THE BANG SOLUTION OF THE COEFFICIENT PROBLEM – 1
Possible introduction (written in common English).

They call themselves Analysts, but what they do
is just manipulating inequalities !...
(ascribed to Arnold)

Let T be a measure space with probability measure μ. Let $\psi_j : T \to \mathbb{R}$ be an orthonormal system of functions in $L^2(T)$, i.e. $||\psi_j||_2 = 1$ for every j and $(\psi_i, \psi_j) \stackrel{\text{def}}{=} \int_T \psi_i \psi_j d\mu = 0$ for every $i \neq j$. Suppose that X is some other space of functions on T such that $||f||_2 \leq ||f||_X$ for every $f \in X$ (the main example to keep in mind is $X = L^\infty(T)$). The question we are going to discuss is whether we can say anything nontrivial about the decay of the Fourier coefficients $(f, \psi_j) = \int_T f \psi_j d\mu$ of functions $f \in X$.

Of course, we always have the Bessel inequality:

$$\sum_j (f, \psi_j)^2 \leq ||f||_2^2 \leq ||f||_X^2.$$

Definition

We will call the space X large (with respect to the system $\{\psi_j\}$) if for every sequence $\{a_j\}$ of positive numbers satisfying $\sum_j a_j^2 = 1$ there exists a function $F \in X$ such that $|(F, \psi_j)| \geq a_j$ for every j.

Roughly speaking, this definition means that the Bessel inequality is the only thing about the decay we may say for sure.

Note that if X is large, we can find a solution F of the coefficient problem $|(F, \psi_j)| \geq a_j$ with uniformly bounded norm (i.e. such that $||F||_X \leq C$ where $C > 0$ does not depend on a_j). Indeed, otherwise for every $k = 1, 2, \ldots$ one can choose a sequence $\{a_j^{(k)}\}$ with $\sum_j (a_j^{(k)})^2 = 1$ and such that there is no function $F \in X$ with $||F||_X \leq 4^k$ satisfying $(F, \psi_j)| \geq a_j^{(k)}$ for all j. But then for the sequence $a_j \stackrel{\text{def}}{=} \sum_{k=0}^\infty 2^{-k} a_j^{(k)}$ there is no $F \in X$ solving the coefficient problem $|(F, \psi_j)| \geq a_j$ (because $a_j \geq 2^{-k} a_j^{(k)}$ for every k and thereby we should have $||F||_X \geq 2^k$ for every k which is impossible), but still $\sum_j a_j^2 \leq 1$.

Which spaces are large and which are not ? There exists the following

Conjecture:

X is large iff there is a constant $c > 0$ such that for every sequence $\{b_j\}$ of positive numbers satisfying $\sum_j b_j^2 = 1$ one can find $f \in X$ such that

$$||f||_X = 1 \quad \textit{and} \quad \sum_j b_j |(f, \psi_j)| \geq c.$$

The conjecture means that the set of sequences $\{ \{|(f, \psi_j)|\} : f \in X, ||f||_X \leq 1\}$ (which is star-like with respect to the origin, but by no means convex) can be treated up to a certain extent as a convex set and one may switch from the problem to its dual and back.

In most cases when X is known to be large (including those we will add below to the existing list) it is very easy to see that the element f solving this dual problem exists. An interesting exception is the space $X = U(T)$ of the functions defined on the unit circumference $\mathbb{T}$ with uniform convergent Fourier series with respect to the standard base $\{z^k\}_{k=-\infty}^{\infty}$ (the norm in $U(\mathbb{T})$ is defined by $||f||_{U(\mathbb{T})} \stackrel{\text{def}}{=} \sup_{n,m\in\mathbb{Z}} ||S_{n,m}f(z)||_\infty$ where $S_{n,m}f(z) \stackrel{\text{def}}{=} \sum_{k=n}^{m} \hat{f}(k)z^k$). It is known that $U(T)$ is large with respect to any orthonormal system $\{\psi_j\}$ satisfying $\sup_j ||\psi_j||_\infty < +\infty$, say, but to check solvability of the dual problem for this space is only a very little bit easier then to solve the coefficient problem itself.

Despite the conjecture above is quite old, there has been found no evidence of that the fact so general should hold. So, a lot of people now (me among them) believe that it is false. Meanwhile de-Leeuw, Kahane, Katznelson and others invented quite powerful technique allowing to check that X is large for many particular spaces X. Their idea was to consider the random sum $f_\varepsilon \stackrel{\text{def}}{=} \sum_j \varepsilon_j a_j \psi_j$ (where $\varepsilon_j = 1$ or $\varepsilon_j = -1$ with probability $\frac{1}{2}$ each) and to show that with positive probability this sum "almost belongs" to X. The exact meaning of this phrase is that for every $A > 0$ it is possible to decompose f_ε into the sum $f_\varepsilon = g_A + h_A$ such that $||g_A||_X \le A$, $||h_A||_2 \le \Delta(A)$ and $\Delta(A)$ decays fast enough as $A \to +\infty$. Then they proceeded by "minor corrections" as follows:

Suppose that we have a function $F \in X$ of norm $||F||_X \le M$ which solves the coefficient problem for all j but those in some exceptional set J and that $\sum_{j\in J} a_j^2 \le b^2$ (for instance, we may start with $F \equiv 0$, $M = 0$, J which is the whole index set and $b = 1$). Let us choose the signs ε_j in such a way that $f_\varepsilon = \sum_{j\in J} \varepsilon_j a_j \psi_j$ almost belongs to X. Note that the function $\widetilde{F} \stackrel{\text{def}}{=} F + 2f_\varepsilon$ solves the coefficient problem $|(\widetilde{F}, \psi_j)| \ge a_j$ for every j but, unfortunately, does not belong to X in general. Nevertheless, it lies not very far from X: as for every $A > 0$ we can decompose f_ε into the sum $g_A + h_A$ where $||g_A||_X \le Ab$ and $||h_A||_2 \le \Delta(A)b$ (the factor b is due to the condition $\sum_{j\in J} a_j^2 \le b^2$, not 1, as in the definition of $\Delta(A)$ above), the functuon $\widetilde{F}$ can be decomposed into the sum of $F_A \stackrel{\text{def}}{=} F + g_A$ and h_A. When we take F_A instead of $\widetilde{F}$, some of the Fourier coefficients get spoiled but since the L^2-norm of the difference h_A is small, we cannot substantially spoil too many of them. Namely, let $c \in (0, 1)$ and let $J' = \{j : |(F_A, \psi_j)| < (1 - c)a_j\}$. Then for every $j \in J'$ we have $|(h_A, \psi_j)| \ge ca_j$ and therefore

$$\sum_{j\in J'} a_j^2 \le c^{-2}||h_A||_2^2 \le \frac{\Delta(A)^2 b^2}{c^2} \stackrel{\text{def}}{=} {b'}^2.$$

