Does Order Matter

T. W. Körner

DPMMS
16 Mill Lane
Cambridge
`twk@pmms.cam.ac.uk`

ABSTRACT. I describe the work of Olevskiĭ, Tao and others on the rearrangement of orthogonal series. It turns out that arbitrary rearrangements produce trouble for all orthogonal and wavelet methods, that decreasing rearrangements produce trouble for Fourier series, but that wavelet expansions continue to work well under decreasing rearrangement.

1. Introduction

In this paper we shall move without comment between the circle $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, the closed interval $[0,1]$ and the half closed interval $[0,1)$ as seems most convenient. When our results deal with almost everywhere behaviour this does not create any problems, when we want everywhere convergence readers may have to produce their own argument to deal with the point 1.

What do we mean when we consider

$$\sum_{u=-\infty}^{\infty} \hat{f}(u) \exp 2\pi i u t\,?$$

Traditionally we take

$$\lim_{N\to\infty} \sum_{u=-N}^{N} \hat{f}(u) \exp 2\pi i u t$$

and the only variations that we allow concern the mode of convergence (pointwise, L^p, etc). A sign that this may be too narrow an approach appears when we consider the two dimensional case of a function $f : \mathbb{T}^2 \to \mathbb{C}$. The obvious way to proceed is to take

$$\sum_{(u,v)\in\mathbb{Z}^2} \hat{f}(u,v) \exp 2\pi i(ut+vs) = \lim_{N\to\infty} \sum_{(u,v)\in\Gamma(N)} \hat{f}(u,v) \exp 2\pi i(ut+vs)$$

with the $\Gamma(N)$ finite subsets of $\mathbb{Z}^2$ such that $\Gamma(1) \subseteq \Gamma(2) \subseteq \dots$ and $\bigcup_{N=1}^{\infty} \Gamma(N) = \mathbb{Z}^2$. However, it is well known that, even when f is quite well behaved, different choices of the sequence $\Gamma(N)$ give rise to different behaviour.

The question of 'correct order' also arises, even in the one dimensional case, from signal processing. If we seek to store and reconstruct a function $f : \mathbb{T} \to \mathbb{C}$ by using its Fourier coefficients it is natural to to use them in decreasing order of magnitude and to consider

$$\sum_{|\hat{f}(u)| \geqslant \eta} \hat{f}(u) \exp 2\pi i u t$$

rather than

$$\sum_{|u| \leqslant N} \hat{f}(u) \exp i u t.$$

Any easy optimism about rearrangements is quenched by the following result.

THEOREM 1. *There exists a real $f \in L^2(\mathbb{T})$ and a bijection $\sigma : \mathbb{Z} \to \mathbb{Z}$ such that*

$$\limsup_{N\to\infty} \left| \sum_{u=-N}^{N} \hat{f}(\sigma(u)) \exp i\sigma(u)t \right| = \infty.$$

for almost all $t \in \mathbb{T}$.

This theorem was first stated by Kolmogorov. A proof of Kolmogorov's statement was sketched by Zahorskiĭ, given in detail by Ulyanov and much simplified by Olevskiĭ. The reader who consults [7] will find an excellent bibliography.

Pólya says that sometimes the easiest way to prove a result is to generalise it and prove the generalisation. By the time Kolmogorov's theorem reached Ulyanov and Olevskiĭ it had taken the following form.

THEOREM 2. *Let $\phi_1, \phi_2, \phi_3, \dots$ form a complete orthonormal system in $L^2([0,1))$. Then there exists a real $f \in L^2(\mathbb{T})$ and a bijection $\sigma : \mathbb{N} \to \mathbb{N}$ such that*

$$\limsup_{N\to\infty} \left| \sum_{u=0}^{N} \hat{f}(\sigma(u)) \phi_{\sigma(u)}(t) \right| = \infty.$$

for almost all $t \in [0,1)$.

Here as usual we write

$$\hat{f}(\phi_r) = \langle f, \phi_r \rangle = \int_0^1 f(t)\phi_r(t)\, dt.$$

In fact Olevskiĭ improved Theorem 2 by replacing $f \in L^2(\mathbb{T})$ by f continuous. We shall prove a result which includes this as Theorem 5.

In the years since Olevskiĭ published his result, more general systems than those of complete orthonormal type have assumed practical importance.

DEFINITION 3. *We say that*

$$\phi_1, \phi_2, \phi_3, \dots$$

form a Riesz basis for $L^2([0,1))$ *if the linear span of the* ϕ_n *is dense in* L^2 *and there exists an* A *with* $A \geqslant 1$ *such that*

$$A^{-1}\sum_{n=0}^{\infty}|a_n|^2 \leqslant \left\|\sum_{n=0}^{\infty}a_n\phi_n\right\|_2^2 \leqslant A\sum_{n=0}^{\infty}|a_n|^2.$$

We call A the Riesz constant of the system. Easy functional analysis reveals the following lemma.

LEMMA 4. *Let* ϕ_1, ϕ_2, ϕ_3, … *form a Riesz basis for* $L^2([0,1))$. *Then there exists a unique sequence* ψ_1, ψ_2, ψ_3, … *of bounded* L^2 *norm such that* $\sum_{n=0}^{\infty}|\langle\psi_n, f\rangle|^2 < \infty$ *and*

$$f = \sum_{n=0}^{\infty}\langle\psi_n, f\rangle\phi_n$$

for every $f \in L^2$. *If* $f = \sum_{n=0}^{\infty}a_n\phi_n$ *then* $a_n = \langle\psi_n, f\rangle$ *for all* n.

We write $\hat{f}(n) = \langle\psi_n, f\rangle$.

As Olevskiĭ indicates very clearly his method can be extended to Riesz bases.

THEOREM 5. *Let* ϕ_1, ϕ_2, ϕ_3, … *form a Riesz basis for* $L^2([0,1])$. *Then there exists a real continuous* f *and a bijection* $\sigma : \mathbb{N} \to \mathbb{N}$ *such that*

$$\limsup_{N\to\infty}\left|\sum_{u=0}^{N}\hat{f}(\sigma(u))\phi_{\sigma(u)}(t)\right| = \infty.$$

for almost all $t \in [0,1]$.

If we ask for which complete orthonormal system Theorem 2 is least plausible one answer would be the Haar system. Set

$$\Lambda(N) = \{(r,s) \in \mathbb{Z}^2 : 0 \leqslant s \leqslant 2^r - 1 \text{ and } N \geqslant r \geqslant 0\}$$

and $\Lambda = \bigcup_{N\geqslant 0}\Lambda(N)$. Consider intervals of the form

$$E(r,s) = [s2^{-r}, (s+1)2^{-r})$$

where $(r,s) \in \Lambda$. We define

$$\begin{aligned}
\chi_{r,s}(t) &= 1 && \text{if } t \in E(r+1, 2s)\\
\chi_{r,s}(t) &= -1 && \text{if } t \in E(r+1, 2s+1)\\
\chi_{r,s}(t) &= 0 && \text{otherwise}
\end{aligned}$$

where, again, $(r,s) \in \Lambda$. We call the $\chi_{r,s}$ together with the function $1 = \chi_{-1,0}$ the Haar system. We call the $T_{r,s} = 2^{r/2}\chi_{r,s}$ together with the function 1 the normalised Haar system.

It is well known that the normalised Haar system is a complete orthonormal system. If we put the standard order

$$(r,s) \ll (r',s') \text{ if } r' > r \text{ or } r = r' \text{ and } s' \geqslant s$$

on the Haar system it enjoys remarkable convergence properties. For example if f is continuous

$$\sum_{(u,v)\ll(r,s)} \hat{f}(u,v)T_{u,v} \to f$$

uniformly as we allow (r,s) to increase so as to exhaust the system.

