin{aligned} \|P(f - T_g f)\|_Y &\leq \limsup_{n \to \infty} \|A_n (f - T_g f)\|_Y \\ &\leq \limsup_{n \to \infty} |F_n|^{-1} \left\| \int_{F_n} T_{h^{-1}} f \, dm_L(h) - \int_{F_n} T_{h^{-1} g} f \, dm_L(h) \right\|_Y \\ &= \limsup_{n \to \infty} |F_n|^{-1} \left\| \int_{F_n} T_{h^{-1}} f \, dm_L(h) - \int_{g^{-1} F_n} T_{h'^{-1}} f \, dm_L(h') \right\|_Y \\ &\leq \limsup_{n \to \infty} A \frac{|F_n \triangle g^{-1} F_n|}{|F_n|} \cdot \|f\|_Y = 0. \end{aligned}$$

Therefore, by considering linear combinations of elements $f - L_g f$ and using the continuity of P, one observes that $L_0 \subseteq \ker(P)$. If, on the other hand, $Pf = 0$, then by the strong convergence $A_n \to P$, the zero element of Y must be contained in C_f and thus, there are finite convex sums $s_m = \sum_{i=1}^{K_m} \alpha_i^{(m)} T_{g_i^{(m)}} f$ such that $\|s_m\|_Y \overset{m\to\infty}{\to} 0$. Therefore,

$$\begin{aligned} \lim_{m\to\infty} \left\| \sum_{i=1}^{K_m} \alpha_i^{(m)} (f - T_{g_i^{(m)}} f) - f \right\|_Y &= \lim_{m\to\infty} \left\| \sum_{i=1}^{K_m} \alpha_i^{m} T_{g_i^{(m)}} f \right\|_Y \\ &= 0. \end{aligned}$$

We conclude that $f \in L_0$, thus $\ker(P) = L_0$ and the proof of statement (iv) is finished.

(iv) $\Rightarrow$ (v) Assume that $0 \neq h \in \mathrm{Fix}(T_G^*)$. Then for all $g \in G$ and all $f \in Y$

$$\langle f, h\rangle_{Y,Y^*} = \langle f, T_g^* h\rangle_{Y,Y^*} = \langle T_g f, h\rangle_{Y,Y^*},$$

so that $\langle f - T_g f, h\rangle_{Y,Y^*} = 0$. Hence, for all $f' \in \ker(P) = L_0$, we must have

$$\langle f', h\rangle_{Y,Y^*} = 0. \tag{4.6}$$

Since $h \neq 0$, there must be some $f^* \in Y$ such that $\langle f^*, h\rangle_{Y,Y^*} \neq 0$. But P is a bounded projection which implies that we can write $Y = \mathrm{ran}(P) + \ker(P)$. So with (4.6) we can assume without loss of generality that $f^* \in \mathrm{ran}(P) = \mathrm{Fix}(T_G)$.

(v) $\Rightarrow$ (vi) Define $Z := \mathrm{Fix}(T_G) + L_0$. We claim first that this sum is direct. So let $z := \sum_i \alpha_i (f_i - T_{g_i} f_i)$, where the sum is finite, the α_i are scalar coefficients and $f_i \in Y$, $g_i \in G$ for all i. Then

$$\begin{aligned} \|A_n z\|_Y &= \left\| |F_n|^{-1} \left[\int_{F_n} T_{h^{-1}} \Big(\sum_i \alpha_i f_i\Big)\, dm_L(h) - \int_{F_n} T_{h^{-1}} \Big(\sum_i \alpha_i T_{g_i} f_i\Big)\, dm_L(h) \right] \right\|_Y \\ &\leq \sum_i |\alpha_i| |F_n|^{-1} \left\| \int_{F_n} T_{h^{-1}} f_i\, dm_L(h) - \int_{g_i^{-1} F_n} T_{h'^{-1}} f_i\, dm_L(h') \right\|_Y \\ &\leq A \sum_i |\alpha_i| \frac{|F_n \triangle g_i^{-1} F_n|}{|F_n|} \|f_i\|_Y \;\to 0 \end{aligned}$$

as $n \to \infty$ by the Følner property. Since the operators T_g are uniformly bounded, we conclude that $A_n z \to 0$ in norm for all $z \in L_0$. If z is also in $\mathrm{Fix}(T_G)$, we have $A_n z = z$ for all $n \in \mathbb{N}$ such that $A_n z \to z$ in norm as $n \to \infty$. By uniqueness of the limit, we conclude that $z = 0$.
With this direct sum, as well as with the closedness of the spaces $\mathrm{Fix}(T_G)$ and L_0, it is not hard to see that Z is closed.
Assume now that there is some $h \in Y^*$ vanishing on Z, i.e. $\langle z, h\rangle_{Y,Y^*} = 0$ for all $z \in Z$. In particular, this then holds for all $z' \in L_0$, which is equivalent to

$$\begin{aligned} & \langle f - T_g f, h\rangle_{Y,Y^*} = 0 \;\forall\, f \in Y,\, g \in G \\ \Leftrightarrow\; & \langle f, (I - T_g^*) h\rangle_{Y,Y^*} = 0 \;\forall\, f \in Y,\, g \in G \\ \Leftrightarrow\; & (I - T_g^*) h = 0 \;\forall\, g \in G \\ \Leftrightarrow\; & h \in \mathrm{Fix}(T_G^*). \end{aligned}$$

Moreover, h has to vanish on $\mathrm{Fix}(T_G)$ which - by assumption - separates $\mathrm{Fix}(T_G^*)$. Since $h \in \mathrm{Fix}(T_G^*)$, we must have $h = 0$ on the whole space Y. By the Hahn-Banach Theorem ([44], Corollary III.1.9), this implies that Z is dense in Y. But Z was closed, so we finally arrive at $Z = Y$.

(vi) $\Rightarrow$ (i) By assumption, we can rewrite each $f \in Y$ uniquely as $f = u + v$, where $u \in \mathrm{Fix}(T_G)$ and $v \in L_0$. We have to show that $u \in C_f$.
To do so, we define first the functions $\Phi_{f,h}(g) := \langle T_{g^{-1}} f, h\rangle_{Y,Y^*}$ for $f \in Y$ and $h \in Y^*$. Since for all $g \in G$,

$$|\Phi_{f,h}(g)| \le \|T_{g^{-1}} f\|_Y \cdot \|h\|_{Y^*} \le A\|f\|_Y \cdot \|h\|_{Y^*},$$

we obtain that $\Phi_{f,h} \in L^\infty(G)$. Furthermore, for all $f \in Y$, $h \in Y^*$ and every $g_0 \in G$,

$$\begin{aligned} \Phi_{f-T_{g_0}f,h}(g) &= \langle T_{g^{-1}} f, h\rangle_{Y,Y^*} - \langle T_{g^{-1}g_0} f, h\rangle_{Y,Y^*} \\ &= \Phi_{f,h}(g) - \Phi_{f,h}(g_0^{-1} g) \\ &= \Phi_{f,h}(g) - L_{g_0}\Phi_{f,h}(g). \end{aligned} \tag{4.7}$$

Take now some left-invariant mean m on $L^\infty(G)$. We claim that $m(\Phi_{v,h}) = 0$ for each $h \in Y^*$. Since $v \in L_0$, for an arbitrary $\varepsilon > 0$, there are finitely many $f_i \in Y$, $\alpha_i \in \mathbb{C}, g_i \in G$ so that with

$$\psi := \sum_i \alpha_i (f_i - T_{g_i} f_i),$$

we have $\|v - \psi\|_Y < \varepsilon$. It follows then from (4.7) that $m(\Phi_{\psi,h}) = 0$. Hence, for all $h \in Y^*$,

$$\begin{aligned} |m(\Phi_{v,h})| &\le \|\Phi_{v,h} - \Phi_{\psi,h}\|_{L^\infty(G)} \\ &\le A\,\|v - \psi\|_Y \cdot \|h\|_{Y^*} < \varepsilon\, A\, \|h\|_{Y^*}. \end{aligned}$$

With $\varepsilon \to 0$ we conclude that indeed, $m(\Phi_{v,h}) = 0$. Assume now that $u \notin C_f$. Since the latter set is closed and convex, by the Hahn-Banach-Theorem (cf. [44], VIII 2.12), we can find some $h \in Y^*$ as well as some $\varepsilon_0 > 0$ such that for all $f' \in C_f$

$$\mathrm{Re}\,\langle u, h\rangle_{Y,Y^*} \ge \varepsilon_0 + \mathrm{Re}\,\langle f', h\rangle_{Y,Y^*}.$$

Using $m(\Phi_{u,h}) = \langle u, h\rangle$ and $m(\Phi_{v,h}) = 0$, we thus obtain

$$\begin{aligned} \mathrm{Re}\, m(\Phi_{f,h}) &= \mathrm{Re}\, m(\Phi_{u,h}) + \mathrm{Re}\, m(\Phi_{v,h}) \\ &= \mathrm{Re}\, m(\langle u, h\rangle_{Y,Y^*} \cdot \mathbb{1}) \\ &= m(\mathrm{Re}\,\langle u, h\rangle_{Y,Y^*} \cdot \mathbb{1}) \\ &\overset{\text{pos. of m}}{>} m(\mathrm{Re}\,\Phi_{f,h}) \\ &= \mathrm{Re}\, m(\Phi_{f,h}), \end{aligned}$$

which is a contradiction. Hence, $u \in C_f$ as claimed.
To see uniqueness, take some $u \neq u' \in \mathrm{Fix}(T_G)$. Since Y can be written as a direct sum of $\mathrm{Fix}(T_G)$ and L_0, the decomposition $f = u + v$ is unique, which implies that $f - u' \notin L_0$. If we assume that $u' \in C_f$, then it could be approximated in norm by finite convex sums $\sum_i \alpha_i T_{g_i} f$, where $g_i \in G$. But this would imply that $f - u'$ can be approximated by some expression $\sum_i \alpha_i (f - T_{g_i} f)$ and hence $f - u' \in L_0$. This is a contradiction and there is only one element (namely u) in $\mathrm{Fix}(T_G) \cap C_f$. □

In the following, we assume that a σ-compact, amenable group acts on a σ-finite measure space $(X, \mathcal{B}, \mu)$ by measure preserving transformations and Y will be chosen to be some L^p-space over X, where $1 \leq p < \infty$. To see that this is a special case of the situation in Theorem 4.3, we prove the following.

Proposition 4.4
Assume that a σ-compact group G acts on a σ-finite measure space $(X, \mathcal{B}, \mu)$ by measure preserving transformations. Then G acts on $L^p(X)$ $(1 \leq p < \infty)$ weakly measurably as a uniformly bounded family $\{T_g\}$ of linear operators defined by

$$T_g : L^p(X, \mathcal{B}, \mu) \to L^p(X, \mathcal{B}, \mu) : (T_g f)(x) := f(\pi(g^{-1}, x)) = f(g^{-1} \cdot x),$$

where π denotes the action of G on X as in Definition 3.11.

Proof
We have to check (i)-(iv) from Definition 4.1. For (i), note that it is clear that the T_g are linear operators on $L^p(X)$ and by the fact that G preserves μ,

$$\|T_g f\|^p_{L^p(X)} = \int_X |f(g^{-1}x)|^p \, d\mu(x) = \int_X |f(x)|^p \, d\mu(x) = \|f\|^p_{L^p(X)},$$

so that $\|T_g\| = 1$ for all $g \in G$. Furthermore, it is a straightforward calculation to verify the validity of (ii) and (iii).
We denote the Borel σ-algebra on $\mathbb{C}$ by $\mathcal{B}(\mathbb{C})$. To prove (iv), note that by the measurability of π and the continuity of the inversion, the function $g \mapsto \int_X F(g^{-1}, x) \, d\mu(x)$ is $\mathcal{J}$-$\mathcal{B}(\mathbb{C})$ measurable if $F(g, x)$ is of the form $F(g, x) := f(\pi(g, x_0))h(x)$ for fixed $f \in L^p(X)$, $h \in L^p(X)^*$ and $x_0 \in X$. A simple approximation argument shows that this is also true for $F(g, x) := f(gx)h(x)$ with $f \in L^p(X)$ and $h \in L^p(X)^*$. In light of that, we conclude that for each $f \in L^p(X)$ and every $h \in L^p(X)^*$, the function

$$\Phi_{f,h}(g) := \langle T_g f, h \rangle_{L^p, (L^p)^*}$$

is $\mathcal{J}$-$\mathcal{B}(\mathbb{C})$ measurable as claimed. □

We now establish the mean ergodic theorem in the situation described in the above Proposition 4.4. The key step is the following lemma.

Lemma 4.5
Let $(X, \mathcal{B}, \mu)$ be a σ-finite measure space and $f \in L^p(X, \mathcal{B}, \mu)$ for some $1 < p < \infty$. As usual, a σ-compact, amenable group G acts on X by measure preserving transformations. For $g \in G$, we define the operators T_g as described in Proposition 4.4.
Then the set

$$C_f := \overline{\text{co}}\{T_g f \mid g \in G\}$$

is weakly compact.
The analogous statement for $p = 1$ is true if $\mu(X) < \infty$.

Proof
We prove first that for all $p \geq 1$, the set C_f is bounded. Indeed, by the fact that G preserves the measure μ, one obtains that $\|T_g f\|_{L^p(X)} = \|f\|_{L^p(X)}$ for all $g \in G$. Hence, if h is a convex

combination of translates $T_g f$, we have $\|h\|_{L^p(X)} \leq \|f\|_{L^p(X)}$ and this property also holds if h is the (strong) limit of such convex combinations. So $\|h\|_{L^p(X)} \leq \|f\|_{L^p(X)}$ for all $h \in C_f$. By the reflexivity of the L^p-spaces and the Banach-Alaoglu Theorem, the statement of the lemma follows for the cases $1 < p < \infty$.
Now, if $p = 1$, we will use the fact that a subset $K \subseteq L^1(X)$ is weakly compact if and only if it is bounded and if for any non-increasing sequence $\{E_n\}$ of sets in $\mathcal{B}$ with $\cap_n E_n = \emptyset$ one obtains

$$\lim_{n\to\infty} \int_{E_n} f(x)\, d\mu(x) = 0$$

uniformly for $f \in K$ (cf e.g. [10], Proposition IV.8.9). We have already seen that C_f is bounded for $f \in L^1(X)$.
Moreover, note that for any $\varepsilon > 0$, there is some $\delta > 0$ such that

$$\int_A |f(x)|\, d\mu(x) < \varepsilon$$

whenever $A \in \mathcal{B}$ with $\mu(A) < \delta$. To see this, take some $\tilde{f} \in L^1(X) \cap L^\infty(X)$ with $\|f - \tilde{f}\|_{L^1} < \varepsilon/2$. We now put $\delta := \varepsilon/2\|\tilde{f}\|_{L^\infty}$ and compute

$$\int_A |f(x)|\, d\mu(x) \leq \int_X |f(x) - \tilde{f}(x)|\, d\mu(x) + \int_A |\tilde{f}(x)|\, d\mu(x) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

for all $A \in \mathcal{B}$ with $\mu(A) < \delta$.
So if $\{E_n\}$ is a non-increasing sequence of measurable sets with $\cap_n E_n = \emptyset$, we have with $\mu(X) < \infty$ that $\lim_{n\to\infty} \mu(E_n) = 0$. By the above criterion and the fact that $\mu(g^{-1}E_n) = \mu(E_n)$ for every $n \in \mathbb{N}$ and each $g \in G$, we conclude that indeed

$$\begin{aligned}
\lim_{n\to\infty} \int_{E_n} T_g f(x)\, d\mu(x) &= \lim_{n\to\infty} \int_{E_n} f(g^{-1}x)\, d\mu(x) \\
&= \lim_{n\to\infty} \int_{g^{-1}E_n} f(x)\, d\mu(x) \\
&= 0
\end{aligned}$$

uniformly in g. So by the linearity of the limit, the same is true for convex combinations of elements $T_g f$, $g \in G$. Again, by simple approximation of an arbitrary $h \in C_f$ by convex combinations of left translates $T_g f$ we conclude that

$$\lim_{n\to\infty} \int_{E_n} h(x)\, d\mu(x) = 0$$

uniformly in $h \in C_f$. Thus, the set $C_f \subseteq L^1(X)$ is weakly compact. □

Remark
For $p = \infty$ on an infinite measure space, there is no similar statement.

Putting all our results together, we finally arrive at the following corollary.

Corollary 4.6 (Mean ergodic theorem)
Let G be a σ-compact, amenable group acting on a σ-finite measure space $(X, \mathcal{B}, \mu)$ by measure preserving transformations. Further, let $\{F_n\}$ be a Følner sequence in G and assume that $1 < p < \infty$. Then, for every $f \in L^p(X)$, there is a G-invariant $f^* \in L^p(X)$ such that the ergodic averages $A_n f$, defined as

$$(A_n f)(x) := |F_n|^{-1} \int_{F_n} f(gx)\, dm_L(g)$$

converge to f^* in $L^p(X)$ as $n \to \infty$. In the case $p = 1$, the same statement holds true if $\mu(X) < \infty$.

Proof
Let $f \in L^p(X)$ be given. By Proposition 4.4, G acts on $L^p(X)$ weakly measurably as a uniformly bounded family $\{T_g\}$ of linear operators and we have seen in Lemma 4.5 that the set $C_f := \overline{\text{co}}\{T_g f \mid g \in G\}$ with

$$T_g : L^p(X) \to L^p(X) : (T_g f)(x) := f(g^{-1}x) \quad (g \in G)$$

is weakly compact. This allows us to apply Theorem 4.2 to derive that there is some $f^* \in \text{Fix}(T_G) \cap C_f$. By the Abstract Mean Ergodic Theorem 4.3, this is equivalent to the fact that $A_n \to P$ (strong operator topology), where P is a bounded projection on $\text{Fix}(T_G)$. By uniqueness of the fixed point (again Theorem 4.3), we have $Pf = f^*$ and thus $A_n f \to f^*$ in norm. Since f^* is a fixed point in C_f, it is indeed G-invariant. □

Remark
In his work (cf. [31]), Nagel follows the same path in order to establish the mean ergodic theorem in the more abstract setting of operator semigroups acting on some Banach space Y. He proves analoga (Theorems 1.2 and 1.7) to Theorem 4.3 and concludes by a fixed point argument (cf. [8], Theorem 1) that each amenable semigroup $\mathcal{H}$ of uniformly bounded linear operators on Y with the property that for each $\xi \in Y$, the set $\mathcal{H}\xi$ is weakly relatively compact ($\sigma(Y, Y^*)$-topology), must be mean ergodic. A similar result, dealing with the convergence of ergodic averages as topological nets can be found in [25], Theorem 6.4.1.

One might raise the question whether the mean ergodicity can be established on an infinite measure space also in the case $p = 1$. The answer is negative, as the following example shows.