For $j \notin J'$ the coefficients (F_A, ψ_j) are easy to repair: just take $F' \stackrel{\text{def}}{=} \frac{1}{1-c} F_A$. Thus, having started with the function $F \in X$ of norm $||F||_X \le M$ and with the exceptional set J for which $\sum_{j\in J} a_j^2 \le b^2$, we got a new function F' of norm $||F'||_X \le M' \stackrel{\text{def}}{=} \frac{M+Ab}{1-c}$ and a new exceptional set J' for which $\sum_{j\in J'} a_j^2 \le {b'}^2$ where $b' = \frac{\Delta(A)b}{c}$. If we contrive to get $b' < b$ without essential increase in the norm estimate M, we will be able to iterate the procedure. To have the process convergent, we need to construct three sequences $\{A_k\}$,

$\{b_k\}$ and $\{c_k\}$ ($k = 0, 1, \dots$), for which $b_0 = 1$; $b_{k+1} = \frac{\Delta(A_k)b_k}{c_k}$, $c_k \in (0, 1)$ for every k; $\{b_k\}$ decreases to 0 as $k \to \infty$ and, at last, both series $\sum_{k=0}^{\infty} c_k$ and $\sum_{k=0}^{\infty} A_k b_k$ converge (the last condition comes just from the necessity of having an uniform bound for the norms of the corresponding functions $F_k \in X$, but what it actually implies is $\sum_k ||F_k - F_{k+1}||_X < +\infty$ which is more than enough for the existence of the limit $F = \lim_{k\to\infty} F_k \in X$).

Let's now figure out when such sequences exist. As $c_k = \frac{\Delta(A_k)b_k}{b_{k+1}}$, the problem reduces to the choice of $\{b_k\}$ and $\{A_k\}$ satisfying the conditions $\sum_{k=0}^{\infty} \frac{\Delta(A_k)b_k}{b_{k+1}} < +\infty$ and $\sum_{k=0}^{\infty} A_k b_k < +\infty$. (If we take just arbitrary sequences satisfying these conditions, several first c_k may turn out to be greater than 1, but this can be easily corrected by enlarging finite number of corresponding A_k's). Uniting these two conditions in one

$$\sum_{k=0}^{\infty} b_k \left(\frac{\Delta(A_k)}{b_{k+1}} + A_k \right) < +\infty,$$

we see that we also have almost no freedom in choosing $\{A_k\}$ if $\{b_k\}$ is given: A_k should minimize the expression $\frac{\Delta(A)}{b_{k+1}} + A$. Since $\Delta(A)$ is a decreasing function and since we do not care about constant factors like 2, the best choice comes from solving $\frac{\Delta(A)}{b_{k+1}} = A$. Denoting the solution of the equation $\Delta(A) = At$ by $A(t)$ (for any continuous decreasing function Δ it exists for all $t > 0$ and is a decreasing continuous function of t), we find out that all we need is to construct a sequence $\{b_k\}$ starting with $b_0 = 1$ and decreasing to 0 such that $\sum_k b_k A(b_{k+1}) < +\infty$.

The trivial estimate

$$\sum_k b_k A(b_{k+1}) > \sum_k (b_k - b_{k+1}) A(b_{k+1}) \geq \int_0^1 A(t)dt$$

shows that we should assume at least that $\int_0 A(t)dt < +\infty$. On the other hand, whatever Δ is, $A(t)$ turns out to be quite regular function. Indeed, since $\frac{t}{2}(2A(t)) = tA(t) = \Delta(A(t)) \geq \Delta(A(\frac{t}{2}))$, we have $A(\frac{t}{2}) \leq 2A(t)$ for every $t > 0$.

Thus, taking $b_k = 2^{-k}$, we get

$$\sum_{k=0}^{\infty} b_k A(b_{k+1}) = \sum_{k=0}^{\infty} 2^{-k} A(2^{-k-1}) \leq 4 \int_0^1 A(t)dt.$$

So, the condition $\int_0 A(t)dt < +\infty$ is also sufficient.

What does it mean in terms of the original function Δ ? Changing variable from t to A, we get $\int^{\infty} A d\left(\frac{\Delta(A)}{A}\right) < +\infty$ and it remains only to integrate by parts to obtain the celebratred

<u>De-Leeuw – Kahane – Katznelson theorem:</u>

If $\displaystyle\int^{\infty} \frac{\Delta(A)}{A} dA < +\infty$, *then* X *is large in* $L^2(T)$ *with respect to the system* ψ_j.

For instance, when $X = L^\infty(T)$ and the norms $||\psi_j||_\infty$ are uniformly bounded, we may use the Khinchin theorem to get that $f_\varepsilon \in L^4$, say, with estimate for the norm $||f_\varepsilon||_4$ not depending on the sequence $\{a_j\}$. This gives $\Delta(A) \le CA^{-1}$ which is more than enough to apply the theorem and to conclude that $L^\infty(T)$ is large.

Though quite powerful in general, the de-Leeuw – Kahane – Katznelson approach has two unpleasant restrictions: to proceed their way one should assume at the very least that

1) *X is dense in $L^2(T)$;*
2) *All ψ_j almost belong to X themselves.*

Correspondently, there remained two problems which could not be solved by the de-Leeuw – Kahane – Katznelson approach:

1. The "support" problem:

Let T' be a subset of T of positive measure. Is the space $X = L^p(T') \stackrel{\text{def}}{=} \{f \in L^p(T) : \operatorname{supp} f \subset T'\}$ $(2 \le p \le \infty)$ large or not ?

and

2. The "minimal assumptions" problem:

Let again $X = L^p(T)$. What are the minimal requirements on the system $\{\psi_j\}$ which garantee that X is large with respect to this system ?

