1 Appendix: Ergodic averages along the squares

by **Anthony Quas**[1] and **Máté Wierdl**[2]
University of Memphis
Department of Mathematical Sciences
373 Dunn Hall
Memphis, TN 38152-3240
quasa@msci.memphis.edu, mw@csi.hu

1.1 Enunciation of the result

In this note we want to present a proof of the almost everywhere convergence of the ergodic averages along the sequence of squares.

Theorem 1.1. *Let τ be a measurable, measure preserving transformation of the σ-finite measure space (X, Σ, μ).*

Then, for $f \in L^2$, the averages

$$S_t f(x) = \frac{1}{t} \sum_{n \le t} f(\tau^{n^2} x)$$

converge for almost every $x \in X$.

The theorem is due to J. Bourgain. To keep our presentation as continuous as possible, we present historical remarks, and cite references in the last section, Section 1.6.

1.2 Subsequence lemma

The main idea of the proof is to analyze the Fourier transform $\widehat{S}_t(\alpha) = 1/t \sum_{n \le t} e^{2\pi i n^2 \alpha}$ of the averages. This analysis permits us to replace the averages S_t by other operators that are easier to handle. The replace-ability of the sequence (S_t) by another sequence (A_t) means that we have an inequality of the form

$$\int \sum_t |S_t f - A_t f|^2 < c \int |f|^2. \tag{1.1}$$

[1]A. Quas' research is partially supported by NSF grant #DMS–0200703
[2]M. Wierdl's research is partially supported by NSF grant #DMS–0100577

Now, if somehow we prove that the sequence $(A_t f(x))$ converges for a.e. x, then the above inequality implies, since its left hand side is finite for $f \in L^2$, that the sequence $(S_t f(x))$ converges a.e. as well.

Well, we will not be able to prove an inequality of the type (1.1) exactly. In the real inequality, we will be able to have an inequality where the t runs through a lacunary sequence. But this is quite all right since it is enough to prove the a.e. convergence of the $(S_t f)$ along a lacunary sequence:

Lemma 1.2. *For $\sigma > 1$ denote*

$$I = I_\sigma = \{t \mid t = \sigma^n \text{ for some positive integer } n\}.$$

Suppose that for each fixed $\sigma > 1$, the sequence $(S_t f)_{t \in I}$ converges a.e.

Then the full (S_t) sequence converges a.e.

Proof. We can assume that the function f is nonnegative. For a given t, choose k so that $\sigma^k \leq t < \sigma^{k+1}$. We can then estimate as

$$S_t f(x) \leq \frac{1}{\sigma^k} \sum_{n \leq \sigma^{k+1}} f(\tau^{n^2} x) = \sigma \cdot S_{\sigma^{k+1}} f(x),$$

and similarly, we have $\sigma^{-1} \cdot S_{\sigma^k} f(x) \leq S_t f(x)$. This means that

$$\sigma^{-1} \cdot \lim_k S_{\sigma^k} f(x) \leq \liminf_t S_t f(x) \leq \limsup_t S_t f(x) \leq \sigma \cdot \lim_k S_{\sigma^k} f(x).$$

Choosing now $\sigma_p = 2^{2^{-p}}$, we get that $\lim_k S_{\sigma_p^k} f(x)$ is independent of p for a.e. x, and, by the above estimates, it is equal to $\lim_t S_t f(x)$. □

For the rest of the proof, we fix $\sigma > 1$, and unless we say otherwise, we always assume that $t \in I = I_\sigma$.

Definition 1.3. *If two sequences (A_t) and (B_t) of $L^2 \to L^2$ operators satisfy*

$$\int \sum_t |A_t f - B_t f|^2 < c \int |f|^2; \quad f \in L^2,$$

then we say that (A_t) and (B_t) are equivalent.

1.3 Oscillation and an instructive example

One standard way of proving a.e. convergence for the usual ergodic averages $1/t \sum_{n \le t} f(\tau^n x)$ is to first prove a maximal inequality, and then note that there is a natural dense class for which a.e. convergence holds.

Unfortunately, the second part of this scheme does not work for the averages along the squares, since there is no known class of functions for which it would be easy to prove a.e. convergence of the averages.

Instead, for the squares, we will prove a so called *oscillation inequality*: for any $t(1) < t(2) < \dots$ with $t(k) \in I$, there is a constant c so that we have

$$\int \sum_k \sup_{t(k)<t<t(k+1)} |S_t f - S_{t(k+1)} f|^2 \le c \int f^2. \tag{1.2}$$

We leave it to the reader to verify why an oscillation inequality implies a.e. convergence of the sequence $(S_t f)$. We also leave it to the reader to verify that if two operator sequences (A_t) and (B_t) are equivalent and (A_t) satisfies an oscillation inequality, then so does (B_t).

An important remark is that by the so called *transference principle* of Calderón, it is enough to prove the inequality in (1.2) on the integers $\mathbb{Z}$ which we consider equipped with the counting measure and the right shift. In this case, we have $S_t f(x) = 1/t \sum_{n \le t} f(x + n^2)$.

To see how Fourier analysis can help in proving an oscillation inequality, let us look at a simpler example first: the case of the usual ergodic averages $U_t f(x) = 1/t \sum_{n \le t} f(x + n)$ (by the transference principle, we only need to prove the oscillation inequality on the integers).

Let us assume that we already know the maximal inequality

$$\int_{\mathbb{Z}} \sup_t |U_t f|^2 \le c \cdot \int_{\mathbb{Z}} |f|^2$$

For the Fourier transform $\widehat{U}_t(\alpha) = 1/t \sum_{n \le t} e^{2\pi i n \alpha}$, $\alpha \in (-1/2, 1/2)$ we easily obtain the estimates

$$|\widehat{U}_t(\alpha) - 1| \le c \cdot t \cdot |\alpha|; \tag{1.3}$$

$$|\widehat{U}_t(\alpha)| \le \frac{c}{t \cdot |\alpha|}. \tag{1.4}$$

The first estimate is effective (nontrivial) when $|\alpha| < 1/t$ and it says that $\widehat{U}_t(\alpha)$ is close to 1. The second estimate is effective when $|\alpha| > 1/t$, and it says that then $|\widehat{U}_t(\alpha)|$ is small. In other words, the estimates in (1.3) and (1.4) say that the function $\mathbb{1}_{(-1/t,1/t)}(\alpha)$ captures the "essence" of $\widehat{U}_t(\alpha)$. How? Let us define the operator A_t via its Fourier transform as $\widehat{A}_t(\alpha) = \mathbb{1}_{(-1/t,1/t)}(\alpha)$. The great advantage of the (A_t) is that it is a monotone sequence of projections. We'll see in a minute how this can help. First we claim that the sequences (U_t) and (A_t) are equivalent. To prove this claim, start by observing that