If we can prove Theorem 2 for the Haar system, one is tempted to say, we can surely prove it for any orthonormal system. Olevskiĭ showed that this is indeed the case and we shall see that once we have Theorem 2 for the Haar system we can obtain the rest of the theorem from this particular case. As we might expect on general grounds the key to the Haar system case lies in a finite version of the theorem.

LEMMA 6. *Let $\epsilon > 0$ and $K \geqslant 1$ be given. Then we can find a bijection*

$$\sigma : \{1, 2, 3, \ldots, 2^{N+1} - 1\} \to \Lambda(N)$$

and real $a_{r,s}$ $[(r,s) \in \Lambda(N)]$ such that

$$\left| \sum_{(r,s)\in\Lambda(N)} a_{r,s}\chi_{r,s}(t) \right| \leqslant 1$$

for all $t \in [0,1)$ but

$$\max_{1\leqslant k\leqslant 2^{N+1}-1} \left| \sum_{j=1}^{k} a_{\sigma(j)}\chi_{\sigma(j)}(t) \right| \geqslant K$$

for all $t \notin E$ where E is a set of measure at most ϵ.

The next section is devoted to the proof of this key lemma. Once it is understood the rest is relatively routine.

2. Olevskiĭ's Lemma

We shall need two simple results, the first combinatorial and the second probabilistic.

LEMMA 7. *If $a_1, a_2, \ldots, a_m \geqslant 0$ and $a_{m+1}, a_{m+2}, \ldots, a_N \leqslant 0$ then*

$$\max_{1\leqslant j\leqslant N} \left| \sum_{r=1}^{j} a_r \right| \geqslant \frac{1}{3} \sum_{r=1}^{N} |a_r|.$$

PROOF. Observe that

$$\sum_{r=1}^{N}|a_r| = 2\sum_{r=1}^{m}a_r - \sum_{r=1}^{N}a_r \leqslant 3\max_{1\leqslant j\leqslant N}\left|\sum_{r=1}^{j}a_r\right|.$$

□

LEMMA 8. *Let $X_1, X_2, \dots, X_N$ be independent identically distributed random variables with $\Pr(X_r = 1) = \Pr(X_r = -1) = 1/2$ and let K be an integer with $K \geqslant 1$. Then*

(i) $\Pr\left(\max_{1\leqslant j\leqslant N}\sum_{r=1}^{j}X_r \geqslant K\right) \leqslant 2\Pr\left(\sum_{r=1}^{N}X_r \geqslant K\right)$.

(ii) $\Pr\left(\max_{1\leqslant j\leqslant N}\left|\sum_{r=1}^{j}X_r\right| \geqslant K\right) \leqslant 4\Pr\left(\left|\sum_{r=1}^{N}X_r\right| \geqslant K\right)$.

(iii) $\Pr\left(\max_{1\leqslant j\leqslant N}\left|\sum_{r=1}^{j}X_r\right| \geqslant K\right) \leqslant 2NK^{-2}$.

PROOF. (i) This is the simplest form of the reflection principle. (See e.g. [2] Chapter 3.)

(ii) Use symmetry.

(iii) Use Chebychev's inequality to bound $\Pr\left(\left|\sum_{r=1}^{N}X_r\right| \geqslant K\right)$.

□

Of course, we can get much better estimates in (iii) but the only use I can find for such estimates is to study the Hausdorff dimension of the exceptional set of convergence in Theorem 2 when applied to Haar functions.

By using the standard interpretation of Haar functions in terms of coin tossing, Lemma 8 gives the following result.

LEMMA 9. *If N and K are strictly positive integers then we can find $a_{p,q}$ taking the values 0 or 1 $[(p,q) \in \Lambda(N)]$ and a set E of measure at most $4NK^{-2}$ such that*

$$\left|\sum_{(p,q)\in\Lambda(N)} a_{p,q}\chi_{p,q}(t)\right| \leqslant K$$

for all $t \in [0,1)$, but

$$\sum_{(p,q)\in\Lambda(N)} |a_{p,q}\chi_{p,q}(t)| = N$$

for all $t \notin E$.

PROOF. Consider $[0,1)$ with Lebesgue measure as a probability space. Let

$$X_j(t) = (-1)^{[2^j t]}$$

(where $[2^j t]$ is the integer part of $2^j t$). The X_j satisfy the conditions of Lemma 8. It follows by Lemma 8 (iii) that the set

$$E = \left\{ t : \max_{1 \leqslant j \leqslant N} \left| \sum_{r=1}^{j} X_r \right| \geqslant K \right\}$$

has measure at most $4NK^{-2}$.

To define $a_{p,q}$ we look at the interval $E(p,q) = [p2^{-q}, (p+1)2^{-q})$ and observe that $\sum_{r=1}^{j} X_r(t)$ is constant on $E(p,q)$ for all $1 \leqslant j \leqslant q-1$. We may thus define $a_{p,q} = 1$ if

$$\max_{1 \leqslant j \leqslant N} \left| \sum_{r=1}^{j} X_r(t) \right| \leqslant K - 1$$

for $t \in E(p,q)$, and $a_{p,q} = 0$ otherwise.

If we now set

$$Y(t) = \sum_{(p,q) \in \Lambda(N)} a_{p,q} \chi_{p,q}(t)$$

then Y is the random variable defined by

$$Y(t) = \sum_{r=1}^{N} X_r(t)$$

if $|\max_{1 \leqslant j \leqslant N} \sum_{r=1}^{j} X_r(t)| \leqslant K$ but

$$Y(t) = \sum_{r=1}^{j(t)} X_r(t)$$

where $j(t)$ is the smallest j with $|\sum_{r=1}^{j(t)} X_r(t)| = K$. In more vivid terms, toss a fair coin keeping track of the difference between the number of heads and tails thrown. If this ever takes the value K or $-K$ stop and record the value as Y. If, after N throws this has not happened, record the value after N throws as Y. By definition $|Y(t)| \leqslant K$ and if $t \notin E$ (so we complete the N throws)

$$\sum_{(p,q) \in \Lambda(N)} |a_{p,q} \chi_{p,q}(t)| = N$$

as required. □

Now, instead of taking the sequence of heads and tails as chance presents them, we wish to take all the heads first and then all the tails. To do this we introduce the Olevskiĭ order on Λ

$$(p', q') \succeq (p, q) \text{ if } (2p'+1)2^{-q'} \geqslant (2p+1)2^{-q}.$$

The Olevskiĭ order is more intricate than may at first appear and the reader should try ordering the (p,q) with $(p,q)\in\Lambda(N)$ for $N=5$ or $N=6$. Observe that $(p',q')\succeq\chi(p,q)$ if the mid-point of $\operatorname{supp}\chi_{p',q'}$ is to the right of (i.e. greater than or equal to) the mid-point of $\operatorname{supp}\chi_{p,q}$. We observe that $\chi_{p,q}(t)\geqslant 0$ if t is strictly to the left of (i.e. strictly less than) the mid-point of $\operatorname{supp}\chi_{p,q}$ and $\chi_{p,q}(t)\leqslant 0$ if t is to the right of (i.e. greater than or equal to) the mid-point of $\operatorname{supp}\chi_{p,q}$. It follows that for each $t\in[0,1)$ there is a $\lambda(t)\in\Lambda(N)$ such that

$$\chi_{p,q}(t)\leqslant 0 \text{ for} \lambda\succeq(p,q),\ \chi_{p,q}(t)\geqslant 0 \text{ for} (p,q)\succeq\chi_\lambda.$$

Using this observation we easily arrive at the result we require.