Example 4.7
Consider the space $X = \mathbb{R}$, endowed with the Borel σ-algebra $\mathcal{B}(\mathbb{R})$ and the Lebesgue measure $\mathcal{L}$. Furthermore, let $G = \mathbb{Z}$, i.e. G acts on $L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{L})$ by the measure preserving transformations $(T_k f)(x) := f(x-k)$. We take $f := \mathbb{1}_{[0,1]}$ as the characteristic function of the unit interval and use the Følner sequence $\{F_n\}$ defined by $F_n := \{-(n-1), -(n-2), \ldots, 0\}$

for $n \in \mathbb{N}$. Then, since m_L is the counting measure,

$$\begin{aligned} A_n f(x) := |F_n|^{-1} \int_{F_n} (T_{k^{-1}} f)(x)\, dm_L(k) &= \frac{1}{n} \cdot \sum_{k=0}^{n-1} \mathbb{1}_{[0,1]}(x-k) \\ &= \frac{1}{n} \cdot \sum_{k=0}^{n-1} \mathbb{1}_{[k,k+1]}(x) \\ &= \frac{1}{n} \cdot \mathbb{1}_{[0,n]}(x). \end{aligned}$$

It is clear that $A_n f \to 0$ pointwise, but

$$\|A_n f\|_{L^1(\mathbb{R})} = \frac{1}{n} \int_{\mathbb{R}} \mathbb{1}_{[0,n]}(x)\, dx = 1$$

for every $n \in \mathbb{N}$ and thus it is not possible that $A_n f \to 0$ in norm. It follows from this that without weak compactness of the set C_f, we cannot expect mean ergodicity of the action of G on X.
One could also argue as follows: the space $\mathrm{Fix}(T_G)$ consists of all periodic (length of period equal to one) equivalence classes in $L^1(\mathbb{R})$ and thus, $\mathrm{Fix}(T_G) = \{0\}$. So, it is clear that $\mathrm{Fix}(T_G)$ cannot separate $\mathrm{Fix}(T_G^*)$ which contains all constant functions on $\mathbb{R}$ and hence has dimension greater or equal than one. By Theorem 4.3, this is equivalent to the fact that A_n does not converge in the strong operator topology.

Although most of the results presented in this chapter have been well-known for several years, there are also recent developments in this area. We mention the work of Ghaffari (cf. [16], Theorem 1) who proved a mean ergodic theorem in the following setting. Take a locally compact Hausdorff semitopological semigroup S which acts as a family of uniformly bounded linear operators on some complex Banach space Y. Assume further that for each $\xi \in Y$, the map $s \mapsto \xi s$ is continuous (with respect to the natural topologies on S and on Y). Denote by $\mathcal{M}(S)$ the Banach algebra of all bounded, regular Borel measures on S with total variation norm and convolution $*$ as multiplication and let $\mathcal{M}_0(S)$ be the semigroup of all probability measures in $\mathcal{M}(S)$. In this context, amenability means that there is some asymptotically $\mathcal{M}_0(S)$-invariant net μ_α in $\mathcal{M}_0(S)$, i.e. $\|\mu * \mu_\alpha - \mu_\alpha\| \xrightarrow{\alpha} 0$ for every $\mu \in \mathcal{M}_0(S)$. Now define for each α the averages (expectations)

$$E_\alpha(\xi) = \int_S \xi s\, d\mu_\alpha(s),$$

where $\xi \in Y$. Then the following theorem holds.

Theorem 4.8 (Ghaffari, 2007)
Assume the situation outlined above. If one takes $\xi \in Y$ *such that the set* $C_\xi := \overline{\mathrm{co}}\{\xi s \mid s \in S\}$ *is weakly compact, then* $E_\alpha(\xi)$ *converges to a* S*-fixed point in* C_ξ*; this fixed point is unique in* C_ξ *and therefore it is independent of the choice of the net* μ_α*.*
If C_ξ *is weakly compact for all* $\xi \in Y$*, then with* $Y_1 := \mathrm{Fix}(S) := \{\xi \in Y \mid \xi s = \xi \text{ for all } s\}$ *and* $Y_2 := \overline{lin}\{\xi - \xi s \mid s \in S, \xi \in Y\}$*, one obtains* $Y = Y_1 \oplus Y_2$ *and* E_α *converges strongly to the projection of* Y *onto* Y_1*.*

Proof
See [16], Theorem 1. □

5 From the maximal inequality to pointwise convergence

We now explain the idea of a so called maximal ergodic theorem which describes a well trodden path in the world of individual ergodic theorems. Indeed, the proofs of all pointwise results mentioned in the introduction of this thesis (Chapter 1) are based on that conception. The crux is the verification of a so-called *L^p-maximal inequality* (cf. Definition 5.2) for the *maximal function* as determined by Definition 5.1. The main Theorem 5.3 of this chapter shows that in the usual setting, the validity of an L^p-maximal inequality for the maximal function is sufficient for pointwise almost everywhere convergence of the abstract ergodic averages along a Følner sequence. Using the Mean Ergodic Theorem 4.6, the pointwise almost everywhere convergence can be verified by elementary methods on a dense subspace of $L^p(X)$ in the case $p > 1$. The *Banach principle* (Lemma 5.4) shows that given an L^p-maximal inequality, the set of all functions with pointwise almost everywhere convergent ergodic averages is closed, such that the Pointwise Convergence Theorem 5.3 must hold on the whole space $L^p(X)$.
As in the proof of the Mean Ergodic Theorem 4.6, the case $p = 1$ requires some special treatment. Assuming here that the maximal function satisfies an L^1- *and* an L^2-maximal inequality, we can prove the Individual Ergodic Theorem 5.3 by means of an interpolation argument.

Definition 5.1
Let $(X, \mathcal{B}, \mu)$ be a σ-finite measure space, $1 \leq p < \infty$ and $\{A_n\}$ be a bounded sequence of linear operators on $L^p(X)$. For $n \in \mathbb{N}$ and $f \in L^p(X)$, define the function $M_n f$ as

$$(M_n f)(x) := \max_{1 \leq j \leq n} \Big|(A_j f)(x)\Big|, \quad x \in X.$$

The maximal function Mf w.r.t. $\{T_n\}$ is then defined by

$$(Mf)(x) := \sup_{n \in \mathbb{N}} \Big|(A_n f)(x)\Big|, \quad x \in X.$$

Remark
We will also call M a maximal *operator* on $L^p(X)$ with respect to $\{A_n\}$. Note that M is sublinear, but not necessarily linear. Also, for $f \in L^p(X)$, Mf might not be p-integrable: turning back to our Example 4.7 with $p = 1$, $X = \mathbb{R}$ and $f = \mathbb{1}_{[0,1]}$, note that $M_n f := \sum_{j=1}^{n} j^{-1} \mathbb{1}_{[j-1,j]}$. Thus, we have $Mf = \sum_{j=1}^{\infty} j^{-1} \mathbb{1}_{[j-1,j]}$ and one obtains

$$\|Mf\|_{L^1(X)} = \sum_{j=1}^{\infty} \frac{1}{j} = \infty.$$

In the following, we will always consider M as the maximal operator with respect to the sequence $\{A_n\}$ of the ergodic averages, interpreted as contractions on $L^p(X)$. To see that this makes sense, use Propostition 4.4 and note that since T_g preserves μ for every $g \in G$,

$$\begin{aligned} \|A_n f\|_{L^p(X)} &\leq |F_n|^{-1} \int_{F_n} \|T_{h^{-1}} f\|_{L^p(X)} \, dm_L(h) \\ &= \|f\|_{L^p(X)}. \end{aligned}$$

Definition 5.2 (L^p-maximal inequality)
Let G be some σ-compact amenable group which acts on a σ-finite measure space by measure preserving transformations. Further, let $1 \leq p < \infty$ be given. For a Følner sequence $\{F_n\}$, consider the ergodic averages $\{A_n\}$ as a sequence of contractions on $L^p(X)$. We say that the corresponding maximal operator M satisfies an L^p-maximal inequality resp. is of weak type (p,p) if there is a constant $C > 0$ such that for all $f \in L^p(X)$ and every $\lambda > 0$

$$\mu(\{x \,|\, (Mf)(x) > \lambda\}) \leq \frac{C}{\lambda^p} \|f\|^p_{L^p(X)}.$$

Remark
If we can show an L^p-maximal inequality, we also say that the sequence $\{A_n\}$ satisfies a so-called *dominated* ergodic theorem.

We now prove the main theorem of this section stating that the validity of an L^p-maximal inequality implies the pointwise ergodic theorem.

Theorem 5.3
Let $(X, \mathcal{B}, \mu)$ be some σ-finite measure space and let G be some σ-compact amenable group acting on it by measure preserving transformations. Further, let $f \in L^p(X, \mathcal{B}, \mu)$ for some $1 < p < \infty$. For a Følner sequence $\{F_n\}$, define the maximal functions $M_n f$ and Mf with respect to the ergodic averages $\{A_n\}$ as in Definition 5.1.
If there is some constant $C > 0$ which only depends on p and the sequence $\{F_n\}$ such that for any $\lambda > 0$, the inequality

$$\mu(\{x \,|\, Mf(x) > \lambda\}) \leq \frac{C}{\lambda^p} \|f\|^p_{L^p(X)}$$

holds, then there is some G-invariant $f^* \in L^p(X)$ such that

$$A_n f(x) = |F_n|^{-1} \int_{F_n} f(gx) \, dm_L(g) \to f^*(x) \text{ for } n \to \infty \text{ and a.e. } x \in X.$$

In particular, if the maximal operator M is of weak type (p,p), then the individual ergodic theorem holds for the whole space $L^p(X)$.
In the case $p = 1$, the individual ergodic theorem holds on $L^1(X)$ if M is of weak type $(1,1)$ and of weak type $(2,2)$.

It turns out that the conclusion of Theorem 5.3 can be proven by rather standard methods on a dense subspace of $L^p(X)$. Since the pointwise almost everywhere convergence is not induced by some topology, we will need an additional tool to pass from the dense space to its closure (which is of course the whole space $L^p(X)$). Here, the maximum inequality will be used.

Lemma 5.4 (Banach's principle)
With the same assumptions as in Theorem 5.3, the set

$$K := \{f \in L^p(X) \,|\, (A_n f)_{n\in\mathbb{N}} \text{ is a.e. convergent}\}$$

is a closed subspace of $L^p(X, \mathcal{B}, \mu)$.

Proof (of Lemma 5.4)
We follow the path of [14], Proposition 10.9. For every $n \in \mathbb{N}$, the operator A_n is linear on $L^p(X)$ and by linearity of the (a.e.)-limit, K is a subspace. To see that it is closed, we choose $f \in L^p(X)$ and $g \in K$. Note that for each $k, l \in \mathbb{N}$, we have by the triangle inequality

$$|A_k f - A_l f| \leq |A_k(f-g)| + |A_k g - A_l g| + |A_l(g-f)| \leq 2 \cdot M(f-g) + |A_k g - A_l g| \quad (5.1)$$

in the pointwise sense. The fact that $g \in K$ implies that $(A_n g)$ converges almost everywhere. By a simple calculation, this can be reformulated as

$$\limsup_{k,l\to\infty} |A_k g(x) - A_l g(x)| = \inf_{n\in\mathbb{N}} \sup_{k,l\geq n} |A_k g(x) - A_l g(x)| = 0$$

for almost every $x \in X$. So taking the limsup in (5.1), we derive

$$h := \limsup_{k,l\to\infty} |A_k f - A_l f| \leq 2 \cdot M(f-g)$$

almost everywhere. In light of that, the set of all $x \in X$ with $h(x) > 2\lambda$ is contained in the set of all $x \in X$ with $M(f-g)(x) > \lambda$ and by the maximal inequality

$$\mu(\{x \,|\, h(x) > 2\lambda\}) \leq \mu(\{x \,|\, M(f-g)(x) > \lambda\}) \leq \frac{C}{\lambda^p} \|f-g\|^p_{L^p(X)}. \quad (5.2)$$

Now we choose $f \in \overline{K}$. Then for any $\varepsilon > 0$ we can find some $g \in K$ such that $\|f-g\|_{L^p(X)} < \varepsilon$. Using (5.2) and sending $\varepsilon \to 0$, we conclude that $\mu(\{x \,|\, h(x) > 2\lambda\}) = 0$ for all $\lambda > 0$. But this implies that $h = 0$ a.e., which in turn means by the definition of h that $(A_n f)$ is a.e. convergent. Thus, $f \in K$ and the lemma is proven. □

We now turn to the proof of Theorem 5.3. It will be necessary to treat the cases $p > 1$ and $p = 1$ separately.

Proof (of Theorem 5.3)
Case $p > 1$: Let $f \in L^p(X)$ be given. As we have shown in the general Mean Ergodic Theorem 4.6, the averages $(A_n f)$ converge *in norm* to some G-invariant $f^* \in L^p(X)$. So *if* this convergence holds also pointwise almost everywhere, the limit must be the same f^*. So define the operator $Pf := f^*$ on $L^p(X)$. It follows from Theorem 4.3 that P is a projection on the fixed space $\mathrm{Fix}(T_G)$ and that we can write

$$L^p = \mathrm{ran}(P) \oplus \mathrm{ran}(I-P) = \mathrm{Fix}(T_G) \oplus L_0,$$

where I is the identity operator and L_0 denotes the (strong) closure of the linear hull of linear combinations $f - L_g f$ with $f \in L^p(X)$ and $g \in G$.
For $f \in \mathrm{Fix}(T_G)$, we note that $T_g f = f$ for all $g \in G$ such that fixing g and taking a representant of f, we have $f(g^{-1}x) = f(x)$ except for a null set $N \subseteq X$. It then follows from Fubini's Theorem that for each $x \in X \setminus N$, we must have $f(g^{-1}x) = f(x)$ for m_L-almost every $g \in G$ and this shows that $A_n f = f$ μ-almost everywhere. Consequently, for $f \in \mathrm{Fix}(T_G)$ the ergodic averages converge trivially since they μ-essentially reproduce f.

In light of that, it is sufficient to consider the case when $f \in L_0$, i.e. $Pf = 0$. Moreover, by Banach's principle (5.4), we can restrict our attention to the norm dense subspace (cf. [44], Lemma II.4.1)

$$L_0^* := \mathrm{lin}\{f - T_g f \mid f \in L^p(X) \cap L^\infty(X)\,, g \in G\}$$

of L_0. So let $h \in L_0^*$ be of the form $f - T_{g_0} f$ for some $f \in L^p(X) \cap L^\infty(X)$ and some $g_0 \in G$. Then by the left invariance of the Haar measure

$$\begin{aligned}(A_n h)(x) = |F_n|^{-1} \int_{F_n} h(gx)\, dm_L(g) &= |F_n|^{-1} \int_{F_n} [f(gx) - f(g_0^{-1} gx)]\, dm_L(g) \\ &= |F_n|^{-1} \Big(\int_{F_n} f(gx)\, dm_L(g) - \int_{g_0^{-1} F_n} f(gx)\, dm_L(g) \Big).\end{aligned}$$

Taking absolute values, this reduces to

$$|(A_n h)(x)| \le |F_n|^{-1} \int_{g_0^{-1} F_n \triangle F_n} |f(gx)|\, dm_L(g) \le |F_n|^{-1} |F_n \triangle g_0^{-1} F_n| \cdot \|f\|_{L^\infty(X)},$$

and this expression converges to 0 uniformly in x since $\{F_n\}$ is a Følner sequence. Clearly, the same holds true for linear combinations of such functions. Hence, putting our results together, we have shown that for every $f \in L^p(X)$ with $p > 1$, the ergodic averages $(A_n f)_n$ converge to $f^* := Pf$ μ-almost everywhere.

Case $p = 1$: We start by showing that the operator P on $L^2(X)$, defined as in the first case above is also a well-defined positive contraction on $L^1(X)$. To do this, consider first some $f \in L^1(X) \cap L^\infty(X)$. It follows that then also $f \in L^2(X)$. Thus, by the general Mean Ergodic Theorem 4.6, $A_n f \to f^*$ in L^2-norm and $Pf := f^*$ is well defined on $L^2(X)$ (see the case $p > 1$ above). Moreover, it follows from Hölder's inequality that for every measurable set $B \subseteq X$ with finite μ-measure,

$$\begin{aligned}\lim_{n\to\infty} \Big| \int_B \Big(A_n f(x) - Pf(x) \Big)\, d\mu(x) \Big| &\le \lim_{n\to\infty} \int_B \Big| A_n f(x) - Pf(x) \Big|\, d\mu(x) \\ &\le \mu(B)^{1/2} \lim_{n\to\infty} \|A_n f - Pf\|_{L^2(X)} = 0.\end{aligned}$$

Hence, for all $B \in \mathcal{B}$ with $\mu(B) < \infty$, one obtains that for any sequence $\varepsilon_k \to 0$, there is a sequence $(n_k)_k \subseteq \mathbb{N}$ with

$$\begin{aligned}
\Big|\int_B Pf(x)\,d\mu(x)\Big| &\leq \varepsilon_k + \int_B |A_{n_k} f(x)|\,d\mu(x) \\
&= \varepsilon_k + |F_{n_k}|^{-1} \int_B \Big| \int_{F_{n_k}} f(gx)\,dm_L(g)\Big|\,d\mu(x) \\
&\overset{\text{Fubini}}{\leq} \varepsilon_k + |F_{n_k}|^{-1} \int_{F_{n_k}} \int_B |f(gx)|\,d\mu(x)\,dm_L(g) \\
&\leq \varepsilon_k + |F_{n_k}|^{-1} \int_{F_{n_k}} \int_X |f(gx)|\,d\mu(x)\,dm_L(g) = \varepsilon_k + \|f\|_{L^1(X)},
\end{aligned}$$

where the last equality is due to the fact that the action of G on $(X, \mathcal{B}, \mu)$ preserves the measure. So, in particular, the latter inequality holds for non-negative equivalence classes $f \in L^1(X) \cap L^\infty(X)$, which with $\varepsilon_k \to 0$ implies

$$\int_B P|f|(x)\,d\mu(x) \leq \|f\|_{L^1(X)} \tag{5.3}$$

for each measurable $B \subseteq X$ of finite measure. Note that by the positivity of the operators A_n, P is also positive and thus $|Pf| \leq P|f|$ for $f \in L^1(X) \cap L^\infty(X)$. Using this fact and applying inequality (5.3) to an increasing sequence of measurable sets $\{B_n\}$ of finite measure such that $\cup_n B_n = X$ (which exists by the σ-finiteness of μ), we yield by the monotone convergence theorem

$$\|Pf\|_{L^1(X)} = \int_X |Pf|\,d\mu \leq \int_X P|f|\,d\mu \overset{(5.3)}{\leq} \|f\|_{L^1(X)} \tag{5.4}$$

for all $f \in L^1(X) \cap L^\infty(X)$. But the latter space is dense in $L^1(X)$ (cf. [44], Lemma II.4.1) so that by (5.4) the definition of P can be extended to a contraction on the whole space $L^1(X)$. Since the L^1-convergence preserves the sign, P is a positive contraction on $L^1(X)$. We claim further that the following statements hold true even for $p = 1$.