$$\widehat{U_t f}(\alpha) = \widehat{U}_t(\alpha) \cdot \widehat{f}(\alpha); \quad \widehat{A_t f}(\alpha) = \widehat{A}_t(\alpha) \cdot \widehat{f}(\alpha),$$

and then estimate, using Parseval's formula, as

$$\begin{aligned}
\int_{\mathbb{Z}} \sum_{t\in I} |A_t f - U_t f|^2 &= \int_{-1/2}^{1/2} \sum_{t\in I} |\widehat{A_t}(\alpha) - \widehat{U}_t(\alpha)|^2 \cdot |\widehat{f}(\alpha)|^2 \, d\alpha \\
&\leq \int_{-1/2}^{1/2} |\widehat{f}(\alpha)|^2 \, d\alpha \cdot \sup_{\alpha} \sum_{t\in I} |\widehat{U}_t(\alpha) - \widehat{A}_t(\alpha)|^2 \\
&= \int_{\mathbb{Z}} f^2 \cdot \sup_{\alpha} \sum_{t\in I} |\widehat{U}_t(\alpha) - \widehat{A}_t(\alpha)|^2
\end{aligned}$$

It follows that it is enough to prove the inequality

$$\sup_{\alpha} \sum_{t\in I} |\widehat{U}_t(\alpha) - \widehat{A}_t(\alpha)|^2 < \infty.$$

To see this, for a fixed α, divide the summation on t into two parts, $t < |\alpha|^{-1}$ and $t > |\alpha|^{-1}$. For the case $t < |\alpha|^{-1}$, use the estimate in (1.3) and in case $t > |\alpha|^{-1}$ use the estimate in (1.4). In both cases, we end up with a geometric progression with quotient $1/\sigma$.

Since (U_t) and (A_t) are equivalent and (U_t) satisfies a maximal inequality, the operators A_t also satisfy a maximal inequality. But then the sequence $(A_t f(x))$ satisfies an oscillation inequality. To see this, first note that if $t(k) \leq t \leq t(k+1)$ then $A_t f(x) - A_{t(k+1)} f(x) = A_t \left(A_{t(k)} f(x) - A_{t(k+1)} f(x)\right)$. It follows, that

$$\begin{aligned}
\int_{\mathbb{Z}} \sup_{t(k)<t<t(k+1)} |A_t f - A_{t(k+1)} f|^2 &= \int_{\mathbb{Z}} \sup_{t} \left|A_t \left(A_{t(k)} f - A_{t(k+1)} f\right)\right|^2 \\
&\leq c \cdot \int_{\mathbb{Z}} \left|A_{t(k)} f - A_{t(k+1)} f\right|^2,
\end{aligned}$$

since the sequence (A_t) satisfies a maximal inequality. But now the oscillation inequality follows from the inequality

$$\int_{\mathbb{Z}} \sum_k |A_{t(k)}f - A_{t(k+1)}f|^2 \leq \int_{\mathbb{Z}} f^2.$$

This inequality, in turn, follows by examining the Fourier transform of the left hand side.

Now the punchline is that the ergodic averages (U_t) also satisfy the oscillation inequality since (U_t) and (A_t) are equivalent.

Let us summarize the scheme above: the maximal inequality for (U_t) implies a maximal inequality for the (A_t) since the two sequences are equivalent. But the (A_t), being a monotone sequence of projections, satisfy an oscillation inequality. But then, again appealing to the equivalence of the two sequences, the (U_t) satisfies an oscillation inequality.

What we have learned is that if a sequence of operators $(B_t f)$ satisfies a maximal inequality, and it is equivalent to a monotone sequence (A_t) of projections, then $(B_t f)$ satisfies an oscillation inequality..

In the remaining sections we will see that the scheme of proving an oscillation inequality for the averages along the squares (S_t) is similar, and ultimately it will be reduced to proving a maximal inequality for a monotone sequence of projections.

1.4 Periodic systems and the circle method

The difference between the usual ergodic averages and the averages along squares is that the squares are not uniformly distributed in residue classes. Indeed, for example no number of the form $3n-1$ is a square. This property of the squares is captured well in the behavior of the Fourier transform, $\widehat{S}_t(\alpha) = 1/t\sum_{n\leq t} e^{2\pi i n^2\alpha}$: for a typical rational $\alpha = b/q$, $\lim_{t\to\infty}\widehat{S}_t(\alpha)$ is nonzero (while it would be 0 if the squares were uniformly distributed $\mod q$).

We need some estimates on the Fourier transform $\widehat{S}_t(\alpha)$. Since we will often deal with the function $e^{2\pi i\beta}$, we introduce the notation $e(\beta) = e^{2\pi i\beta}$. Also, the estimates for the Fourier transform $\widehat{S}_t(\alpha)$ are simpler if instead of the averages $\frac{1}{t}\sum_{n\leq t}\tau^{n^2}f(x)$ we consider the weighted averages $1/t\sum_{n^2\leq t}(2n-1)\tau^{n^2}f(x)$. The weight $2n-1$ is motivated by $n^2-(n-1)^2 = 2n-1$. Everything we said about the averages along the squares applies equally well to these new weighted averages. Furthermore, it is an exercise in summation by parts to show that the a.e. convergence of the weighted and non weighted averages is equivalent.

So from now on, we use the notation

$$S_t f(x) = \frac{1}{t} \sum_{n^2 \le t} (2n-1) \tau^{n^2} f(x).$$

Let $\widehat{\Lambda}(\alpha) = \lim_t \widehat{S}_t(\alpha)$. By Weyl's theorem, $\widehat{\Lambda}(\alpha) = 0$ for irrational α and for rational $\alpha = b/q$, if b/q is in reduced terms, we have the estimate

$$|\widehat{\Lambda}(b/q)| \le \frac{c}{q^{1/2}} \tag{1.5}$$

This inequality tells us that while the squares are not uniformly distributed in residue classes $\mod q$, at least they try to be: $\widehat{\Lambda}_t(b/q) \to 0$ as $q \to \infty$.