LEMMA 10. *Suppose that N and K are strictly positive integers and $a_{p,q}$ and E are as in Lemma 9. Then*

$$\max_{\lambda\in\Lambda(N)}\left|\sum_{\lambda\succeq(p,q)}a_{p,q}\chi_{p,q}(t)\right|\geqslant N/3$$

for all $t\notin E$.

PROOF. By Lemma 7 and the last sentence of the previous paragraph

$$\max_{\lambda\in\Lambda(N)}\left|\sum\lambda\succeq(p,q)a_{p,q}\chi_{p,q}(t)\right|\geqslant\frac{1}{3}\sum_{(p,q)\in\Lambda(N)}|a_{p,q}\chi_{p,q}(t)|=N/3$$

for all $t\notin E$. □

THEOREM 11. *Let $\epsilon>0$ and $\kappa>1$ be given. Then we can find an $N\geqslant 1$, $1\geqslant b>0$, $b_{p,q}$ taking the values 0 or b $[(p,q)\in\Lambda(N)]$ and a set E such that*

(i) $|\sum_{(p,q)\in\Lambda(N)}b_{p,q}\chi_{p,q}(t)|\leqslant 1$ *for all* $t\in[0,1)$,

(ii) $\max_{\lambda\in\Lambda(N)}|\sum_{\lambda\succeq(p,q)}a_{p,q}\chi_{p,q}(t)|\geqslant\kappa$ *for all* $t\notin E$

(iii) $|E|<\epsilon$.

PROOF. Choose an integer M with $M>4\epsilon^{-1}$ and $M^2\geqslant 3\kappa$. Set $N=M^5$, $K=M^3$ and choose $a_{p,q}$ and E as in Lemma 9. If we now put $b_{p,q}=a_{p,q}/K$, $b=1/K$ then all the conclusions of the lemma with the exception of condition (ii) follow at once from Lemma 9. Condition (ii) itself follows from Lemma 10. □

Standard 'rolling hump' (or 'condensation of singularities') methods now give the following result.

EXERCISE 12. *Let χ_n $[n\in\mathbb{N}]$ be the Haar system enumerated in some way. Then there exists a real $f\in L^\infty([0,1))$ and a bijection $\sigma:\mathbb{N}\to\mathbb{N}$ such that*

$$\limsup_{N\to\infty}\left|\sum_{u=0}^{N}\hat f(\sigma(u))\chi_{\sigma(u)}(t)\right|=\infty.$$

for almost all $t\in[0,1)$.

We leave it as an exercise for the reader because we shall prove stronger results. However these results will require extra complications in their proofs and the reader may find the more complex proofs easier to follow if he or she has worked through an easier case.

3. Extension To General Spaces

We have not yet exhausted the strength of Pólya's dictum. What happens if we seek to extend Theorem 2 to more general measure spaces? A moment's reflection brings to mind the the fact that, so far as measure theory is concerned, all nice measure spaces which are not obviously different are the same. Recall that a measure space $(X, \mathcal{F}, \mu)$ with positive measure μ is *non-atomic* if given $E \in \mathcal{F}$ with $\mu(E) > 0$ we can find $F \in \mathcal{F}$ with $F \subseteq E$ and $\mu(E) > \mu(F) > 0$. The following result is typical (see e.g. theorem 9, page 327 of [8]).

THEOREM 13. *Let X be a complete separable metric space equipped with $\mathcal{B}_X$ its σ-algebra of Borel sets and μ a non-atomic measure on $\mathcal{B}_X$. If $I = [0,1]$ is the unit interval with its usual metric, $\mathcal{B}_I$ its σ-algebra of Borel sets, and m is the usual Lebesgue measure then there exists a bijective map $F : I \to X$ such that F and F^{-1} carry Borel sets to Borel sets of the same measure.*

We shall not use Theorem 13 but we shall use the lemma which underlies its proof and the proof of results like it.

LEMMA 14. *Let $(X, \mathcal{F}, \mu)$ be a non-atomic probability space. If $E \in \mathcal{F}$ and $1 \geqslant \alpha \geqslant 0$ then we can find $F \in \mathcal{F}$ with $E \supseteq F$ and $\mu(F) = \alpha\mu(E)$.*

Lemma 17 follows easily from a lemma of Saks given as Lemma 7 of section IV.9.8 of Dunford and Schwartz [1]. There is a discussion of isomorphism theorems in Chapter VIII of Halmos's *Measure Theory* [3]. However before readers rush off to inspect the wilder shores of measure theory they should note that all they will learn is that they could have stayed at home, since $[0,1]$ with Lebesgue measure is the type of all non-atomic measure spaces.

In view of the preceding discussion, it is natural to aim at the following generalisation of Theorem 2. (The extension of Definition 3 to general probability spaces is obvious.)

THEOREM 15. *Let $(X, \mathcal{F}, \mu)$ be a probability space with μ non-atomic. Let $\phi_1, \phi_2, \dots$ be a Riesz basis in $L^2(X)$. Then there exists a real $f \in L^2(X)$ and a bijection $\sigma : \mathbb{N} \to \mathbb{N}$ such that*

$$\limsup_{N\to\infty} \left| \sum_{u=0}^{N} \hat{f}(\sigma(u))\phi_{\sigma(u)}(t) \right| = \infty.$$

for almost all $t \in X$.

The result is clearly false if μ is not non-atomic. Suppose $E \in \mathcal{F}$ is an atom, that is $\mu(E) > 0$ and if $F \in \mathcal{F}$ with $F \subseteq E$ then $\mu(F) = \mu(E)$ or $\mu(F) = 0$. Then if $g_n, g \in L^2(X)$ and $||g_n - g||_2 \to 0$ it follows that $g_n(t) \to g(t)$ for μ-almost all $t \in E$.

It is not hard to obtain Theorem 15 from Theorem 2 by the standard argument used to prove results like Theorem 13. However, it is more interesting to ask how we might tackle Theorem 15 directly and, in particular, what we are to make of Theorem 11 in the general context. If we do so we are rewarded with a key insight — Theorem 11 is a combinatorial theorem.

THEOREM 16. *Let $(X, \mathcal{F}, \mu)$ be a probability space. Let $\epsilon > 0$ and $\kappa > 1$ be given. Then we can find an $N \geqslant 1$, $1 \geqslant b > 0$ and $b_{p,q}$ taking the values 0 or b $[(p,q) \in \Lambda(N)]$ with the following property.*

Suppose that we have a collection of sets $E_{n,r} \in \mathcal{F}$ such that

(A) $E(2r-1, n+1) \cap E(2r, n+1) = \emptyset$*,* $E(2r-1, n+1) \cup E(2r, n+1) = E(r, n)$ *for all* $(r, n) \in \Lambda_N$*.*

(B) $|E(r,n)| = 2^{-n}$ *for all* $(r,n) \in \Lambda_{N+1}$*.*

Let us set

$$\begin{aligned} H_{r,n}(t) &= 1 && \text{for } t \in E(2r-1, n+1), \\ H_{r,n}(t) &= -1 && \text{for } t \in E(2r, n+1), \\ H_{r,n}(t) &= 0 && \text{otherwise.} \end{aligned}$$

whenever $(r,n) \in \Lambda_N$*. Then*

(i) $|\sum_{(p,q)\in\Lambda(N)} b_{p,q} H_{p,q}(t)| \leqslant 1$ *for all* $t \in [0,1)$*,*

and there is a set $E \in \mathcal{F}$ *such that*

(ii) $\max_{\lambda\in\Lambda(N)} |\sum_{\lambda\succeq(p,q)} b_{p,q} H_{p,q}(t)| \geqslant \kappa$ *for all* $t \notin E$

(iii) $|E| < \epsilon$*.*

PROOF. This is just Theorem 11. □

To see why this is indeed an insight let us return to the concrete case of Lebesgue measure on $[0,1)$ and consider the most important Riesz basis of all, the exponentials $e_n(t) = \exp(2\pi i n t)$. If we try to convert Theorem 11 into a result on Fourier series we run into the problem that different Haar functions do not 'occupy different parts of the frequency spectrum'. (Indeed the Fourier coefficients of the Haar functions $\chi_{p,q}$ with fixed p are all of same amplitude differing only in phase.) However, if we 'shuffle $[0,1)$' the $E_{p,q}$ can be chosen so that the $H_{p,q}$ are, for practical purposes, in 'different parts of the frequency spectrum'. That is, although we do not have the ideal outcome in which at most one of the $\hat{H}_{p,q}(j)$ is non-zero for each j, we can arrange that at most one of the $\hat{H}_{p,q}(j)$ is large for each j.

