Interlacing polynomials, restricted invertibility and multi-paving

Summer School in Operator Theory

Part I

July 9, 2019

Restricted invertibility

The restricted invertibility principle of Bourgain and Tzafriri states that if A is an $n \times n$ matrix whose columns Ae_j have Euclidean norm equal to 1 then there exists $\sigma \subset [n]$ of cardinality $|\sigma| \geqslant cn/\|A\|_2^2$ such that the restriction A_σ of A to $\mathrm{span}\{e_j : j \in \sigma\}$ is well-invertible.

Restricted invertibility

The restricted invertibility principle of Bourgain and Tzafriri states that if A is an $n \times n$ matrix whose columns Ae_j have Euclidean norm equal to 1 then there exists $\sigma \subset [n]$ of cardinality $|\sigma| \geqslant cn/\|A\|_2^2$ such that the restriction A_σ of A to $\mathrm{span}\{e_j : j \in \sigma\}$ is well-invertible.

Bourgain-Tzafriri

There exist absolute constants $\delta, \gamma > 0$ such that if $A : \ell_2^n \longrightarrow \ell_2^n$ is a linear operator with $|Ae_j| = 1$ for all $j = 1, \ldots, n$ then one may find a subset $\sigma \subseteq [n]$ of cardinality $|\sigma| \geqslant \delta n/\|A\|_2^2$ such that

$$\Big|\sum_{j\in\sigma} t_j A e_j\Big|^2 \geqslant \gamma \sum_{j\in\sigma} |t_j|^2$$

for any choice of scalars $\{t_j\}_{j\in\sigma}$.

The restricted invertibility principle of Bourgain and Tzafriri states that if A is an $n \times n$ matrix whose columns Ae_j have Euclidean norm equal to 1 then there exists $\sigma \subset [n]$ of cardinality $|\sigma| \geqslant cn/\|A\|_2^2$ such that the restriction A_σ of A to $\mathrm{span}\{e_j : j \in \sigma\}$ is well-invertible.

Bourgain-Tzafriri

There exist absolute constants $\delta, \gamma > 0$ such that if $A : \ell_2^n \longrightarrow \ell_2^n$ is a linear operator with $|Ae_j| = 1$ for all $j = 1, \ldots, n$ then one may find a subset $\sigma \subseteq [n]$ of cardinality $|\sigma| \geqslant \delta n/\|A\|_2^2$ such that

$$\Big| \sum_{j\in\sigma} t_j A e_j \Big|^2 \geqslant \gamma \sum_{j\in\sigma} |t_j|^2$$

for any choice of scalars $\{t_j\}_{j\in\sigma}$.

- Recall that the stable rank of A is defined by $\mathrm{srank}(A) := \|A\|_{\mathrm{HS}}^2/\|A\|_2^2$.

Restricted invertibility

The restricted invertibility principle of Bourgain and Tzafriri states that if A is an $n \times n$ matrix whose columns Ae_j have Euclidean norm equal to 1 then there exists $\sigma \subset [n]$ of cardinality $|\sigma| \geqslant cn/\|A\|_2^2$ such that the restriction A_σ of A to $\mathrm{span}\{e_j : j \in \sigma\}$ is well-invertible.

Bourgain-Tzafriri

There exist absolute constants $\delta, \gamma > 0$ such that if $A : \ell_2^n \longrightarrow \ell_2^n$ is a linear operator with $|Ae_j| = 1$ for all $j = 1, \ldots, n$ then one may find a subset $\sigma \subseteq [n]$ of cardinality $|\sigma| \geqslant \delta n/\|A\|_2^2$ such that

$$\Big|\sum_{j\in\sigma} t_j A e_j\Big|^2 \geqslant \gamma \sum_{j\in\sigma} |t_j|^2$$

for any choice of scalars $\{t_j\}_{j\in\sigma}$.

- Recall that the stable rank of A is defined by $\mathrm{srank}(A) := \|A\|_{\mathrm{HS}}^2/\|A\|_2^2$.
- Assuming that $|Ae_j| = 1$ for all $j = 1, \ldots, n$ we have $\|A\|_{\mathrm{HS}}^2 = n$, therefore the cardinality of the set σ provided by this theorem is proportional to the stable rank of A.

Restricted invertibility

- Note also that if A_σ is the restriction of A to $\mathrm{span}\{e_j : j \in \sigma\}$ then the theorem is equivalent to the fact that $s_{\min}(A_\sigma) \geqslant \gamma$, where

$$s_{\min}(T) = \lambda_{\min}(\sqrt{T^t T}) = \lambda_{\min}(\sqrt{T T^t})$$

denotes the smallest singular number of an operator T.

Restricted invertibility

- Note also that if A_σ is the restriction of A to $\mathrm{span}\{e_j : j \in \sigma\}$ then the theorem is equivalent to the fact that $s_{\min}(A_\sigma) \geqslant \gamma$, where

$$s_{\min}(T) = \lambda_{\min}(\sqrt{T^t T}) = \lambda_{\min}(\sqrt{T T^t})$$

  denotes the smallest singular number of an operator T.
- The original proof uses random selection, the Sauer-Shelah lemma and Grothendieck's factorization theorem.

Restricted invertibility

- Note also that if A_σ is the restriction of A to $\mathrm{span}\{e_j : j \in \sigma\}$ then the theorem is equivalent to the fact that $s_{\min}(A_\sigma) \geqslant \gamma$, where

$$s_{\min}(T) = \lambda_{\min}(\sqrt{T^t T}) = \lambda_{\min}(\sqrt{T T^t})$$

  denotes the smallest singular number of an operator T.
- The original proof uses random selection, the Sauer-Shelah lemma and Grothendieck's factorization theorem.

Vershynin generalized this result as follows.

Vershynin

Let $I_n = \sum_{j=1}^m u_j u_j^t$ be an arbitrary decomposition of the identity and $A : \ell_2^n \to \ell_2^n$ be a linear operator. Then, for any $\epsilon \in (0,1)$ one can find $\sigma \subset [m]$ of cardinality $|\sigma| \geqslant (1-\epsilon)\,\mathrm{srank}(A)$ such that for any choice of scalars $(t_j)_{j\in\sigma}$,

$$\Big| \sum_{j\in\sigma} t_j \frac{Au_j}{|Au_j|} \Big| \geqslant c(\epsilon) \Big(\sum_{j\in\sigma} t_j^2 \Big)^{1/2},$$

where $c(\varepsilon) > 0$ is a constant depending only on ϵ.

Restricted invertibility

- Note that if $|Ae_j| = 1$ for all j then, applying Vershynin's theorem for the standard decomposition $I_n = \sum_{j=1}^n e_j e_j^t$ we recover the Bourgain-Tzafriri theorem.

Restricted invertibility

- Note that if $|Ae_j| = 1$ for all j then, applying Vershynin's theorem for the standard decomposition $I_n = \sum_{j=1}^n e_j e_j^t$ we recover the Bourgain-Tzafriri theorem.
- Moreover, we may now find $\sigma \subseteq [n]$ of cardinality greater than $(1-\epsilon)n/\|A\|_2^2$ *for any* $\epsilon \in (0,1)$ so that

$$\Big|\sum_{j\in\sigma} t_j A e_j\Big|^2 \geqslant \gamma \sum_{j\in\sigma} |t_j|^2$$

will hold true, of course with a constant $\gamma = \gamma(\epsilon)$ depending on ϵ.

Restricted invertibility

- Note that if $|Ae_j| = 1$ for all j then, applying Vershynin's theorem for the standard decomposition $I_n = \sum_{j=1}^{n} e_j e_j^t$ we recover the Bourgain-Tzafriri theorem.
- Moreover, we may now find $\sigma \subseteq [n]$ of cardinality greater than $(1-\epsilon)n/\|A\|_2^2$ *for any* $\epsilon \in (0,1)$ so that
$$\Big|\sum_{j\in\sigma} t_j A e_j\Big|^2 \geqslant \gamma \sum_{j\in\sigma} |t_j|^2$$
will hold true, of course with a constant $\gamma = \gamma(\epsilon)$ depending on ϵ.