The first question was open even for $p = 2$, $T = \mathbb{T}$, (as usual, we denote by $\mathbb{T}$ the unit circumference with the Haar measure on it) and the standard base $\psi_j = z^j$ $(j \in \mathbb{Z})$! Of course, the density of X in $L^2(T)$ can be restored if we completely forget about $T \setminus T'$ and consider T' as the whole space (with the measure $d\mu' \stackrel{\text{def}}{=} \frac{d\mu}{\mu(T')}$). Then we should also forget about the the original system ψ_j and consider the functions $\psi'_j \stackrel{\text{def}}{=} \sqrt{\mu(T')}\psi_j|_E$. (the renormalization which preserves L^2-norm) instead. Unfortunately the system $\{\psi'_j\}$ is no longer orthogonal. In general it can be just an arbitrary system of functions satisfying $||\sum_j c_j\psi'_j||_{L^2(T',\mu')} \le \left(\sum_j c_j^2\right)^{\frac{1}{2}}$. And it is also fatal: first, now we cannot say that the coefficients (f_ε, ψ'_j) of the sum $f_\varepsilon = \sum_j \varepsilon_j a_j \psi_j$ are just $\varepsilon_j a_j$ and second, we no longer have a possibility to correct one coefficient without substantial influencing others (just because ψ'_j may be even linearly dependent now) and it seems that the entire idea of "minor corrections" fails.

As to the second problem, the de-Leeuw – Kahane – Katznelson technique can be applied only if we assume a priori something like $\psi_j \in L^2 \log^{2+\delta} L$, but there is no evident reason for a condition of this kind to be necessary (and actually it is not).

What are we going to do ?

Consider the sets $S_j \stackrel{\text{def}}{=} \{f \in L^2(T) : |(f,\psi_j)| < a_j\}$. Obviously, X is large iff the unit ball $B_X \stackrel{\text{def}}{=} \{f \in X : ||f||_X \le 1\}$ of the space X cannot be covered by sets S_j with arbitrarily small $\sum_j a_j^2$. Since in general B_X is just a convex set, symmetric with respect to the origin and since from the geometric point of view S_j is a strip of width $2a_j$, the whole thing resembles a lot the famous "plank problem" posed by Tarski in the very beginning of the century and solved by Bang in 1950s:

Plank problem:

Given a convex set $B \in \mathbb{R}^2$ (or $\mathbb{R}^n$, for Bang's solution it is all the same), *is it possible to cover it by several strips of total width less than the width of* B *?* (The width of a convex body is defined as the width of the narrowest strip containing it).

The conjectured answer was "no", but it took about 40 years to prove that. The proof, when found, was ... 2 pages long and required from the reader only basic knowlege of elementary geometry. For reader's convenience it is included into this paper as an appendix.

Of course, to solve the coefficient problem as it was stated above you cannot just apply the result (because the width of B_X is 0 unless $X = L^2(T)$ and because we need to estimate from below the sum $\sum a_j^2$, not $\sum a_j$), but it turns out that a minor modification of the proof is enough. (So minor that I actually even do not pretend to be an author of the next two sections; rather I act there like a shadow that enters and goes over many strange places which completely eliminate the attention of his master just passing by).

We will deal with both the "support" and the "minimal assumptions" problems simultaneously, so, from the very beginning we will assume that the inequality

$$\left\|\sum_j c_j\psi_j\right\|_2 \leq \left(\sum_j c_j^2\right)^{\frac{1}{2}}$$

is the only a priori information about the system ψ_j we have.

THE BANG SOLUTION OF THE COEFFICIENT PROBLEM - 2
Between L^2 and L^∞ (written in Analysis language).

Let T be a measure space with probability measure μ. Let, as before, $\{\psi_j\}$ be a system of functions satisfying $||\sum_j c_j\psi_j||_2 \leq \left(\sum c_j^2\right)^{\frac{1}{2}}$ for every $c_j \in \mathbb{R}$.

The question we are going to solve below is when $X = L^p(T)$ $(2 \leq p \leq \infty)$ is large with respect to the system $\{\psi_j\}$. Also we will give a good estimate for the norm of the function F, solving the coefficient problem.

Trivial necessary condition.

In general it is $||\psi_j||_{X^*} \geq \beta > 0$. For the case $X = L^p(T)$ this means

$$\left(\int_T |\psi_j|^q\right)^{\frac{1}{q}} \geq \beta > 0 \qquad \text{for every } j \qquad (B_p)$$

where q is the exponent conjugate to p, i.e. $\frac{1}{p} + \frac{1}{q} = 1$.

Our aim is to show that (B_p) is also sufficient. More precisely, we have the following

Theorem:

If the system $\{\psi_j\}$ satisfies (B_p), then for every sequence of positive numbers a_j satisfying $\sum_j a_j^2 = 1$ we can find a function $F \in L^p(T)$, such that

$$||F||_p \leq \left(\frac{3\pi}{2}\right)^{1-\frac{2}{p}} \beta^{-2} \leq 5\beta^{-2} \quad \text{and} \quad \left|\int_T F\psi_j d\mu\right| \geq a_j \text{ for every } j.$$

Remark:

1) Without additional assumptions on ψ_j the estimate for the norm is almost the best possible one for small β. Indeed, consider $T = \{1, 2\}$ with the measure μ defined by $\mu\{1\} = \mu\{2\} = \frac{1}{2}$. Then $L^p(T)$ is merely $\mathbb{R}^2$ with the norm $||(x_1, x_2)||_p = \left(\frac{|x_1|^p+|x_2|^p}{2}\right)^{\frac{1}{p}}$. Let n be a large positive integer. Put $\psi_j \stackrel{\text{def}}{=} \sqrt{\frac{2}{n}}(\cos\frac{\pi j}{n}, \sin\frac{\pi j}{n})$ $(j = 1, \ldots, n)$. Note that $||\psi_j||_2 = \sqrt{\frac{1}{n}}$ and therefore

$$\left\|\sum_j c_j\psi_j\right\|_2 \le \sqrt{\frac{1}{n}}\sum_j |c_j| \le \left(\sum c_j^2\right)^{\frac{1}{2}}.$$

On the other hand, for every $1 \le q \le 2$

$$||\psi_j||_q \ge \sqrt{\frac{1}{2n}} \stackrel{\text{def}}{=} \beta.$$

Let now $a_j = \sqrt{\frac{1}{n}}$ $(j = 1, \ldots, n)$. Notice that for every $j = 1, \ldots, n-1$ we have $||\psi_j - \psi_{j+1}||_2 \le \sqrt{\frac{1}{n}}\frac{\pi}{n}$ and thereby for every $p \ge 2$ we get $\left|\int_T F\psi_j d\mu - \int_T F\psi_{j+1} d\mu\right| \le \sqrt{\frac{1}{n}}\frac{\pi}{n}||F||_p$ If the signs of $\int_T F\psi_j d\mu$ and $\int_T F\psi_{j+1} d\mu$ are different for some j, then, due to the previous inequality, at least one of them doesn't exceed $\sqrt{\frac{1}{n}}\frac{\pi}{2n}||F||_p$ and therefore

$$||F||_p \ge \frac{2n}{\pi} = \frac{1}{\pi}\beta^{-2}.$$

If all the signs are the same, we get the same result from $||\psi_1 + \psi_n||_2 \le \sqrt{\frac{1}{n}}\frac{\pi}{n}$.