(i) P is a projection on $L^1(X)$,

(ii) $\operatorname{ran}(P) \subseteq \operatorname{Fix}(T_G)$,

(iii) Define $L_0^* := \overline{\operatorname{lin}}^{\|\cdot\|_{L^2}}\{h - T_g h \,|\, h \in L^2(X),\, g \in G\}$ and $L_1^* := L_0^* \cap L^1(X)$. Then L_1^* is L^1-dense in $\ker(P)$.

Note that if we can verify the claims (i)-(iii), similar arguments as in the case $p > 1$ can be used to prove the individual ergodic theorem. Since P is a projection by (i), we can write each $f \in L^1(X)$ as $f = f_1 + f_2$ with $f_1 := Pf \in \operatorname{ran}(P)$ and $f_2 := f - Pf \in \ker(P)$. By claim (ii), f_1 is a fixed function and therefore, $A_n f_1 = f_1$ for all $n \in \mathbb{N}$ (see above) and this implies that $A_n f_1$ converges pointwise almost everywhere to Pf. Claim (iii) shows that f_2 can be approximated in $L^1(X)$ by representants in L_1^*. Since we assumed the maximal operator M to be of weak type $(2,2)$, we conclude from the case $p = 2$ (see above) that the ergodic averages $A_n h$ converge to zero almost everywhere for every $h \in L_1^*$. But this space is L^1-dense in $\ker(P)$ and therefore, $\operatorname{Fix}(T_G) + L_1^*$ is dense in $L^1(X)$. The fact that the maximal operator is also of weak type $(1,1)$ allows us to apply Banach's Principle, Lemma 5.4, which implies that the pointwise ergodic theorem is valid on the whole space $L^1(X)$.

So let us prove the claims (i)-(iii).

To see (i), pick a function $f \in L^1(X)$ and find some sequence $\{f_n\} \subseteq L^1(X) \cap L^\infty(X) \subseteq L^2(X)$ such that $f_n \xrightarrow{n} f$ in $L^1(X)$. It follows from the Mean Ergodic Theorem 4.3 and Proposition 4.4 for $p = 2$ that $P^2 f_n = P f_n$ for all $n \in \mathbb{N}$. Since P is a contraction this implies $P^2 f = Pf$ and since f was arbitrary, P is a projection on $L^1(X)$.

For (ii), assume first that $f \in \operatorname{ran}(P)$, i.e. there is some $h \in L^1(X)$ such that $Ph = f$. Approximating h in $L^1(X)$ by some sequence $\{h_n\} \subseteq L^1(X) \cap L^\infty(X)$ we have by the Abstract Mean Ergodic Theorem 4.3 and Proposition 4.4 in the case $p = 2$ that $T_g(Ph_n) = Ph_n$ and using the continuity of P and T_g we observe that

$$\begin{aligned} \|T_g f - f\|_{L^1(X)} &= \|T_g Ph - Ph\|_{L^1(X)} \\ &\leq \limsup_{n\to\infty} \|T_g Ph - T_g Ph_n\|_{L^1(X)} + \limsup_{n\to\infty} \|Ph_n - Ph\|_{L^1(X)} = 0 \end{aligned}$$

for all $g \in G$, which means that $f \in \operatorname{Fix}(T_G)$.

We finally prove (iii). To do so, assume that $Pf = 0$ for $f \in L^1(X)$. As usual, we choose some sequence $\{f_n\} \subseteq L^1(X) \cap L^\infty(X)$ converging to f in $L^1(X)$. By the Mean Ergodic Theorem 4.3 and Proposition 4.4 in the case $p = 2$ we have $L_1^* \subseteq \ker(P)$. By the same arguments, we can write

$$f_n = f_n^{(1)} + f_n^{(2)}$$

with $f_n^{(1)}, f_n^{(2)} \in L^2(X)$ for all $n \in \mathbb{N}$. Noting that $Pf_n^{(2)} = 0$ we see that $f_n^{(2)} \in L_0^*$ for all $n \in \mathbb{N}$. Moreover, $f_n^{(1)} = Pf_n$ and by continuity of P we have $f_n^{(1)} \in L^1(X)$ for every $n \in \mathbb{N}$. But this implies that $f_n^{(2)} \in L^1(X)$ as well such that $f_n^{(2)} \in L_1^*$ for every $n \in \mathbb{N}$. Since P is a contraction and $Pf = 0$, we see that $f_n^{(1)} \xrightarrow{n} 0$ in $L^1(X)$ and therefore, the expression

$$\|f - f_n^{(2)}\|_{L^1(X)} \leq \|f - f_n\|_{L^1(X)} + \|f_n^{(1)}\|_{L^1(X)}$$

tends to zero as n tends to infinity. We conclude that each $f \in \ker(P)$ can be approximated in $L^1(X)$ by some sequence in L_1^*. This finishes the proof. □

6 The transfer principle

We have seen in the previous chapter that for establishing a pointwise ergodic theorem, it is sufficient to verify the validity of an L^p-maximal inequality, see Theorem 5.3. To do so, we use the *transfer principle* and prove a so-called *transfer inequality* (see Theorem 6.4). This method is a well developed technique allowing us to restrict our attention to the natural action $G \times G \to G$ of group multiplication instead of considering the action of G on the σ-finite measure space $(X, \mathcal{B}, \mu)$. Thus, the information *whether* an action of some amenable group G with Følner sequence $\{F_n\}$ on a measure space gives rise to pointwise a.e. convergence of the corresponding ergodic averages $A_n f$ is encoded in the intrinsic structures of the group and the sequence. This idea dates back to the first proofs of the classic pointwise ergodic theorem and was explicitly formulized by Calderon in [5]. Few years later, Emerson extended and generalized these results in [11] by proving the following theorem.

Theorem 6.1 (Transfer principle, Emerson 1974)
Let G be a σ-compact amenable group. Then for every Følner sequence $\{F_n\}$, the expression

$$(Sf)(g_0) := \sup_{n \in \mathbb{N}} \left| |F_n|^{-1} \int_{F_n} f(gg_0)\, dm_L(g) \right|$$

is a well defined sublinear operator S on $L^1_{loc}(G)$.
Suppose further that G acts on a σ-finite measure space $(X, \mathcal{B}, \mu)$ by measure preserving transformations. If S is of weak type (p, p) for some $1 \leq p < \infty$, then so is also the operator M on $L^p(X)$, defined as

$$(Mf)(x) := \sup_{n \in \mathbb{N}} \left| |F_n|^{-1} \int_{F_n} f(gx)\, dm_L(g) \right|.$$

Proof
See [11], Corollary 1. □

Hence, with the transfer principle, we can derive an L^p-maximal inequality by checking if the operator S on $L^1_{loc}(G)$ is of weak type (p, p); no information about the action $G \times X \to X$ is necessary. In the same paper, Emerson used this technique to derive the individual ergodic theorem along Følner sequences satisfying the Tempelman condition (3.2).

Theorem 6.2 (Individual ergodic theorem, Emerson 1974)
Let G be a σ-compact group acting on a σ-finite measure space $(X, \mathcal{B}, \mu)$ by measure preserving transformations. Further, let $\{F_n\}$ be a Følner sequence satisfying the Tempelman condition. If now $1 < p < \infty$ and if $f \in L^p(X)$, then there is a G-invariant $f^* \in L^p(X)$

such that the ergodic averages $A_n f$, defined as in Theorem 4.3, converge to f^* pointwise μ-almost everywhere. In addition to that we have $f^* = Pf$, where P is the mean ergodic projection on $L^p(X)$.
In the case $p = 1$ this result also holds if $\mu(X) < \infty$.

Proof
See [11], Theorem 4'. □

Remark
Combining Emerson's results (see Theorem 6.1) with the elaborations in Chapter 5, one can observe easily that the pointwise ergodic theorem must be valid on a (possibly non-finite) measure space also in the case $p = 1$.

However, the machinery given by Theorem 6.1 is based on quite involved ideas and is established by very sophisticated and technical calculations. In his work, Lindenstrauss uses the transfer principle in a more subtle way, showing only the things which are indispensable for the proof of the individual ergodic theorem. None the less, one can use Lindenstrauss' results in [29] to verify the conditions in Theorem 6.1, which provides a slightly (but not really) different proof of the pointwise ergodic theorem. We will do this in Corollary 8.8.
So let us describe Lindenstrauss' version of the transfer principle. He only consideres the case $p = 1$ in [29]; we provide an extension to the cases $1 < p < \infty$ on a σ-finite measure space. We start with the following lemma:

Lemma 6.3
Let G be a σ-compact amenable group with a Følner sequence $\{F_n\}$. Let $k \in \mathbb{N}$ and choose $\varepsilon > 0$. Then, for $\tilde{F}_k := \cup_{i=1}^k F_i$ there is some $n_k \in \mathbb{N}$ such that for $\overline{F}_k := \tilde{F}_k F_{n_k}$ we have

$$|\overline{F}_k| \leq (1+\varepsilon)|F_{n_k}|.$$

Proof
Let $\varepsilon > 0$ and $k \in \mathbb{N}$. As a finite union of compact sets, the set $\tilde{F}_k$ is compact as well. Hence, by the Følner property (3.5) there must be some element F_{n_k} of the Følner sequence such that

$$|\overline{F}_k| - |F_{n_k}| \leq |F_{n_k} \triangle \tilde{F}_k F_{n_k}| \leq \varepsilon |F_{n_k}|,$$

which proves the lemma. □

The key step to establish a sufficient condition for the validity of an L^p-maximal inequality is the following theorem.

Theorem 6.4 (Transfer principle, Lindenstrauss 2001)
Let G be a σ-compact, amenable group acting on a σ-finite measure space $(X, \mathcal{B}, \mu)$ by measure preserving transformations. Further, let $\varepsilon = 1$ and for each $k \in \mathbb{N}$, choose the sets $\tilde{F}_k, F_{n_k}$ and $\overline{F}_k$ as in Lemma 6.3, where $\{F_n\}$ is a Følner sequence of G.
Then the maximal operator $M = \sup_{n\in\mathbb{N}} A_n$ satisfies an L^p-maximal inequality for

$1 \le p < \infty$ if there is a constant $c > 0$ such that for any $f \in L^p(X)$, for each $k \in \mathbb{N}$, for every $\lambda > 0$ and for almost every $x \in X$, we have the transfer inequality

$$\left|\bigcup_{j=1}^{k} B_j\right| \le c\lambda^{-p} \int\limits_{\overline{F}_k} |f(gx)|^p \, dm_L(g), \tag{6.1}$$

where $B_j := B_j(x) := \{g \in F_{n_k} \mid |(A_j f)(gx)| \ge \lambda\}$ for $1 \le j \le k$.

Proof

Let $c > 0$ be such a constant and take $f \in L^p(X)$. We first fix $k \in \mathbb{N}$, $\lambda > 0$ as well as some $x \in X$ for which the function $f_x(g) := f(gx)$ is p-integrable over $\overline{F}_k$ (Note that by Fubini's theorem, the function $f_x(\cdot)$ is locally p-integrable on G for almost every $x \in X$). Then set

$$(M_k f)(x) := \max_{1 \le j \le k} |(A_j f)(x)|$$

as in Definition 5.1 and

$$D_k := \{x \in X \mid (M_k f)(x) \ge \lambda\}.$$

We see that

$$\begin{aligned} a \in \bigcup_{j=1}^{k} B_j \quad &\Leftrightarrow \quad a \in F_{n_k} \wedge \Big(\exists\, 1 \le j \le k : |(A_j f)(ax)| \ge \lambda\Big) \\ &\Leftrightarrow \quad a \in F_{n_k} \wedge (M_k f)(ax) \ge \lambda \\ &\Leftrightarrow \quad a \in F_{n_k} \wedge ax \in D_k, \end{aligned}$$

so that

$$\left|\bigcup_{j=1}^{k} B_j\right| = \int_{F_{n_k}} \mathbb{1}_{D_k}(gx) \, dm_L(g). \tag{6.2}$$

Note further that $\mu(D_k) < \infty$ since otherwise, we could find some $1 \le j_0 \le k$ such that $\|A_{j_0} f\|_{L^p} = \infty$, which clearly is a contradiction.

Hence, we can apply Fubini's Theorem and by using the fact that the action of G on X is measure preserving, we compute

$$\begin{aligned} \mu(D_k) \quad &= \quad \int_X \mathbb{1}_{D_k}(x) \, d\mu(x) \\ &= \quad \int_{F_{n_k}} |F_{n_k}|^{-1} \int_X \mathbb{1}_{D_k}(gx) \, d\mu(x) \, dm_L(g) \\ &\overset{\text{Fubini}}{=} \quad \int_X |F_{n_k}|^{-1} \int_{F_{n_k}} \mathbb{1}_{D_k}(gx) \, dm_L(g) \, d\mu(x). \end{aligned}$$

Using equality (6.2) and the assumption (6.1), one obtains further

$$\begin{aligned} \mu(D_k) \quad &\le \quad |F_{n_k}|^{-1} \int_X c\lambda^{-p} \int_{\overline{F}_k} |f(gx)|^p \, dm_L(g) \, d\mu(x) \\ &\overset{\text{Fubini}}{=} \quad c\lambda^{-p} |F_{n_k}|^{-1} \int_{\overline{F}_k} \int_X |f(gx)|^p \, d\mu(x) \, dm_L(g) \\ &\le \quad c\lambda^{-p} \frac{|\overline{F}_k|}{|F_{n_k}|} \|f\|^p_{L^p(X)}. \end{aligned} \tag{6.3}$$

As we have chosen the sets F_{n_k} and $\overline{F}_k$ according to Lemma 6.3 with $\varepsilon = 1$, it holds true that

$$\frac{|\overline{F}_k|}{|F_{n_k}|} \leq 2,$$

which with inequality (6.3) yields

$$\mu(D_k) \leq 2c\lambda^{-p}\|f\|^p_{L^p(X)}. \tag{6.4}$$

Finally, we note that the inequality $(Mf)(x) > \lambda$ implies that there must be some $k \in \mathbb{N}$ such that $(A_k f)(x) \geq \lambda$ which means that

$$\mu(\{x \,|\, (Mf)(x) > \lambda\}) \quad \leq \quad \mu(\cup_{k=1}^{\infty} D_k). \tag{6.5}$$

Now k was arbitrarily chosen and the sets $\{D_k\}$ are increasing, hence by taking the limit in (6.4) one finally obtains with inequality (6.5)

$$\begin{aligned} \mu(\{x \,|\, (Mf)(x) > \lambda\}) &\leq \lim_{k\to\infty} \mu(D_k) \\ &\leq \frac{2c}{\lambda^p}\|f\|^p_{L^p(X)}, \end{aligned}$$

which gives the desired L^p-maximal inequality. □

7 Decomposition of the group - deterministic case

By the previous two chapters, in order to prove a general pointwise ergodic theorem for σ-compact amenable groups one can make use of the transfer principle and verify inequality (6.1). We will do this for countable amenable groups in Theorem 7.1 by deterministic methods which are due to WEISS (see [43]). The key step is the proof of an abstract combinatorial lemma (7.2).

We approach the verification of the transfer inequality by the following simple observation: fix $f \in L^p(X)$ ($1 \leq p < \infty$), $\lambda > 0$ and $x \in X$ as in Theorem 6.4. Moreover, recall that the transfer inequality (6.1) involves sets $B_j \subseteq G$ of specific $b \in G$ such that $|A_j f(bx)| \geq \lambda$ (as usual A_j denotes the j^{th} ergodic average). So if $\{F_n\}$ is the corresponding Følner sequence in G, then by definition of B_j and Hölder's inequality ($q := p/(p-1)$, i.e. $p - p/q = 1$ for $1 < p < \infty$) we obtain for each $b \in B_j$

$$\begin{aligned}
\int_{F_j b} |f(gx)|^p \, dm_L(g) &\geq |F_j b|^{-p/q} \left(\int_{F_j b} |f(gx)| \, dm_L(g) \right)^p \\
&= |F_j b|^{-p/q} \left(\Delta(b) \int_{F_j} |f(g(bx))| \, dm_L(g) \right)^p \\
&= |F_j b|^{-p/q} \cdot [\Delta(b)]^p \cdot [|F_j| \cdot (A_j |f|)(bx)]^p \\
&\geq |F_j b|^{-p/q} \cdot [\Delta(b) \cdot |F_j|]^p \cdot \lambda^p \\
&= \lambda^p \cdot |F_j b|^{p-p/q} \\
&= \lambda^p \cdot |F_j b|, \qquad (7.1)
\end{aligned}$$

where Δ stands for the modular function $\Delta : G \to (0, \infty)$, defined so that for $g \in G$,

$$m_L(B \cdot g) = \Delta(g) \cdot m_L(B)$$

for all measurable sets $B \subseteq G$ (cf. [9], Section 1.4). As Δ is also a group homomorphism, this implies for all integrable functions f that

$$\int_X f(hg) \, dm_L(h) = \Delta(g)^{-1} \cdot \int_X f(h) \, dm_L(h).$$

Note that (7.1) is trivial if $p = 1$. Thus, to verify the condition of Theorem 6.4, one might try to cover the set $\overline{F}_k$ by right-translates $F_j b$ of Følner elements F_j with $b \in B_j$. Then, if we assure that

(1) two distinct translates have small overlaps (where the word 'small' has to be specified) and

(2) the total mass of these translates exceeds the left-hand side of (6.1),

we can use inequality (7.1) in an obvious manner to derive the transfer inequality (6.1). The classical covering lemmas of this kind show the existence of a countable *maximal* collection of *pairwise disjoint* translates $F_j b, b \in B_j$ which satisfies the above properties (1)(clear) and (2). Recall that pointwise ergodic theorems do in general not hold along arbitrary Følner sequences $\{F_n\}$ (cf. Theorem 3.13) and so far, we have not considered any restrictions on $\{F_n\}$. Indeed, one needs the growth condition on $\{F_n\}$ exactly at this point of our argumentation. Therefore, an additional requirement on the considered Følner sequence is indispensable for the validity of a general covering (resp. disjointification-)lemma. The most popular approach is the use of the Tempelman condition (3.2), as demonstrated e.g. in [11], [33], [34] and [43]. While WIENER applies a disjointification lemma in his proof of the classical $\mathbb{Z}$-case (cf. [45]), CALDERON uses assumptions on the Følner sequence which are stronger than amenability (see [4]).
Moreover, all these covering results can be obtained for arbitrary sequences of sets obeying some growth restriction; the sequence does *not* have to be a Følner sequence, i.e. as long as we work with some well-conditioned sequence, the statements in this and the next chapter may even hold true for non-amenable groups.