Spielman and Srivastava gave a generalization, in the spirit of Vershynin's theorem, with optimal dependence on ϵ.

Spielman-Srivastava

Let $\epsilon \in (0,1)$ and $u_1, \ldots, u_m \in \mathbb{R}^n$ such that $I_n = \sum_{j=1}^{m} u_j u_j^t$. Let $A : \ell_2^n \to \ell_2^n$ be a linear operator. We can find $\sigma \subseteq [m]$ of cardinality $|\sigma| \geqslant \lfloor (1-\epsilon)^2 \mathrm{srank}(A) \rfloor$ such that the set $\{Au_j : j \in \sigma\}$ is linearly independent and

$$\lambda_{\min}\Big(\sum_{j\in\sigma} (Au_j)(Au_j)^t\Big) \geqslant \epsilon^2 \frac{\|A\|_{\mathrm{HS}}^2}{m},$$

where the smallest eigenvalue $\lambda_{\min}$ is computed on the subspace $\mathrm{span}\{Au_j : j \in \sigma\}$.

Restricted invertibility

- The inequality

$$\lambda_{\min}\Big(\sum_{j\in\sigma}(Au_j)(Au_j)^t\Big) \geqslant \epsilon^2\,\frac{\|A\|_{\mathrm{HS}}^2}{m},$$

is equivalent to the fact that, for any choice of scalars $(t_j)_{j\in\sigma}$,

$$\Big|\sum_{j\in\sigma} t_j A u_j\Big| \geqslant \epsilon\,\frac{\|A\|_{\mathrm{HS}}}{\sqrt{m}}\Big(\sum_{j\in\sigma} t_j^2\Big)^{1/2}.$$

Restricted invertibility

- The inequality

$$\lambda_{\min}\Big(\sum_{j\in\sigma}(Au_j)(Au_j)^t\Big) \geqslant \epsilon^2\,\frac{\|A\|_{\mathrm{HS}}^2}{m},$$

is equivalent to the fact that, for any choice of scalars $(t_j)_{j\in\sigma}$,

$$\Big|\sum_{j\in\sigma} t_j A u_j\Big| \geqslant \epsilon\,\frac{\|A\|_{\mathrm{HS}}}{\sqrt{m}}\Big(\sum_{j\in\sigma} t_j^2\Big)^{1/2}.$$

- The Bourgain-Tzafriri theorem follows with constants $\gamma(\epsilon)=\epsilon^2\ \ \delta(\epsilon)=(1-\epsilon)^2$; consider the standard decomposition of the identity $I_n=\sum_{i=1}^n e_j e_j^t$, where $\{e_j : j=1,\dots,n\}$ and recall that $\|A\|_{\mathrm{HS}}^2=n$.

Strong restricted invertibility

Strong B-T

There exists an absolute constant $\gamma > 0$ with the following property: for every $B > 0$ there exists $r = r(B) \in \mathbb{N}$ such that if $A : \ell_2^n \to \ell_2^n$ is a linear operator with $\|A\| \leqslant B$ and $|Ae_j| = 1$ for all $i = 1, \ldots, n$, then we may find a partition $\{\sigma_k\}_{k=1}^r$ of $[n]$ such that for every $k = 1, \ldots, r$ and any choice of real coefficients $(t_j)_{j\in\sigma_k}$,

$$\Big|\sum_{j\in\sigma_k} t_j Ae_j\Big|^2 \geqslant \gamma \sum_{j\in\sigma_k} |t_j|^2.$$

- This statement is called the *strong Bourgain-Tzafriri conjecture*.

Strong restricted invertibility

Strong B-T

There exists an absolute constant $\gamma > 0$ with the following property: for every $B > 0$ there exists $r = r(B) \in \mathbb{N}$ such that if $A : \ell_2^n \to \ell_2^n$ is a linear operator with $\|A\| \leqslant B$ and $|Ae_j| = 1$ for all $i = 1, \ldots, n$, then we may find a partition $\{\sigma_k\}_{k=1}^r$ of $[n]$ such that for every $k = 1, \ldots, r$ and any choice of real coefficients $(t_j)_{j\in\sigma_k}$,

$$\Big|\sum_{j\in\sigma_k} t_j A e_j\Big|^2 \geqslant \gamma \sum_{j\in\sigma_k} |t_j|^2.$$

- This statement is called the *strong Bourgain-Tzafriri conjecture*.
- It is now a theorem, after the proof of the *paving conjecture* by Marcus, Spielman and Srivastava.

Strong restricted invertibility

Strong B-T

There exists an absolute constant $\gamma > 0$ with the following property: for every $B > 0$ there exists $r = r(B) \in \mathbb{N}$ such that if $A : \ell_2^n \to \ell_2^n$ is a linear operator with $\|A\| \leqslant B$ and $|Ae_j| = 1$ for all $i = 1, \ldots, n$, then we may find a partition $\{\sigma_k\}_{k=1}^r$ of $[n]$ such that for every $k = 1, \ldots, r$ and any choice of real coefficients $(t_j)_{j\in\sigma_k}$,

$$\Big|\sum_{j\in\sigma_k} t_j A e_j\Big|^2 \geqslant \gamma \sum_{j\in\sigma_k} |t_j|^2.$$

- This statement is called the *strong Bourgain-Tzafriri conjecture*.
- It is now a theorem, after the proof of the *paving conjecture* by Marcus, Spielman and Srivastava.
- In fact, the paving theorem is equivalent to the statement above, and also provides an affirmative answer to the Kadison-Singer problem.

Paving theorem

Paving

For every $\varepsilon > 0$ there exists $r = r(\varepsilon) \in \mathbb{N}$ such that: if $S : \ell_2^n \to \ell_2^n$ is a linear operator with diagonal $D(S) = 0$ then we may find a partition $\{\sigma_k\}_{k=1}^r$ of $[n]$ such that for every $k = 1, \ldots, r$

$$\|R_{\sigma_k} S R_{\sigma_k}\| \leqslant \varepsilon \|S\|.$$

Paving theorem

Paving

For every $\varepsilon > 0$ there exists $r = r(\varepsilon) \in \mathbb{N}$ such that: if $S : \ell_2^n \to \ell_2^n$ is a linear operator with diagonal $D(S) = 0$ then we may find a partition $\{\sigma_k\}_{k=1}^r$ of $[n]$ such that for every $k = 1, \dots, r$

$$\|R_{\sigma_k} S R_{\sigma_k}\| \leqslant \varepsilon \|S\|.$$

- Here, R_σ is the restriction map (orthogonal projection) onto the subspace $\mathrm{span}\{e_j : j \in \sigma\}$.

Paving theorem

Paving

For every $\varepsilon > 0$ there exists $r = r(\varepsilon) \in \mathbb{N}$ such that: if $S : \ell_2^n \to \ell_2^n$ is a linear operator with diagonal $D(S) = 0$ then we may find a partition $\{\sigma_k\}_{k=1}^r$ of $[n]$ such that for every $k = 1, \dots, r$

$$\|R_{\sigma_k} S R_{\sigma_k}\| \leqslant \varepsilon \|S\|.$$

- Here, R_σ is the restriction map (orthogonal projection) onto the subspace $\mathrm{span}\{e_j : j \in \sigma\}$.
- The delicate point is that r should depend only on ε and not on n.

Paving implies restricted invertibility

- Let $A : \ell_2^n \to \ell_2^n$ satisfy $|Ae_j| = 1$ for all $i \leqslant n$. Then, $B = A^t A$ has diagonal $D(B) = (1, \ldots, 1)$ and $\|B\| = \|A^t A\| = \|A\|^2$.