2) Another curious observation is that all the conditions (B_p) are equivalent if we do not care about the exact value of β. So, if any of the spaces $L^p(T)$ is large, then all of them are !

Proof of the theorem:

Consider all sign corteges $\varepsilon = \{\varepsilon_j\}$ where $\varepsilon_j = +1$ or -1. For each of them put $f_\varepsilon \stackrel{\text{def}}{=} \sum_j \varepsilon_j a_j \psi_j$. Regardless of the choice of the signs we have $f_\varepsilon \in L^2(T)$ and $||f_\varepsilon||_2 \le 1$.

Let now $\Phi : \mathbb{R} \to \mathbb{R}_+$ be any C^2-smooth even function satisfying $\Phi(0) = \Phi'(0) = 0$ and $0 < \Phi''(x) \le 1$ for every $x \in \mathbb{R}$.

It is easy to check that the integral $I(f) \stackrel{\text{def}}{=} \int_T \Phi(f)d\mu$ is well-defined and continuous in $L^2(T)$. As the family $\{f_\varepsilon\}$ is compact in the topology of $L^2(T)$, one can find a cortege $\bar\varepsilon$ for which $I(f_\varepsilon)$ attains its maximal value. Let $f = f_{\bar\varepsilon}$. Fix now some j and consider the function f_j which is obtained from $f_{\bar\varepsilon}$ by replacing $\bar\varepsilon_j a_j\psi_j$ by $-\bar\varepsilon_j a_j\psi_j$ in the corresponding sum. So $f_j = f - 2\bar\varepsilon_j a_j\psi_j$.

We have

$$0 \le \int_T \Phi(f)d\mu - \int_T \Phi(f_j)d\mu = \int_T \Phi'(f)(f-f_j)d\mu + \frac{1}{2}\int_T \Phi''(g)(f-f_j)^2 d\mu.$$

where $g(t)$ lies between $f(t)$ and $f_j(t)$ for every $t \in T$. Recalling what $f - f_j$ is, we get from here

$$\bar{\varepsilon}_j \int_T \Phi'(f)\psi_j d\mu \ge a_j \int_T \Phi''(g)\psi_j^2 d\mu.$$

If we contrive to choose $\Phi(x)$ in such a way that $\Phi'(f) \in L^p(T)$ for every $f \in L^2(T)$ and that the integral on the right can be estimated from below by some constant depending on β only, it will remain only to put $F \stackrel{\text{def}}{=} A\Phi'(f)$ with constant A large enough.

Case $p = 2$.

Here the choice is easy: just put $\Phi(x) = \frac{x^2}{2}$. Then $\Phi'(f) = f$ and $\Phi''(g) \equiv 1$, so the integral reduces to $\int_T \psi_j^2 \ge \beta^2$. Thus, the function $F \stackrel{\text{def}}{=} \beta^{-2} f$ is what we are looking for.

General case.

Let $\Phi(x)$ be the solution of the equation $\Phi''(x) = (1+x^2)^{\frac{2}{p}-1}$ for which $\Phi(0) = \Phi'(0) = 0$. As $p \ge 2$, we really have $\Phi''(x) \le 1$ for every $x \in \mathbb{R}$.

Note that

$$|\Phi'(x)| \le \int_0^{|x|} (1+s^2)^{\frac{2}{p}-1} ds \le \left(\int_0^{|x|} ds\right)^{\frac{2}{p}} \left(\int_0^{|x|} (1+s^2)^{-1} ds\right)^{1-\frac{2}{p}} \le \left(\frac{\pi}{2}\right)^{1-\frac{2}{p}} |x|^{\frac{2}{p}}.$$

Thus,

$$||\Phi'(f)||_p \le \left(\frac{\pi}{2}\right)^{1-\frac{2}{p}} \left(\int_T |f|^2 d\mu\right)^{\frac{2}{p}} \le \left(\frac{\pi}{2}\right)^{1-\frac{2}{p}}.$$

On the other hand,

$$\left(\int_T (1+g^2)d\mu\right)^{1-\frac{q}{2}} \left(\int_T (1+g^2)^{\frac{2}{p}-1}\psi_j^2 d\mu\right)^{\frac{q}{2}} \ge \int_T |\psi_j|^q$$

for $1 - \frac{q}{2} + (\frac{2}{p} - 1)\frac{q}{2} = 0$.

As $g^2 \le f^2 + f_j^2$, we get $\int_T (1+g^2)d\mu \le 3$ and therefore

$$\int_T \Phi''(g)\psi_j^2 d\mu = \int_T (1+g^2)^{\frac{2}{p}-1}\psi_j^2 d\mu \ge 3^{1-\frac{2}{q}} \left(\int_T |\psi_j|^q\right)^{\frac{2}{q}} \ge 3^{1-\frac{2}{q}}\beta^{-2} = 3^{\frac{2}{p}-1}\beta^{-2}.$$

So, $F = 3^{1-\frac{2}{p}}\beta^{-2}\Phi'(f)$ is what we need.

THE BANG SOLUTION OF THE COEFFICIENT PROBLEM – 3

Beyond L^∞ (written in Banach space geometry language)

Our main aim here will be to emphasize the geometric charachter of the problem and of our approach to it. Every time when we can make the geometry clearer by some loss in constants, we will do it. So, when doing L^∞-case again, we will get the estimate $\approx \beta^{-3}$ for the norm instead of the best possible β^{-2}. The reader may think himself over what should be changed to gain exactly the same result as before.