As pointed out in Chapter 3, Følner sequences with Tempelman condition do not necessarily exist in amenable groups. This is different for the Shulman condition (cf. Lemma 3.10). Hence, it is more appropriate to work with tempered sequences. A first result involving a Følner sequence with Shulman condition is due to eponym SHULMAN who managed to prove the pointwise convergence theorem in the L^2-case (cf. [39]). An elementary proof of the transfer inequality for discrete, σ-compact (and hence countable) groups, which we will present below, was given by WEISS (cf. [43], p. 251-254). The case for second countable, σ-compact amenable groups had been open for almost three decades, until one of WEISS' graduate students, LINDENSTRAUSS, solved the problem in 2001 by treating possible coverings as outcome of random variables (cf. [29], Lemma 2.1).
We now describe WEISS' proof of inequality (6.1) for a countable (discrete) amenable group G. In contrast to the techniques developed by LINDENSTRAUSS (see next chapter below), he builds his line of argumentation on deterministic and combinatorial arguments; no randomness is involved.

Theorem 7.1 (Transfer inequality in the countable case, cf. [43])
Let G be a countable ($|G| > 10$), discrete group with a tempered sequence $\{F_n\}$ (with constant C) of finite subsets of G. Further, G acts on a σ-finite measure space $(X, \mathcal{B}, \mu)$ by measure preserving transformations. For $1 \le p < \infty$, let $f \in L^p(X)$ be given. Moreover, we fix $N \in \mathbb{N}$, $\lambda > 0$ and choose $x \in X$ as in the proof of Theorem 6.4.
Then, if F is some compact (finite) subset of G and if $\{B_j\}_{j=1}^N$ is a finite sequence of finite, pairwise disjoint subsets in G such that $F_j B_j \subseteq F$ and $B_j \subseteq \{g \in G \,|\, |(A_j f)(gx)| \ge \lambda\}$ for all $1 \le j \le N$, then

$$\left|\bigcup_{j=1}^N B_j\right| = \sum_{j=1}^N |B_j| \le \frac{6(C+1)}{\lambda^p} \cdot \sum_{g \in F} |f(gx)|^p. \tag{7.2}$$

The proof of this theorem is mainly based on the following abstract combinatorial lemma.

Lemma 7.2 (Abstract combinatorial lemma, cf. [43])
Let F be a finite set. Furthermore, let a finite sequence $\{V_i\}_{i=1}^m$ of subsets of F be given which all have the same size, i.e. $|V_i| = |V_1| =: v < \infty$ for all $1 \leq j \leq m$. We further assume that $v \geq 10$ and that a positive measure ϕ is defined on F such that

(i) there exists some $\theta > 0$, such that $\phi(V_i) \geq \theta v$ for all $1 \leq i \leq m$,

(ii) $\sum_{i=1}^m \mathbb{1}_{V_i}(g) \leq v$ for all $g \in F$.

Then there is a subcollection $\{V_i \,|\, i \in I \subseteq \{1, \ldots, m\}\}$ satisfying

(a) $\phi(\bigcup_{i \in I} V_i) \geq \frac{1}{3}\theta \cdot m$,

(b) $|I| \cdot v \leq \frac{3}{\theta}\phi(\bigcup_{i \in I} V_i)$.

Proof
We give a proof in form of an algorithm. As a first step, we set $i(1) = 1$. Assume further that $i(1), \ldots, i(k)$ $(k \geq 1)$ are chosen and define $I_k := \{i(1), \ldots, i(k)\}$. If for all $l > i(k)$,

$$\phi\left(V_l \setminus \bigcup_{j=1}^{k} V_{i(j)}\right) < \frac{1}{2}\phi(V_l), \qquad \text{(case A)}$$

then set $I := I_k = \{i(1), \ldots, i(k)\}$. Otherwise choose $i(k) < l \leq m$ as the least integer between $i(k)$ and m such that

$$\phi\left(V_l \setminus \bigcup_{j=1}^{k} V_{i(j)}\right) \geq \frac{1}{2}\phi(V_l), \qquad \text{(case B)}$$

set $I_{k+1} := I_k \cup \{l\}$ and repeat this procedure until we arrive in case (A) which returns the resulting set I.
Again, we distinguish two cases.

(I) $k = |I| \geq m/v$;
For claim (b), note that

$$\begin{aligned}
\phi\left(\bigcup_{j=1}^{k} V_{i(j)}\right) &= \phi\left(\bigcup_{j=1}^{k}\left(V_{i(j)} \setminus \bigcup_{s=1}^{j-1} V_{i(s)}\right)\right) \\
&= \sum_{j=1}^{k} \phi\left(V_{i(j)} \setminus \bigcup_{s=1}^{j-1} V_{i(s)}\right) \\
&\overset{\text{Def. I}}{\geq} \frac{1}{2}\sum_{j=1}^{k} \phi(V_{i(j)}) \\
&\overset{\text{(i)}}{\geq} \frac{1}{2}k\theta v \geq \frac{\theta}{3}|I| \cdot v.
\end{aligned}$$

Continuing this calculation and using the fact that $|I| \geq m/v$, we arrive at

$$\phi\left(\bigcup_{j=1}^{k} V_{i(j)}\right) \geq \frac{\theta}{3} \cdot m,$$

which is claim (a).

(II) $k < m/v$;
We set $\overline{I} := \{1, \ldots, m\} \setminus I$ and $U := \cup_{i \in I} V_i$. Then all V_i corresponding to elements in $\overline{I}$ have most of their mass in U, i.e. for $i_0 \in \overline{I}$, one has

$$\begin{aligned} \phi(V_{i_0}) - \phi(V_{i_0} \cap U) &= \phi\left(V_{i_0} \setminus U\right) \\ &\leq \phi\left(V_{i_0} \setminus \bigcup_{j: i(j) < i_0} V_{i(j)}\right) \\ &\overset{\text{Def. } \overline{I}}{<} \frac{1}{2}\phi(V_{i_0}), \end{aligned}$$

so that

$$\frac{1}{2}\phi(V_{i_0}) < \phi(V_{i_0} \cap U).$$

By summation over all $i \in \overline{I}$ and by property (ii),

$$\begin{aligned} \frac{1}{2}\sum_{i \in \overline{I}} \phi(V_i) &\leq \sum_{i=1}^{m} \int_F \mathbb{1}_{V_i}(g) \cdot \mathbb{1}_U(g)\, d\phi(g) \\ &\overset{\text{(ii)}}{\leq} v \cdot \int_F \mathbb{1}_U(g)\, d\phi(g) \\ &= v \cdot \phi(U). \end{aligned}$$

Therefore,

$$\begin{aligned} \phi\left(\bigcup_{i \in I} V_i\right) = \phi(U) &\geq \frac{1}{2v} \sum_{i \in \overline{I}} \phi(V_i) \\ &\overset{\text{(i)}}{\geq} \frac{1}{2v} |\overline{I}| \theta v \\ &= \frac{\theta}{2} \cdot (m - k) = \frac{\theta}{2} m \cdot (1 - \frac{k}{m}) \\ &\overset{\text{(II)}}{>} \frac{1}{2}(1 - \frac{1}{v}) \cdot \theta m \\ &\overset{v \geq 10}{\geq} \frac{9}{20}\theta m \geq \frac{\theta}{3} m, \end{aligned}$$

which gives (a). By assumption, $m > |I| \cdot v$ and thus, claim (b) is proven as well.

To conclude, the statements (a) and (b) hold in each case and the proof of the lemma is finished. □

Proof (of Theorem 7.1)
As in the case of Lemma 7.2, we give an algorithm as a proof. Note first that $\phi(A) := \sum_{g\in A} |f(gx)|^p$ defines a positive measure on Borel subsets of F. Since G is unimodular (i.e. the modular function is the constant one function), we have $|F_j g| = |F_j|$ for all $g \in G$ and each $1 \le j \le N$. It follows also from the unimodularity of G that $\lim_{n\to\infty} |F_n| = |G| > 10$ such that without loss of generality, we can assume that $v_j := |F_j| \ge 10$ for all $1 \le j \le N$. Consider the collection $\{F_j b \,|\, b \in B_j\}$ for some fixed $1 \le j \le N$. So with $m_j := |B_j|$, $B_j = \{b_1^{(j)}, \dots, b_{m_j}^{(j)}\}$, we can define $V_i^{(j)} := F_j b_i^{(j)}$ for $1 \le i \le m_j$. It follows from inequality (7.1) that $\phi(V_i^{(j)}) \ge \lambda^p |V_i^{(j)}|$ for all $1 \le i \le m_j$. In addition to that, since all the $V_i^{(j)}$ are translates of the set F_j, each $g \in F$ is contained in at most $|F_j|$ different translates. Hence, $\sum_{i=1}^{m_j} \mathbb{1}_{V_i^{(j)}}(g) \le v_j$ for $g \in F$.

Therefore, for all $1 \le j \le N$, Lemma 7.2 is applicable to (subcollections of) $\{V_i^{(j)}\}_{i=1}^{m_j}$ with $v = v_j \ge 10$, $m = m_j$, $\theta = \lambda^p$ and ϕ defined as above. We will make use of this fact in the algorithm below.

(I) Set $j = N$;
Set $\vartheta_N := 1$ and apply Lemma 7.2 to the collection $\{F_N b \,|\, b \in B_N\} = \{V_i^{(N)}\}_{i=1}^{m_N}$. Hence we find $I_N \subseteq B_N$ with

$$|B_N| = m_N \le \frac{3}{\lambda^p}\phi\left(\bigcup_{d\in I_N} F_N d\right) = \frac{3}{\lambda^p}\phi(F_N I_N);$$

$$|I_N|\cdot|F_N| = |I_N|\cdot v_N \le \frac{3}{\lambda^p}\phi\left(\bigcup_{d\in I_N} F_N d\right) = \frac{3}{\lambda^p}\phi(F_N I_N).$$

(II) Replace j by $(j-1)$; if $j = 0$ then go to step (IV).

(III) Set $\overline{B}_j := \{b \in B_j \,|\, F_j b \cap \bigcup_{i=j+1}^{n} F_i I_i = \emptyset\}$ and distinguish the following two cases (A) and (B):

(A) $|\overline{B}_j| \ge |B_j|/2$;
Put $\vartheta_j := 1$ and apply Lemma 7.2 to $\overline{B}_j$ to construct $I_j \subseteq \overline{B}_j \subseteq B_j$ such that

$$|\overline{B}_j| \le \frac{3}{\lambda^p}\phi(F_j I_j);$$

$$|B_j| \le \frac{6}{\lambda^p}\phi(F_j I_j);$$

$$|I_j|\cdot|F_j| \le \frac{3}{\lambda^p}\phi(F_j I_j).$$

Return to step (II).

(B) $|B_j \setminus \overline{B}_j| \ge |B_j|/2$;
Put $\vartheta_j := 0$ and note that $b \in B_j \setminus \overline{B}_j$ implies $b \in F_j^{-1} F_{i_0} I_{i_0}$ for some $i_0 > j$. Thus,

$$B_j \setminus \overline{B}_j \subseteq \bigcup_{i>j} F_j^{-1} F_i I_i.$$

We simply record this fact, set $I_j = \emptyset$ and return to step (II).

(IV) We summarize our results:
in the cases where $\vartheta_j = 1$ occured, one obtains by disjointness of the sets $F_j I_j$

$$\sum_{j:\vartheta_j=1} |B_j| \overset{\text{step III (A)}}{\leq} \frac{6}{\lambda^p} \sum_{j:\vartheta_j=1} \phi(F_j I_j) \leq \frac{6}{\lambda^p} \phi(F).$$

For the cases $\vartheta_j = 0$, using the disjointness of the sequence $\{B_j\}$ and the fact that the sequence $\{F_j\}_{j=1}^N$ is tempered, we compute

$$\begin{aligned}
\sum_{j:\vartheta_j=0} |B_j| &\leq 2\left|\bigcup_{j:\vartheta_j=0} B_j \setminus \overline{B}_j\right| \leq 2\left|\bigcup_{j=1}^{N-1} \bigcup_{i>j} F_j^{-1} F_i I_i\right| \\
&= 2\left|\bigcup_{i=2}^{N} \bigcup_{j<i} F_j^{-1} F_i I_i\right| \leq 2 \sum_{i=2,\vartheta_i=1}^{N} \left|\bigcup_{j<i} F_j^{-1} F_i I_i\right| \\
&\leq 2 \sum_{i=2,\vartheta_i=1}^{N} \left|\bigcup_{j<i} F_j^{-1} F_i\right| \cdot |I_i| \leq 2 \sum_{i=2,\vartheta_i=1}^{N} C|F_i| \cdot |I_i| \\
&\overset{\text{step III (A)}}{\leq} 2 \sum_{i=2,\vartheta_i=1}^{N} \frac{3C}{\lambda^p} \phi(F_i I_i) \leq \frac{6C}{\lambda^p} \phi(F),
\end{aligned}$$

where in the last inequality, we exploited the disjointness of the sets $F_i I_i$.

We conclude that

$$\begin{aligned}
\sum_{j=1}^{N} |B_j| &= \sum_{j=1,\vartheta_j=0}^{N} |B_j| + \sum_{j=1,\vartheta_j=1}^{N} |B_j| \\
&\leq \frac{6C}{\lambda^p} \phi(F) + \frac{6}{\lambda^p} \phi(F) \\
&= \frac{6(C+1)}{\lambda^p} \sum_{g \in F} |f(gx)|^p,
\end{aligned}$$

which proves the transfer inequality (7.2). □

With this result, the individual ergodic theorem for countable, discrete groups along tempered Følner sequences is easily established.

Corollary 7.3 (Pointwise ergodic theorem for countable, discrete groups)
Let G be a countable, discrete group which acts on a σ-finite measure space $(X, \mathcal{B}, \mu)$ by measure preserving transformations. Further, let $\{F_n\}$ be a tempered Følner sequence with constant C. Then for any $f \in L^p(X)$, $1 \leq p < \infty$, the ergodic averages

$$(A_n f)(x) := \frac{1}{|F_n|} \sum_{g \in F_n} f(gx)$$

converge pointwise a.e. to some G-invariant $f^* \in L^p(X)$.

Proof
The proof is an application of the transfer principle, Chapter 6. For finite groups, the ergodic theorem is trivial, hence assume $|G| \geq 10$. By the amenability of G, we find for each $k \in \mathbb{N}$ finite sets $\tilde{F}_k$, F_{n_k} and $\overline{F}_k$ as in Lemma 6.3 with $\varepsilon = 1$. Further, take some $f \in L^p(X)$ as well as some $\lambda > 0$ and fix $x \in X$ as in the proof of Theorem 6.4. We define $B_j := \{g \in F_{n_k} \,|\, |(A_j f)(gx)| \geq \lambda\}$ for $1 \leq j \leq k$, as well as

$$\begin{aligned} \overline{B}_k &:= B_k, \\ \overline{B}_j &:= \{g \in F_{n_k} \setminus (\cup_{i>j} B_i) \,|\, |(A_j f)(gx)| \geq \lambda\}, \quad (k-1 \geq j \geq 1). \end{aligned}$$

Then $\overline{B}_j \subseteq B_j \subseteq \{g \in G \,|\, |(A_j f)(gx)| \geq \lambda\}$ and $F_j \overline{B}_j \subseteq F_j B_j \subseteq \overline{F}_k$ for every $1 \leq j \leq k$. Note further that the sets $\overline{B}_j$ are pairwise disjoint and

$$\left| \bigcup_{j=1}^{k} B_j \right| = \left| \bigcup_{j=1}^{k} \overline{B}_j \right| = \sum_{j=1}^{k} |\overline{B}_j|. \tag{7.3}$$

Hence we apply Theorem 7.1 with $F = \overline{F}_k$ to obtain with equality (7.3)

$$\left| \bigcup_{j=1}^{k} B_j \right| \leq \frac{6(C+1)}{\lambda^p} \int_{\overline{F}_k} |f(gx)|^p \, dm_L(g),$$

where $m_L(\cdot)$ is the counting measure on G. Since C depends only on the Følner sequence, it follows from the transfer principle (Theorem 6.4) that the maximal operator M is of weak type (p,p), hence satisfies an L^p-maximal inequality. The pointwise convergence now follows from Theorem 5.3. □

8 Lindenstrauss' decompositions

This section is devoted to the presentation of Elon Lindenstrauss' celebrated proof of the transfer inequality (6.1) for most σ-compact amenable groups (cf. [29]). The argumentation is built on the *Decomposition Lemma* (Theorem 8.1) which is valid for all second countable groups. Note that amenability of the group is *not* required. In the proof, it will be necessary to distinguish the cases ('G is discrete') and ('G is not discrete'). The latter situation is more involved because one has to work with the theory of Poisson point processes on locally compact groups (see Definition 8.4). However, the differences are rather technical (compare e.g. the technical Lemmas 8.2 and 8.7), i.e. the random algorithms determining the subcollection $\mathcal{Z}(\omega)$ of $\mathcal{F}$ are nearly identical in both cases. Note that a priori, it is not clear that Poisson point processes exist on locally compact groups. As Theorem 8.5 shows, this is indeed the case. We will not present the proof of the existence theorem in whole detail, but give an outline of the main ideas.
Finally, we present proofs of the pointwise ergodic theorems for σ-compact amenable groups in Corollary 8.8, where we use Lindenstrauss' and Emerson's versions of the transfer principle (cf. Theorems 6.4 and 6.1) respectively.

We proceed as follows. Assume as above that we are given some compact set $F \subseteq G$, a tempered sequence $\{F_j\}_{j=1}^N$ of compact sets and sets $\{B_j\}_{j=1}^N$ such that $F_jB_j \subseteq F$ for all $1 \leq j \leq N$. Unlike before, the selection of the right-translates F_jb is now based on the outcome of some random variable $\mathcal{Z}$ which will be defined on a carefully chosen probability space $(\Omega, \mathcal{H}, \mathbb{P})$. We will see that on average, the resulting subcollections of $\mathcal{F} := \{F_jb \,|\, 1 \leq j \leq N,\, b \in B_j\}$ satisfy the 'nice' properties leading to the transfer inequality in a similar way as in Chapter 7.