Since there seems no further advantage in considering general measure spaces we shall stay within $[0,1)$ and $[0,1]$. However, readers who wish to work more generally will find that they only need the following simple consequence of Lemma 14.

LEMMA 17. *Let $(X, \mathcal{F}, \mu)$ be a non-atomic probability space. If $F \in \mathcal{F}$ and $\mu(F) > 0$ we can find a sequence e_j of orthogonal functions and sets $F_j \in \mathcal{F}$ with $F_j \subseteq F$ such that*

(i) $e_j(t) = 1$ *for* $t \in F_j$, $e_j(t) = -1$ *for* $t \in F \setminus F_j$, $e_j(t) = 0$ *otherwise,*
(ii) $\mu(F_j) = \mu(F)/2$.

PROOF. Set $E(0,0) = F$. By repeated use of Lemma 14 with $\alpha = 1/2$ we can find $E_{n,r} \in \mathcal{F}$ such that

(A) $E(2r-1, n+1) \cap E(2r, n+1) = \emptyset$, $E(2r-1, n+1) \cup E(2r, n+1) = E(r,n)$ for all $1 \leqslant r \leqslant 2^n$,
(B) $|E(r,n)| = 2^{-n}|F|$ for all $1 \leqslant r \leqslant 2^n$.

Now set $F_j = \bigcup_{r=1}^{2^j} E(2r-1, j+1)$ and define e_j as in condition (i). □

4. Extension to general Riesz bases

In this section we work in $[0,1]$ with Lebesgue measure. In order to put into effect the programme sketched at the end of the last section we need a sequence of easy lemmas. It may be helpful for the reader to keep in mind as examples both the 'well behaved' orthonormal system of exponentials and some other system ϕ_n where $\phi_n \notin L^\infty$.

LEMMA 18. *Let* $\phi_1, \phi_2, \ldots$ *be a Riesz basis. If* $e_1, e_2, \ldots$ *form an orthonormal sequence then*

$$\hat{e}_k(j) \to 0$$

as $k \to \infty$ *for each fixed* j.

PROOF. Referring back to Lemma 4 we see that $\hat{e}_k(j) = \langle \psi_j, e_k \rangle$ and the result is obvious. □

LEMMA 19. *Let* $\phi_1, \phi_2, \ldots$ *be a Riesz basis. Suppose that* $\delta > 0$, $M \geqslant 0$ *and that* F *is a set of strictly positive measure. Then we can find* E *a measurable subset of* F *with* $|E| = |F|/2$ *and* $M' > M$ *such that if we set* $H(t) = 1$ *for* $t \in E$, $H(t) = -1$ *for* $t \in F \setminus E$ *and* $H(t) = 0$ *otherwise then*

$$\sum_{j=1}^{M} |\hat{H}(j)|^2 + \sum_{j=M'}^{\infty} |\hat{H}(j)|^2 < \delta.$$

PROOF. Combining the results of Lemma 17 and Lemma 18 we see that we can find E a measurable subset of F with $|E| = |F|/2$ such that, if H is defined as stated,

$$\sum_{j=1}^{M} |\hat{H}(j)|^2 < \delta/2.$$

Since $H \in L^2$ there exists an $M' > M$ such that $\sum_{j=M'}^{\infty} |\hat{H}(j)|^2 < \delta/2$, so we are done. □

We now have our basic construction.

THEOREM 20. *Let ϕ_1, ϕ_2, … be a Riesz basis. Suppose that $\delta > 0$ and $N, M'(0,0) \geqslant 1$ are given. Then we can find integers $M(r,n) < M'(r,n)$ with $M(r,n) > M'(0,0)$—together with a collection of measurable sets $E(n,r)$ such that*

(A) $E(2r-1,n+1) \cap E(2r,n+1) = \emptyset$, $E(2r-1,n+1) \cup E(2r,n+1) = E(r,n)$ *for all* $(r,n) \in \Lambda_N$,
(B) $|E(r,n)| = 2^{-n}$ *for all* $(r,n) \in \Lambda_{N+1}$,
(C) *if we write* $\Delta(r,n) = \{k : M(r,n) \leqslant k \leqslant M'(r,n)\}$ *then* $\Delta(r,n) \cap \Delta(s,m) = \emptyset$ *whenever* $(r,n) \neq (s,m)$,

and such that, if we set

$$\begin{aligned} H_{r,n}(t) &= 1 && \text{for } t \in E(2r-1,n+1), \\ H_{r,n}(t) &= -1 && \text{for } t \in E(2r,n+1), \\ H_{r,n}(t) &= 0 && \text{otherwise.} \end{aligned}$$

whenever $(r,n) \in \Lambda_N$ *then*

$$\sum_{j \notin \Delta(r,n)} |\hat{H}_{r,n}(j)|^2 < \delta.$$

PROOF. Use Lemma 19 repeatedly. □

THEOREM 21. *Let $\epsilon > 0$ and $\kappa > 1$ be given and take $N \geqslant 1$, b and $b_{p,q}$ $[(p,q) \in \Lambda(N)]$ be as in Theorem 16. Suppose that ϕ_1, ϕ_2, … is a Riesz basis, $\delta > 0$ and $M'(0,0) \geqslant 1$ are given and that $E(n,r)$, $\Delta(r,n)$ and $H_{r,n}$ are constructed as in Theorem 20. Then if we set*

$$f(t) = \sum_{(p,q) \in \Lambda(N)} b_{p,q} H_{p,q}(t)$$

we have

(i) $||f||_\infty \leqslant 1$.

Further, provided only that δ is small enough, there is a measurable set F such that

(ii) $\max_{\lambda \in \Lambda(N)} |\sum_{\lambda \succeq (p,q)} \sum_{j \in \Delta(p,q)} |\hat{f}(j)\phi_j(t)| \geqslant \kappa - 1$ *for all* $t \notin F$,
(iii) $|F| < 2\epsilon$.