Now the so called *circle method* of Hardy and Littlewood tells us about the structure of $\widehat{S}_t(\alpha)$. Let us introduce the notations $P(t) = t^{1/3}$, $Q(t) = 2t/P(t) = 2t^{2/3}$. According to the circle method, we have the following estimates

$$\left|\widehat{S}_t(\alpha) - \widehat{\Lambda}(b/q) \cdot \widehat{U}_t(\alpha - b/q)\right| \le c \cdot t^{-1/6}; \quad q \le P(t), \quad |\alpha - b/q| < 1/Q(t) \tag{1.6}$$

$$\left|\widehat{S}_t(\alpha)\right| < c \cdot t^{-1/6}, \quad \text{otherwise,} \tag{1.7}$$

where recall that U_t denotes the usual ergodic averages so $\widehat{U}_t(\beta) = 1/t \sum_{n \le t} e(n\alpha)$. In other words, the estimate above tells us that $\widehat{S}_t(\alpha)$ is close to $\widehat{\Lambda}(b/q) \cdot \widehat{U}_t(\alpha - b/q)$ if α is close to a rational point b/q with small denominator, and otherwise $|\widehat{S}_t(\alpha)|$ is small.

Given these estimates, it is easy to see that the sequence (S_t) is equivalent to the sequence (A_t) defined by its Fourier transform as

$$\widehat{A}_t(\alpha) = \sum_{\substack{b/q \\ q \le P(t)}} \widehat{\Lambda}(b/q) \cdot \widehat{U}_t(\alpha - b/q) \cdot \mathbb{1}_{(-1/Q(t), 1/Q(t))}(\alpha - b/q)$$

It remains to prove an oscillation inequality for the A_t. To do this, first we group those b/q for which q is of similar size:

$$E_p = \{b/q \mid 2^p \le q < 2^{p+1}\}$$

By the estimates in (1.5), we have

$$\sup_{b/q \in E_p} |\widehat{\Lambda}(b/q)| \leq c \cdot 2^{-p/2}. \tag{1.8}$$

Note also that if $b/q \in E_p$ then the term $\widehat{\Lambda}(b/q)$ occurs in the definition of A_t only when $t > 2^{3p}$. Define the operator $A_{p,t}$ by its Fourier transform as

$$\widehat{A}_{p,t}(\alpha) = \sum_{b/q \in E_p} \widehat{\Lambda}(b/q) \cdot \widehat{U}_t(\alpha - b/q) \cdot \mathbb{1}_{(-1/Q(t), 1/Q(t))}(\alpha - b/q), \quad t > 2^{3p}.$$

Using the triangle inequality for the summation in p, we see that an oscillation inequality for (A_t) would follow from the inequality

$$\int_{\mathbb{Z}} \sum_k \sup_{t(k) < t < t(k+1)} |A_{p,t}f - A_{p,t(k+1)}f|^2 \leq c \cdot \frac{p^2}{2^p} \cdot \int_{\mathbb{Z}} f^2 \tag{1.9}$$

We have learned in the previous section, Section 1.3, that it is useful to try work with projections. As a step, we introduce the operators $B_{p,t}$ defined via

$$\widehat{B}_{p,t}(\alpha) = \sum_{b/q \in E_p} \widehat{\Lambda}(b/q) \cdot \mathbb{1}_{(-1/t, 1/t)}(\alpha - b/q), \quad t > 2^{3p}.$$

Note that for each α there is at most one $b/q \in E_p$ so that $\mathbb{1}_{(-1/t,1/t)}(\alpha - b/q) \neq 0$ or $\mathbb{1}_{(-1/Q(t),1/Q(t))}(\alpha - b/q) \neq 0$ for some $t > 2^{3p}$. Hence, using the estimates in (1.3), (1.4), and (1.8), we get

$$|\widehat{A}_{p,t}(\alpha) - \widehat{B}_{p,t}(\alpha)| \leq c \cdot 2^{-p/2} \cdot \min\{t|\alpha|, (t|\alpha|)^{-1}\}; \quad t > 2^{3p}.$$

It follows that we can replace the $(A_{p,t})$ by the $(B_{p,t})$:

$$\int_{\mathbb{Z}} \sum_{t > 2^{3p}} |A_{p,t}f - B_{p,t}f|^2 \leq c \cdot 2^{-p} \cdot \int_{\mathbb{Z}} f^2$$

In order to prove the required oscillation inequality for the $B_{p,t}$, we make one more reduction. Namely, we claim that defining $C_{p,t}$ by

$$\widehat{C}_{p,t}(\alpha) = \sum_{b/q \in E_p} \mathbb{1}_{(-1/t,1/t)}(\alpha - b/q), \quad t > 2^{3p}.$$

(so $\widehat{C}_{p,t}$ is just $B_{p,t}$ without the multipliers $\widehat{\Lambda}(b/q)$), we need to prove

$$\int_{\mathbb{Z}} \sum_k \sup_{t(k)<t<t(k+1)} |C_{p,t} - C_{p,t(k+1)}|^2 \leq c \cdot p^2 \cdot \int_{\mathbb{Z}} f^2. \tag{1.10}$$

We leave the proof of this implication to the reader with the hint to replace the function f by g defined by its Fourier transform as

$$\widehat{g}(\alpha) = \sum_{b/q \in E_p} \widehat{\Lambda}(b/q) \cdot \mathbb{1}_{(-2^{-3p}, 2^{-3p})}(\alpha - b/q) \cdot \widehat{f}(\alpha).$$

Indeed, then $B_{p,t} f(x) = C_{p,t} g(x)$ and $\int_{\mathbb{Z}} g^2 \leq c \cdot 2^{-p} \int_{\mathbb{Z}} f^2$ by (1.8).

Now, the $C_{p,t}$ form a monotone (in t) sequence of projections, and hence they will satisfy the oscillation inequality in (1.10) once they satisfy the maximal inequality

$$\int_{\mathbb{Z}} \sup_{t > 2^{3p}} |C_{p,t}|^2 \leq c \cdot p^2 \cdot \int_{\mathbb{Z}} f^2. \tag{1.11}$$

To encourage the reader, we emphasize that our only remaining task is to prove the inequality in (1.11) above.