Paving implies restricted invertibility

- Let $A : \ell_2^n \to \ell_2^n$ satisfy $|Ae_j| = 1$ for all $i \leqslant n$. Then, $B = A^t A$ has diagonal $D(B) = (1, \dots, 1)$ and $\|B\| = \|A^t A\| = \|A\|^2$.
- Write $B = I_n + S$. Then, $D(S) = 0$ and applying the paving theorem we may find $r = r(\|A\|)$ and a partition $\{\sigma_k\}_{k=1}^r$ of $[n]$ such that for every $k = 1, \dots, r$

$$\|R_{\sigma_k}(B - I_n)R_{\sigma_k}\| \leqslant \varepsilon.$$

Paving implies restricted invertibility

- Let $A : \ell_2^n \to \ell_2^n$ satisfy $|Ae_j| = 1$ for all $i \leqslant n$. Then, $B = A^t A$ has diagonal $D(B) = (1, \dots, 1)$ and $\|B\| = \|A^t A\| = \|A\|^2$.
- Write $B = I_n + S$. Then, $D(S) = 0$ and applying the paving theorem we may find $r = r(\|A\|)$ and a partition $\{\sigma_k\}_{k=1}^r$ of $[n]$ such that for every $k = 1, \dots, r$

$$\|R_{\sigma_k}(B - I_n)R_{\sigma_k}\| \leqslant \varepsilon.$$

- It follows that for every $x \in \ell_2^n$ and every $k \leqslant r$ we have

$$\begin{aligned} \langle BR_{\sigma_k}x, R_{\sigma_k}x\rangle &= \langle R_{\sigma_k}BR_{\sigma_k}x, R_{\sigma_k}x\rangle \\ &= \langle R_{\sigma_k}x, R_{\sigma_k}x\rangle - \langle R_{\sigma_k}(I_n - B)R_{\sigma_k}x, R_{\sigma_k}x\rangle \\ &\geqslant |R_{\sigma_k}x|^2 - |R_{\sigma_k}(I_n - B)R_{\sigma_k}x|\,|R_{\sigma_k}x| \\ &\geqslant |R_{\sigma_k}x|^2 - \varepsilon|R_{\sigma_k}x|^2 = (1-\varepsilon)|R_{\sigma_k}x|^2. \end{aligned}$$

Paving implies restricted invertibility

- Let $A : \ell_2^n \to \ell_2^n$ satisfy $|Ae_j| = 1$ for all $i \leqslant n$. Then, $B = A^t A$ has diagonal $D(B) = (1, \ldots, 1)$ and $\|B\| = \|A^t A\| = \|A\|^2$.
- Write $B = I_n + S$. Then, $D(S) = 0$ and applying the paving theorem we may find $r = r(\|A\|)$ and a partition $\{\sigma_k\}_{k=1}^r$ of $[n]$ such that for every $k = 1, \ldots, r$

$$\|R_{\sigma_k}(B - I_n)R_{\sigma_k}\| \leqslant \varepsilon.$$

- It follows that for every $x \in \ell_2^n$ and every $k \leqslant r$ we have

$$\begin{aligned}
\langle BR_{\sigma_k}x, R_{\sigma_k}x\rangle &= \langle R_{\sigma_k}BR_{\sigma_k}x, R_{\sigma_k}x\rangle \\
&= \langle R_{\sigma_k}x, R_{\sigma_k}x\rangle - \langle R_{\sigma_k}(I_n - B)R_{\sigma_k}x, R_{\sigma_k}x\rangle \\
&\geqslant |R_{\sigma_k}x|^2 - |R_{\sigma_k}(I_n - B)R_{\sigma_k}x|\,|R_{\sigma_k}x| \\
&\geqslant |R_{\sigma_k}x|^2 - \varepsilon|R_{\sigma_k}x|^2 = (1 - \varepsilon)|R_{\sigma_k}x|^2.
\end{aligned}$$

- This verifies the strong Bourgain-Tzafriri conjecture.

The case of ℓ_1^n

Schechtman

Let $S : \ell_1^n \to \ell_1^n$ with $D(S) = 0$. For every $\varepsilon > 0$ there exists $\sigma \subset [n]$ with $|\sigma| \geqslant \frac{\varepsilon n}{2\|S\|}$ such that

$$\|R_\sigma S R_\sigma\| \leqslant \varepsilon.$$

Schechtman

Let $S : \ell_1^n \to \ell_1^n$ with $D(S) = 0$. For every $\varepsilon > 0$ there exists $\sigma \subset [n]$ with $|\sigma| \geqslant \frac{\varepsilon n}{2\|S\|}$ such that

$$\|R_\sigma S R_\sigma\| \leqslant \varepsilon.$$

- If $S : \mathbb{R}^n \to \mathbb{R}^n$ is a linear operator and $a_{ij} = \langle Se_j, e_i \rangle$ then

$$\|S : \ell_1^n \to \ell_1^n\| = \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} |a_{ij}|$$

and

$$\|S : \ell_\infty^n \to \ell_\infty^n\| = \max_{1 \leqslant j \leqslant n} \sum_{i=1}^{n} |a_{ij}|.$$

Schechtman

Let $S : \ell_1^n \to \ell_1^n$ with $D(S) = 0$. For every $\varepsilon > 0$ there exists $\sigma \subset [n]$ with $|\sigma| \geqslant \frac{\varepsilon n}{2\|S\|}$ such that

$$\|R_\sigma S R_\sigma\| \leqslant \varepsilon.$$

- If $S : \mathbb{R}^n \to \mathbb{R}^n$ is a linear operator and $a_{ij} = \langle Se_j, e_i \rangle$ then

$$\|S : \ell_1^n \to \ell_1^n\| = \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} |a_{ij}|$$

and

$$\|S : \ell_\infty^n \to \ell_\infty^n\| = \max_{1 \leqslant j \leqslant n} \sum_{i=1}^{n} |a_{ij}|.$$

- Therefore, we may replace S by S', which is defined by $\langle S' e_j, e_i \rangle = |a_{ij}|$, and assume that all a_{ij} are non-negative.

The case of ℓ_1^n

Paving in ℓ_1^n

Let $A = (a_{ij})_{i,j=1}^n$ be an $n \times n$ matrix such that:

- $a_{ij} \geqslant 0$ and $a_{ii} = 0$.
- For all $i \leqslant n$, $\sum_{j=1}^n a_{ij} \leqslant 1$.

Then, for every $\varepsilon > 0$ there exists $\sigma \subset [n]$ with $|\sigma| \geqslant \frac{\varepsilon n}{2}$ such that: for every $i \in \sigma$, $\sum_{j\in\sigma} a_{ij} \leqslant \varepsilon$.

Paving in ℓ_1^n

Let $A = (a_{ij})_{i,j=1}^n$ be an $n \times n$ matrix such that:

- $a_{ij} \geqslant 0$ and $a_{ii} = 0$.
- For all $i \leqslant n$, $\sum_{j=1}^n a_{ij} \leqslant 1$.

Then, for every $\varepsilon > 0$ there exists $\sigma \subset [n]$ with $|\sigma| \geqslant \frac{\varepsilon n}{2}$ such that: for every $i \in \sigma$, $\sum_{j\in\sigma} a_{ij} \leqslant \varepsilon$.

- In fact, we have something stronger.

The case of ℓ_1^n

Paving in ℓ_1^n

Let $A = (a_{ij})_{i,j=1}^n$ be an $n \times n$ matrix such that:

- $a_{ij} \geqslant 0$ and $a_{ii} = 0$.
- For all $i \leqslant n$, $\sum_{j=1}^n a_{ij} \leqslant 1$.