Let H be a Hilbert space over $\mathbb{R}$, B be a closed convex set, containing the origin (from this point we will call such sets "standard"). Denote by $P_B f$ the nearest to f (in the metrics of H) element of B. We get a projection $P_B : H \to B$ (unfortunately not linear in general).

Lemma 1.

For every $f', f'' \in H$ we have $||P_B f' - P_B f''|| \le ||f' - f''||$.

Proof:

As B is convex, $(f' - P_B f', g - P_B f') \le 0$ for every $g \in B$. In particular, $(f' - P_B f', P_B f'' - P_B f') \le 0$. Analogously, $(f'' - P_B f'', P_B f' - P_B f'') \le 0$. Thus,

$$\begin{aligned}
&(f' - f'', P_B f' - P_B f'') \\
&= ||P_B f' - P_B f''||^2 - (f' - P_B f', P_B f'' - P_B f') - (f'' - P_B f'', P_B f' - P_B f'') \\
&\ge ||P_B f' - P_B f''||^2.
\end{aligned}$$

The statement of Lemma 1 follows from here immediately.

Actually, this all means just the following: if you have four points X, Y, Z, T in space and the angles XYZ and YZT are not less than $\frac{\pi}{2}$, then $|YZ| \le |XT|$.

Thus, $P_B f$ depends continuously on f (actually we proved that $P_B \in Lip1$). Let's now show that it also depends continuously on B.

Lemma 2.

Let $B_H = \{x \in H : ||x|| \le 1\}$ be the unit ball in H. Let $\delta > 0$, $f \in H$. If two standard sets B and B' satisfy $P_{B'} f \in B + \delta B_H$ and $P_B f \in B' + \delta B_H$ (i.e. $P_{B'} f$ is not more than δ distant from B and vice versa), then

$$||P_B f - P_{B'} f|| \le \sqrt{2||f||\delta}$$

Proof

Let $a = ||f - P_B f||$, $b = ||f - P_{B'} f||$. Suppose, for definiteness, that $a \ge b$. If $a = 0$, then $P_B f = P_{B'} f = f$ and there is nothing to prove. Otherwise f and B lie in different half-spaces separated by the hyperplane orthogonal to $f - P_B f$ and containing the point $P_B f$. Let H_+ be the half-space containing f and H_- be the half-space, containing B. As $a \ge b$, we have $P_{B'} f \in H_+$. Let $u = ||P_B f - P_{B'} f||$, v be the distance from $P_{B'} f$ to H_-. The Pithagorean theorem gives $v^2 - (a - v)^2 = u^2 - b^2$, or, what is the same,

$u^2 = b^2 - a^2 + 2av$. As v doesn't exceed the distance from $P_{B'}f$ to B, which is not more than δ, and as $b \le a$, we get $u \le \sqrt{2a\delta}$. Since $0 \in B$, we have $a = ||f - P_B f|| \le ||f - 0|| = ||f||$ and the conclusion of Lemma 2 follows immediately.

The Bang functional Φ_B.

Let B be a standard set in H. Consider the functional $\Phi_B : H \to \mathbb{R}$ defined by

$$\Phi_B(f) = ||f||^2 - ||f - P_B f||^2 = 2(f, P_B f) - ||P_B f||^2 \qquad (f \in H).$$

The functional Φ_B has the following remarkable

Property:

$$\Phi(f) - \Phi(f') \le 2(P_B f, f - f') - ||P_B f - P_B f'||^2.$$

Proof:

$$\begin{aligned}\Phi(f) - \Phi(f') &= 2(f, P_B f) - 2(f', P_B f') - ||P_B f||^2 + ||P_B f'||^2 \\ &= 2(P_B f, f - f') + 2(f', P_B f - P_B f') - (P_B f + P_B f', P_B f - P_B f') \\ &= 2(P_B f, f - f') + (2f' - P_B f - P_B f', P_B f - P_B f') \\ &= 2(P_B f, f - f') + 2(f' - P_B f', P_B f - P_B f') - ||P_B f - P_B f'||^2.\end{aligned}$$

It remains only to note that the second term is nonpositive.

Observation:

Let $\{\psi_j\}$ be a system of vectors in H, B be a standard set and, at last, $\{a_j\}$ be a sequence of positive numbers.

Suppose that every cortege $\varepsilon = \{\varepsilon_j\}$ where $\varepsilon_j = 1$ or -1 the series $\sum_j \varepsilon_j a_j \psi_j$ converges in H to some element f_ε.

Consider the cortege $\bar\varepsilon$ for which $\Phi_B(f_\varepsilon)$ attains its maximal value (as Φ_B is continuous and as convergence of all the series $\sum_j \varepsilon_j a_j \psi_j$ implies their uniform convergence in H with respect to the choice of signs ε_j, the cortege $\bar\varepsilon$ exists and can be constructed by diagonal process, say). Let $f = f_{\bar\varepsilon}$ be the corresponding element of H. Fix some j consider the series in which $\bar\varepsilon_j$ is replaced by $-\bar\varepsilon_j$. Let $f'_j = f - 2\bar\varepsilon_j a_j \psi_j$ be the sum of this series. Then

$$|(P_B f, \psi_j)| \ge \frac{||P_B f - P_B f'_j||^2}{4a_j}$$

Proof:

We have $0 \le \Phi_B(f) - \Phi_B(f'_j) \le 2(P_B f, f - f'_j) - ||P_B f - P_B f'_j||^2$ which results in $2(P_B f, f - f'_j) \ge ||P_B f - P_B f'_j||^2$ and, after recalling that $f - f'_j = 2\bar\varepsilon_j a_j \psi_j$, in

$$\bar\varepsilon_j (P_B f, \psi_j) \ge \frac{||P_B f - P_B f'_j||^2}{4a_j}.$$

It remains to note that $|\bar{\varepsilon}_j| = 1$.

Let now $X = L^\infty(T)$, $H = L^2(T)$ where T is some measure space with probability measure μ. Let ψ_j be a system of vectors in H satisfying $||\sum_j c_j\psi_j|| \leq \sqrt{\sum_j c_j^2}$ for every cortege of coefficients $c_j \in \mathbb{R}$ and obvious necessary condition $\int_T |\psi_j| d\mu \geq \beta > 0$. At last, let $a_j > 0$ satisfy $\sum_j a_j^2 = 1$.