1.5 The main inequality

Since the least common multiple of the denominators of rational numbers in the set E_p is not greater than 2^{cp2^p} and the distance between two elements of E_p is at least 2^{-2p}, the estimate in (1.11) follows from the following result

Theorem 1.4. *Let $0 < \delta < 1/2$ and $e(\alpha_1), e(\alpha_2), \ldots, e(\alpha_J)$ be distinct complex Q-th roots of unities with $|\alpha_i - \alpha_j| > \delta$ for $i \neq j$. We assume that $\delta^{-1} \leq Q$. Define the projections R_t by*

$$\widehat{R}_t(\alpha) = \sum_{j \leq J} \mathbb{1}_{(-1/t, 1/t)}(\alpha - \alpha_j).$$

Then we have, with an absolute constant c,

$$\int_{\mathbb{Z}} \sup_{t \geq \delta^{-1}} |R_t f|^2 \leq c \cdot (\log \log Q)^2 \cdot \int_{\mathbb{Z}} |f|^2.$$

We restrict the the range on t to $t \geq \delta^{-1}$, because then the sum making up R_t contains pairwise orthogonal elements—as a result on the separation hypothesis $|\alpha_i - \alpha_j| > \delta$.

Proof. Two essentially different techniques will be used to handle the supremum. The first technique will handle the range $\delta^{-1} \leq t < Q^4$, and the other technique will handle the remaining $t > Q^4$ range.

Let us start with proving the inequality

$$\int_{\mathbb{Z}} \sup_{\delta^{-1} \leq t \leq Q^4} |R_t f|^2 \leq c \cdot (\log \log Q)^2 \cdot \int_{\mathbb{Z}} |f|^2. \tag{1.12}$$

We can assume that Q^4 is a power of σ, say $Q^4 = \sigma^S$, and then the range $\delta^{-1} \leq t \leq Q^4$ can be rewritten as $c \log \delta^{-1} \leq s \leq S$, where we take log with base σ. Introduce the monotone sequence of projections $P_s = R_{\sigma^{S-s}}$, $s \leq S - c \log \delta^{-1}$. All follows from

$$\int_{\mathbb{Z}} \sup_{s \leq S - c \log \delta^{-1}} |P_s f|^2 \leq c \cdot \log^2 S \cdot \int_{\mathbb{Z}} |f|^2.$$

It is clearly enough to show the inequality for dyadic $S - c \log \delta^{-1}$:

$$\int_{\mathbb{Z}} \sup_{s \leq 2^M} |P_s f|^2 \leq c \cdot M^2 \cdot \int_{\mathbb{Z}} |f|^2.$$

For each integer $m \leq M$ consider the sets

$$H_m = \left\{ P_{(d+1)\cdot 2^m} - P_{d \cdot 2^m} \mid d = 0, 1, \ldots, 2^{M-m} - 1 \right\}.$$

If the dyadic expansion of s is $s = \sum_{m \leq M} \epsilon_m \cdot 2^m$, where ϵ_m is 0 or 1, then for some $X_m \in H_m$, $P_s = \sum_{m \leq M} \epsilon_m \cdot X_m$. It follows that

$$|P_s f(x)|^2 \leq M \cdot \sum_{m \leq M} |X_m f(x)|^2.$$

For each m, we have

$$|X_m f(x)|^2 \leq \sum_{d \leq 2^{M-m}} |P_{(d+1)\cdot 2^m} f(x) - P_{d \cdot 2^m} f(x)|^2,$$

hence

$$\begin{aligned}\int_{\mathbb{Z}} \sup_{s\le 2^M} |P_s f|^2 &\le M\cdot \int_{\mathbb{Z}} \sum_{m\le M} \sum_{d\le 2^{M-m}} |P_{(d+1)\cdot 2^m} f(x) - P_{d\cdot 2^m} f(x)|^2 \\ &\le M\cdot \sum_{m\le M} \sum_{s\le 2^M} \int_{\mathbb{Z}} |P_{s+1} f - P_s f|^2 \\ &\le M^2\cdot 2\cdot \int_{\mathbb{Z}} |f|^2.\end{aligned}$$

Let us now handle the remaining range for t. We want to prove

$$\int_{\mathbb{Z}} \sup_{t>Q^4} |R_t f|^2 \le c\cdot \int_{\mathbb{Z}} |f|^2. \tag{1.13}$$

It seems best if we replace the operators R_t by the operators

$$A_t f(x) = \frac{1}{t} \sum_{n\le t} \sum_{j\le J} e(n\alpha_j) f(x+n).$$

This replacement is possible if we prove the following two inequalities

$$\int_{\mathbb{Z}} \sum_{t>\delta^{-2}} |A_t f - R_t f|^2 \le c\cdot \int_{\mathbb{Z}} |f|^2. \tag{1.14}$$

and

$$\int_{\mathbb{Z}} \sup_{t>Q^4} |A_t f|^2 \le c\cdot \int_{\mathbb{Z}} |f|^2. \tag{1.15}$$

Let us start with proving (1.14). By Parseval's formula, we need to prove

$$\sup_{\alpha} \sum_{t} |\widehat{A}_t(\alpha) - \widehat{R}_t(\alpha)|^2 < \infty.$$

Fix α. Without loss of generality we can assume that of the α_j, the point α_1 is closest to α. Possibly dividing the sum on j into two and reindexing them, we also assume that $\alpha_1 < \alpha_2 < \cdots < \alpha_J$. Using the separation hypothesis $|\alpha_i - \alpha_j| > \delta$ for $i \ne j$, we have that $|\alpha - \alpha_j| > (j-1)\delta$ for $j > 1$.

For $t \le 1/|\alpha-\alpha_1|$ we can thus estimate (recall that $\widehat{U}_t(\beta) = 1/t\sum_{n\le t} e(n\beta)$) as

$$\begin{aligned}
|\widehat{A}_t(\alpha) - \widehat{R}_t(\alpha)| &\le |\widehat{U}_t(\alpha-\alpha_1)| + \sum_{2\le j\le J} |\widehat{U}_t(\alpha-\alpha_j)| \text{ by (1.3) and (1.4)} \\
&\le c\cdot\left(t|\alpha-\alpha_1| + \sum_{2\le j\le J}\frac{1}{t(j-1)\delta}\right) \\
&\le c\cdot (t|\alpha-\alpha_1| + \log J/(\delta t)) \\
&\le c\cdot \big(t|\alpha-\alpha_1| + \delta^{-2}/t\big)
\end{aligned}$$

where we used in the last estimate that $J \le \delta^{-1}$. Summing this estimate over $t \in I$ with $\delta^{-2} \le t \le 1/|\alpha-\alpha_1|$ we get a finite bound independent of α.