PROOF. Conclusion (i) is just conclusion (i) of Theorem 20. To obtain (ii) and (iii) we proceed as follows. Observe that

$$\begin{aligned}
&\left\| \sum_{\lambda\succeq(p,q)} \sum_{j\in\Delta(p,q)} \hat{f}(j)\phi_j - \sum_{\lambda\succeq(p,q)} b_{p,q}H_{p,q} \right\|_2 \\
&\qquad \leqslant \sum_{\lambda\succeq(p,q)} |b_{p,q}| \sum_{j\in\Delta(p,q)} \left\| H_{p,q} - \sum_{j\in\Delta(p,q)} \hat{f}(j)\phi_j \right\|_2 \\
&\qquad \leqslant \sum_{\lambda\succeq(p,q)} \sum_{j\in\Delta(p,q)} \left\| H_{p,q} - \sum_{j\in\Delta(p,q)} \hat{f}(j)\phi_j \right\|_2 .
\end{aligned}$$

But, for each $(p,q) \in \Lambda(N)$,

$$\begin{aligned}
&\left\| H_{p,q} - \sum_{j\in\Delta(p,q)} \hat{f}(j)\phi_j \right\|_2 \\
&\quad \leqslant \left\| \sum_{j\notin\Delta(p,q)} \hat{H}_{p,q}(j)\phi_j \right\|_2 + \sum_{(r,s)\neq(p,q)} \left\| \sum_{j\in\Delta(p,q)} \hat{H}_{r,s}(j)\phi_j \right\|_2 \\
&\quad \leqslant \left\| \sum_{j\notin\Delta(p,q)} \hat{H}_{p,q}(j)\phi_j \right\|_2 + \sum_{(r,s)\neq(p,q)} \left\| \sum_{j\notin\Delta(r,s)} \hat{H}_{r,s}(j)\phi_j \right\|_2 \\
&\quad = \sum_{(r,s)\in\Lambda(N)} \left\| \sum_{j\notin\Delta(r,s)} \hat{H}_{r,s}(j)\phi_j \right\|_2 .
\end{aligned}$$

But, writing A for the Riesz constant of the basis, Theorem 20 tells us that

$$\left\| \sum_{j\notin\Delta(r,s)} \hat{H}_{r,s}(j)\phi_j \right\|_2^2 \leqslant A \sum_{j\notin\Delta(r,s)} |\hat{H}_{r,s}(j)|^2 < A\delta.$$

Thus, retracing our steps,

$$\left\| H_{p,q} - \sum_{j\in\Delta(p,q)} \hat{f}(j)\phi_j \right\|_2 < 2^N (A\delta)^{1/2}$$

and

$$\left\| \sum_{\lambda \succeq (p,q)} \sum_{j \in \Delta(p,q)} \hat{f}(j)\phi_j - \sum_{\lambda \succeq (p,q)} b_{p,q} H_{p,q} \right\|_2 < 2^{2N}(A\delta)^{1/2}$$

for all $\lambda \in \Lambda(N)$.

If we now write

$$F_\lambda = \left\{ t : \left| \sum_{\lambda \succeq (p,q)} \sum_{j \in \Delta(p,q)} \hat{f}(j)\phi_j - \sum_{\lambda \succeq (p,q)} b_{p,q} H_{p,q} \right| \geqslant 1 \right\},$$

then Chebychev's inequality and the last inequality of the preceding paragraph tell us that

$$|F_\lambda| < 2^{4N} A\delta.$$

Thus if we set $F = E \cup \bigcup_{\lambda \in \Lambda(N)} F_\lambda$ conclusion (ii) follows from conclusion (ii) Theorem 20 whilst conclusion (iii) Theorem 20 tells us that

$$|F| < \epsilon + 2^{5N} A\delta,$$

and (iii) holds provided only that δ is small enough. □

Theorems 20 and 21 give us what we want. However, we have accumulated a fair amount of notation in the course of the construction which we can now jettison to provide a simpler conclusion.

THEOREM 22. *Let $\phi_1, \phi_2, \ldots$ be a Riesz basis. Given $K \geqslant 1$ we can find an integer $M(K) \geqslant 1$ with the following properties. Given any integer $m \geqslant 1$ and any $\eta > 0$ we can find a function $f \in L^\infty([0,1])$, an integer $m' > m$, a bijection $\sigma : \mathbb{N} \to \mathbb{N}$ with $\sigma(r) = r$ for $1 \leqslant r \leqslant m$ and for $r \geqslant m'$, integers $m \leqslant p(1) \leqslant p(2) \leqslant p(3) \ldots p(M) \leqslant m'$ and a measurable set E such that*

(i) $||f||_\infty \leqslant 1$,
(ii) $\sum_{j=1}^m |\hat{f}(j)|^2 \leqslant \eta$,
(iii) $\max_{1 \leqslant k \leqslant M} |\sum_{j=1}^{p(k)} \hat{f}(\sigma(j))\phi_{\sigma(j)}| \geqslant K$ *for all* $t \in E$,
(iv) $|E| > 1 - K^{-1}$.

The distance from $L^\infty([0,1])$ to $C([0,1])$ is usually not very great. The present case is no exception.

THEOREM 23. *In Theorem 22 we may take f continuous.*

PROOF. This is entirely routine. Let the $M(K)$ of our new Theorem 23 be chosen to be the $M(2K+2)$ of the old Theorem 22. Then, by Theorem 22, given any $m \geqslant 1$ and any $\eta > 0$ we can find a function $f \in L^\infty([0,1))$, an integer $m' > m$, a bijection $\sigma : \mathbb{N} \to \mathbb{N}$ with $\sigma(r) = r$ for $1 \leqslant r \leqslant m$ and for $r \geqslant m'$, integers $m \leqslant p(1) \leqslant p(2) \leqslant p(3) \ldots p(M) \leqslant m'$ and a measurable set E' such that

(i) $'\ ||f||_\infty \leqslant 1$,
(ii) $'\ \sum_{j=1}^{m} |\hat{f}(j)|^2 \leqslant \eta$,
(iii) $'\ \max_{1\leqslant k\leqslant M} |\sum_{j=1}^{p(k)} \hat{f}(\sigma(j))\phi_{\sigma(j)}(t)| \geqslant K+1$ for all $t \in E$,
(iv) $'\ |E'| > 1 - K^{-1}/2$.

Let $\delta > 0$ be a small number to be determined and choose $f \in C([0,1))$ such that f satisfies condition (i) and $||f-g||_2 \leqslant \delta/A^2$. Automatically

$$\left(\sum_{j=1}^{\infty} |\hat{f}(j) - \hat{g}(j)|^2\right)^{1/2} \leqslant \delta/A$$

so condition (ii) is satisfied provided only that we choose δ small enough.

Next we observe that

$$\begin{aligned}\left\|\sum_{j=1}^{p(k)} \hat{f}(j)\phi_j - \sum_{j=1}^{p(k)} \hat{g}(j)\phi_j\right\|_2 &\leqslant A\left(\sum_{j=1}^{p(k)} |\hat{f}(j) - \hat{g}(j)|^2\right)^{1/2} \\ &\leqslant A\left(\sum_{j=1}^{\infty} |\hat{f}(j) - \hat{g}(j)|^2\right)^{1/2} \leqslant \delta.\end{aligned}$$

Thus writing

$$E_k = \left\{t \in [0,1) : \left\|\sum_{j=1}^{p(k)} \hat{f}(j)\phi_j - \sum_{j=1}^{p(k)} \hat{g}(j)\phi_j\right\| \geqslant 1\right\},$$

we have, by Chebychev's inequality,

$$|E_k| \leqslant \left\|\sum_{j=1}^{p(k)} \hat{f}(j)\phi_j - \sum_{j=1}^{p(k)} \hat{g}(j)\phi_j\right\|_2 = \delta.$$

Thus, setting $E = E' \setminus \bigcup_{k=1}^{N}$, we see that (iii) holds automatically and

$$|E| \geqslant |E'| - \sum_{k=1}^{N} |E_k| \geqslant (1-\eta/2) - N\delta$$

so that (iv) holds provided only that we choose δ small enough. □

The remainder of the construction follows a standard rolling hump (condensation of singularities) pattern using Chebychev's inequality in the same way as in the two previous proofs. In view of the amount of notation involved readers will probably prefer to do it themselves rather than follow my proof.