Theorem 8.1 (Decomposition Lemma, Lindenstrauss 2001)
Let G be a second countable group, $N \in \mathbb{N}$ and assume that $\{F_j\}_{j=1}^N$ is a finite sequence of tempered compact sets in G with constant $C > 0$. Further, let $0 < \delta \leq 1$ be an arbitrary positive number.
If $F \subseteq G$ is a compact set and if there are Borel-measurable sets $\{B_j\}_{j=1}^N$ such that $F_jB_j \subseteq F$ for all $1 \leq j \leq N$, then there is some probability space $(\Omega, \mathcal{H}, \mathbb{P})$ as well as a random variable $\mathcal{Z}$ on Ω taking values $\mathcal{Z}(\omega)$ in the set of all subcollections of $\mathcal{F} := \{F_jb \,|\, 1 \leq j \leq N,\, b \in B_j\}$ such that the counting function

$$\Lambda : \Omega \times F \to \mathbb{N}_0 : \Lambda(\omega, g) := \sum_{B \in \mathcal{Z}(\omega)} \mathbb{1}_B(g)$$

has the following properties:

(1) $\mathcal{Z}(\omega)$ is a finite set almost surely and Λ is a measurable function on $\Omega \times F$.

(2) For all $g \in F$,

$$\mathbb{E}(\Lambda_g(\cdot) \mid \Lambda_g(\cdot) \geq 1) \leq (1+\delta),$$

where $\mathbb{E}(\Lambda(\cdot, g) \mid L)$ *with* $L \in \mathcal{H}$ *stands for the expectation of the counting function* $\Lambda_g := \Lambda(\cdot, g) : \Omega \to \mathbb{N}_0$ *with respect to the probability measure* $\mathbb{P}$ *under the condition that the event* L *has been realized.*

(3) Moreover, if we define $\gamma(\delta, C) := \delta(1+C\delta)^{-1}$*, then*

$$\mathbb{E}\left(\int_F \Lambda_g(\cdot)\, dm_L(g)\right) = \mathbb{E}\left(\sum_{B \in \mathcal{Z}(\cdot)} |B|\right) \geq \gamma(\delta, C)\left|\bigcup_{j=1}^{N} B_j\right|.$$

Remark
Statement (2) stands for the almost disjointness of the outcomes $\mathcal{Z}(\omega)$ on average. We emphasize that this property does not only mean that the expectation of the counting function Λ_g is controlled by some bound which is slightly greater than one. It also says that even if we know already that $\Lambda(\omega, g) \geq 1$, we can still expect that $\Lambda(\omega, g) \leq (1+\delta)$. In light of that, there cannot exist 'significantly' many $\omega \in \Omega$ with $\Lambda(\omega, g) > 1$ for some fixed $g \in F$. Hence, for every $g \in F$ and for 'most' subcollections $\mathcal{Z}(\omega)$ ($\omega \in \Omega$), there is at most one translate $F_j b$ in $\mathcal{Z}(\omega)$ containing g.
The inequality in (3) makes sure that the total mass of right translates $F_j b$ in the collections $\mathcal{Z}(\omega)$ is on average big enough to beat the left-hand side of inequality (6.1).
As already pointed out in Chapter 7, these two properties are the crucial ingredients for the proof of the transfer inequality.

We start with the case that G is discrete. In the proof, we will need the following elementary probabilistic lemma.

Lemma 8.2
Let $\{Z_i\}$ *be a sequence of independent, identically distributed random variables which take the value 1 with probability* p *and the value 0 with probability* $(1-p)$ *(we say that the* Z_i *have Bernoulli distribution with parameter* p*). For* $n \in \mathbb{N}$*, define the random variable* $S_n := \sum_{i=1}^{n} Z_i$*. Then*

$$\mathbb{E}(S_n \mid S_n \geq 1) \leq 1 + (n-1)p$$

for every $n \in \mathbb{N}$.

Proof
Note that the claim is trivial for $p \in \{0, 1\}$. Hence we can assume that $0 < p < 1$. Then

$$\begin{aligned}(1-p)^{-n+1} = \left(\frac{1}{1-p}\right)^{n-1} &= \left(\sum_{k=0}^{\infty} p^k\right)^{n-1} \\ &\geq (1+p)^{n-1} \geq 1 + (n-1)p,\end{aligned}$$

where the last inequality is the Bernoulli inequality. Multiplying by $(1-p)^n$ gives

$$\begin{aligned}(1-p) &\geq (1-p)^n(1+(n-1)p);\\ 1+(n-1)p-(1-p)^n(1+(n-1)p) &\geq np;\\ (1+(n-1)p)(1-(1-p)^n) &\geq np.\end{aligned}$$

We divide by $(1-(1-p)^n)$ to obtain

$$\begin{aligned}1+(n-1)p &\geq \frac{np}{1-(1-p)^n} = \frac{\mathbb{E}(S_n)}{1-\mathbb{P}[S_n=0]}\\ &= \frac{\sum_{k=0}^n k\cdot\mathbb{P}[S_n=k]}{\mathbb{P}[S_n\geq 1]} = \sum_{k=1}^n k\cdot\frac{\mathbb{P}[S_n=k, S_n\geq 1]}{\mathbb{P}[S_n\geq 1]}\\ &= \sum_{k=0}^n k\cdot\mathbb{P}[S_n=k\,|\,S_n\geq 1] = \mathbb{E}(S_n\,|\,S_n\geq 1).\end{aligned}$$

We now prove the Decomposition Lemma for discrete groups. □

Proof (of Theorem 8.1, *G discrete*)
Since G is second countable and discrete, it must also be at most countable. Hence, we can put

$$\Omega := \{0,1\}^{N\times|G|}$$

as the set of all $(N\times|G|)$-matrices with entries 0 or 1. With its powerset $\mathcal{H}$, we extend Ω to a measurable space. For every pair (j,b), $1\leq j\leq N$, $b\in G$ we define pairwise independent Bernoulli random variables $Z_{j,b}$ with parameter $p_{j,b} := p_j := \delta/|F_j|$. These random variables induce the product measure $\mathbb{P}$ on Ω, i.e. for a set $M\subseteq\{1,\dots,N\}$ and some finite set $F\subseteq G$, we have for $(l_{j,b})_{j\in M,b\in F}\in\{0,1\}^{|M|\times|F|}$ that

$$\mathbb{P}[\{Z_{j,b}=l_{j,b}\}, j\in M,\, b\in F] = \prod_{j\in M}\prod_{b\in F}\Big(\chi_{[l_{j,b}=1]}\cdot p_j + \chi_{[l_{j,b}=0]}\cdot(1-p_j)\Big).$$

We now construct the random variable $\mathcal{Z}(\cdot)$ by the following algorithm. Let $\omega\in\Omega$ be given.

1. Put $j := N$ and $B_i^{(N+1)}(\omega) := B_i$ for $1\leq i\leq N$.
2. Set

$$\Sigma_j(\omega) := \{b\in B_j^{(j+1)}(\omega)\,|\,\omega_{j,b}=1\}$$

   and

$$\mathcal{Z}_j(\omega) := \{F_j b\,|\,b\in\Sigma_j(\omega)\}.$$

3. For all $i<j$, remove from $B_i^{(j+1)}(\omega)$ those elements b with $F_i b\cap(\cup_{k\geq j}\mathcal{Z}_k(\omega))\neq\emptyset$ and obtain the sets $B_i^{(j)}(\omega)$ for $i<j$.
4. As long as $j\neq 1$, replace j by $(j-1)$ and return to Step 2.

5. Put

$$\mathcal{Z}(\omega) := \bigcup_{j=1}^{N} \mathcal{Z}_j(\omega).$$

The randomness of the algorithm comes into play in Step 2. Every right-translate $F_j b$ in $\mathcal{F}$ corresponds to a pair (j, b) with $1 \le j \le N$ and $b \in B_j \subseteq G$ which means that for each such set, we flip some coin showing 'head' with probability p_j. Hence if $\omega_{j,b} = 1$ (i.e. the coin shows 'head'), the corresponding translate becomes an element of the resulting subcollection $\mathcal{Z}(\omega)$ unless it does not intersect translates $F_k b\,(k > j)$ which have been chosen as an element of $\mathcal{Z}(\omega)$ in a previous step.
Note also that the random variables $B_i^{(j)}$ for $1 \le i < j \le N$ can be interpreted as the set of all $b \in B_i$ for which $F_i b \in \mathcal{Z}(\omega)$ is still possible given $(\Sigma_i(\omega))_{i \ge j}$ and steps $N, N-1, \ldots, j$. Furthermore, the outcomes of the $B_i^{(j)}$ as well as of the random variables $\Sigma_j\,(1 \le j \le N)$ depend on Φ_j, the smallest σ-algebra of subsets of Ω generated by the pairwise independent random variables $Z_{k,a}$ for $k \ge j$ and all $a \in G$. Thus we observe by stochastic independence that if for $2 \le j \le N$, the 'value' of the information function

$$\overline{\Phi}_j(\omega) := (\Sigma_k(\omega), B_i^{(k)}(\omega), \mathcal{Z}_k(\omega))_{j \le k \le N,\, 1 \le i < k}$$

on Ω (which is based on sets in Φ_j) has been determined, the random collection $\bigcup_{i=1}^{j-1} \mathcal{Z}_i(\omega)$ has exactly the same distribution as the distribution one obtains by applying the algorithm on the tempered sequence $\{F_i\}_{i=1}^{j-1}$ and the translation sets $\{B_i^{(j)}(\omega)\}_{i=1}^{j-1}$. In the following, we will refer to this essential fact as the so-called *recursive property* of the algorithm. For the sake of completeness, we define the value $\overline{\Phi}_{N+1}(\omega)$ for all $\omega \in \Omega$ as

$$\overline{\Phi}_{N+1}(\omega) = (B_j)_{1 \le j \le N}.$$

To prove properties (1) - (3), we work with the counting functions of the $\mathcal{Z}_j(\omega)$ defined as

$$\Lambda^{(j)}(\omega, g) = \sum_{B \in \mathcal{Z}_j(\omega)} \mathbb{1}_B(g)$$

on $\Omega \times F$. Since $\mathcal{H}$ is the powerset of Ω and since G is discrete, the functions $\Lambda^{(j)}$ are measurable for $1 \le j \le N$. By disjointness of the collections $\mathcal{Z}_j(\omega)$ and $\mathcal{Z}_k(\omega)$ for $j \ne k$, the events $I_j := [\Lambda^{(j)}(\cdot, g) \ge 1]$ and $I_k := [\Lambda^{(k)}(\cdot, g) \ge 1]$ are mutually exclusive for $g \in F$. It follows that the supports of the functions $\Lambda_g^{(j)}(\omega) := \Lambda^{(j)}(\omega, g)$ on Ω are pairwise disjoint for fixed $g \in F$. Therefore,

$$\Lambda_g(\omega) = \sum_{j=1}^{N} \Lambda_g^{(j)}(\omega)$$

for all $\omega \in \Omega$ and

$$\begin{aligned}
\mathbb{E}(\Lambda_g \,|\, \Lambda_g \geq 1) &= \sum_{j=1}^{N} \mathbb{E}(\Lambda_g^{(j)} \,|\, \Lambda_g \geq 1) \\
&= \sum_{j=1}^{N} \Big(\mathbb{P}[\Lambda_g \geq 1]\Big)^{-1} \int_{[\Lambda_g \geq 1]} \Lambda_g^{(j)}(\omega)\, d\mathbb{P}(\omega) \\
&= \sum_{j=1}^{N} \frac{\mathbb{P}[\Lambda_g^{(j)} \geq 1]}{\mathbb{P}[\Lambda_g \geq 1]} \cdot \frac{\int_{[\Lambda_g^{(j)} \geq 1]} \Lambda_g^{(j)}(\omega)\, d\mathbb{P}(\omega)}{\mathbb{P}[\Lambda_g^{(j)} \geq 1]} \\
&= \sum_{j=1}^{N} \alpha_j \cdot \mathbb{E}(\Lambda_g^{(j)} \,|\, \Lambda_g^{(j)} \geq 1),
\end{aligned}$$

where $\alpha_j := P[\Lambda_g^{(j)} \geq 1]/P[\Lambda_g \geq 1]$ and $\sum_{j=1}^{N} \alpha_j = 1$. In light of that, to prove (2), it is sufficient to show

$$\mathbb{E}(\Lambda_g^{(j)} \,|\, \Lambda_g^{(j)} \geq 1) \leq (1+\delta)$$

for $1 \leq j \leq N$. So take an arbitrary j and assume that the information $\overline{\Phi}_{j+1}(\omega)$ is given for $\omega \in \Omega$. This implies that we know $B_j^{(j+1)}(\omega)$, i.e. the translation set for the $(N-j+1)^{\text{th}}$ run of the algorithm and we obtain

$$\Lambda_g^{(j)}(\omega) = \sum_{b \in B_j^{(j+1)}(\omega),\, g \in F_j b} Z_{j,b}(\omega) = \sum_{b \in B_j^{(j+1)}(\omega) \cap F_j^{-1} g} Z_{j,b}(\omega).$$

Hence, $\Lambda_g^{(j)}$ is a sum of $|B_j^{(j+1)}(\omega) \cap F_j^{-1} g| \leq |F_j|$ independent random variables with values in $\{0,1\}$ and with parameter p_j. By the tower property of the conditional expectation, we have

$$\mathbb{E}(\Lambda_g^{(j)} \,|\, \Lambda_g^{(j)} \geq 1) = \mathbb{E}\Big(\mathbb{E}(\Lambda_g^{(j)} \,|\, \Lambda_g^{(j)} \geq 1,\, \Phi_{j+1})(\cdot) \,\Big|\, \Lambda_g^{(j)} \geq 1\Big).$$

The fact that we condition on the σ-algebra Φ_{j+1} determining the information $\overline{\Phi}_{j+1}(\omega)$ allows us to consider the set $B_j^{(j+1)}(\omega)$ as fixed and with Lemma 8.2, we can estimate the inner expectation as

$$\begin{aligned}
\mathbb{E}(\Lambda_g^{(j)} \,|\, \Lambda_g^{(j)} \geq 1 \,:\, \Phi_{j+1})(\omega) &\leq 1 + |B_j^{(j+1)}(\omega) \cap F_j^{-1} g| \cdot p_j \\
&\leq 1 + |F_j| \cdot p_j = 1 + \delta.
\end{aligned}$$

This proves (2) and also shows with Fubini's Theorem that

$$\begin{aligned}
\mathbb{E}\Big(\sum_{B \in \mathcal{Z}(\cdot)} |B|\Big) &= \int_F \mathbb{E}(\Lambda_g(\cdot))\, dm_L(g) \\
&\leq (1+\delta)\,|F| < \infty.
\end{aligned}$$

Since we have $|F_j b| \geq \gamma := \min_{1 \leq j \leq N} |F_j| > 0$ for $1 \leq j \leq N$ and $b \in B_j$, this inequality can only hold true if $\mathcal{Z}(\omega)$ is finite almost surely. Thus the proof of statement (1) is completed

as well.
We show claim (3) for some constant $\gamma = \gamma(\delta, C)$ satisfying also $\gamma < \min\{\delta, C^{-1}\}$. The recursive property of the algorithm as well as the tempered condition on the sequence $\{F_j\}_{j=1}^N$ allow us to use induction on N.
For $N = 1$, we obtain

$$\begin{aligned} \mathbb{E}\Big(\sum_{g\in F} \Lambda_g\Big) &= \mathbb{E}\Big(\sum_{B\in\mathcal{Z}_N(\cdot)} |B|\Big) = \sum_{b\in B_N} |F_N b|\,\mathbb{E}(Z_{N,b}) \\ &= |F_N|\,|B_N|\,p_N = \delta\,|B_N|. \end{aligned} \tag{8.1}$$

Now assume that $N > 1$. With the recursive property of the algorithm we apply the method on $\{F_j\}_{j=1}^{N-1}$ as well as on $\{B_j^{(N)}(\omega)\}_{j=1}^{N-1}$ determined by the σ-algebra Φ_N. As before, we have to condition the expectation on Φ_N to consider the input information $\overline{\Phi}_N(\omega)$ as fixed for $\omega \in \Omega$. By the disjointness of the subcollections $\mathcal{Z}_j(\omega)$ for $1 \le j \le N$, we obtain

$$\mathbb{E}\Big(\sum_{B\in\mathcal{Z}(\cdot)} |B| \,\Big|\, \Phi_N\Big)(\omega) = \sum_{B\in\mathcal{Z}_N(\omega)} |B| + \mathbb{E}\Big(\sum_{B\in\cup_{j=1}^{N-1}\mathcal{Z}_j(\cdot)} |B| \,\Big|\, \Phi_N\Big)(\omega). \tag{8.2}$$

The induction hypothesis and the recursive property of the algorithm make sure that

$$\mathbb{E}\Big(\sum_{B\in\cup_{j=1}^{N-1}\mathcal{Z}_j(\cdot)} |B| \,\Big|\, \Phi_N\Big)(\omega) \ge \gamma \Big|\bigcup_{j=1}^{N-1} B_j^{(N)}(\omega)\Big| \tag{8.3}$$

with $\gamma < \min\{\delta, C^{-1}\}$. We know from the third step of the algorithm that the sets $B_j^{(N)}(\omega)$ arise from the sets B_j from which we remove those elements b with $F_j b \cap F_N \Sigma_N(\omega) \neq \emptyset$. Thus, for $j < N$, one obtains $B_j^{(N)}(\omega) = B_j \setminus F_j^{-1} F_N \Sigma_N(\omega)$ such that

$$\bigcup_{j=1}^{N-1} B_j^{(N)}(\omega) \supseteq \bigcup_{j=1}^{N-1} B_j \setminus \left(\bigcup_{j=1}^{N-1} F_j^{-1} F_N\right) \Sigma_N(\omega). \tag{8.4}$$

Since $\sum_{B\in\mathcal{Z}_N(\omega)} |B| = |F_N| \cdot |\Sigma_N(\omega)|$ and by the fact that the sequence $\{F_j\}_{j=1}^N$ is tempered, we obtain from the inequalities (8.2), (8.3) and (8.4) that

$$\mathbb{E}\Big(\sum_{B\in\mathcal{Z}(\cdot)} |B| \Big| \Phi_N\Big)(\omega) \ge |F_N| \cdot |\Sigma_N(\omega)| + \gamma \left(\left|\bigcup_{j=1}^{N-1} B_j\right| - C|F_N||\Sigma_N(\omega)|\right).$$

Taking expectations and using that $C\gamma < 1$, we have with (8.1) that

$$\mathbb{E}\Big(\sum_{B\in\mathcal{Z}(\cdot)} |B|\Big) \ge \delta|B_N| + \gamma \left(\left|\bigcup_{j=1}^{N-1} B_j\right| - C\delta|B_N|\right).$$

One readily verifies that

$$\delta - C\delta \cdot \gamma = \gamma \Leftrightarrow \gamma = \frac{\delta}{1 + C\delta}.$$

It follows that

$$\mathbb{E}\Big(\sum_{B\in\mathcal{Z}(\cdot)} |B|\Big) \geq \gamma\left(\left|\bigcup_{j=1}^{N-1} B_j\right| + |B_N|\right)$$

with the explicit constant $\gamma(\delta, C) = \delta\,(1 + C\delta)^{-1} < \min\{\delta, C^{-1}\}$.
This finishes the proof. □