Then for every cortege ε we have the corresponding series convergent in $L^2(T)$ to the sum f_ε of norm $||f_\varepsilon||_H \leq 1$. Let $s > 0$, $B = sB_{L^\infty(T)}$. The projection $P_B f$ is easy to compute for every $f \in H$. Namely,

$$(P_B f)(t) = \begin{cases} -s, \text{ if } f(t) \leq -s; \\ f(t), \text{ if } -s \leq f(t) \leq s; \\ s, \text{ if } f(t) \geq s \end{cases} \qquad (t \in T).$$

Thus, (for the same f and f' as above), $||P_B f - P_B f'||^2 \geq \int_E |f(t) - f'(t)|^2 d\mu(t) = 4a_j^2 \int_E \psi_j^2 d\mu$, where $E = \{t \in T : |f(t)| \leq s, |f'(t)| \leq s\}$. The measure of the complement $T \setminus E$ of the set E can be estimated from above by $\frac{2}{s^2}$ (just from Tschebyshev inequality). On the other hand, if $\mu(T \setminus E) \leq \frac{\beta^2}{4}$, then

$$\int_{T\setminus E} |\psi_j| d\mu \leq \mu(T \setminus E)^{\frac{1}{2}} (\int_T \psi_j^2 d\mu)^{\frac{1}{2}} \leq \frac{\beta}{2}$$

and thereby

$$\int_E \psi_j^2 d\mu \geq \left(\int_E |\psi_j| d\mu \right)^2 \geq \frac{\beta^2}{4}$$

Thus, if we put $s = \frac{4}{\beta}$, say, we will have

$$|(P_B f, \psi_j)| \geq \frac{\beta^2}{4} a_j$$

for every j. Then the function $F = \frac{4}{\beta^2} P_B f$ has L^∞-norm $||F||_{L^\infty(T)} \leq \frac{16}{\beta^3}$ and its Fourier coefficients with respect to ψ_j are greater than a_j in absolute value. So, $L^\infty(T)$ is large in $L^2(T)$.

From boundedness to continuity.

If T is a topological space and the space $C(T)$ of bounded continuous functions (with the same norm as in L^∞) is dense in $L^2(T)$, there is a temptation to improve the constructed function g with large Fourier coefficients to a continuous one.

Indeed, if the set $\{\psi_j\}$ is finite, all we need is to take the new function $F' \in C(T)$ sufficiently close to F in L^2-norm. We can do it for any finite system with uniform estimate for the norm, the space $C(T)$ is complete and therefore ... nothing follows ! One of many strange things connected with the problem is that nobody can prove or disprove the result that general (namely, it is not known whether it is always enough to check that X is large with respect to all finite subsystems of the system $\{\psi_j\}$ (which you may assume even exactly orthogonal) with uniform estimate for the norm to conclude that X is large).

So, instead of referring to some general principle, we have to work by hand. Let us consider again the standard set $B = sB_{L^\infty(T)}$ with $s = \frac{4}{\beta}$ as before, the sequence of numbers $a_j > 0$ satisfying $\sum_j a_j^2 = 1$ and the corresponding family $\{f_\varepsilon = \sum_j \varepsilon_j a_j \psi_j : \varepsilon_j = 1 \text{ or } -1\}$. Note that the family $\{f_\varepsilon\}$ is compact and thereby the set $\{P_B f_\varepsilon\}$ is compact as well (both in topology of $L^2(T)$, of course). As $sB_{C(T)}$ is dense in $B = sB_{L^\infty(T)}$ (again, in L^2-sense), we can find for every $\delta > 0$ a finite set of elements $g_1, \dots, g_{N(\delta)} \in sB_{C(T)}$ such that for every cortege ε the element $P_B f_\varepsilon$ is not more than δ distant from one of g_m. Let $B(\delta)$ be the convex hull of the origin and the points g_m. Then $B(\delta)$ is a standard set for every $\delta > 0$ and $B(\delta) \subset sB_{C(T)} \subset B$ (the main advantage of $B(\delta)$ compared to $sB_{C(T)}$ is that $B(\delta)$ is closed in topology of $L^2(T)$). Besides, every element $P_B f_\varepsilon$ is not more than δ distant from $B(\delta)$. Fix some ε and consider again the pair of functions $f = f_\varepsilon$ and $f'_j = f - 2\varepsilon_j a_j \psi_j$. We have

$$||P_{B(\delta)} f - P_{B(\delta)} f'_j|| \geq ||P_B f - P_B f'_j|| - ||P_{B(\delta)} f - P_B f|| - ||P_{B(\delta)} f'_j - P_B f'_j||.$$

But, as we saw above, the first term on the right is at least βa_j while lemma 2 together with the estimates $||f||, ||f'_j|| \leq 1$ implies that both other tems do not exceed $\sqrt{2\delta}$. Thus, if $\delta \leq \frac{\beta^2 a_j^2}{32}$, we still have

$$||P_{B(\delta)} f - P_{B(\delta)} f'_j|| \geq \frac{\beta}{2} a_j.$$

Let us now consider some sequence $\delta_k \to 0$ and define the functional $\Phi : H \to \mathbb{R}$ to maximize by

$$\Phi(f) = \sum_{k=1}^{\infty} \frac{1}{k(k+1)} \Phi_{B(\delta_k)}(f).$$

Again, let $\bar{\varepsilon}$ be the cortege for which Φ attains its maximal value. For the functions $f = f_{\bar{\varepsilon}}$ and $f'_j = f - 2\bar{\varepsilon}_j a_j \psi_j$ we get

$$0 \leq \Phi(f) - \Phi(f'_j) \leq \sum_{k=1}^{\infty} \frac{1}{k(k+1)} [2(P_{B(\delta_k)} f, f - f'_j) - ||P_{B(\delta)} f - P_{B(\delta)} f'_j||^2]$$

Thus, for the function $F = \sum_k \frac{1}{k(k+1)} P_{B(\delta)} f$ (which clearly is in $C(T)$) we have

$$\bar{\varepsilon}_j (F, \psi_j) \geq \frac{1}{4a_j} \sum_{k=1}^{\infty} \frac{1}{k(k+1)} ||P_{B(\delta)} f - P_{B(\delta)} f'_j||^2.$$

But

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)} ||P_{B(\delta)} f - P_{B(\delta)} f'_j||^2 \geq \sum_{k \,:\, \delta_k \leq \frac{\beta^2 a_j^2}{32}} \frac{1}{k(k+1)} \frac{\beta^2}{4} a_j^2 = \frac{\beta^2 a_j^2}{4k(a_j)},$$

where for every $a > 0$ the index $k(a)$ is defined as the first index k for which $\delta_k \le \frac{\beta^2 a_j^2}{32}$.