For $t > 1/|\alpha-\alpha_1|$, we have

$$|\widehat{A}_t(\alpha) - \widehat{R}_t(\alpha)| \le \sum_{1\le j\le J} |\widehat{U}_t(\alpha-\alpha_j)| \le c\cdot\frac{\delta^{-2}}{t},$$

which, upon summing over the full range $\delta^{-2} < t$, again gives a finite bound independent of α.

Let us single out a consequence of inequality (1.14): there is a constant c so that

$$\int_{\mathbb{Z}} |A_t f|^2 \le c\cdot\int_{\mathbb{Z}} |f|^2; \quad t > \delta^{-2}. \tag{1.16}$$

Our only remaining task is to prove inequality (1.15).

For a given t, let q be the largest integer so that $qQ^2 \le t$. Note that $q \ge Q^2$ since $t > Q^4$. We can estimate as

$$\begin{aligned}
&\Big|\sum_{n\le t}\sum_{j\le J} e(n\alpha_j) f(x+n)\Big| \\
&\quad\le \Big|\sum_{n\le qQ^2}\sum_{j\le J} e(n\alpha_j) f(x+n)\Big| + \Big|\sum_{qQ^2<n\le t}\sum_{j\le J} e(n\alpha_j) f(x+n)\Big|.
\end{aligned}\tag{1.17}$$

We estimate the second term on the right trivially as

$$\Big|\sum_{qQ^2<n\le t}\sum_{j\le J} e(n\alpha_j) f(x+n)\Big| \le J\cdot\sum_{qQ^2<n\le (q+1)Q^2}\big|f(x+n)\big|.$$

With this, we have

$$\begin{aligned}
&\sup_{t>Q^4}\Big(\frac{1}{t}\Big|\sum_{qQ^2<n\le t}\sum_{j\le J} e(n\alpha_j)f(x+n)\Big|\Big)^2 \\
&\quad\le \sup_{q\ge Q^2}\Big(\frac{J}{qQ^2}\cdot\sum_{qQ^2<n\le(q+1)Q^2}|f(x+n)|\Big)^2 \text{ by Cauchy's inequality} \\
&\quad\le \sup_{q\ge Q^2}\frac{J^2\cdot Q^2}{qQ^2}\cdot\frac{\sum_{qQ^2<n\le(q+1)Q^2}|f(x+n)|^2}{qQ^2} \\
&\quad\le \sum_{q\ge Q^2}\frac{J^2}{q^2}\cdot\frac{1}{Q^2}\sum_{qQ^2<n\le(q+1)Q^2}|f(x+n)|^2
\end{aligned}$$

Integrating the last line, we obtain the bound

$$\sum_{q\ge Q^2}\frac{J^2}{q^2}\cdot\int_{\mathbb{Z}}|f|^2 \le c\cdot\frac{J^2}{Q^2}\int_{\mathbb{Z}}|f|^2 \le c\cdot\int_{\mathbb{Z}}|f|^2.$$

since $J \le Q$.

Let us now handle the first term on the right of (1.17). Since $e(\alpha_j)$ satisfies $e((mQ^2+h)\alpha_j) = e(h\alpha_j)$ (this is the first and last time we use that the $e(\alpha_j)$ are Q-th roots of unities), we can write, defining $Tg(x) = g(x+Q^2)$,

$$\Big|\frac{1}{t}\sum_{n\le qQ^2}\sum_{j\le J} e(n\alpha_j)f(x+n)\Big| \le \Big|\frac{1}{q}\sum_{m\le q} T^m\frac{1}{Q^2}\sum_{h\le Q^2}\sum_{j\le J} e(h\alpha_j)f(x+h)\Big|.$$

By the ergodic maximal inequality, applied to T, the ℓ^2 norm of our maximal operator is bounded by the ℓ^2 norm of

$$\frac{1}{Q^2}\sum_{h\le Q^2}\sum_{j\le J} e(h\alpha_j)f(x+h).$$

But the estimate in (1.14) says, the ℓ^2 norm of the above is bounded independently of Q since $Q^2 > \delta^{-2}$ by assumption. □

1.6 Notes

More details More details and references can be found in [RosW]. In particular, the circle method and the transference principle are described in complete details—though no proof of the main inequality of Bourgain, Theorem 1.4, is given. The inequalities (1.6) and (1.7) appear as (4.23) and (4.24) in [RosW].

Theorem 1.1 The result is due to Bourgain ([Bou1]). He later extended the result to $f \in L^p$, $p > 1$; cf [Bou2]. The case $p = 1$ is the most outstanding unsolved problem in this subject.

Idea of proof The basic structure of the proof is that of Bourgain's ([Bou2]) but we used ideas from Lacey's paper [La] as well—not to mention some personal communication with M. Lacey.

Other sequences The sequence of primes is discussed in [Wi]. But we'd like to emphasize that the L^2 theory of the primes is identical to the case of the squares. The only difference is in the estimates in (1.6) and (1.7).

A characterization of sequences which are good for the pointwise and mean ergodic theorems can be found in [BoQW].

References

[A1] J. Auslander, On the proximal relation in topological dynamics, *Proc. Amer. Math. Soc.* **11** (1960), 890–895.

[Ba] S. Banach, Sur le problème de la mesure, *Fund. Math.*, **4**, (1923), 7–33.

[BaT] S. Banach and A. Tarski, Sur la décomposition des ensembles de points en parties respectivement congruentes, *Fund. Math.* **6** (1924) 244–277.

[BB] D. Berend and V. Bergelson, Jointly ergodic measure-preserving transformations, *Israel J. Math.* **49** (1984), no. 4, 307–314.

[Be1] V. Bergelson, Sets of recurrence of $\mathbb{Z}^m$-actions, and properties of sets of differences in $\mathbb{Z}^m$, *J. London Math Soc. (2)* **31** (1985), 295–304.

[Be2] V. Bergelson, Weakly mixing PET, *Erg. Th. and Dynam. Sys.* **7** (1987), 337–349.

[Be3] V. Bergelson, Ergodic Ramsey theory - an update, *Ergodic theory of $\mathbb{Z}^d$ actions (Warwick, 1993-1994)*, 273–296, London Math. Soc. Lecture Note Ser., 228, Cambridge Univ. Press, Cambridge, 1996.