It is easy to put the bricks of Theorem 23 together.

LEMMA 24. *Let $\phi_1, \phi_2, \ldots$ be a Riesz basis with constant A. Let $m'(0) = 1$ We can construct inductively positive integers $M(n)$, $m(n)$, $m'(n)$ with $m'(n-1) < m(n)$, continuous functions f_n, bijections $\sigma_n : \mathbb{N} \to \mathbb{N}$ with $\sigma_n(r) = r$ for $1 \leqslant r \leqslant m'(n-1)$ and for $m'(n) \leqslant r$, integers $m \leqslant p(1,n) \leqslant p(2,n) \leqslant p(3,n) \ldots p(n, M(n)) \leqslant m'$ and a measurable set E such that*

(i) $_n$ $||f_n||_\infty \leqslant 2^{-n}$,
(ii) $_n$ $\sum_{r=1}^{m(n)} |\hat{f}_n(r)|^2 \leqslant A^{-1} 2^{-4n-4} m(n)^2$,
(iii) $_n$ $\max_{1 \leqslant k \leqslant M(n)} |\sum_{j=1}^{p(k,n)} \hat{f}_n(\sigma(j)) \phi_{\sigma_n(j)}(t)| \geqslant 2^n$ *for all* $t \in E$,
(iv) $_n$ $|E_n| > 1 - 2^{-n}$,
(v) $_n$ $\sum_{j=1}^{n} \sum_{r=m'(n)}^{\infty} |\hat{f}_j(r)|^2 \leqslant A^{-1} 2^{-2n-4} M(n+1)^{-2}$.

PROOF. Conditions (i)$_n$ to (iv)$_n$ come directly from Theorem 23. The key point is that Theorem 23 allows us to define $M(n+1)$ before we define $m'(n)$ in such a way as to satisfy condition (v)$_n$. □

PROOF OF THEOREM 5. As I said above, this is routine. It is easy to check that $\sigma(r) = \sigma_n(r)$ for $m'(n-1) \leqslant r \leqslant m'(n)$ gives a well defined bijection $\sigma : \mathbb{N} \to \mathbb{N}$. By the conditions (i)$_n$ $g_n = \sum_{r=1}^{n} f_r$ converges uniformly to a continuous function f as $n \to \infty$. Trivially

$$\begin{aligned} \left| \sum_{j=1}^{p(k,n)} \hat{f}(\sigma(j)) \phi_{\sigma(j)}(t) - \sum_{j=1}^{p(k,n)} \hat{f}_n(\sigma(j)) \phi_{\sigma(j)}(t) \right| &\leqslant ||g_{n-1}||_\infty + |G_{k,n}(t)| \\ &\leqslant 1 + |G_{k,n}(t)|, \end{aligned}$$

where

$$G_{k,n}(t) = \sum_{r=1}^{n-1} \sum_{j=p(k,n)+1}^{\infty} \hat{f}_r(\sigma(j)) \phi_{\sigma(j)}(t) + \sum_{r=n+1}^{\infty} \sum_{j=1}^{p(k,n)} \hat{f}_r(\sigma(j)) \phi_{\sigma(j)}(t).$$

Using the properties of Riesz bases together with conditions (ii)$_{(r)}$ and (v)$_{(r)}$ we have

$$
\begin{aligned}
&\|G_{k,n}\|_2 \\
&\leqslant \sum_{r=1}^{n-1}\left\|\sum_{j=p(k,n)+1}^{\infty}\hat{f}_r(\sigma(j))\phi_{\sigma(j)}\right\|_2 + \sum_{r=n+1}^{\infty}\left\|\sum_{j=1}^{p(k,n)}\hat{f}_r(\sigma(j))\phi_{\sigma(j)}\right\|_2 \\
&\leqslant A^{1/2}\sum_{r=1}^{n-1}\left(\sum_{j=p(k,n)+1}^{\infty}|\hat{f}_r(\sigma(j))|^2\right)^{1/2} + A^{1/2}\sum_{r=n+1}^{\infty}\left(\sum_{j=1}^{p(k,n)}|\hat{f}_r(\sigma(j))|^2\right)^{1/2} \\
&\leqslant A^{1/2}\sum_{r=1}^{n-1}\left(\sum_{j=m'(n-1)}^{\infty}|\hat{f}_r(\sigma(j))|^2\right)^{1/2} + A^{1/2}\sum_{r=n+1}^{\infty}\left(\sum_{j=1}^{m(n)}|\hat{f}_r(\sigma(j))|^2\right)^{1/2} \\
&\leqslant A^{1/2}\sum_{r=1}^{n-1}\left(\sum_{j=m'(n-1)}^{\infty}|\hat{f}_r(\sigma(j))|^2\right)^{1/2} + A^{1/2}\sum_{r=n+1}^{\infty}\left(\sum_{j=1}^{m(r-1)}|\hat{f}_r(\sigma(j))|^2\right)^{1/2} \\
&\leqslant 2^{-n-1}M(n)^{-1} + \sum_{r=n+1}^{\infty}2^{-2r-4}m(r)^{-1} \leqslant 2^{-n}M(n)^{-1}.
\end{aligned}
$$

Thus writing

$$E(k,n) = \{t : |G_{k,n}(t)| \geqslant 1\},$$

we have, by Chebychev's theorem, $|E(k,n)| \leqslant 2^{-n}M(n)^{-1}$ and, if we set

$$E'(n) = E(n)\bigcup_{k=1}^{M(n)}E(k,n),$$

we have, by $(iv)_n$,

$$|E'(n)| \leqslant 2^{-n+1}$$

and by (iii)$_n$

$$\max_{1\leqslant k\leqslant M(n)}\left|\sum_{j=1}^{p(k,n)}\hat{f}(\sigma(j))\phi_{\sigma(j)}(t)\right| \geqslant 2^n - 2$$

for all $t \in E'(n)$. The theorem follows. □

The reader may readily check that Theorem 15 may be obtained by very similar arguments as can the following direct generalisation of Theorem 5

THEOREM 25. *Let (X,τ) be a compact Hausdorff space. Let $(X,\mathcal{F},\mu)$ be a regular non-atomic probability space. Let ϕ_1, ϕ_2, … form a Riesz basis for $L^2(\mu)$. Then there exists* *a real continuous f and a bijection $\sigma : \mathbb{N} \to \mathbb{N}$ such that*

$$\limsup_{N\to\infty} \left| \sum_{u=0}^{N} \hat{f}(\sigma(u))\phi_{\sigma(u)}(t) \right| = \infty.$$

for almost all $t \in X$.

5. Further Questions

It is natural to ask if our results can be improved so as to replace 'divergence almost everywhere' by 'divergence everywhere'. The obvious answer is no, since if we take a complete orthonormal system ϕ_n on $[0, 1)$ and define $\tilde{\phi}_n$ by $\tilde{\phi}_n(t) = \phi_n(t)$ for $t \neq 0$, $\tilde{\phi}_n(0) = 0$ the result remains a complete orthonormal system but any linear combination of a finite set of $\tilde{\phi}_n$ will take the value 0 at zero.

The "unfair" example just given means that we must reformulate our question and suggests that, at least initially, we should consider particular complete orthonormal systems. In the case of Fourier series we recall the remarkable theorem of Kahane and Katznelson that every set of measure zero is a set on which the Fourier sum of a continuous function diverges (see e.g. [4], Chapter II, Section 3) and consider the following sequence of lemmas.