This results in

$$|(F, \psi_j)| \ge \frac{\beta^2}{16} \frac{a_j}{k(a_j)}$$

for every j. As the sequence δ_k can tend to 0 as fast as we desire, $k(a)$ is just an arbitrary decreasing finction which has values $1, 2, 3, \ldots$ and tends to infinity as $a \to 0$. Therefore the sequence $\{\frac{a_j}{k(a_j)}\}$ is nothing but an arbitrary sequence b_j of positive numbers satisfying $\sum_j b_j^2 < 1$. Thus, $C(T)$ is large and we have done.

How about $U(\mathbb{T})$?

For a while I thought that I knew how to get the result that the space $U(\mathbb{T})$ (i.e. the space of functions defined on the unit circumference $\mathbb{T}$ for which their Fourier series with respect to the standard base z^j converges uniformly with norm defined as the supremum of L^∞-norms of partial sums) is large under the same assumptions about ψ_j as before, but then I discovered that I made quite rough (though very hard to notice) mistake in my reasoning. Nevertheless, the "simple" part of what I was doing has survived and gives the following result which is only a little weaker:

Proposition:

Suppose that a sequence of numbers $a_j \in (0, \frac{1}{2})$, say, satisfies the condition

$$\sum_j a_j^2 |\log a_j|^\gamma < \infty \qquad \text{with some } \gamma > 2.$$

Then there is a function $F \in U(\mathbb{T})$ for which $|(F, \psi_j)| \ge a_j$ for every j.

Proof:

Let us do the same as for the case $X = C(T)$. The main difference is that now we can approximate an element $P_B f_\varepsilon \in B = sB_{L^\infty(\mathbb{T})}$ by an element $g \in U(\mathbb{T})$ only if we allow the norm $||g||_{U(\mathbb{T})}$ to grow to infinity. Fortunately, we do not need it to grow very fast. The best possible estimate is given by the famous

Kislyakov correction theorem:

If $||f||_{L^\infty(\mathbb{T})} \le 1$, then for every $\delta \in (0, \frac{1}{2})$, say, there exists $g \in U(\mathbb{T})$ such that $||f - g||_{L^2(\mathbb{T})} \le \delta$ while $||g||_{U(\mathbb{T})} \le C \log \frac{1}{\delta}$ ($C > 0$ is some absolute constant).

We can now repeat the construction of standard sets $B(\delta)$ using the Kislyakov theorem instead of trivial approximation of bounded functions by continuous ones. For every $\delta > 0$ we see that $B(\delta) \subset B + \delta B_{L^2(\mathbb{T})}$ and the $U(\mathbb{T})$-norms of all elements in $B(\delta)$ do not exceed $Cs \log \frac{s}{\delta}$. It allows to repeat all the estimates except that for the norm of the element F constructed. Now, to get the series $\sum_k \frac{1}{k(k+1)} P_{B(\delta)} f$ convergent in $U(\mathbb{T})$ we have to demand $\sum_k \frac{1}{k(k+1)} \log \frac{1}{\delta_k} < \infty$.

We will just put $\delta_k = \exp\{-k^{\frac{2}{\gamma}}\}$ to provide that (any further advance which we could gain from a better choice is hardly noticable compared to the distance to the desired result that $U(\mathbb{T})$ is large). This gives $k(a) \approx |\log a|^{\frac{\gamma}{2}}$ for small a and we conclude that for every sequence of numbers $a_j \in (0, \frac{1}{2})$ satisfying $\sum_j a_j^2 = 1$ there is a function $F \in U(\mathbb{T})$ such that for every j

$$|(F, \psi_j)| \geq \frac{a_j}{|\log a_j|^{\frac{\gamma}{2}}},$$

which is equivalent to the Proposition to prove.

THE BANG SOLUTION OF THE COEFFICIENT PROBLEM – 4

How Bang solved the plank problem.

Let $B \subset \mathbb{R}^n$ be a compact convex body with smooth boundary covered by finitely many open strips S_j (general case can be easily reduced to this one by standard approximation argument). Let H be the width of B, h_j be the width of S_j. Put $a_j \stackrel{\text{def}}{=} \frac{h_j}{2}$ and consider the vectors ψ_j of length a_j orthogonal to the boundary hyperplanes of the corresponding strips S_j.

Bang's solution consists of two independent steps:

Lemma 1:

If ψ is a vector of length $a < \frac{H}{2}$, then the intersection $(B - \psi) \cap (B + \psi)$ contains a homotetic image of B with coefficient $\frac{H-2a}{H}$

Corollary:

If $\sum_j h_j < H$, then there exists a point $x_0 \in \mathbb{R}^n$ such that for every sign cortege $\varepsilon = \{\varepsilon_j\}$ the point $x_\varepsilon \stackrel{\text{def}}{=} x_0 + \sum_j \varepsilon_j \psi_j \in B$.

Lemma2:

For every $x_0 \in \mathbb{R}^n$ at least one of the points x_ε doesn't belong to any of the strips S_j.

Proof of lemma 1:

Let yz be the longest section of B by a line parallel to ψ. Let ℓ be the length of the interval yz, c be the middle of this interval. Note that the tangent hyperplanes to B at the points y and z are parallel (othewise we might move the line yz a bit and get a longer section). Since ℓ is not less than the distance between these hyperplanes, we conclude that $\ell \geq H$.

As $\psi = \frac{a}{\ell}(z - y)$, for every $x \in B$ we have

$$\frac{2a}{\ell}c + \frac{\ell - 2a}{\ell}x + \psi = \frac{2a}{\ell}\frac{y+z}{2} + \frac{\ell - 2a}{\ell}x + \frac{a}{\ell}(z - y) = \frac{2a}{\ell}z + \frac{\ell - 2a}{\ell}x \in B$$

(because B is convex, $x, z \in B$ and $\frac{2a}{\ell} + \frac{\ell-2a}{\ell} = 1$). Therefore, $\frac{2a}{\ell}c + \frac{\ell-2a}{\ell}x \in B - \psi$. Analogously we get that also $\frac{2a}{\ell}c + \frac{\ell-2a}{\ell}x \in B + \psi$. It remains to note that the point $\frac{2a}{\ell}c + \frac{\ell-2a}{\ell}x$ runs over the image of B under the homotopy with the center c and the coefficient $\frac{\ell-2a}{\ell} \geq \frac{H-2a}{H}$ as x runs over B.