Then, for every $\varepsilon > 0$ there exists $\sigma \subset [n]$ with $|\sigma| \geqslant \frac{\varepsilon n}{2}$ such that: for every $i \in \sigma$, $\sum_{j \in \sigma} a_{ij} \leqslant \varepsilon$.

- In fact, we have something stronger.

Bourgain

Let $A = (a_{ij})_{i,j=1}^n$ be an $n \times n$ matrix such that:

- $a_{ij} \geqslant 0$ and $a_{ii} = 0$.
- For every $i \leqslant n$, $\sum_{j=1}^n a_{ij} \leqslant 1$.

Then, for every $k \in \mathbb{N}$ there exists a partition $\{\sigma_1, \ldots, \sigma_k\}$ of $[n]$ such that: for all $\ell \leqslant k$ and any $i \in \sigma_\ell$,

$$\sum_{j \in \sigma_\ell} a_{ij} \leqslant \frac{2}{k}.$$

The case of ℓ_1^n

This implies paving.

Paving in ℓ_1^n

For any $k \in \mathbb{N}$ and any $S : \ell_1^n \to \ell_1^n$ with $D(S) = 0$ there exists a partition $\{\sigma_1, \ldots, \sigma_k\}$ of $[n]$ such that: for all $\ell \leqslant k$

$$\|R_{\sigma_\ell} S R_{\sigma_\ell}\| \leqslant \frac{2}{k}\|S\|.$$

The case of ℓ_1^n

- The next argument is due to K. Ball.

The case of ℓ_1^n

- The next argument is due to K. Ball. We may assume that $a_{ij} > 0$ if $i \neq j$ and that for every $i \leqslant n$ we have $\sum_{j=1}^{n} a_{ij} = 1$.

The case of ℓ_1^n

- The next argument is due to K. Ball. We may assume that $a_{ij} > 0$ if $i \neq j$ and that for every $i \leqslant n$ we have $\sum_{j=1}^{n} a_{ij} = 1$.
- Then $\rho = 1$ is an eigenvalue of A with right eigenvector $\mathbf{1} = (1, \dots, 1)$. So, there exists non zero $\gamma = (\gamma_1, \dots, \gamma_n)$ with $\gamma A = \gamma$.

The case of ℓ_1^n

- The next argument is due to K. Ball. We may assume that $a_{ij} > 0$ if $i \neq j$ and that for every $i \leqslant n$ we have $\sum_{j=1}^n a_{ij} = 1$.
- Then $\rho = 1$ is an eigenvalue of A with right eigenvector $\mathbf{1} = (1, \dots, 1)$. So, there exists non zero $\gamma = (\gamma_1, \dots, \gamma_n)$ with $\gamma A = \gamma$.

Claim 1

We may assume that $\gamma_i > 0$ for all $i \leqslant n$.

The case of ℓ_1^n

- The next argument is due to K. Ball. We may assume that $a_{ij} > 0$ if $i \neq j$ and that for every $i \leqslant n$ we have $\sum_{j=1}^{n} a_{ij} = 1$.
- Then $\rho = 1$ is an eigenvalue of A with right eigenvector $\mathbf{1} = (1, \ldots, 1)$. So, there exists non zero $\gamma = (\gamma_1, \ldots, \gamma_n)$ with $\gamma A = \gamma$.

Claim 1

We may assume that $\gamma_i > 0$ for all $i \leqslant n$.

- Since $\gamma A = \gamma$ we have $\sum_{i=1}^{n} \gamma_i a_{ij} = \gamma_j$ for all $j \leqslant n$.

The case of ℓ_1^n

- The next argument is due to K. Ball. We may assume that $a_{ij} > 0$ if $i \neq j$ and that for every $i \leqslant n$ we have $\sum_{j=1}^n a_{ij} = 1$.
- Then $\rho = 1$ is an eigenvalue of A with right eigenvector $\mathbf{1} = (1, \ldots, 1)$. So, there exists non zero $\gamma = (\gamma_1, \ldots, \gamma_n)$ with $\gamma A = \gamma$.

Claim 1

We may assume that $\gamma_i > 0$ for all $i \leqslant n$.

- Since $\gamma A = \gamma$ we have $\sum_{i=1}^n \gamma_i a_{ij} = \gamma_j$ for all $j \leqslant n$.
- It follows that

$$\sum_{j=1}^n |\gamma_j| = \sum_{j=1}^n \Big| \sum_{i=1}^n \gamma_i a_{ij} \Big| \leqslant \sum_{j=1}^n \sum_{i=1}^n |\gamma_i| a_{ij} = \sum_{i=1}^n |\gamma_i| \sum_{j=1}^n a_{ij} = \sum_{i=1}^n |\gamma_i|.$$

The case of ℓ_1^n

- The next argument is due to K. Ball. We may assume that $a_{ij} > 0$ if $i \neq j$ and that for every $i \leqslant n$ we have $\sum_{j=1}^n a_{ij} = 1$.
- Then $\rho = 1$ is an eigenvalue of A with right eigenvector $\mathbf{1} = (1, \ldots, 1)$. So, there exists non zero $\gamma = (\gamma_1, \ldots, \gamma_n)$ with $\gamma A = \gamma$.

Claim 1

We may assume that $\gamma_i > 0$ for all $i \leqslant n$.

- Since $\gamma A = \gamma$ we have $\sum_{i=1}^n \gamma_i a_{ij} = \gamma_j$ for all $j \leqslant n$.
- It follows that

$$\sum_{j=1}^n |\gamma_j| = \sum_{j=1}^n \Big| \sum_{i=1}^n \gamma_i a_{ij} \Big| \leqslant \sum_{j=1}^n \sum_{i=1}^n |\gamma_i| a_{ij} = \sum_{i=1}^n |\gamma_i| \sum_{j=1}^n a_{ij} = \sum_{i=1}^n |\gamma_i|.$$

- Since we have equality everywhere, all γ_j have the same sign. So, we may assume that $\gamma_j \geqslant 0$ for all $j \leqslant n$. We shall show that $\gamma_j > 0$ for all $j \leqslant n$.

- The next argument is due to K. Ball. We may assume that $a_{ij} > 0$ if $i \neq j$ and that for every $i \leqslant n$ we have $\sum_{j=1}^n a_{ij} = 1$.
- Then $\rho = 1$ is an eigenvalue of A with right eigenvector $\mathbf{1} = (1, \ldots, 1)$. So, there exists non zero $\gamma = (\gamma_1, \ldots, \gamma_n)$ with $\gamma A = \gamma$.

Claim 1

We may assume that $\gamma_i > 0$ for all $i \leqslant n$.

- Since $\gamma A = \gamma$ we have $\sum_{i=1}^n \gamma_i a_{ij} = \gamma_j$ for all $j \leqslant n$.
- It follows that

$$\sum_{j=1}^n |\gamma_j| = \sum_{j=1}^n \Big| \sum_{i=1}^n \gamma_i a_{ij} \Big| \leqslant \sum_{j=1}^n \sum_{i=1}^n |\gamma_i| a_{ij} = \sum_{i=1}^n |\gamma_i| \sum_{j=1}^n a_{ij} = \sum_{i=1}^n |\gamma_i|.$$

- Since we have equality everywhere, all γ_j have the same sign. So, we may assume that $\gamma_j \geqslant 0$ for all $j \leqslant n$. We shall show that $\gamma_j > 0$ for all $j \leqslant n$.
- Suppose that $\gamma_j = 0$ for some j. Since $a_{ij} > 0$ if $i \neq j$, from the equation $\sum_{i=1}^n \gamma_i a_{ij} = \gamma_j$ we get $\gamma = 0$, a contradiction.