[Be4] V. Bergelson, The multifarious Poincare recurrence theorem, *Descriptive set theory and dynamical systems (Marseille-Luminy, 1996)*, 31–57, London Math Soc. Lecture Note Ser. 277, Cambridge Univ. Press, Cambridge, 2000.

[Be5] V. Bergelson, Ergodic theory and diophantine problems, *Topics in symbolic dynamics and applications (Temuco, 1997)*, 167-205, London Math Soc. Lecture Note Ser. 279, Cambridge Univ. Press, Cambridge, 2000.

[Be6] V. Bergelson, Minimal idempotents and ergodic Ramsey theory, *Topics in dynamics and ergodic theory, 8–39*, London Math. Soc. Lecture Note Ser., 310, Cambridge Univ. Press, Cambridge, 2003.

[Be7] V. Bergelson, Multiplicatively large sets and ergodic Ramsey theory, to appear in Israel J. Math.

[BeBB] V. Bergelson, M. Boshernitzan, and J. Bourgain, Some results on nonlinear recurrence, *J. Anal. Math.* **62** (1994), 29–46.

[BeBH] V. Bergelson, A. Blass and N. Hindman, Partition theorems for spaces of variable words, *Proc. London Math. Soc. (3)* **68** (1994), no. 3, 449–476.

[BeFHK] V. Bergelson, H. Furstenberg, N. Hindman, and Y. Katznelson, An algebraic proof of van der Waerden's theorem, *Enseign. Math. (2)* **35** (1989), no. 3-4, 209–215.

[BeFM] V. Bergelson, H. Furstenberg, and R. McCutcheon, IP-sets and polynomial recurrence, *Ergodic Theory and Dynam. Systems* **16** (1996), no. 5, 963–974.

[BeG] V. Bergelson and A. Gorodnik, Weakly mixing group actions: a brief survey and an example, *Modern Dynamical Systems and Applications, B. Hasselblatt, M. Brin, Y. Pesin, eds.*, 3–25, Cambridge University Press, New York, 2004.

[BeH1] V. Bergelson and N. Hindman, Nonmetrizable topological dynamics and Ramsey theory, *Trans. AMS.* **320** (1990), 293–320.

[BeH2] V. Bergelson and N. Hindman, Some topological semicommutative van der Waerden type theorems and their combinatorial consequences, *J. London Math. Soc. (2)* **45** (1992), no. 3, 385–403.

[BeH3] V. Bergelson and N. Hindman, Additive and multiplicative Ramsey theorems in N—some elementary results, *Combin. Probab. Comput.* **2** (1993), no. 3, 221–241.

[BeH4] V. Bergelson and N. Hindman, On IP* sets and central sets, *Combinatorica* **14** (1994), no. 3, 269–277.

[BeL1] V. Bergelson and A. Leibman, Polynomial extensions of van der Waerden's and Szemerédi's theorems, *Journal of AMS* **9** (1996), 725 –753.

[BeL2] V. Bergelson and A. Leibman, Set-polynomials and polynomial extension of the Hales-Jewett theorem. *Ann. of Math. (2)* **150** (1999), no. 1, 33–75.

[BeL3] V. Bergelson and A. Leibman, A nilpotent Roth theorem. *Invent. Math.* **147** (2002), no. 2, 429–470.

[BeL4] V. Bergelson and A. Leibman, Topological multiple recurrence for polynomial configurations in nilpotent groups. *Adv. Math.* **175** (2003), no. 2, 271–296.

[BeLM] V. Bergelson, A. Leibman and R. McCutcheon, Polynomial Szemerédi theorems for countable modules over integral domains and finite fields, to appear in *J. d'Analyse Math.*

[BeM1] V. Bergelson and R. McCutcheon, Uniformity in the polynomial Szemerédi theorem. *Ergodic theory of $\mathbb{Z}^d$ actions (Warwick, 1993-1994)*, 273–296, London Math. Soc. Lecture Note Ser., 228, Cambridge Univ. Press, Cambridge, 1996.

[BeM2] V. Bergelson and R. McCutcheon, An ergodic IP Szemerédi theorem. *Mem. Amer. Math. Soc.* **146** (2000), no. 695, viii+106 pp.

[BeM3] V. Bergelson and R. McCutcheon, Recurrence for semigroup actions and a non-commutative Schur theorem. *Topological dynamics and applications (Minneapolis, MN, 1995)*, 205–222, Contemp. Math., 215, Amer. Math. Soc., Providence, RI, 1998.

[BeMZ] V. Bergelson, R. McCutcheon and Q. Zhang, A Roth theorem for amenable groups, *Amer. J. Math.* **119** (1997), no. 6, 1173–1211.

[BeR1] V. Bergelson and J. Rosenblatt, Mixing actions of groups, *Illinois J. Math.* **32** (1988), no. 1, 65–80.

[BeR2] V. Bergelson and J. Rosenblatt, Joint ergodicity for group actions, *Ergodic Theory Dynam. Systems*, **8** (1988), no. 3, 351–364.

[BeS] V. Bergelson and D. Shapiro, Multiplicative subgroups of finite index in a ring. *Proc. Amer. Math. Soc.* **116** (1992), no. 4, 885–896.

[Bi] P. Billingsley, *Probability and Measure*, Wiley, New York, 1986.

[BPT] A. Blaszczyk, S. Plewik, S. Turek, Topological multidimensional van der Waerden theorem. *Comment. Math. Univ. Carolin.* **30** (1989), no. 4, 783–787.

[BoQW] M. Boshernitzan, G. Kolesnik, A. Quas and M. Wierdl, Ergodic averaging sequences, 33 pages, to appear in *J. d'Analyse Math.*, available as `http://www.csi.hu/mw/hardy_ergav.pdf` or `http://www.csi.hu/mw/hardy_ergav.dvi`

[Bou1] J. Bourgain, On the maximal ergodic theorem for certain subsets of the integers, *Israel J. Math.* **61** (1988), 39–72.

[Bou2] J. Bourgain, Pointwise ergodic theorems for arithmetic sets (appendix: The return time theorem), *Publ. Math. I.H.E.S.* **69** (1989), 5–45.

[Br] T. C. Brown, An interesting combinatorial method in the theory of locally finite semigroups. *Pacific J. Math.* **36** (1971) 285–289.

[Ca] T.J. Carlson, Some unifying principles in Ramsey theory, *Discrete Math.* **68** (1988), 117–169.