Proof of the corollary:

By induction in the number of vectors ψ_j we find that the intersection $\cap_\varepsilon(B - \sum_j \varepsilon_j \psi_j)$ contains a honotopic image of B with the coefficient $\frac{H-\sum_j h_j}{H} > 0$ and therefore is not empty. Every point of this intersection can be taken as x_0.

Remark to the step 1:

We didn't use this part of Bang's solution in our reasoning at all because in our case we had the body B symmetric with respect to the origin which allows just to put $x_0 = 0$. Also, it is this part where the distinction between $\sum_j a_j$ in the plank problem and $\sum_j a_j^2$ in the coefficient problem comes from: in the plank problem we deal with completely arbitrary vectors ψ_j and therefore the triangle inequality is the only estimate for the norm we can use, while in the coefficient problem we had our vecrors mutually orthogonal and might apply the Pithagorean theorem instead.

Proof of lemma 2:

Let p_k be any point of the "middle" hyperplane of the strip S_k. Then

$$S_k = \{x \in \mathbb{R}^n : |(x - p_k, \psi_k)| < a_j^2.$$

As adding ψ_j to a point means pushing the point toward the "positive" side of the strip S_j while adding $-\psi_j$ means pushing it in the opposite direction, it is natural to try to find the point x_ε which lies from the positive or from the negative side of S_k corresponding to whether $\varepsilon_k = 1$ or -1. This gives the system

$$\varepsilon_k(x_\varepsilon - p_k, \psi_k) \geq a_k^2 \qquad (k = 1, 2, \dots)$$

of inequalities to solve. Recalling that $x_\varepsilon = x_0 + \sum_j \varepsilon_j \psi_j$, we get

$$\varepsilon_k(x_0 - p_k, \psi_k) + \sum_j \varepsilon_k \varepsilon_j (\psi_k, \psi_j) \geq a_k^2 \qquad (k = 1, 2, \dots).$$

Note that when $j = k$, we have the term $\varepsilon_k^2 |\psi_k|^2 = a_k^2$ on the left. Cancelling it out, we get left with

$$\varepsilon_k(x_0 - p_k, \psi_k) + \sum_{j: j \neq k} \varepsilon_k \varepsilon_j (\psi_k, \psi_j) \geq 0 \qquad (k = 1, 2, \dots).$$

Let $\Phi(\varepsilon) \stackrel{\text{def}}{=} \sum_k \varepsilon_k(x_0 - p_k, \psi_k) + \sum_{j,k: j<k} \varepsilon_k \varepsilon_j (\psi_k, \psi_j)$. Note that the inequality above is exactly $\frac{1}{2}(\Phi(\varepsilon) - \Phi(\varepsilon')) \geq 0$ where $\varepsilon_j' = \varepsilon_j$ for every $j \neq k$ and $\varepsilon_k' = -\varepsilon_k$. Thus, if we take the sign cortege $\bar{\varepsilon}$ for which the function $\Phi(\varepsilon)$ attains its maximal value, the corresponding point $x_{\bar{\varepsilon}}$ will lye in none of the strips. That's all !

THE BANG SOLUTION OF THE COEFFICIENT PROBLEM – 5
Concluding remarks (written a year later).

This paper was written about a year ago. I intentionally delayed with its publication hoping that someone would be able to combine Bang's approach with some other idea (say, with the classical Kahane–Katznelson–de-Leeuw technique) and to get the result that $U(\mathbb{T})$ is large with respect to a system $\{\psi_j\}$ satisfying the same a priori assumption $\|\sum_j a_j\psi_j\|_2 \le \left(\sum_j a_j^2\right)^{1/2}$ and the trivial necessary condition $\int_{\mathbb{T}} |\psi_j| d\mu \ge \beta > 0$. Despite the paper has got quite famous during that year and despite all my own efforts, this has not happened yet. Moreover, even a simpler question about "largeness" of $X = H^\infty(\mathbb{D})$ in $H = H^2(\mathbb{D})$ remains unanswered.

Still I've got several interesting remarks from various people. Two most interesting of them were the following.

First, Keith Ball pointed out to me that the "coefficient problem" can be solved completely if one is interested in majorizing a ℓ^1 sequence $\{a_j\}$ instead of ℓ^2 that. Namely, in 1991 he proved the following

Theorem.

Let X be any Banach space, $\psi_j \in X^*$ be any family of bounded linear functionals on X such that $\|\psi_j\|_{X^*} = 1$ for every j. Then for any sequence $\{a_j\}$ of positive numbers satisfying $\sum_j a_j < 1$ there exists $f \in X$ such that $\|f\|_X \le 1$ and $|(f, \psi_j)| \ge a_j$ for every j.

The proof of the theorem can be found in [3].

The second remark is due to Francoise Lust-Piquard. She noticed that the Bang approach works also in a "non-commutative" setting. Among other results her forthcoming paper [4] contains the following

Proposition.

Let $\{a_{ij}\}_{i,j=1}^\infty$ be a matrix with positive entries such that

$$\sum_{j=1}^\infty a_{ij}^2 \le 1 \text{ for every } i \quad \text{and} \quad \sum_{i=1}^\infty a_{ij}^2 \le 1 \text{ for every } j.$$

Then there exists a bounded operator $A : \ell^2 \to \ell^2$ of norm $\|A\| \le 10$, say, such that its matrix satisfies $|A_{ij}| \ge a_{ij}$ for every i, j.

The above condition is obviously necessary for existence of such an operator of norm 1.

Literature cited

[1] Bang T., A solution of the "plank problem." Proc. Amer. Math. Soc. 2, (1951). 990–993

[2] de Leeuw, K; Katznelson, Y; Kahane, J.-P. Sur les coefficients de Fourier des fonctions continues. C. R. Acad. Sci. Paris Sr. A-B 285 (1977), no. 16, A1001–A1003

[3] Ball K. The plank problem for symmetric bodies. Invent. Math. 104 (1991), no. 3, 535–543

[4] Lust-Piquard, F. On the coefficient problem: a version of the Kahane-Katznelson-De Leeuw theorem for spaces of matrices. to appear in Journal of Functional Analysis.