The case of ℓ_1^n

- Let $k \geqslant 2$ and for any partition $\Delta = \{\delta_1, \dots, \delta_k\}$ of $[n]$ define

$$f(\Delta) = \sum_{\ell=1}^{k} \sum_{i,j \in \delta_\ell} \gamma_i a_{ij}.$$

The case of ℓ_1^n

- Let $k \geqslant 2$ and for any partition $\Delta = \{\delta_1, \ldots, \delta_k\}$ of $[n]$ define

$$f(\Delta) = \sum_{\ell=1}^{k} \sum_{i,j \in \delta_\ell} \gamma_i a_{ij}.$$

- Choose a partition $\Sigma = \{\sigma_1, \ldots, \sigma_k\}$ that minimizes $f(\Delta)$ over all such partitions.

- Let $k \geqslant 2$ and for any partition $\Delta = \{\delta_1, \ldots, \delta_k\}$ of $[n]$ define

$$f(\Delta) = \sum_{\ell=1}^{k} \sum_{i,j \in \delta_\ell} \gamma_i a_{ij}.$$

- Choose a partition $\Sigma = \{\sigma_1, \ldots, \sigma_k\}$ that minimizes $f(\Delta)$ over all such partitions.

Claim 2

For all $\ell \leqslant k$ and $i \in \sigma_\ell$ we have $\sum_{j \in \sigma_\ell} a_{ij} \leqslant \frac{2}{k}$.

- Assuming the contrary, there exists $r \in \sigma_1$ such that $\theta := \sum_{j \in \sigma_1} a_{rj} > \frac{2}{k}$.

- Let $k \geqslant 2$ and for any partition $\Delta = \{\delta_1, \dots, \delta_k\}$ of $[n]$ define

$$f(\Delta) = \sum_{\ell=1}^{k} \sum_{i,j \in \delta_\ell} \gamma_i a_{ij}.$$

- Choose a partition $\Sigma = \{\sigma_1, \dots, \sigma_k\}$ that minimizes $f(\Delta)$ over all such partitions.

Claim 2

For all $\ell \leqslant k$ and $i \in \sigma_\ell$ we have $\sum_{j \in \sigma_\ell} a_{ij} \leqslant \frac{2}{k}$.

- Assuming the contrary, there exists $r \in \sigma_1$ such that $\theta := \sum_{j \in \sigma_1} a_{rj} > \frac{2}{k}$.
- We define $(k-1)$ new partitions $\Sigma^2, \dots, \Sigma^k$ as follows: for every $s = 2, \dots, k$ we define $\Sigma^s = \{\sigma_1^s, \dots, \sigma_k^s\}$ where

$$\sigma_1^s = \sigma_1 \setminus \{r\}, \;\; \sigma_s^s = \sigma_s \cup \{r\} \;\text{ and }\; \sigma_\ell^s = \sigma_\ell \;\text{ if }\; \ell \neq 1, s.$$

The case of ℓ_1^n

- Observe that

$$f(\Sigma) - f(\Sigma^s) = \sum_{i,j\in\sigma_1} \gamma_i a_{ij} + \sum_{i,j\in\sigma_s} \gamma_i a_{ij} - \sum_{i,j\in\sigma_1\setminus\{r\}} \gamma_i a_{ij} - \sum_{i,j\in\sigma_s\cup\{r\}} \gamma_i a_{ij}$$
$$= \gamma_r \sum_{j\in\sigma_1} a_{rj} + \sum_{i\in\sigma_1} \gamma_i a_{ir} - \gamma_r \sum_{j\in\sigma_s} a_{rj} - \sum_{i\in\sigma_s} \gamma_i a_{ir}.$$

The case of ℓ_1^n

- Observe that

$$f(\Sigma) - f(\Sigma^s) = \sum_{i,j\in\sigma_1} \gamma_i a_{ij} + \sum_{i,j\in\sigma_s} \gamma_i a_{ij} - \sum_{i,j\in\sigma_1\setminus\{r\}} \gamma_i a_{ij} - \sum_{i,j\in\sigma_s\cup\{r\}} \gamma_i a_{ij}$$
$$= \gamma_r \sum_{j\in\sigma_1} a_{rj} + \sum_{i\in\sigma_1} \gamma_i a_{ir} - \gamma_r \sum_{j\in\sigma_s} a_{rj} - \sum_{i\in\sigma_s} \gamma_i a_{ir}.$$

- Adding these equations we see that

$$\sum_{s=2}^{k} \left(f(\Sigma) - f(\Sigma^s)\right) = (k-1)\gamma_r \sum_{j\in\sigma_1} a_{rj} + (k-1) \sum_{i\in\sigma_1} \gamma_i a_{ir} - \gamma_r \sum_{j\notin\sigma_1} a_{rj} - \sum_{i\notin\sigma_1} \gamma_i a_{ir}$$
$$\geqslant (k-1)\gamma_r \sum_{j\in\sigma_1} a_{rj} - \gamma_r \sum_{j\notin\sigma_1} a_{rj} - \sum_{i=1}^{n} \gamma_i a_{ir}$$
$$= (k-1)\gamma_r\theta - \gamma_r(1-\theta) - \gamma_r$$
$$= \gamma_r(k\theta - 2) > 0.$$

The case of ℓ_1^n

- Observe that

$$\begin{aligned} f(\Sigma) - f(\Sigma^s) &= \sum_{i,j\in\sigma_1} \gamma_i a_{ij} + \sum_{i,j\in\sigma_s} \gamma_i a_{ij} - \sum_{i,j\in\sigma_1\setminus\{r\}} \gamma_i a_{ij} - \sum_{i,j\in\sigma_s\cup\{r\}} \gamma_i a_{ij} \\ &= \gamma_r \sum_{j\in\sigma_1} a_{rj} + \sum_{i\in\sigma_1} \gamma_i a_{ir} - \gamma_r \sum_{j\in\sigma_s} a_{rj} - \sum_{i\in\sigma_s} \gamma_i a_{ir}. \end{aligned}$$

- Adding these equations we see that

$$\begin{aligned} \sum_{s=2}^{k} \left(f(\Sigma) - f(\Sigma^s)\right) &= (k-1)\gamma_r \sum_{j\in\sigma_1} a_{rj} + (k-1) \sum_{i\in\sigma_1} \gamma_i a_{ir} - \gamma_r \sum_{j\notin\sigma_1} a_{rj} - \sum_{i\notin\sigma_1} \gamma_i a_{ir} \\ &\geqslant (k-1)\gamma_r \sum_{j\in\sigma_1} a_{rj} - \gamma_r \sum_{j\notin\sigma_1} a_{rj} - \sum_{i=1}^{n} \gamma_i a_{ir} \\ &= (k-1)\gamma_r\theta - \gamma_r(1-\theta) - \gamma_r \\ &= \gamma_r(k\theta - 2) > 0. \end{aligned}$$

- Therefore, for some $s \in \{2, \ldots, k\}$ we must have $f(\Sigma) > f(\Sigma^s)$, a contradiction.

Random selection

Let $A : \ell_2^n \to \ell_2^n$ such that $|Ae_i| = 1$ for all $i \leqslant n$. There exists $\sigma_1 \subset [n]$ with $|\sigma_1| \geqslant c_1 n / \|A\|^2$ such that, for all $i \in \sigma_1$,

$$\left|P_{\langle Ae_j : j \in \sigma_1 \setminus \{i\} \rangle}(Ae_i)\right| < \frac{1}{\sqrt{2}}.$$

- We fix $\delta \in (0,1)$ and consider independent 0-1 random variables $\xi_1, \ldots, \xi_n$ with $\mathbb{E}(\xi_i) = \delta$ on a probability space (Ω, Σ, μ).