Example 8.3
To illustrate the method presented in the proof of the Decomposition Lemma, we give a concrete example for $N = 3$. Figure 8.1 shows the group G as well as the compact set $F \subseteq G$. The (tempered) sequence $\{F_j\}_{j=1}^3$ is displayed by a square F_1, a triangle F_2 and a circle F_3 respectively. Further, we are given translation sets $B_1 := \{b_k^{(1)}\}_{k=1}^4$, $B_2 := \{b_l^{(2)}\}_{l=1}^4$ and $B_3 := \{b_m^{(3)}\}_{m=1}^3$ such that the squares $S_k := F_1 b_k^{(1)}$, $1 \leq k \leq 4$, the triangles $T_l := F_2 b_l^{(2)}$, $1 \leq l \leq 4$ as well as the circles $C_m := F_3 b_m^{(3)}$, $1 \leq m \leq 3$ are all contained in F and form the collection $\mathcal{F}$. Assume now for the corresponding random variables $Z_{j,b}$, we obtain the realizations

$$\{Z_{j,b^{(j)}}(\omega)\}_{1\leq j\leq 3,\, b^{(j)}\in B_j} = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & \end{pmatrix}.$$

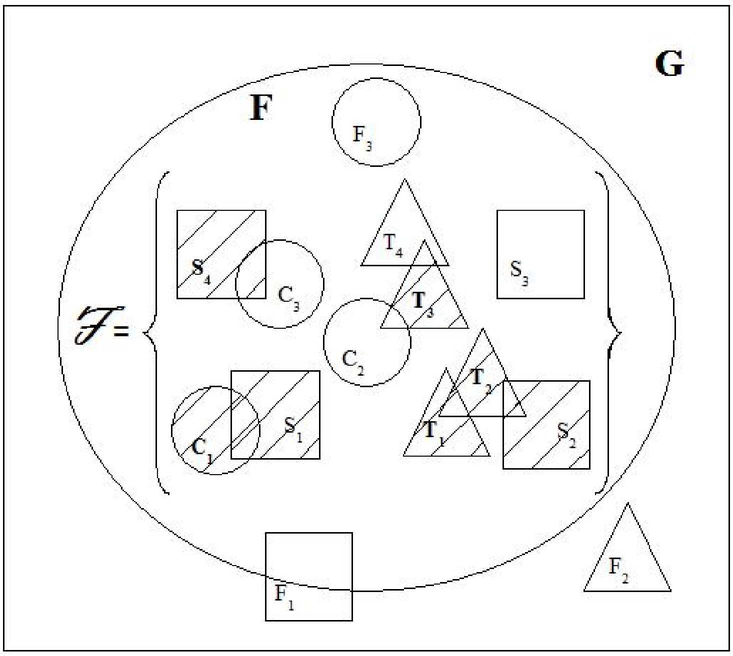

Figure 8.1: Decomposition Algorithm for $N = 3$

The shaded elements in Figure 8.1 are those $F_j b$ with $Z_{j,b}(\omega) = 1$, i.e. the (unfair) coin shows 'head' for the realization $\omega \in \Omega$. In our concrete case, these are the squares S_1, S_2, S_4 as well as the triangles T_1, T_2, T_3 and the circle C_1.
Now we can construct the subcollection $\mathcal{Z}(\omega)$ of $\mathcal{F}$. We start with $j = N = 3$ which means that we consider circles. Since C_1 is the only shaded circle, we obtain $\mathcal{Z}_3(\omega) = \{C_1\}$ by Step 2 of the algorithm. We see further that $C_1 \cap S_1 \neq \emptyset$ which means that the square S_1 will not be contained in the final collection $\mathcal{Z}(\omega)$ by Step 3.
In the second run of the algorithm, we have $j = 2$ and we draw our attention to the shaded triangles. None of those intersects $\cup \mathcal{Z}_3(\omega) = C_1$ non-trivially which means by Step 2 that $\mathcal{Z}_2(\omega) = \{T_1, T_2, T_3\}$. Since $T_2 \cap S_2 \neq \emptyset$, the square S_2 is removed from the cadidate list for the final collection $\mathcal{Z}(\omega)$ by Step 3 of our method.
Finally, we turn to the shaded squares in the third run ($j = 1$). The squares S_1 and S_2 are marked but have already been excluded from the final collection in the previous runs. In light of that, we arrive at $\mathcal{Z}_1(\omega) = S_4$ again by Step 2 of the algorithm. Since $j = 1$, we miss out the Steps 3 and 4 and Step 5 returns $\mathcal{Z}(\omega) = \mathcal{Z}_1(\omega) \cup \mathcal{Z}_2(\omega) \cup \mathcal{Z}_3(\omega) = \{S_4, T_1, T_2, T_3, C_1\}$ as final collection (see elements marked in bold face in Figure 8.1).

We now turn to the case that G is not discrete. As mentioned above, we will make use of Poisson point processes on locally compact groups. For that, one considers the measurable space $(G, \mathcal{J})$, where $\mathcal{J}$ is the Borel σ-algebra on G. Further, we denote by Ω the collection of all locally finite subsets of G (i.e. subsets such that their intersection with each compact set in G is a finite set) and we endow this set with the smallest σ-algebra $\mathcal{H}$ such that for all $B \in \mathcal{J}$, the function

$$N(\cdot)[B] : (\Omega, \mathcal{H}) \to (\mathbb{N}_0 \cup \{\infty\}, \mathcal{P}(\mathbb{N}_0 \cup \{\infty\})) : N(\Upsilon)[B] := \#(\Upsilon \cap B)$$

is measurable, where $\# A$ stands for the number of elements contained in A and $\mathcal{P}$ denotes the power set. Hence, for each $B \in \mathcal{J}$ we have found a random variable mapping locally finite subsets Υ of G to the cardinality of $\Upsilon \cap B$.
With that, we can introduce the notion of a Poisson point process.

Definition 8.4
Let ν be a σ-finite, non-atomic measure on the measurable space $(G, \mathcal{J})$, i.e. $\nu(\{g\}) = 0$ for all $g \in G$. A probability measure $\mathbb{P}$ on the measurable space $(\Omega, \mathcal{H})$ defined above is called a Poisson point process with intensity (measure) ν if

(i) for all $B \in \mathcal{J}$ with $0 < \nu(B) < \infty$, the random variable $N(\cdot)[B]$ (on Ω) has Poisson distribution with parameter $\nu(B)$, i.e.

$$\mathbb{P}[N(\cdot)[B] = k] = e^{-\nu(B)} \frac{\nu(B)^k}{k!}$$

for all $k \in \mathbb{N}$;

(ii) the fact that for some $n \in \mathbb{N}$, $\{B_j\}_{j=1}^n$ is a family of pairwise disjoint elements in $\mathcal{J}$ implies that the random variables $\{N(\cdot)[B_j]\}_{j=1}^n$ are independent, i.e.

$$\mathbb{P}[N(\cdot)[B_j] = k_j,\, 1 \le j \le n] = \prod_{j=1}^n \mathbb{P}[N(\cdot)[B_j] = k_j]$$

for all $k = \{k_j\}_{j=1}^n \in (\mathbb{N}_0 \cup \{\infty\})^n$.

Remark
By definition of the random variables $N(\cdot)[B]$ for $B \in \mathcal{J}$, the assumption on ν to be non-atomic is inevitable. Namely, if we assume that there is some $g \in G$ such that $\nu(\{g\}) > 0$, then property (i) of the Poisson point process implies that $N(\cdot)[\{g\}]$ has Poisson distribution with parameter $\nu(\{g\})$. But then we arrive at

$$0 < \mathbb{P}[N(\cdot)[\{g\}] \geq 2] = \mathbb{P}[\{\Upsilon \mid \#(\Upsilon \cap \{g\}) \geq 2\}] = 0,$$

which is a contradiction.

Note that so far, it is not clear whether such a probability measure $\mathbb{P}$ exists on the measurable space $(\Omega, \mathcal{H})$. However, *if* it exists, then the probability measure is (up to distribution) uniquely defined by the Poisson random variables $N(\cdot)[B]$ for $B \in \mathcal{J}$. The reason for this is that the intensity measure ν can be written as

$$\nu : \mathcal{J} \to [0, \infty] : B \mapsto \nu(B) = \mathbb{E}[N(\cdot)[B]] = \int_\Omega N(\cdot)[B]\, d\mathbb{P}$$

and that the Poisson point process is determined uniquely by the measure ν (cf. [38], Section 3.6).
We will not prove the existence of a Poisson point process on $(\Omega, \mathcal{H})$, but give a rough outline of the construction. For a detailed discussion, the interested reader may refer to [24], Chapter 2.

Theorem 8.5 (Existence theorem for Poisson point processes)
Let G be a σ-compact group and denote by $\mathcal{J}$ its Borel σ-algebra. Further, let Ω and $\mathcal{H}$ be given as above. Then if ν is a non-atomic, σ-finite measure on $(G, \mathcal{J})$, there is a Poisson point process $\mathbb{P}$ on $(\Omega, \mathcal{H})$ with intensity ν.

We give a sketch of the proof which is based on [24], Chapter 2. For this, we need the notion of independent Poisson point processes.

Definition 8.6 (Independence of Poisson point processes)
Let $(G, \mathcal{J})$ be a measurable space and assume that Ω and $\mathcal{H}$ are chosen as above. A (possibly infinite) family $\{\mathbb{P}_n\}_{n \in M}$, $M \subseteq \mathbb{N}$ of Poisson point processes on $(\Omega, \mathcal{H})$ is called independent if for every finite set $M' \subseteq M$ and all sets $B \in \mathcal{J}$, the corresponding random variables $\{N_m(\cdot)[B]\}_{m \in M'}$ are independent.

PROOF (OF THEOREM 8.5, SKETCH)
The proof consists of three major parts.

(1) The first step is the verification of a *Disjointness Lemma* showing that if $\mathbb{P}_1$ and $\mathbb{P}_2$ are two independent Poisson point processes on $(\Omega, \mathcal{H})$ with intensities ν_1 and ν_2 and if $B \in \mathcal{J}$ with $\nu_1(B), \nu_2(B) < \infty$, then the locally finite random sets Υ_1 and Υ_2 realized by the processes are almost surely disjoint on B.

To see this, we must treat $\mathbb{P}_1$ and $\mathbb{P}_2$ as processes on the set B. So we first endow the collection $\overline{\Omega} := (B \times B)^f$ consisting of all finite subsets of $B \times B$ with the coarsest σ-algebra $\overline{\mathcal{H}}$ making the map

$$\overline{\Omega} \to \mathbb{N} : \overline{\Upsilon} := (\Upsilon_1, \Upsilon_2) \mapsto \#(\overline{\Upsilon} \cap \overline{C})$$

measurable for all sets $\overline{C} \subseteq B \times B \in \mathcal{J} \otimes \mathcal{J}$, where $\mathcal{J} \otimes \mathcal{J}$ stands for the product σ-algebra of $\mathcal{J}$ with itself. Then one shows that the mapping

$$\pi : B^f \times B^f \to (B \times B)^f : (\Upsilon_1, \Upsilon_2) \mapsto \Upsilon_1 \times \Upsilon_2$$

is measurable with respect to $\mathcal{H}_0 \times \mathcal{H}_0$ on $B^f \times B^f$, where $\mathcal{H}_0$ is the smallest σ-algebra on B^f making the map $\Upsilon \mapsto \#(\Upsilon \cap C)$ measurable for all measurable $C \subseteq B$. As a consequence, the map π induces on $(\overline{\Omega}, \overline{\mathcal{H}})$ the joint probability measure $\overline{\mathbb{P}}$ of the processes $\mathbb{P}_1$ and $\mathbb{P}_2$. Note that the event $E_B := \{(\Upsilon_1, \Upsilon_2) \,|\, \Upsilon_1 \cap \Upsilon_2 \cap B = \emptyset\}$ can likewise be expressed as $E_B := \{\overline{\Upsilon} \,|\, \#(\overline{\Upsilon} \cap D_B) = 0\}$, where D_B is the (measurable!) diagonal set in $B \times B$. By the notion of the image measure,

$$\overline{\mathbb{P}}(E_B) = (P_1 \times P_2)(\pi^{-1}(E_B)),$$

where P_1 and P_2 are the distributions of the independent random variables $N_1(\cdot)[B]$ and $N_2(\cdot)[B]$. A short calculation using Fubini's Theorem shows that

$$(P_1 \times P_2)(\pi^{-1}(E_B)) = 1$$

and therefore, the outcomes Υ_1 and Υ_2 of the processes are $\overline{\mathbb{P}}$-almost surely disjoint random subsets in B.

(2) Next, using the Disjointness Lemma, one proves the so-called *Superposition Theorem*. It says that if $\{\mathbb{P}_i\}$ is a sequence of independent Poisson point processes on $(\Omega, \mathcal{H})$ with intensities $\{\nu_i\}$, then there is a Poisson point process $\mathbb{P}$ on $(\Omega, \mathcal{H})$ with intensity measure $\nu := \sum_{i=1}^{\infty} \nu_i$.
In a canonical manner, $\mathbb{P}$ will be the probability distribution of subsets in G represented as

$$\Upsilon := \bigcup_{i=1}^{\infty} \Upsilon_i,$$

where the random sets Υ_i are distributed according to $\mathbb{P}_i$.
By the Disjointness Lemma, the random sets are essentially pairwise disjoint in each set $B \in \mathcal{J}$ of finite measure. In light of that, the number of points of some $\Upsilon \subseteq \Omega$ in B is given by the random variable

$$N(\cdot)[B] := \sum_{i=1}^{\infty} N_i(\cdot)[B].$$

Since the countable sum of independent Poisson random variables is again a Poisson variable with mean equal to the value of the convergent (!) series of the parameters of the sequence (one way to see that is to use generating functions) we conclude that

for every $B \in \mathcal{J}$ with $0 < \nu(B) < \infty$, the random variable $N(\cdot)[B]$ has indeed Poisson distribution with mean $\nu(B)$.
The independence of a sequence $\{N(\cdot)[B_j]\}_{j=1}^n$ for disjoint sets B_j in $\mathcal{J}$ is now easily established. By the fact that all $\mathbb{P}_i$ are Poisson point processes, the family $\{N_i(\cdot)[B_j]\}_{j=1}^n$ must be independent for all $i \in \mathbb{N}$. Moreover, the independence of the $\mathbb{P}_i$ guarantees that the double array $\{N_i(\cdot)[B_j]\}_{i\in\mathbb{N},\,1\leq j\leq n}$ consists of independent random variables, which is sufficient. The Kolmogorov Existence Theorem (cf. [40], Theorem 12.8) makes sure that there extists indeed a measure $\mathbb{P}$ on $(\Omega, \mathcal{H})$ having the desired properties.

(3) Finally, one gives the *construction of the Poisson point process* on $(\Omega, \mathcal{H})$. Since the measure ν is σ-finite, there is a sequence $\{G_i\}$ of disjoint measurable subsets of G such that $\cup_{i=1}^{\infty} G_i = G$ and $0 < \nu(G_i) < \infty$ for all $i \in \mathbb{N}$. We define measures $\nu_i(\cdot) := \nu(\,\cdot \cap G_i)$ on $(G, \mathcal{J})$ for every $i \in \mathbb{N}$. Now fix $i_0 \in \mathbb{N}$ and denote by L an arbitrary Poisson distributed random variable with mean $\nu_{i_0}(G) := \nu(G_{i_0})$ on a suitable measurable space $(\Omega_L, \mathcal{H}_L)$. Assume further that $\{Z_j\}$ is a sequence of identically and independently distributed, G-valued random variables on a probability space $(\Omega_Z, \mathcal{H}_Z, \mathbb{P}_Z)$ which are also independent of L and which all have distribution

$$\mathbb{P}_Z := \nu_{i_0}(\cdot)/\nu_{i_0}(G) = \nu(\,\cdot\, \cap G_{i_0})/\nu(G_{i_0}).$$

Then, for each $\tau \in \Omega_L$ and every $\omega_j \in \Omega_Z$, a random subset of G is determined by some $\mathcal{H}$-measurable (choose $\mathcal{H}_L$ and $\mathcal{H}_Z$ appropriately) map

$$\Upsilon_{i_0}(\tau, \{\omega_j\}) := \{Z_j(\omega_j) \,|\, 1 \leq j \leq L(\tau)\}.$$

Further, define

$$N_{i_0}(\Upsilon_{i_0})[B] := \#\{\Upsilon_{i_0} \cap B\}$$

on $(\Omega, \mathcal{H})$ for $B \in \mathcal{J}$. Let $\{B_j\}_{j=1}^n$ be a disjoint family of measurable sets in G and set $B_0 := G \setminus \cup_{j=1}^n B_j$. By the choice of the distributions of the Z_j, we observe that given the fact that $\#\Upsilon_{i_0} = m \in \mathbb{N}$, the random variables $\{N_{i_0}(B_j)\}_{j=0}^n$ obey a multinomial distribution, i.e.

$$\mathbb{P}(N_{i_0}(\cdot)[B_j] = m_j,\, 1 \leq j \leq n \,|\, \#\Upsilon_{i_0} = m) = \frac{m!}{\prod_{j=0}^n m_j!} \cdot \prod_{j=0}^n \left(\frac{\nu_{i_0}(B_j)}{\nu_{i_0}(G)}\right)^{m_j},$$

where $m_0 := m - \sum_{j=1}^n m_j \geq 0$ (The probability is zero if $m_0 < 0$).
Using the definition of conditional probability and the fact that L has Poisson distribution with mean $\nu_{i_0}(G)$, one finally arrives at

$$\mathbb{P}(N_{i_0}(\cdot)[B_j] = m_j,\, 1 \leq j \leq m) = \prod_{j=0}^n \exp(-\nu_{i_0}(B_j))\, \frac{\nu_{i_0}(B_j)^{m_j}}{m_j!},$$

which shows with a summing argument that each random variable $N_{i_0}(\cdot)[B_j]$ is indeed distributed according to the Poisson distribution with mean $\nu_{i_0}(B_j) = \nu(B_j \cap G_{i_0})$ and that the family $\{N_{i_0}(\cdot)[B_j]\}_{j=1}^n$ is independent. We conclude that the image measure $\mathbb{P}_{i_0}$ of the mapping $\Upsilon_{i_0}(\tau, \{\omega_j\})$ is a Poisson point process on $(\Omega, \mathcal{H})$ with intensity measure

ν_{i_0}.
Since i_0 was arbitrary, we can construct for every $i \in \mathbb{N}$ a process $\mathbb{P}_i$ in this manner. It is clear that $\nu = \sum_{i=1}^{\infty} \nu_i$ and it follows from the disjointness of the G_i that the processes $\mathbb{P}_i$ are independent of each other. Hence, the Superposition Theorem tells us that there must be a Poisson point process $\mathbb{P}$ with intensity ν on $(\Omega, \mathcal{H})$. □

For the proof of Theorem 8.1 for non-discrete amenable groups we need an elementary lemma which is similar to Lemma 8.2.