LEMMA 26. *Given any $\epsilon > 0$, any $K > 1$ and any integer $N \geqslant 1$ we can find a trigonometric polynomial P, a set E which is the union of a finite set of intervals and a bijection $\sigma : \mathbb{N} \to \mathbb{Z}$ such that*

(i) $||P||_\infty \leqslant 1$,
(ii) $\hat{P}(n) = 0$ *for all* $n \leqslant N$,
(iii) $\max_{k\geqslant 0} |\sum_{j\leqslant k} \hat{P}(\sigma(j)) \exp(i\sigma(j)t)| \geqslant K$ *for all* $t \notin E$.
(iv) $|E| < \epsilon$.

PROOF. Apply de la Vallée Poussin summation to Theorem 23 (with the exponentials $\exp(2\pi int)$ as the Riesz basis) to obtain a P satisfying all the conditions except possibly (ii). To obtain (ii) it suffices to replace $P(t)$ by $\exp(iMt)P(t)$ with M a suitable large positive integer. □

LEMMA 27 (Kahane and Katznelson). *Given any $K > 1$ there exists an $\epsilon(K)$ with the following property. Given any set E which is the union of a finite set of intervals and has $|E| < \epsilon(K)$ and any integer $N \geqslant 1$ we can find a trigonometric polynomial P such that*

(i) $||P||_\infty \leqslant 1$,
(ii) $\hat{P}(n) = 0$ *for all* $n \leqslant N$,
(iii) $\max_{k\geqslant 0} |\sum_{j\leqslant k} \hat{P}(j) \exp(ijt)| \geqslant K$ *for all* $t \notin E$.

PROOF. See [4], Chapter II, Section 3. □

Combining the last two lemmas we obtain the following result.

LEMMA 28. *Given any* $\epsilon > 0$, *any* $K > 1$ *and any integer* $N \geqslant 1$ *we can find a trigonometric polynomial* P *and a bijection* $\sigma : \mathbb{N} \to \mathbb{Z}$ *such that*

(i) $||P||_\infty \leqslant 1$,

(ii) $\hat{P}(n) = 0$ *for all* $n \leqslant N$,

(iii) $\max_{k\geqslant 0} |\sum_{j\leqslant k} \hat{P}(\sigma(j)) \exp(2\pi i\sigma(j)t)| \geqslant K$ *for all* $t \in \mathbb{T}$.

It is now easy to prove the desired result.

THEOREM 29. *We can find a continuous function* $f : \mathbb{T} \to \mathbb{C}$ *with* $\hat{f}(n) = 0$ *for* $n < 0$ *and a bijection* $\sigma : \mathbb{N} \to \mathbb{N}$ *such that*

$$\sup_{k\geqslant 0} |\sum_{j\leqslant k} \hat{f}(\sigma(j)) \exp(2\pi ijt)| = \infty$$

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

If instead of considering complex valued functions and the orthonormal system $\exp 2\pi int$, we wish to consider real valued functions and the orthonormal system formed by $\sin 2\pi nt$ and $\cos 2\pi nt$ then we can replace the P of Lemma 28 by $\Re P + \Im P$ and first sum the $\cos 2\pi\sigma(j)t$ terms and then the $\sin 2\pi\sigma(j)t$ terms. The existence of a continuous function whose rearranged Fourier series diverges everywhere was first proved by L. V. Taĭkov [10] using Olevskiĭ's theorem in a different way.

So far as I know the general question remains open. In particular we may ask the following question.

QUESTION 30. *Consider the Haar system on* $[0, 1)$ *ordered in some way. Does there exist a continuous function* $f : [0, 1] \to \mathbb{C}$ *and a bijection* $\sigma : \mathbb{N} \to \mathbb{N}$ *such that*

$$\sup_{k\geqslant 0} \left| \sum_{j=0}^{k} a_{\sigma(j)} \chi_{\sigma(j)}(t) \right| = \infty$$

for all $t \in [0, 1)$*?*

We remark that a similar proof to the one above (replacing the non-trivial lemma of Kahane and Katznelson by a trivial parallel lemma) gives the following.

LEMMA 31. *Consider the Haar system on* $[0, 1)$ *ordered in some way. There exists a real* $f \in L^2$ *and a bijection*

$$\sigma : \mathbb{N} \to \mathbb{N}$$

such that

$$\sup_{k\geqslant 1} \left| \sum_{j=1}^{k} \hat{f}(\chi_{\sigma(j)}) \chi_{\sigma(j)}(t) \right| = \infty$$

for all $t \in [0, 1)$.

The method of proof applies to any 'well behaved' system (though the 'unfair example' with which we began the section shows that it can not apply to all Riesz bases).

6. Hard summation

The reader with more practical interests will observe that striking as Kolmogorov's theorem and its generalisations may be, they do not answer the question in the case of any particular rearrangement. Let us return to the question we started with but apply it to a general othonormal system $\phi_1, \phi_2, \dots$. If $f \in L^2$ we consider

$$S_\delta(t) = \sum_{|\hat{f}(n)| \geqslant \delta} \hat{f}(n)\phi_n(t)$$

where $\hat{f}(n) = \langle f, \phi_n \rangle$ is the usual Fourier coefficient for the given system. This method of forming a sum is called 'hard summation' and, as we said at the beginning of the talk, is a very natural one to use.

(The reader may well ask what 'soft summation' is. Here we consider something like

$$\sigma_\delta \sum_{|\hat{f}(n)| \geqslant \delta} \hat{f}(n)\phi_n(t) + \sum_{\delta \geqslant |\hat{f}(n)| \geqslant \delta/2} \frac{|\hat{f}(n)| - \delta/2}{\delta/2} \hat{f}(n)\phi_n(t).$$

I suspect that the idea of soft summation derives from the use of Cesàro sums and similar filtering techniques to improve the behaviour of Fourier sums. I also suspect that the analogy is false since I know of no reason why soft summation should behave better than hard summation. However, this is merely opinion, and like all opinion liable to be disproved by fact.)

In [5] I constructed the following example.

THEOREM 32. *There exists a $f \in L^2(\mathbb{T})$ such that*

$$\limsup_{\eta \to 0+} \left| \sum_{|\hat{f}(u)| \geqslant \eta} \hat{f}(u) \exp 2\pi i u t \right| = \infty.$$

for almost all $t \in \mathbb{T}$.

Later in [6] I showed that we could take f continuous.

It therefore came as a great surprise to me to learn that Tao had proved the following theorem

THEOREM 33. *hard summation works for well behaved (rapidly decreasing) wavelets.*

Still more surprising was the simple nature of his proof. In order to show how it works I will prove it in a special case where the argument is particularly clean.

THEOREM 34. *Consider the Haar system on* $\mathbb{T}$*. If* $f : \mathbb{T} \to \mathbb{C}$ *is continuous then*

$$\sum_{|\hat{f}(\chi)| \geqslant \eta} \hat{f}(\chi)\chi \to f$$

uniformly on $\mathbb{T}$*.*

[The exceptionally good behaviour of the Haar system means that just as in some sense it is the *hardest* to prove a Kolmogorov type theorem for, so it is in some sense the *easiest* to prove a Tao type theorem for. However, I think all the essential ideas can be seen even in this simple special case.]

The reader should notice that we now use a different normalisation for the Haar functions since we want an *orthonormal* system in which $\langle \chi, \chi \rangle = 1$. We shall return to this point when we consider Lemma 37.