The Bourgain-Tzafriri argument

Random selection

Let $A : \ell_2^n \to \ell_2^n$ such that $|Ae_i| = 1$ for all $i \leqslant n$. There exists $\sigma_1 \subset [n]$ with $|\sigma_1| \geqslant c_1 n/\|A\|^2$ such that, for all $i \in \sigma_1$,

$$\left|P_{\langle Ae_j : j \in \sigma_1 \setminus \{i\}\rangle}(Ae_i)\right| < \frac{1}{\sqrt{2}}.$$

- We fix $\delta \in (0,1)$ and consider independent 0-1 random variables $\xi_1, \ldots, \xi_n$ with $\mathbb{E}(\xi_i) = \delta$ on a probability space (Ω, Σ, μ).
- For each $\omega \in \Omega$ we set $\sigma(\omega) = \{i \leqslant n : \xi_i(\omega) = 1\}$

The Bourgain-Tzafriri argument

Random selection

Let $A : \ell_2^n \to \ell_2^n$ such that $|Ae_i| = 1$ for all $i \leqslant n$. There exists $\sigma_1 \subset [n]$ with $|\sigma_1| \geqslant c_1 n/\|A\|^2$ such that, for all $i \in \sigma_1$,

$$\left|P_{\langle Ae_j : j \in \sigma_1 \setminus \{i\}\rangle}(Ae_i)\right| < \frac{1}{\sqrt{2}}.$$

- We fix $\delta \in (0,1)$ and consider independent 0-1 random variables $\xi_1, \ldots, \xi_n$ with $\mathbb{E}(\xi_i) = \delta$ on a probability space (Ω, Σ, μ).
- For each $\omega \in \Omega$ we set $\sigma(\omega) = \{i \leqslant n : \xi_i(\omega) = 1\}$
- Note that $|\sigma(\omega)| = \sum_{i=1}^n \xi_i(\omega)$.

Random selection

Let $A : \ell_2^n \to \ell_2^n$ such that $|Ae_i| = 1$ for all $i \leqslant n$. There exists $\sigma_1 \subset [n]$ with $|\sigma_1| \geqslant c_1 n/\|A\|^2$ such that, for all $i \in \sigma_1$,

$$\left|P_{\langle Ae_j : j \in \sigma_1 \setminus \{i\}\rangle}(Ae_i)\right| < \frac{1}{\sqrt{2}}.$$

- We fix $\delta \in (0,1)$ and consider independent 0-1 random variables $\xi_1, \ldots, \xi_n$ with $\mathbb{E}(\xi_i) = \delta$ on a probability space (Ω, Σ, μ).
- For each $\omega \in \Omega$ we set $\sigma(\omega) = \{i \leqslant n : \xi_i(\omega) = 1\}$
- Note that $|\sigma(\omega)| = \sum_{i=1}^n \xi_i(\omega)$.
- From Bernstein's inequality we have

$$\mu\Big(\Big\{\omega : |\sigma(\omega)| < \frac{\delta n}{2}\Big\}\Big) \leqslant \exp(-\delta n/10).$$

The Bourgain-Tzafriri argument

- By the independence of the ξ_i's we get

$$\begin{aligned}
\mathbb{E}\Big(\sum_{i=1}^{n}\xi_i(\omega)\big|P_{\langle\xi_j(\omega)Ae_j:j\neq i\rangle}(Ae_i)\big|^2\Big) &= \sum_{i=1}^{n}(\mathbb{E}\xi_i(\omega))\,\mathbb{E}\big|P_{\langle\xi_j(\omega)Ae_j:j\neq i\rangle}(Ae_i)\big|^2 \\
&= \delta\,\mathbb{E}\Big(\sum_{i=1}^{n}\big|P_{\langle\xi_j(\omega)Ae_j:j\neq i\rangle}(Ae_i)\big|^2\Big) \leqslant \delta\,\mathbb{E}\Big(\sum_{i=1}^{n}\big|P_{\langle\xi_j(\omega)Ae_j:j\leqslant n\rangle}(Ae_i)\big|^2\Big) \\
&= \delta\,\mathbb{E}\big\|P_{\langle\xi_j(\omega)Ae_j:j\leqslant n\rangle}A\big\|_{HS}^2 \leqslant \delta\|A\|^2\,\mathbb{E}\big\|P_{\langle\xi_j(\omega)Ae_j:j\leqslant n\rangle}\big\|_{HS}^2 \\
&= \delta\|A\|^2\,\mathbb{E}\,[\dim(\langle\xi_j(\omega)Ae_j : j\leqslant n\rangle)] \leqslant \delta\|A\|^2\,\mathbb{E}\Big(\sum_{j=1}^{n}\xi_i(\omega)\Big) \\
&= \delta^2 n\|A\|^2.
\end{aligned}$$

The Bourgain-Tzafriri argument

- By the independence of the ξ_i's we get

$$\mathbb{E}\Big(\sum_{i=1}^{n}\xi_i(\omega)\Big|P_{\langle\xi_j(\omega)Ae_j:j\neq i\rangle}(Ae_i)\Big|^2\Big)=\sum_{i=1}^{n}(\mathbb{E}\xi_i(\omega))\;\mathbb{E}\Big|P_{\langle\xi_j(\omega)Ae_j:j\neq i\rangle}(Ae_i)\Big|^2$$

$$=\delta\;\mathbb{E}\Big(\sum_{i=1}^{n}\Big|P_{\langle\xi_j(\omega)Ae_j:j\neq i\rangle}(Ae_i)\Big|^2\Big)\leqslant\delta\;\mathbb{E}\Big(\sum_{i=1}^{n}\Big|P_{\langle\xi_j(\omega)Ae_j:j\leqslant n\rangle}(Ae_i)\Big|^2\Big)$$

$$=\delta\;\mathbb{E}\Big\|P_{\langle\xi_j(\omega)Ae_j:j\leqslant n\rangle}A\Big\|_{HS}^2\leqslant\delta\|A\|^2\;\mathbb{E}\Big\|P_{\langle\xi_j(\omega)Ae_j:j\leqslant n\rangle}\Big\|_{HS}^2$$

$$=\delta\|A\|^2\;\mathbb{E}\;[\dim(\langle\xi_j(\omega)Ae_j:j\leqslant n\rangle)]\leqslant\delta\|A\|^2\;\mathbb{E}\Big(\sum_{j=1}^{n}\xi_i(\omega)\Big)$$

$$=\delta^2 n\|A\|^2.$$

- From Markov's inequality, with probability greater than or equal to $1/2$ we get

$$\sum_{i=1}^{n}\xi_i(\omega)\Big|P_{\langle\xi_j(\omega)Ae_j:j\neq i\rangle}(Ae_i)\Big|^2\leqslant 2\delta^2 n\|A\|^2.$$

The Bourgain-Tzafriri argument

- So, we may find $\omega_0 \in \Omega$ such that $\sigma_0 := \sigma(\omega_0)$ satisfies $|\sigma_0| \geqslant \delta n/2$

$$\sum_{i\in\sigma_0} \left|P_{\langle Ae_j : j\in\sigma_0\setminus\{i\}\rangle}(Ae_i)\right|^2 \leqslant 2\delta^2 n\|A\|^2.$$

The Bourgain-Tzafriri argument

- So, we may find $\omega_0 \in \Omega$ such that $\sigma_0 := \sigma(\omega_0)$ satisfies $|\sigma_0| \geqslant \delta n/2$

$$\sum_{i\in\sigma_0} \left|P_{\langle Ae_j : j\in\sigma_0\setminus\{i\}\rangle}(Ae_i)\right|^2 \leqslant 2\delta^2 n\|A\|^2.$$

- We define

$$\tau := \{i \in \sigma_0 : |P_{\langle Ae_j : j\in\sigma_0\setminus\{i\}\rangle}(Ae_i)| > 4\|A\|\sqrt{\delta}\}.$$