[CL1] J.P. Conze and E. Lesigne, Théorèmes ergodiques pour des mesures diagonales, *Bull. Soc. Math. France* **112** (1984) no. 2, 143–175.

[CL2] J.P. Conze and E. Lesigne, Sun un théorème ergodique pour des mesures diagonales, *Probabilités*, 1–31, Publ. Inst. Rech. Math. Rennes, 1987–1, Univ. Rennes I, Rennes, 1988.

[CL3] J.P. Conze and E. Lesigne, Sun un théorème ergodique pour des mesures diagonales, *C. R. Acad. Sci. Paris Sér. I Math.* **306** (1988) no. 12, 491–493.

[CN] W. Comfort and S. Negrepontis, *The Theory of Ultrafilters*, Springer-Verlag, Berlin and New York, 1974.

[CY] P. Civin and B. Yood, The second conjugate space of a Banach algebra as an algebra, *Pac. J. Math.* **11** (1961), 847–870.

[D1] L. E. Dickson, On the congruence $x^n + y^n = z^n \equiv 0 \pmod p$. *Journal of Reine Angew. Math.* **135** (1909) p.134–141.

[D2] L. E. Dickson, *History of the Theory of Numbers, vol. II (Diophantine analysis)*, Chelsea, New York, 1971.

[E1] R. Ellis, Distal transformation groups, *Pac. J. Math.* **8** (1958), 401–405.

[E2] R. Ellis, A semigroup associated with a transformation group, *Trans. AMS*, **94** (1960), 272–281.

[ET] P. Erdős and P. Turán, On some sequences of integers, *J. London Math. Soc.* **11** (1936), 261–264.

[Fø] E. Følner, On groups with full Banach mean values, *Math. Scand.* **3** (1955), 243–254.

[F1] H. Furstenberg, The structure of distal flows, *Amer. J. Math.* **85** (1963), 477–515.

[F2] H. Furstenberg, Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions, *J. d'Analyse Math.* **31** (1977), 204–256.

[F3] H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton University Press, 1981.

[F4] H. Furstenberg, Poincaré recurrence and number theory, *Bull. AMS (New Series)* **5** (1981), 211–234.

[FK1] H. Furstenberg and Y. Katznelson, An ergodic Szemerédi theorem for commuting transformations, *J. d'Analyse Math.* **34** (1978), 275–291.

[FK2] H. Furstenberg and Y. Katznelson, An ergodic Szemerédi theorem for IP-systems and combinatorial theory, *J. d'Analyse Math.* **45** (1985), 117–168.

[FK3] H. Furstenberg and Y. Katznelson, Idempotents in compact semigroups and Ramsey theory, *Israel J. Math.* **68** (1990), 257–270.

[FK4] H. Furstenberg and Y. Katznelson, A density version of the Hales-Jewett theorem, *J. d'Analyse Math.* **57** (1991), 64–119.

[FKO] H. Furstenberg, Y. Katznelson, D. Ornstein, The ergodic theoretical proof of Szemerédi's theorem, *Bull. Amer. Math. Soc. (N.S.)* **7** (1982), no. 3, 527–552.

[FW1] H. Furstenberg and B. Weiss, Topological dynamics and combinatorial number theory, *J. d'Analyse Math.* **34** (1978), 61–85.

[FW2] H. Furstenberg and B. Weiss, A mean ergodic theorem for $\frac{1}{N}\sum_{n=1}^{N} f(T^n x)g(T^{n^2}x)$, *Convergence in ergodic theory and probability (Columbus, OH, 1993)*, 193–227, Ohio State Univ. Math. Res. Inst. Publ., 5, de Gruyter, Berlin, 1996.

[GJ] L. Gillman and M. Jerison, *Rings of continuous functions*, Springer-Verlag, New York, 1976.

[GraRS] R. Graham, B. Rothschild and J. Spencer, *Ramsey Theory*, Wiley, New York, 1980.

[Gre] F. Greenleaf, *Invariant Means on Topological Groups*, von Nostrand, New York, 1969.

[HaJ] A.W. Hales and R.I. Jewett, Regularity and positional games, *Trans. AMS* **106** (1963), 222–229.

[Hal] P. Halmos, *Lectures on Ergodic Theory*, Mathematical Society of Japan, Tokyo, 1956.

[HalN] P. Halmos and J. von Neumann, Operator methods in classical mechanics, II, *Ann. of Math.* **43** (1942), 332–350.

[HaL] G. Hardy and J. Littlewood, Some problems of Diophantine approximation, *Acta Mathematica*, **37** (1914), 155–191.

[Hau] F. Hausdorff, *Grundzüge der Mengenlehre*, Verlag von Veit, Leipzig, 1914. Reprinted by Chelsea, New York, 1949.

[Hel] H. Helson, *Harmonic Analysis*, Addison-Wesley, Reading, MA, 1983.

[HeR] E. Hewitt and K. Ross, *Abstract Harmonic Analysis I*, Springer-Verlag, 1963.

[Hil] D. Hilbert, Über die Irreduzibilität ganzer rationaler Funktionen mit ganzzahligen Koeffizienten, *J. Math.* **110** (1892), 104–129.

[Hin1] N. Hindman, Finite sums from sequences within cells of a partition of $\mathbb{N}$, *J. Combinatorial Theory* (Series A), **17** (1974), 1–11.

[Hin2] N. Hindman, Partitions and sums and products of integers, *Trans. AMS*, **247** (1979), 227–245.

[HinS] N. Hindman and D. Strauss, *Algebra in the Stone-Čech Compactification. Theory and Applications*, de Gruyter Expositions in Mathematics, 27. Walter de Gruyter & Co., Berlin, 1998. xiv+485 pp.

[Hopf] E. Hopf, *Ergodentheorie,* Chelsea, New York, 1948.

[HoK1] B. Host and B. Kra, Convergence of Conze-Lesigne averages, *Ergodic Theory Dynam. Systems* **21** (2001), no. 2, 493–509.

[HoK2] B. Host and B. Kra, Nonconventional ergodic averages and nilmanifolds, to appear in *Annals of Math.*

[HoK3] B. Host and B. Kra, Convergence of polynomial ergodic averages, to appear in *Israel J. of Math.*

[KM] T. Kamae and M. Mendès-France, van der Corput's difference theorem, *Israel J. Math.* **31** (1978), no. 3-4, 335–342.