Lemma 8.7
Assume that $(G, \mathcal{J})$ is a Borel measurable space endowed with a σ-finite measure ν without atoms. Moreover, let $\mathbb{P}$ be a Poisson point process on the space Ω of all locally finite subsets of G with the canonical σ-algebra $\mathcal{H}$. Then for any measurable $B \in \mathcal{J}$ with $\nu(B) < \infty$ there is a random variable $\Upsilon_B(\cdot)$ with distribution $\mathbb{P}_B$ realizing the process on B, i.e. it takes values in the set $\overline{\Omega}$ of all locally finite subsets of B with trace σ-algebra $\overline{\mathcal{H}} := \mathcal{H} \cap \overline{\Omega} := \{H \cap \overline{\Omega} \mid H \in \mathcal{H}\}$ and

$$\mathbb{E}(\# \Upsilon_B(\cdot) \mid \# \Upsilon_B(\cdot) \geq 1) \leq 1 + \nu(B).$$

Alternatively, $\mathbb{P}_B$ can be interpreted as a Poisson point process on $(\Omega, \mathcal{H})$ with intensity measure $\nu_B := \nu(\cdot \cap B)$ (restriction property).

Proof
We define the measure ν_B as

$$\nu_B(C) := \nu(C \cap B)$$

for all $C \in \overline{\mathcal{H}}$. We have seen in the construction that $\mathbb{P}$ is the distribution of some $\mathcal{H}$-measurable random variable Υ with values in Ω. Therefore, we set

$$\Upsilon_B(\cdot) := \Upsilon(\cdot) \cap B.$$

It is clear that Υ_B is $\overline{\mathcal{H}}$-measurable and that it has the probability distribution $\mathbb{P}_B$ induced by $\mathbb{P}$. By checking the definitions, one observes that $\mathbb{P}_B$ determines a Poisson point process on $(\overline{\Omega}, \overline{\mathcal{H}})$ with intensity measure ν_B.
Thus,

$$\mathbb{P}_B(\# \Upsilon_B = m) = \exp(-\nu(B)) \cdot \frac{(\nu(B))^m}{m!}$$

for all $m \in \mathbb{N}$ and

$$\begin{aligned} \mathbb{E}(\# \Upsilon_B(\cdot) \mid \# \Upsilon_B(\cdot) \geq 1) &= \frac{\sum_{m \geq 1} m \, \mathbb{P}_B(\# \Upsilon_B = m)}{\sum_{m \geq 1} \mathbb{P}_B(\# \Upsilon_B = m)} = \frac{\mathbb{E}(\# \Upsilon_B(\cdot))}{1 - \mathbb{P}_B(\Upsilon_B = \emptyset)} \\ &= \frac{\nu(B)}{1 - \exp(-\nu(B))} \leq \frac{\nu(B)}{1 - (1 + \nu(B))^{-1}} \\ &= \frac{\nu(B)}{\nu(B)(1 + \nu(B))^{-1}} = 1 + \nu(B), \end{aligned}$$

where the inequality is due to the classical inequality $e^{-x} \leq (1+x)^{-1}$ for $x \geq 0$.
By extending $\mathbb{P}_B$ to $\mathbb{P}'_B$ on the whole space $(\Omega, \mathcal{H})$, where we put $\mathbb{P}'_B(C) = 0$ for all $C \notin \overline{\mathcal{H}}$ we see that $\mathbb{P}'_B$ determines a Poisson point process on $(\Omega, \mathcal{H})$ with intensity ν_B. This finishes the proof. □

Now, we have all tools to complete the proof of Theorem 8.1.

Proof (of Theorem 8.1, *G non-discrete*)
We put

$$\Omega := \left\{ \omega = (\Upsilon_j)_{j=1}^N \,\middle|\, \Upsilon_j \subseteq G \text{ locally finite for } 1 \leq j \leq N \right\}$$

and endow this set with the σ-algebra $\mathcal{H}$ which is the N-product of $\mathcal{H}'$, the coarsest σ-algebra on the set Ω' of all locally finite sets in G making the map

$$N'(\cdot)[B] : \Omega' \to \mathbb{N} \cup \{\infty\} : \Upsilon \mapsto \#(\Upsilon \cap B)$$

measurable for all $B \in \mathcal{J}$. Following Theorem 8.5, we now establish independent Poisson point processes $\mathbb{P}_j$ on the space $(\Omega', \mathcal{H}')$ with intensities $\nu_j := \alpha_j m_R(\cdot)$, where $m_R(\cdot)$ is the *right* Haar measure on G and $\alpha_j := \delta/|F_j|$ for $1 \leq j \leq N$. Hence, the corresponding product measure $\mathbb{P}$, defined on the measurable space $(\Omega, \mathcal{H})$ can be seen as a product of N independent Poisson point processes on $(\Omega', \mathcal{H}')$.
Note that since G is also σ-compact by Proposition 3.4 and as $m_R(\cdot)$ is a Radon measure, the intensity measures ν_j are σ-finite. Moreover, the fact that G is not discrete guarantees that they are non-atomic (see e.g. [9], Proposition 1.4.4). In light of that, the intensities satisfy at least all technical requirements. The reasons for which these measures are also the 'right' choices in the sense that they lead to success will become clear below.
So let us formulate the algorithm for the construction of the subcollection $\mathcal{Z}(\omega)$ for $\omega = (\Upsilon_j)_{j=1}^N \in \Omega$.

1. Put $j := N$ and $B_i^{(N+1)}(\omega) := B_i$ for $1 \leq i \leq N$.
2. Set $\Sigma_j(\omega) := \Upsilon_j \cap B_j^{(j+1)}(\omega)$ and
$$\mathcal{Z}_j(\omega) := \{F_j b \,|\, b \in \Sigma_j(\omega)\}.$$
3. For all $i < j$, remove from $B_i^{(j+1)}(\omega)$ those elements b with $F_j b \cap (\cup_{k \geq j} \mathcal{Z}_k(\omega)) \neq \emptyset$ to obtain the sets $B_i^{(j)}(\omega)$ for $i < j$.
4. As long as $j \neq 1$, replace j by $(j-1)$ and return to Step 2.
5. Return
$$\mathcal{Z}(\omega) := \bigcup_{j=1}^N \mathcal{Z}_j(\omega).$$

As mentioned above, the algorithm is similar to the corresponding one in the discrete case. The possible candidates for the final collection $\mathcal{Z}(\omega)$ are determined by random sets generated by the the Poisson point process. Note that for every $1 \leq j \leq N$ and every $a \in F_j$, the continuity of the group multiplication in G guarantees that $a\overline{B}_j = \overline{aB_j} \subseteq F$ and therefore, the closure of B_j is compact. It follows that $\alpha_j m_R(B_j) < \infty$ such that Υ_j is finite almost surely by the proof of Lemma 8.7.
Further, we show that the counting function $\Lambda : \Omega \times G \to \mathbb{N}_0 \cup \{\infty\} : (\omega, g) \mapsto \sum_{j=1}^{N} \Lambda^{(j)}(\omega, g)$ with

$$\Lambda^{(j)} : \Omega \times G \to \mathbb{N}_0 \cup \{\infty\} : \Lambda^{(j)}(\omega, g) := \sum_{B \in \mathcal{Z}_j(\omega)} \mathbb{1}_B(g)$$

is measurable with respect to the product σ-algebra $\mathcal{H} \otimes \mathcal{J}$. To see this, note that it is sufficient to deal with the single functions $\Lambda_g^{(j)}$ since they all have disjoint support on F. We assume first that $j = N = 1$ and that $\Lambda_g(\omega) = \Lambda_g^{(N)}(\omega) \geq 1$. This implies $g \in F_N B_N$ and we calculate for $M \geq 1$

$$\begin{aligned}
[\Lambda = M] = [\Lambda^{(N)} = M] &= \{(\Upsilon_N, g) \,|\, \#(B_N \cap \Upsilon_N \cap F_N^{-1} g) = M\} \\
&= \{(\Upsilon_N, g) \,|\, N^{'}(\Upsilon_N)[F_N^{-1} g] = M\} \\
&= \{(\Upsilon_N, g) \,|\, N^{'}(\Upsilon_N g^{-1})[F_N^{-1}] = M\}.
\end{aligned}$$

Consider the map

$$\eta : \Omega^{'} \times G \to \Omega^{'} : (H, g) \mapsto H g^{-1}.$$

For $A \in \mathcal{H}^{'}$, we have $\eta^{-1}(A) = \{(g, Ag) \,|\, g \in G\}$. Since $\{g\} \times \{Ag\} \in \mathcal{H}^{'} \otimes \mathcal{J}$ for all g and since G is a second countable, hence separable topological group, it follows that η is $(\mathcal{H}^{'} \otimes \mathcal{J})$-$\mathcal{H}^{'}$-measurable and by the measurability of the counting map we conclude that $[\Lambda \geq M] \in \mathcal{H}^{'} \otimes \mathcal{J}$ for $M \geq 1$. Considering complements, we observe that the same must be true for $M = 0$.
If $N > j \geq 1$ and if the maps $\Lambda^{(i)}$ are measurable for $j < i \leq N$, we have $g \in F_j B_j^{(j+1)}(\omega)$ and therefore

$$\begin{aligned}
[\Lambda^{(j)} = M] &= \{(\omega, g) \,|\, \#(B_j^{(j+1)}(\omega) \cap \Upsilon_j \cap F_j^{-1} g) = M\} \\
&= \{(\omega, g) \,|\, N^{'}(\Upsilon_j g^{-1})[F_j^{-1}] = M]\}
\end{aligned}$$

for $M \geq 1$. We remark that then also $\Lambda^{(i)} = 0$ for $i > j$ by the disjointness of the $\mathcal{Z}_i(\cdot)$ such that by the independence of the Poisson point processes, it follows in the same manner as above that

$$[\Lambda^{(j)} = M] = \underbrace{\Omega^{'} \times \cdots \times \Omega^{'}}_{j-1} \times \mathcal{S}$$

with

$$\mathcal{S} \in \underbrace{\mathcal{H}^{'} \otimes \cdots \otimes \mathcal{H}^{'}}_{N-j+1} \otimes \mathcal{J}.$$

So we have indeed that $[\Lambda^{(j)} = M] \in \mathcal{H} \otimes \mathcal{J}$. Again, for $M = 0$, we just consider complements.
Consequently, the counting function Λ is $\mathcal{H} \otimes \mathcal{J}$-measurable as the sum of the $\Lambda^{(j)}$.

For the proof of the second statement of Theorem 8.1, as in the discrete case, it is sufficient to show that for all $1 \leq j \leq N$ and every $\omega \in \Omega$, we have for every $g \in F$ that

$$\mathbb{E}(\Lambda_g^{(j)} \,|\, \Lambda_g^{(j)} \geq 1 : \Phi_{j+1})(\omega) \leq (1+\delta),$$

where Φ_{j+1} is the smallest σ-algebra generated by the random variables $\Upsilon_i(\cdot)\ (i > j)$ induced by the processes $\mathbb{P}_i\ (i > j)$ and $\Lambda_g^{(j)}(\omega) := \Lambda^{(j)}(g, \omega)$.
To do so, recall that for each $g \in F$ and every $1 \leq j \leq N$, we have

$$\Lambda_g^{(j)}(\omega) = M \iff N'(\Upsilon_j(\omega))[B_j^{(j+1)}(\omega) \cap F_j^{-1} g] = M.$$

Of course this argumentation only makes sense when $M \geq 1$. For $M = 0$ we simply have $\Upsilon_j(\omega) \cap B_j^{(j+1)}(\omega) \cap F_j^{-1} g = \emptyset$. By the choice of the α_j, we obtain

$$\alpha_j\, m_R(B_j^{(j+1)}(\omega) \cap F_j^{-1} g) \leq \frac{\delta}{|F_j|} \cdot m_R(F_j^{-1} g) = \frac{\delta}{|F_j|} \cdot |g^{-1} F_j| = \delta$$

for all $1 \leq j \leq N$. Using Lemma 8.7 with $\nu = \alpha_j\, m_R$ and $B = B_j^{(j+1)}(\omega) \cap F_j^{-1} g$ we finally arrive at

$$\begin{aligned} \mathbb{E}(\Lambda_g^{(j)} \,|\, \Lambda_g^{(j)} \geq 1 : \Phi_{j+1})(\omega) &= \mathbb{E}(\#\, \Upsilon_B(\cdot) \,|\, \#\, \Upsilon_B(\cdot) \geq 1) \\ &\leq 1 + \alpha_j\, m_R(B_j^{(j+1)}(\omega) \cap F_j^{-1} g) \\ &\leq (1+\delta), \end{aligned}$$

which completes the proof of statement (2).
So let us turn to the proof of the claim (3) of Theorem 8.1. If we fix $N \in \mathbb{N}$, then by the definition of the modular function, we obtain

$$\begin{aligned} \mathbb{E}\Big(\int_F \Lambda_g(\cdot)\, dm_L(g) \Big) &= \mathbb{E}\Big(\sum_{B \in \mathcal{Z}_N(\cdot)} |B| \Big) \\ &= \mathbb{E}\Big(\sum_{b \in \Sigma_N(\cdot)} |F_N b| \Big) \\ &= \mathbb{E}\Big(\sum_{b \in \Sigma_N(\cdot)} \Delta(b) \Big) |F_N|. \end{aligned} \tag{8.5}$$

Further, if $\{C_i\}_{i=1}^{L}$ is a disjoint family of measurable sets in G with $m_R(C_i) < \infty$ and if $\gamma_i \in \mathbb{C}$, $(1 \leq i \leq L)$, then for the function

$$h(g) := \sum_{i=1}^{L} \gamma_i\, \mathbb{1}_{C_i}(g) \in L^\infty(G),$$

one has

$$\begin{aligned}\mathbb{E}\Big(\sum_{b\in\Upsilon_N(\cdot)} h(b)\Big) &= \mathbb{E}\Big(\sum_{b\in\Upsilon_N(\cdot)}\sum_{i=1}^{L}\gamma_i\,\mathbb{1}_{C_i}(b)\Big)\\ &= \mathbb{E}\Big(\sum_{i=1}^{L}\gamma_i\cdot N'(\Upsilon_N(\cdot))[C_i]\Big)\\ &= \sum_{i=1}^{L}\gamma_i\cdot\alpha_N\, m_R(C_i)\\ &= \alpha_N\int_G h(g)\,dm_R(g) \qquad (8.6)\end{aligned}$$

by the first property of a Poisson point process. Hence, if $h\in L^\infty(G)$ is a non-negative function with support of finite m_R-measure, it can be approximated pointwise and monotonically by non-negative simple functions of the above form and by the Monotone Convergence Theorem, this implies that equality (8.6) is valid for all functions $h\in L^\infty(G)$ with $m_R(\mathrm{supp}(h))<\infty$. Moreover, applying this fact to the function $h(g):=\mathbb{1}_{B_N}(g)\,\Delta(g)$ (note that h has support of finite measure since $m_R(B_N)<\infty$, see above) and using equality (8.5), we arrive at

$$\begin{aligned}\mathbb{E}\Big(\int_F\Lambda_g(\cdot)\,dm_L(g)\Big) &= \mathbb{E}\Big(\sum_{b\in B_N\cap\Upsilon_N(\cdot)}\Delta(b)\Big)\,|F_N|\\ &= \mathbb{E}\Big(\sum_{b\in\Upsilon_N(\cdot)}\mathbb{1}_{B_N}(b)\,\Delta(b)\Big)\,|F_N|\\ &= \alpha_N\,|F_N|\int_{B_N}\Delta(g)\,dm_R(g).\end{aligned}$$

Since $\Delta(g)\,dm_R(g)=dm_L(g)$ by the definition of the modular function and since $\alpha_N\,|F_N|=\delta$ by the choice of α_N, we finally conclude

$$\mathbb{E}\Big(\int_F\Lambda_g(\cdot)\,dm_L(g)\Big)=\mathbb{E}(|\Sigma_N(\omega)|\cdot|F_N|)=\delta\,|B_N|.$$

Thus, statement (3) of Theorem 8.1 is verified for $N=1$. Again, we exploit the recursive property of the algorithm (same arguments as in the discrete case) and proceed from here by induction on N. In *exactly* the same manner as above, we can derive equality (8.2), inequality (8.3) and inclusion (8.4). As before, we combine these results with the assumption that $C\gamma<1$ and with the fact that $\{F_j\}_{j=1}^N$ is a tempered sequence such that by taking expectations, we obtain indeed that

$$\mathbb{E}\Big(\Lambda_g(\cdot)\,dm_L(g)\Big)=\mathbb{E}\left(\sum_{B\in\mathcal{Z}(\cdot)}|B|\right)\geq\delta(1+C\delta)^{-1}\left|\bigcup_{j=1}^{N}B_j\right|.$$

This shows statement (3) of the Decomposition Lemma and thus finishes the proof. □

Finally, we are able to prove the main theorem of this thesis. We give two slightly different proofs.

Corollary 8.8 (Pointwise ergodic theorem)
Let G be a second countable (and hence σ-compact), amenable group G acting on some σ-finite measure space $(X, \mathcal{B}, \mu)$ by measure preserving transformations. Further, let $1 \leq p < \infty$ and assume that $\{F_n\}$ is a tempered Følner sequence of G with constant $C > 0$. Then the pointwise convergence theorem holds, i.e. for every $f \in L^p(X)$ there is some G-invariant $f^* \in L^p(X)$ such that

$$A_n f(x) := |F_n|^{-1} \int_{F_n} f(gx)\, dm_L(g) \to f^*(x)$$

for μ-almost every $x \in X$ as $n \to \infty$.