The Bourgain-Tzafriri argument

- So, we may find $\omega_0 \in \Omega$ such that $\sigma_0 := \sigma(\omega_0)$ satisfies $|\sigma_0| \geqslant \delta n/2$

$$\sum_{i \in \sigma_0} \left| P_{\langle Ae_j : j \in \sigma_0 \setminus \{i\} \rangle}(Ae_i) \right|^2 \leqslant 2\delta^2 n \|A\|^2.$$

- We define

$$\tau := \{ i \in \sigma_0 : |P_{\langle Ae_j : j \in \sigma_0 \setminus \{i\} \rangle}(Ae_i)| > 4\|A\|\sqrt{\delta} \}.$$

- Applying Markov's inequality once again we get

$$|\tau|(16\delta \|A\|^2) \leqslant 2\delta^2 n \|A\|^2,$$

which gives

$$|\tau| \leqslant \delta n/8.$$

The Bourgain-Tzafriri argument

- So, we may find $\omega_0 \in \Omega$ such that $\sigma_0 := \sigma(\omega_0)$ satisfies $|\sigma_0| \geqslant \delta n/2$

$$\sum_{i\in\sigma_0} \left|P_{\langle Ae_j : j\in\sigma_0\setminus\{i\}\rangle}(Ae_i)\right|^2 \leqslant 2\delta^2 n\|A\|^2.$$

- We define

$$\tau := \{i \in \sigma_0 : |P_{\langle Ae_j : j\in\sigma_0\setminus\{i\}\rangle}(Ae_i)| > 4\|A\|\sqrt{\delta}\}.$$

- Applying Markov's inequality once again we get

$$|\tau|(16\delta\|A\|^2) \leqslant 2\delta^2 n\|A\|^2,$$

which gives

$$|\tau| \leqslant \delta n/8.$$

- If we set $\sigma_1 = \sigma_0 \setminus \tau$, then $|\sigma_1| \geqslant 3\delta n/8$ and, for all $i \in \sigma_1$,

$$|P_{\langle Ae_j : j\in\sigma_1\setminus\{i\}\rangle}(Ae_i)| \leqslant |P_{\langle Ae_j : j\in\sigma_0\setminus\{i\}\rangle}(Ae_i)| \leqslant 4\|A\|\sqrt{\delta}.$$

The Bourgain-Tzafriri argument

- So, we may find $\omega_0 \in \Omega$ such that $\sigma_0 := \sigma(\omega_0)$ satisfies $|\sigma_0| \geqslant \delta n/2$

$$\sum_{i\in\sigma_0} \left|P_{\langle Ae_j : j\in\sigma_0\setminus\{i\}\rangle}(Ae_i)\right|^2 \leqslant 2\delta^2 n\|A\|^2.$$

- We define

$$\tau := \{i \in \sigma_0 : |P_{\langle Ae_j : j\in\sigma_0\setminus\{i\}\rangle}(Ae_i)| > 4\|A\|\sqrt{\delta}\}.$$

- Applying Markov's inequality once again we get

$$|\tau|(16\delta\|A\|^2) \leqslant 2\delta^2 n\|A\|^2,$$

which gives

$$|\tau| \leqslant \delta n/8.$$

- If we set $\sigma_1 = \sigma_0 \setminus \tau$, then $|\sigma_1| \geqslant 3\delta n/8$ and, for all $i \in \sigma_1$,

$$|P_{\langle Ae_j : j\in\sigma_1\setminus\{i\}\rangle}(Ae_i)| \leqslant |P_{\langle Ae_j : j\in\sigma_0\setminus\{i\}\rangle}(Ae_i)| \leqslant 4\|A\|\sqrt{\delta}.$$

- Choosing $\delta = 1/(32\|A\|^2)$ we get the result.

Sauer-Shelah

Let D be a subset of $E_2^n = \{-1,1\}^n$ and $k \leqslant n$. If

$$|D| > \sum_{j=0}^{k-1} \binom{n}{j},$$

then there exists $\sigma \subset [n]$ with cardinality $|\sigma| = k$ such that $R_\sigma(D) = E_2^\sigma = \{-1,1\}^\sigma$, where R_σ is the restriction to the coordinates of σ.

The Bourgain-Tzafriri argument

Combinatorial step

Let $A : \ell_2^n \to \ell_2^n$ such that $|Ae_i| = 1$ for all $i \leqslant n$. There exists $\sigma_2 \subset [n]$ with $|\sigma_2| \geqslant c_2 n/\|A\|^2$ such that, for any choice of coefficients $(t_i)_{i\in\sigma_2}$,

$$\Big|\sum_{i\in\sigma_2} t_i A e_i\Big| \geqslant c_2 \frac{\sum_{i\in\sigma_2}|t_i|}{\sqrt{|\sigma_2|}}.$$

- We had found σ_1 with $|\sigma_1| \geqslant c_1 n/\|A\|^2$ such that, for all $i \in \sigma_1$,

$$\big|P_{\langle Ae_j : j\in\sigma_1\setminus\{i\}\rangle}(Ae_i)\big| < \frac{1}{\sqrt{2}}.$$

Combinatorial step

Let $A : \ell_2^n \to \ell_2^n$ such that $|Ae_i| = 1$ for all $i \leqslant n$. There exists $\sigma_2 \subset [n]$ with $|\sigma_2| \geqslant c_2 n/\|A\|^2$ such that, for any choice of coefficients $(t_i)_{i\in\sigma_2}$,

$$\Big|\sum_{i\in\sigma_2} t_i Ae_i\Big| \geqslant c_2 \frac{\sum_{i\in\sigma_2}|t_i|}{\sqrt{|\sigma_2|}}.$$

- We had found σ_1 with $|\sigma_1| \geqslant c_1 n/\|A\|^2$ such that, for all $i \in \sigma_1$,

$$\big|P_{\langle Ae_j : j\in\sigma_1\setminus\{i\}\rangle}(Ae_i)\big| < \frac{1}{\sqrt{2}}.$$

- For every $i \in \sigma_1$ we define $u_i' = Ae_i - P_{\langle Ae_j : j\in\sigma_1\setminus\{i\}\rangle}(Ae_i)$ and $u_i = u_i'/|u_i'|$. Then, $|u_i| = 1$ and, for all $i \in \sigma_1$,

$$\langle Ae_i, u_i\rangle \geqslant \frac{1}{\sqrt{2}} \ \text{ and } \ \langle Ae_i, u_j\rangle = 0 \quad j \neq i.$$

The Bourgain-Tzafriri argument

Combinatorial step

Let $A : \ell_2^n \to \ell_2^n$ such that $|Ae_i| = 1$ for all $i \leqslant n$. There exists $\sigma_2 \subset [n]$ with $|\sigma_2| \geqslant c_2 n/\|A\|^2$ such that, for any choice of coefficients $(t_i)_{i\in\sigma_2}$,

$$\Big|\sum_{i\in\sigma_2} t_i Ae_i\Big| \geqslant c_2 \frac{\sum_{i\in\sigma_2}|t_i|}{\sqrt{|\sigma_2|}}.$$

- We had found σ_1 with $|\sigma_1| \geqslant c_1 n/\|A\|^2$ such that, for all $i \in \sigma_1$,

$$\big|P_{\langle Ae_j : j\in\sigma_1\setminus\{i\}\rangle}(Ae_i)\big| < \frac{1}{\sqrt{2}}.$$

- For every $i \in \sigma_1$ we define $u_i' = Ae_i - P_{\langle Ae_j : j\in\sigma_1\setminus\{i\}\rangle}(Ae_i)$ and $u_i = u_i'/|u_i'|$. Then, $|u_i| = 1$ and, for all $i \in \sigma_1$,

$$\langle Ae_i, u_i\rangle \geqslant \frac{1}{\sqrt{2}} \ \text{ and } \ \langle Ae_i, u_j\rangle = 0 \quad j \neq i.$$

- By the parallelogram law,

$$\mathrm{Ave}\Big\{\Big|\sum_{j\in\sigma_1}\varepsilon_j u_j\Big|^2 : \varepsilon_j = \pm 1\Big\} = \sum_{j\in\sigma_1}|u_j|^2 = |\sigma_1|.$$

The Bourgain-Tzafriri argument

- Consider the set

$$D := \Big\{ (\varepsilon_j)_{j\in\sigma_1} \in E_2^{\sigma_1} : \Big| \sum_{j\in\sigma_1} \varepsilon_j u_j \Big| \leqslant \sqrt{2|\sigma_1|} \Big\}.$$

The Bourgain-Tzafriri argument

- Consider the set
$$D := \Big\{ (\varepsilon_j)_{j\in\sigma_1} \in E_2^{\sigma_1} : \Big| \sum_{j\in\sigma_1} \varepsilon_j u_j \Big| \leqslant \sqrt{2|\sigma_1|} \Big\}.$$
- From Markov's inequality we have $|D| \geqslant 2^{|\sigma_1|-1}$.