[Kh] A. Y. Khintchine, *Three Pearls of Number Theory*, Graylock Press, Rochester, N.Y., 1952.

[KoN] B.O. Koopman and J. von Neumann, Dynamical systems of continuous spectra, *Proc. Nat. Acad. Sci. U.S.A.*, **18** (1932), 255–263.

[Kre] U. Krengel, Weakly wandering vectors and weakly independent partitions, *Trans. Amer. Math. Soc.* **164** (1972), 199–226.

[Kro] L. Kronecker, Die Periodensysteme von Funktionen Reeller Variablen, *Berliner Sitzungsberichte* (1884), 1071–1080.

[La] M. Lacey, On an inequality due to Bourgain, *Illinois J. Math.* **41** (1997), 231–236.

[Le1] A. Leibman, Multiple recurrence theorem for nilpotent group actions, *Geom. Funct. Anal.* **4** (1994), no. 6, 648–659.

[Le2] A. Leibman, Multiple recurrence theorem for measure preserving actions of a nilpotent group, *Geom. and Funct. Anal.*, **8** (1998), 853–931.

[Le3] A. Leibman, The structure of unitary actions of finitely generated nilpotent groups, *Ergodic Theory Dynam. Systems* **20** (2000), 809–820.

[Le4] A. Leibman, Convergence of multiple ergodic averages along polynomials of several variables, to appear in *Israel J. of Math*

[Li] E. Lindenstrauss, Pointwise theorems for amenable groups, *Invent. Math.* **146** (2001), 259–295.

[McC] R. McCutcheon, *Elemental Methods in Ergodic Ramsey Theory*, Springer Lecture Notes in Mathematics 1722, Springer-Verlag, Berlin, 1999.

[M] K. Milliken, Ramsey's theorem with sums or unions, *J. Combinat. Th. (Series A)* **18** (1975), 276–290.

[Na] I. Namioka, Følner's condition for amenable semi-groups, *Math. Scand.*, **15** (1964), 18–28.

[Ne1] J. von Neumann, Zur algemeinen Theorie des Masses, *Fund. Math.*, **13** 1929, 73–116.

[Ne2] J. von Neumann, Zur Operatorenmethode in der klassischen Mechanik, *Ann. of Math.* **33** (1932), 587–642.

[Nu] K. Numakura, On bicompact semigroups, *Math. J. of Okayama University* **1** (1952), 99–108.

[Ol1] A. Olshanskii, On the question of the existence of an invariant mean on a group, *Uspekhi Mat. Nauk.* **35** (1980), no. 4 (214), 199–200.

[Ol2] A. Olshanskii, An infinite group with subgroups of prime orders, *Izv. Akad. Nauk SSSR Ser. Mat.* **44** (1980), no. 2, 309–321.

[OrW] D. Ornstein, B. Weiss, Entropy and isomorphism theorems for actions of amenable groups, *J. d' Analyse Math.*, **48** (1987), 1–141.

[Pa] A. Patterson, *Amenability*, Amer. Math. Soc., Providence, 1988.

[Pe] K. Petersen, *Ergodic Theory*, Cambridge Univ. Press, Cambridge, 1981.

[Pier] J.-P. Pier, *Amenable Locally Compact Groups*, John Wiley and Sons, 1984.

[Po] H. Poincaré, Sur le problème des trois corps et les équations de la Dynamique, Acta Mathematica **13** (1890), 1–270.

[Rado] R. Rado, Note on combinatorial analysis, *Proc. London Math. Soc.* **48** (1993), 122–160.

[Ram] F. P. Ramsey, On a problem of formal logic, *Proc. London Math. Soc.* **30** (1930), 264–286.

[Ri] P. Ribenboim, *13 Lectures on Fermat's Last Theorem*, Springer-Verlag, New York, 1979.

[RosW] J. Rosenblatt and M. Wierdl, Pointwise ergodic theorems via harmonic analysis, *Ergodic Theory and Its Connection with Harmonic Analysis (K. Petersen and I. Salama, eds.)*, 3–151, London Math. Soc. Lecture Note Ser. 205, Cambridge University Press, Cambridge, 1995.

[Roy] H. L. Royden, *Real analysis*, Macmillan, New York, 1968.

[Sa] A. Sárközy, On difference sets of integers III, *Acta. Math. Acad. Sci. Hungar.*, **31** (1978) 125–149.

[Schm] W. Schmidt, *Equations over Finite Fields: an Elementary Approach*, Springer-Verlag, Berlin ; New York, 1976.

[Schur] I. Schur, Über die Kongruenz $x^m + y^m \equiv z^m \pmod p$, *Jahresbericht der Deutschen Math.-Ver.* **25** (1916), 114–117.

[Sz] E. Szemerédi, On sets of integers containing no k elements in arithmetic progression, *Acta Arith.* **27** (1975), 199–245.

[T] A. Taylor, A canonical partition relation for finite subsets of ω, *J. Combinat. Th. (Series A)* **17** (1974), 1–11.

[vdW1] B. van der Waerden, Beweis einer Baudetschen Vermutung, *Nieuw. Arch. Wisk.* **15** (1927), 212–216.

[vdW2] B. van der Waerden, How the proof of Baudet's conjecture was found, *Studies in Pure Mathematics presented to Richard Rado (1971),* L. Mirsky, ed., Academic Press, London, 251–260.

[Wag] S. Wagon, *The Banach-Tarski Paradox*, Cambridge University Press, Cambridge, 1985.

[Wal] P. Walters, *An Introduction to Ergodic Theory*, Springer-Verlag, New York, 1982.

[Weyl] H. Weyl, Über die Gleichverteilung von Zahlen mod. Eins, *Math. Ann.* **77** (1916), 313–352.

[Wi] M. Wierdl, Pointwise ergodic theorem along the prime numbers, *Israel J. Math.*, **64** (1988), 315–336.

[Zh] Q. Zhang, On convergence of the averages $\frac{1}{N}\sum_{n=1}^{N} f_1(R^n x) f_2(S^n x) f_3(T^n x)$, *Monat. Math.* **122** (1996), 275-300.

[Zie] T. Ziegler, Universal characteristic factors and Furstenberg averages, preprint.

[Zim1] R. Zimmer, Extensions of ergodic group actions, *Illinois J. Math* **20** (1976), no. 3, 373–409.

[Zim2] R. Zimmer, Ergodic actions with generalized discrete spectrum, *Illinois J. Math* **20** (1976), no. 4, 555–588.