Proof
It suffices to prove the transfer inequality (6.1). The theorem follows then from the transfer principle, Theorem 6.4 and Theorem 5.3.
To do so, fix $f \in L^p(X)$, $\lambda > 0$ and $x \in X$. Moreover, we set $\varepsilon = 1$ and for each $k \in \mathbb{N}$, we choose the sets $\tilde{F}_k$, F_{n_k} and $\overline{F}_k$ according to Lemma 6.3. Further, for each $k \in \mathbb{N}$ we apply the Decomposition Lemma, Theorem 8.1 with $\delta = 1$ on the sets $\overline{F}_k$ and $B_j := \{g \in F_{n_k} \,|\, |(A_j f)(gx)| \geq \lambda\}$ which have in fact the property that $F_j B_j \subseteq \overline{F}_k$ for all $1 \leq j \leq k$.
By the inequality (7.1), we obtain

$$\begin{aligned} \mathbb{E}\Big(\int_{\overline{F}_k} \Lambda_g(\cdot)\, |f(gx)|^p \, dm_L(g) \Big) &= \mathbb{E}\Big(\sum_{B \in \mathcal{Z}(\cdot)} \int_B |f(gx)|^p \, dm_L(g) \Big) \\ &\geq \lambda^p \, \mathbb{E}\Big(\sum_{B \in \mathcal{Z}(\cdot)} |B| \Big). \end{aligned}$$

Using the third statement of the Decomposition Lemma, we arrive at

$$\mathbb{E}\Big(\sum_{B \in \mathcal{Z}(\cdot)} |B| \Big) \geq (1 + C)^{-1} \cdot \left| \bigcup_{j=1}^{k} B_j \right|,$$

which implies that

$$\mathbb{E}\Big(\int_{\overline{F}_k} \Lambda_g(\cdot)\, |f(gx)|^p \, dm_L(g) \Big) \geq \lambda^p \, (1 + C)^{-1} \cdot \left| \bigcup_{j=1}^{k} B_j \right|. \tag{8.7}$$

On the other hand, using Fubini's Theorem, it follows from the second statement of the Decomposition Lemma that

$$\begin{aligned} \mathbb{E}\Big(\int_{\overline{F}_k} \Lambda_g(\cdot)\, |f(gx)|^p \, dm_L(g) \Big) &= \int_{\overline{F}_k} \mathbb{E}(\Lambda_g(\cdot))\, |f(gx)|^p \, dm_L(g) \\ &\leq 2 \int_{\overline{F}_k} |f(gx)|^p \, dm_L(g). \end{aligned} \tag{8.8}$$

Therefore, putting the inequalities (8.7) and (8.8) together, we arrive at

$$\left| \bigcup_{j=1}^{k} B_j \right| \leq \frac{2(C+1)}{\lambda^p} \int_{\overline{F}_k} |f(gx)|^p \, dm_L(g)$$

for all $k \in \mathbb{N}$, each $f \in L^p(X)$, every $\lambda > 0$ and almost every $x \in X$.
This proves the desired transfer inequality (6.1). □

With the Decomposition Lemma, one can likewise use EMERSON'S transfer principle (Theorem 6.1) to prove the individual ergodic theorem.

PROOF (POGORZELSKI 2010)
We define the operator S on $L^1_{loc}(G)$ as in Theorem 6.1, i.e. $(Sf)(g_0) := \sup_{n\in\mathbb{N}} |S_n f(g_0)|$, where

$$S_n f(g_0) := |F_n|^{-1} \int_{F_n} f(gg_0)\, dm_L(g)$$

for $f \in L^1_{loc}(G)$. We show that S is of weak type (p,p) which means that there is some $c > 0$ such that

$$m_L(\{g \in G \,|\, (Sf)(g_0) > \lambda\}) \le c\lambda^{-p} \|f\|^p_{L^p(G)}$$

for all $\lambda > 0$ and all $f \in L^p(G)$ (note that $L^p(G) \subseteq L^p_{loc}(G) \subseteq L^1_{loc}(G)$). Since G is σ-compact, one can find some sequence $\{\overline{B}_n\}$ of compact sets with $\cup_n \overline{B}_n = G$ and $\overline{B}_n \subseteq \overline{B}_{n+1}$ for all $n \in \mathbb{N}$. For $n \in \mathbb{N}$, we put further $\overline{F}_n := (\cup_{j=1}^n F_j)\overline{B}_n$. Hence, the sets $\overline{F}_n$ are compact and we have $F_j\overline{B}_j \subseteq \overline{F}_n$ for $n \in \mathbb{N}$ and $1 \le j \le n$.
We now fix $n_0 \in \mathbb{N}$, $f \in L^p(G)$, as well as $\lambda > 0$ and define $B_j := \{g \in \overline{B}_{n_0} \,|\, |(S_j f)(g)| \ge \lambda\}$ for $1 \le j \le n_0$. Similarly as in inequality (7.1), one computes for $g_0 \in B_j$

$$\begin{aligned}
\int_{F_j g_0} |f(g)|^p\, dm_L(g) &\ge |F_j g_0|^{-p/q} \Big(\int_{F_j g_0} |f(g)|\, dm_L(g) \Big)^p \\
&= |F_j g_0|^{-p/q} \Big(\Delta(g_0) \int_{F_j} |f(gg_0)|\, dm_L(g) \Big)^p \\
&\ge |F_j g_0|^{-p/q} (\Delta(g_0)^p \cdot \lambda^p \cdot |F_j|^p) \\
&= \lambda^p \cdot |F_j g_0|^{-p/q} \cdot |F_j g_0|^p \\
&= \lambda^p \cdot |F_j g_0|, \qquad (8.9)
\end{aligned}$$

where $q := p/(p-1)$ for $1 < p < \infty$. Note that this inequality holds trivially if $p = 1$.
We now apply the Decomposition Lemma (Theorem 8.1) to the compact set $\overline{F}_{n_0}$ and the sets B_j, $1 \le j \le n_0$. As usual, we choose $\delta = 1$ and the constant from the Shulman condition of the Følner sequence shall be denoted by C. For $\omega \in \Omega$, we write $\mathcal{Z}(\omega)$ for the resulting subcollection of right-translates from

$$\mathcal{F} := \{F_j g \,|\, 1 \le j \le n_0,\, g \in B_j\},$$

as well as $\Lambda_\cdot(\omega)$ for the corresponding counting function on $\overline{F}_{n_0}$.
By the second property of the decomposition, we obtain with Fubini's Theorem

$$\begin{aligned}
\mathbb{E}\Big(\sum_{B\in\mathcal{Z}(\cdot)} \int_B |f(g)|^p\, dm_L(g) \Big) &= \mathbb{E}\Big(\int_{\overline{F}_{n_0}} \sum_{B\in\mathcal{Z}(\cdot)} \mathbb{1}_B(g)\, |f(g)|^p\, dm_L(g) \Big) \\
&= \mathbb{E}\Big(\int_{\overline{F}_{n_0}} \Lambda_g(\cdot) |f(g)|^p\, dm_L(g) \Big) \\
&= \int_{\overline{F}_{n_0}} \mathbb{E}(\Lambda_g(\cdot))\, |f(g)|^p\, dm_L(g) \\
&\le \int_G 2\, |f(g)|^p\, dm_L(g) \\
&= 2\, \|f\|^p_{L^p(G)}. \qquad (8.10)
\end{aligned}$$

By the monotonicity of the expected value and inequality (8.9), it is true that

$$\mathbb{E}\Big(\sum_{B\in\mathcal{Z}(\cdot)}\int_B |f(g)|^p\, dm_L(g)\Big) \geq \lambda^p \cdot \mathbb{E}\Big(\sum_{B\in\mathcal{Z}(\cdot)}|B|\Big)$$

and we conclude with the third statement of the Decomposition Lemma that

$$\begin{aligned}\lambda^p \cdot (C+1)^{-1} \cdot \Big|\bigcup_{j=1}^{n_0} B_j\Big| &\leq \lambda^p \cdot \mathbb{E}\Big(\sum_{B\in\mathcal{Z}(\cdot)}|B|\Big)\\ &\leq \mathbb{E}\Big(\sum_{B\in\mathcal{Z}(\cdot)}\int_B |f(g)|^p\, dm_L(g)\Big).\end{aligned}$$

Combining this estimate with inequality (8.10) and noting that $n_0 \in \mathbb{N}$ was arbitrary, we see that

$$\Big|\bigcup_{j=1}^{n} B_j\Big| \leq \frac{2(C+1)}{\lambda^p}\cdot \|f\|^p_{L^p(G)} \tag{8.11}$$

for all $n \in \mathbb{N}$. Recall that the sets B_j depend on n, thus we now write $B_j^{(n)}$ instead. We define $D_n := \cup_{j=1}^n B_j^{(n)}$ for $n \in \mathbb{N}$ and claim that

$$\{g \in G \,|\, (Sf)(g) > \lambda\} \subseteq \bigcup_{n=1}^{\infty} D_n.$$

To see this, note that the event $[(Sf)(g) > \lambda]$ for some $g \in G$ implies that there is some $n \in \mathbb{N}$ such that $g \in \overline{B}_n$ and that there is some $l \in \mathbb{N}$ with $|(S_l f)(g)| \geq \lambda$. Since $\{\overline{B}_n\}$ is increasing this implies that $g \in \cup_{j=1}^d B_j^{(d)}$, where $d := \max\{l, n\}$.
By the choice of the sets $\overline{B}_n$, the sequence $\{D_n\}$ is increasing and since the right-hand side of inequality (8.11) does not depend on n we conclude that

$$\begin{aligned} m_L(\{g \in G \,|\, Sf(g) > \lambda\}) &\leq m_L(\cup_{n=1}^\infty D_n)\\ &= \lim_{n\to\infty} m_L\Big(\bigcup_{j=1}^n B_j^{(n)}\Big)\\ &\leq \frac{2(C+1)}{\lambda^p}\cdot \|f\|^p_{L^p(G)}.\end{aligned}$$

Hence, the operator S is indeed of weak type (p,p). By the transfer principle (Theorem 6.1), the maximal operator M with respect to the ergodic averages $\{A_n\}$ satisfies an L^p-maximal inequality. The pointwise ergodic theorem now follows from Theorem 5.3. □

Finally, we compare the methods presented in the Chapters 7 and 8. We have seen that Weiss' version of the transfer inequality (see Theorem 7.1) is based on the Abstract Combinatorial Lemma 7.2. Its proof does not require stochastic arguments but is rather complex from the combinatorial point of view and only works for countable groups. The Decomposition Lemma of Lindenstrauss (Theorem 8.1) based on random methods takes into account the general situation of σ-compact (second countable) groups. The use of the Poisson point processes in the non-discrete case is very elegant and yields a major breakthrough in the theory of pointwise ergodic theorems. By choosing the coverings randomly, Lindenstrauss overcomes the counting difficulties that one might have with selection methods of the Weiss type. Hence, the question about the construction of a deterministic algorithm for the general case is far from being trivial and has not been answered yet. Although the stochastic version for discrete groups needs more theory than the deterministic version, it is very clear and thus very understandable.

9 A short outlook

Besides the fact that the general pointwise ergodic theorem (Corollary 8.8) is a beautiful assertion on its own, it also has various applications.
In his article LINDENSTRAUSS enhances the Decomposition Lemma (Theorem 8.1) significantly (cf. [29], Section 2) and provides a powerful toolbox of combinatorial arguments for problems in entropy theory. For discrete amenable groups, he proved e.g. the existence of the entropy function $h(\mathcal{P})$ of measure preserving, ergodic actions on the measure space X as a function of finite measurable partitions $\mathcal{P}$ of X. More precisely, Theorem 1.3 in this paper describes a general version of the Shannon McMillan Breiman Theorem which states that $h(\mathcal{P})$ is obtained as an L^1- *and* an a.e.- limit of generalized entropy approximants involving rapidly growing, tempered Følner sequences $\{F_n\}$ as well as finite partitions $\mathcal{P}^{F_n}$ as shifts of $\mathcal{P}$ by the elements in F_n.

Further, the individual ergodic theorem can be helpful in mathematical physics. LENZ and STRUNGARU use this statement in their work about diffraction theory to construct averages of specific measures (see [28], Section 4).
Another interesting problem in mathematical physics is the existence of the integrated density of states of an operator which can be interpreted as cumulative distribution function of the eigenvalues. It is known for a couple of years now that Banach space valued ergodic theorems are instrumental for the uniform approximation of the integrated density of states of operators. One can find concrete results by LENZ and STOLLMANN for Delone dynamical systems in [27] as well as by LENZ, SCHWARZENBERGER and VESELIĆ for bounded operators on Cayley graphs of a large class of discrete amenable groups in [26].
As a concrete project, one might attempt to use LINDENSTRAUSS' decompositions (see [29], Section 2) in order to extend the latter result to *all* finitely generated amenable groups.

Bibliography

[1] M. A. Akcoglu, "A Pointwise Ergodic Theorem in L^p-spaces," *Canadian J. Math.*, vol. 27, pp. 1075–1082, (1975).

[2] G. D. Birkhoff, "Proof of the Ergodic Theorem," *Proc. Nat. Acad. Sci.*, vol. 17, pp. 656–660, (1931).

[3] S. Butkevich, "Convergence of Averages in Ergodic Theory," Ph.D. dissertation, Ohio State University, 2001.

[4] A. P. Calderon, "A General Ergodic Theorem," *Ann. Math.*, vol. 58, pp. 182–191, (1953).

[5] ——, "Ergodic Theory and Translation Invariant Operators," *Proc. Nat. Acad. Sci. USA*, vol. 59, pp. 349–353, (1968).

[6] J. B. Conway, *Functions of one complex variable.* Springer Verlag Berlin Heidelberg, 1978.

[7] M. M. Day, "Amenable Semigroups," *Illinois J. Math.*, pp. 509–544, (1957).

[8] ——, "Fixed Point Theorems for Compact Convex Sets," *Illinois J. Math.*, vol. 5, pp. 585–590, (1961).

[9] A. Deitmar and S. Echterhoff, *Principles of Harmonic Analysis.* Springer Verlag Berlin, 2004.

[10] N. Dunford and J. T. Schwarz, *Linear Operators, Part I.* John Wiley and Sons, New York, 1957.

[11] W. R. Emerson, "The Pointwise Ergodic Theorem for Amenable Groups," *American J. Math.*, vol. 96, pp. 472–487, (1974).

[12] W. R. Emerson and F. P. Greenleaf, "Covering Properties and Følner Conditions for Locally Compact Groups," *Math. Zeitschr.*, vol. 102, pp. 370–384, (1967).

[13] F. Erling, "On Groups with Full Banach Mean Value," *Math. Scandinavica*, vol. 3, pp. 243–254, (1955).

[14] B. Farkas, T. Eisner, M. Haase, and R. Nagel, *Ergodic Theory - An Operator Theoretic Approach.* `http://isem.mathematik.tu-darmstadt.de/isem`: 12th Internet Seminar, 2008.

[15] A. Garsia, *Topics in Almost Everywhere Convergence.* Markham Publishing Company Chicago, 1970.

[16] A. Ghaffari, "Ergodic Theory of Amenable Semigroup Actions," *Proc. Math. Sci./Indian Acad. Sci.*, vol. 117, pp. 177–183, (2007).

[17] F. P. Greenleaf, *Invariant Means on Topological Groups.* Van Nostrand Reinhold Companies New York, 1969.

[18] ——, "Ergodic Theorems and the Construction of Summing Sequences in Amenable Locally Compact Groups," *Communications Pure Applied Math.*, vol. 26, pp. 29–46, (1973).

[19] E. Hopf, "The General Temporally Discrete Markov Processes," *J. of Rat. Mechan. Analysis*, vol. 3, (1954).

[20] ——, "On the Ergodic Theorem for Positive Linear Operators," *J. Reine Angewandte Math.*, vol. 205, (1961).

[21] A. Hulanicki, "Means and Følner Conditions on Locally Compact Groups," *Studia Math.*, vol. 27, pp. 87–104, (1960).

[22] A. Ionescu Tulcea, "Ergodic Properties of Isometries in L^p-spaces, $1 < p < \infty$," *Bull. American Math. Soc.*, vol. 70, pp. 366–371, (1964).

[23] J. L. Kelley and I. Namioka, *Linear Topological Spaces.* Springer Verlag Berlin Heidelberg New York, 1963.

[24] J. F. C. Kingman, *Poisson Processes.* Clarendon Press Oxford, 1993.

[25] U. Krengel, *Ergodic Theorems.* Walter de Gruyter Berlin New York, 1985.

[26] D. Lenz, F. Schwarzenberger, and I. Veselić, "A Banach Space Valued Ergodic Theorem and the Integrated Density of States," *Geometricae Dedicata*, (2010).

[27] D. Lenz and P. Stollmann, "An Ergodic Theorem for Delone Dynamical Systems and Existence of the Integrated Density of States," *J. d'Analyse Math.*, vol. 97, pp. 1–23, (2005).

[28] D. Lenz and N. Strungaru, "Pure Point Spectrum for Measure Dynamical Systems on Locally Compact Abelian Groups," *J. Math. Pures Appl.*, vol. 92, pp. 323–341, (2009).

[29] E. Lindenstrauss, "Pointwise Theorems for Amenable Groups," *Invent. Math.*, vol. 146, pp. 259–295, (2001).

[30] M. I. Monastyrsky, "John von Neumann," *Russian Acad. Sci.*, vol. 174, pp. 1271–1280, (2004).

[31] R. Nagel, "Mittelergodische Halbgruppen linearer Operatoren," *Ann. Institut Fourier*, vol. 23, pp. 75–87, (1973).

[32] I. Namioka, "Følner's Condition for Amenable Semigroups," *Math. Scandinavica*, vol. 15, pp. 18–28, (1964).

[33] D. Ornstein and B. Weiss, "The Shannon-McMillan-Breiman Theorem for a Class of Amenable Groups," *Israel J. Math.*, vol. 44, pp. 53–60, (1983).

[34] ——, "Subsequence Theorems for Amenable Groups," *Israel J. Math.*, vol. 79, pp. 113–127, (1992).

[35] A. Rosenthal, "The 1913 papers of Rosenthal and Plancherel - Proof of the Impossibility of Ergodic Gas Systems," *Transp. Theory Statist. Phys.*, vol. 1, pp. 287–311, (1971).

[36] J. J. Rotman, *An Introduction to the Theory of Groups.* Springer Verlag New York, 1995.

[37] V. Runde, *Lectures on Amenability.* Springer Verlag Berlin Heidelberg New York, 2001.

[38] R. Serfozo, *Basics of Applied Stochastic Processes.* Springer Verlag Berlin Heidelberg, 2009.

[39] A. Shulman, "Maximal Ergodic Theorems on Groups," *Dep. Lit. NIINTI*, no. 2184, pp. 1–22, (1988).

[40] Y. G. Sinai and L. B. Koralov, *Theory of Probability and Random Processes.* Springer Verlag Berlin Heidelberg, 2007.

[41] M. Takesaki, *Theory of Operator Algebras III.* Springer Verlag Berlin Heidelberg New York, 2001.

[42] J. von Neumann, "Proof of the quasi-ergodic Hypothesis," *Proc. Nat. Acad. Sci. USA*, vol. 18, pp. 70–82, (1932).

[43] B. Weiss, "Topics in Dynamics and Ergodic Theory - Actions of Amenable Groups," *London Math. Soc. Lecture Notes 310*, pp. 226–262, (2003).

[44] D. Werner, *Funktionalanalysis.* Springer Verlag Berlin Heidelberg New York, 1995.

[45] N. Wiener, "The Ergodic Theorem," *Duke Math. J.*, vol. 5, pp. 1–18, (1939).