Tao's idea is to use one of the key characters in 20th century harmonic analysis — the maximal function. If we write

$$P_N f(t) = \sum_{\mathrm{rank}(\chi) \leqslant N} \hat{f}(\chi)\chi(t)$$

$$S_\eta f(t) = \sum_{|\hat{f}(\chi)| \geqslant \eta} \hat{f}(\chi)\chi(t)$$

then corresponding maximal functions are

$$P^* f(t) = \sup_{N \geqslant 0} |P_N(f)(t)|$$

$$S^* f(t) = \sup_{\eta > 0} |S_\eta(f)(t)|.$$

Since the Haar system is so well behaved P^* is easy to bound.

LEMMA 35. $|P^* f(t)| \leqslant \|f\|_\infty$.

PROOF. Observe that $P_N f(t)$ is piecewise constant with $P_N f(t)$ taking the average value of f on the interval of the form $[q2^{-N}, (q+1)2^{-N})$ to which t belongs. Thus $|P_N f(t)| \leqslant \|f\|_\infty$ for all N and so $|P^* f(t)| \leqslant \|f\|_\infty$. □

In general maximal functions are not so easy to bound and the main work of the proof of Tao's result lies in obtaining a bound for S^*.

THEOREM 36. *There exists a constant* K *such that if* $f : \mathbb{T} \to \mathbb{C}$ *is continuous then*

$$|S^* f(t)| \leqslant K \|f\|_\infty.$$

PROOF OF THEOREM 34 FROM THEOREM 36. Let $\epsilon > 0$. Then, since $P_N f(t) \to f$ uniformly we can find an M such that $\|P_M f - f\|_\infty < \epsilon$. Set $g = f - P_M(f)$. We observe that $\hat{g}(\chi) = 0$ if $\mathrm{rank}\,\chi \leqslant N$ and $\widehat{P_M(f)}(\chi) = 0$ if $\mathrm{rank}\,\chi \geqslant N+1$.

Let $\delta_0 = 1$ if $P_M(f) = 0$ and set

$$\delta_0 = \min\{|P_M\hat{}(f)(\chi)| \,;\, P_M\hat{}(f)(\chi) \neq 0\}$$

otherwise. If $\delta_0 > \delta > 0$ we have

$$S_\delta(f) = P_M(f) + S_\delta(g)$$

and so

$$\|S_\delta(f) - f\|_\infty \leqslant \|P_M(f) - f\|_\infty + \|S_\delta(g)\|_\infty \leqslant \epsilon + S^*(g) \leqslant (K+1)\epsilon$$

using Theorem 36. Since ϵ was arbrary the result follows. □

In the proof of Theorem 36 we look out for ways in which (well behaved) wavelets differ from their Fourier counterparts. The first is that although the Riemann Lebesgue lemma represents the strongest statement we can make about the decrease of Fourier coefficients, much stronger results hold for wavelets.

LEMMA 37. *If $f : [0,1] \to \mathbb{C}$ is continuous and χ has rank n then*

$$|\hat{f}(\chi)| \leqslant \|f\|_\infty 2^{-n/2}$$

PROOF. Direct calculation gives

$$\begin{aligned} |\hat{f}(\chi)| &= \left|\int_{\operatorname{supp}\chi} f(t)\chi(t)\,dt\right| \leqslant \|f\|_\infty \|\chi\|_\infty |\operatorname{supp}\chi| \\ &= \|f\|_\infty 2^{n/2} 2^{-n} = \|f\|_\infty 2^{-n/2}. \end{aligned}$$

Note how the normalisation of χ comes into the calculation. □

The second and best known difference is the localisation property of wavelets. This will play an important role in our estimate of the maximal function $S^* f$.

PROOF OF THEOREM 36. . It is suficicient to show that if $\|f\|_\infty = 1$, and $1 \geqslant \delta > 0$ then $|S_\delta(f)(t)| \leqslant K$ for some fixed K. To this end, choose n so that $2^{-n/2} \geqslant \delta > 2^{-(n+1)}/2$ and observe that by Lemma 37 this means that $|\hat{f}(\chi)| < \delta$ for $rank \chi > n$.

Thus

$$\begin{aligned} |S_\delta(f)(t)| &= \left| P_n(f)(t) - \sum_{\operatorname{rank}\chi \leqslant n,\ |\hat{f}(\chi)|<\delta} \hat{f}(\chi)\chi(t) \right| \\ &\leqslant |P_n(f)(t)| + \Big| \sum_{\operatorname{rank}\chi \leqslant n,\ |\hat{f}(\chi)|<\delta} \hat{f}(\chi)\chi(t) \Big| \\ &\leqslant P^*(f)(t) + \delta \sum_{\operatorname{rank}\chi \leqslant n} |\chi(t)| \\ &\leqslant \|f\|_\infty + \delta \sum_{\operatorname{rank}\chi \leqslant n} |\chi(t)|, \chi(t) \neq 0 \\ &\leqslant 1 + \delta \sum_0^n 2^{r/2} \end{aligned}$$

since exactly one χ of each rank is non-zero at t (this is an extreme version of the localisation property of wavelets).

Doing some simple calculations we have

$$|S_\delta(f)(t)| \leqslant 1 + \delta \sum_0^n 2^{r/2} \leqslant 1 + \frac{\delta 2^{(n+1)/2}}{2^{1/2}-1} \leqslant 1 + \frac{2^{1/2}}{2^{1/2}-1},$$

which is an inequality of the required form (with $K = 1 + 2^{1/2}/(2^{1/2}-1)$). □

To extend beyond continuous functions we need to use the Hardy-Littlewood maximal operator

$$Mf(t) = \sup_{r>0} \frac{1}{2r} \int_{t-r}^{t+r} |f(x)|\, dx.$$

If we do so, we get pointwise convergence almost everywhere for $f \in L^p$ $[1 \leqslant p \leqslant \infty]$ for the Haar system.

We end by stating Tao's general theorem from [9]. He works on $\mathbb{R}$ which is a more useful space than $\mathbb{T}$.

THEOREM 38 (Tao). *We work on $\mathbb{R}$. Suppose ϕ has integral zero, and is bounded and rapidly decreasing. Suppose that the set of functions (wavelets) given by*

$$\phi_{j,k}(x) = 2^{j/2}\phi(2^j(x-k))$$

form an orthonormal system. Then, if $f \in L^p(\mathbb{R})$ $[1 < p < \infty]$, we have

$$\sum_{|\hat{f}(\phi_{j,k})|>\lambda} \hat{f}(\phi_{j,k})\phi_{j,k}(x) \to f(x)$$

almost everywhere as $\lambda \to 0+$.

References

[1] N. Dunford and J. T. Schwartz, *Linear Operators (Part I)* (2nd Ed), Wiley, 1957.

[2] W. Feller, *An Introduction to Probability Theory and its Applications*, (3rd Ed) Wiley, 1968.

[3] P. R. Halmos *Measure Theory*, Van Nostrand, 1950.

[4] Y. Katznelson *An Introduction to Harmonic Analysis* Wiley, 1968.

[5] T. W. Körner, *Divergence of decreasing rearranged Fourier series* Ann. of Math. **144**, no. 1, 167–180 (1996).

[6] T. W. Körner *Decreasing rearranged Fourier series* J. Fourier Anal. Appl. **5** , no. 1, 1–19 (1999).

[7] A. M. Olevskiĭ, *Fourier Series with Respect to General Orthogonal Systems* Springer 1975.

[8] H. L. Royden, *Real Analysis* (2nd Ed), MacMillan, New York, 1968.

[9] T. Tao *On the almost everywhere convergence of wavelet summation methods* Appl. Comput. Harmon. Anal. **3**, no. 4, 384–387 (1996).

[10] L. V. Taĭkov, *The divergence of Fourier series of continuous functions with respect to a rearranged trigonometric system*, Dokl. Akad. Nauk SSSR, **150**, 262–265 (1963).