The Bourgain-Tzafriri argument

- Consider the set

$$D := \Big\{(\varepsilon_j)_{j\in\sigma_1} \in E_2^{\sigma_1} : \Big|\sum_{j\in\sigma_1} \varepsilon_j u_j\Big| \leqslant \sqrt{2|\sigma_1|}\Big\}.$$

- From Markov's inequality we have $|D| \geqslant 2^{|\sigma_1|-1}$.
- By the Sauer-Shelah lemma (with $k \geqslant |\sigma_1|/2$) we find $\sigma_2 \subset \sigma_1$ with $|\sigma_2| \geqslant |\sigma_1|/2 \geqslant (c_1/2)n/\|A\|^2$ such that

$$R_{\sigma_2}(\mathrm{conv}(D)) = \mathrm{conv}\Big(\{-1,1\}^{\sigma_2}\Big) = [-1,1]^{\sigma_2}.$$

The Bourgain-Tzafriri argument

- Consider the set

$$D := \left\{ (\varepsilon_j)_{j\in\sigma_1} \in E_2^{\sigma_1} : \left| \sum_{j\in\sigma_1} \varepsilon_j u_j \right| \leqslant \sqrt{2|\sigma_1|} \right\}.$$

- From Markov's inequality we have $|D| \geqslant 2^{|\sigma_1|-1}$.
- By the Sauer-Shelah lemma (with $k \geqslant |\sigma_1|/2$) we find $\sigma_2 \subset \sigma_1$ with $|\sigma_2| \geqslant |\sigma_1|/2 \geqslant (c_1/2)n/\|A\|^2$ such that

$$R_{\sigma_2}(\mathrm{conv}(D)) = \mathrm{conv}\left(\{-1,1\}^{\sigma_2}\right) = [-1,1]^{\sigma_2}.$$

- By the definition of D, if $(q_j)_{j\in\sigma_1} \in \mathrm{conv}(D)$ then

$$\left| \sum_{j\in\sigma_1} q_j u_j \right| \leqslant \sqrt{2|\sigma_1|} \leqslant 2\sqrt{|\sigma_2|}.$$

The Bourgain-Tzafriri argument

- Let $(t_i)_{i\in\sigma_2}$. We write

$$\frac{1}{\sqrt{2}}\sum_{i\in\sigma_2}|t_i| \leqslant \Big\langle \sum_{i\in\sigma_2} t_i A e_i, \sum_{j\in\sigma_2} \mathrm{sign}(t_j) u_j \Big\rangle.$$

The Bourgain-Tzafriri argument

- Let $(t_i)_{i\in\sigma_2}$. We write

$$\frac{1}{\sqrt{2}}\sum_{i\in\sigma_2}|t_i| \leqslant \Big\langle \sum_{i\in\sigma_2} t_i A e_i, \sum_{j\in\sigma_2} \mathrm{sign}(t_j) u_j \Big\rangle.$$

- Observe that

$$\big(\mathrm{sign}(t_j)\big)_{j\in\sigma_2} \in [-1,1]^{\sigma_2} \subset R_{\sigma_2}(\mathrm{conv}(D)).$$

The Bourgain-Tzafriri argument

- Let $(t_i)_{i\in\sigma_2}$. We write
$$\frac{1}{\sqrt{2}}\sum_{i\in\sigma_2}|t_i| \leqslant \Big\langle \sum_{i\in\sigma_2} t_i A e_i, \sum_{j\in\sigma_2}\mathrm{sign}(t_j)u_j\Big\rangle.$$
- Observe that
$$\big(\mathrm{sign}(t_j)\big)_{j\in\sigma_2} \in [-1,1]^{\sigma_2} \subset R_{\sigma_2}(\mathrm{conv}(D)).$$
- Therefore, we may find $(q_j)_{j\in\sigma_1} \in \mathrm{conv}(D)$ such that
$$q_j = \mathrm{sign}(t_j) \;\text{ if }\; j \in \sigma_2.$$

The Bourgain-Tzafriri argument

- Let $(t_i)_{i\in\sigma_2}$. We write
$$\frac{1}{\sqrt{2}}\sum_{i\in\sigma_2}|t_i| \leqslant \Big\langle \sum_{i\in\sigma_2} t_i Ae_i, \sum_{j\in\sigma_2}\mathrm{sign}(t_j)u_j\Big\rangle.$$
- Observe that
$$\big(\mathrm{sign}(t_j)\big)_{j\in\sigma_2} \in [-1,1]^{\sigma_2} \subset R_{\sigma_2}(\mathrm{conv}(D)).$$
- Therefore, we may find $(q_j)_{j\in\sigma_1} \in \mathrm{conv}(D)$ such that
$$q_j = \mathrm{sign}(t_j) \text{ if } j\in\sigma_2.$$
- Since $\langle Ae_i, u_j\rangle = 0$ whenever $i\in\sigma_2$ and $j\in\sigma_1\setminus\sigma_2$, we have
$$\Big\langle \sum_{i\in\sigma_2} t_i Ae_i, \sum_{j\in\sigma_2}\mathrm{sign}(t_j)u_j\Big\rangle = \Big\langle \sum_{i\in\sigma_2} t_i Ae_i, \sum_{j\in\sigma_1} q_j u_j\Big\rangle.$$

- Let $(t_i)_{i\in\sigma_2}$. We write

$$\frac{1}{\sqrt{2}}\sum_{i\in\sigma_2}|t_i| \leqslant \Big\langle \sum_{i\in\sigma_2} t_i Ae_i, \sum_{j\in\sigma_2}\mathrm{sign}(t_j)u_j\Big\rangle.$$

- Observe that

$$\big(\mathrm{sign}(t_j)\big)_{j\in\sigma_2} \in [-1,1]^{\sigma_2} \subset R_{\sigma_2}(\mathrm{conv}(D)).$$

- Therefore, we may find $(q_j)_{j\in\sigma_1} \in \mathrm{conv}(D)$ such that

$$q_j = \mathrm{sign}(t_j) \ \text{ if } \ j \in \sigma_2.$$

- Since $\langle Ae_i, u_j\rangle = 0$ whenever $i\in\sigma_2$ and $j\in\sigma_1\setminus\sigma_2$, we have

$$\Big\langle \sum_{i\in\sigma_2} t_i Ae_i, \sum_{j\in\sigma_2}\mathrm{sign}(t_j)u_j\Big\rangle = \Big\langle \sum_{i\in\sigma_2} t_i Ae_i, \sum_{j\in\sigma_1} q_j u_j\Big\rangle.$$

- From the Cauchy-Schwarz inequality,

$$\frac{1}{\sqrt{2}}\sum_{i\in\sigma_2}|t_i| \leqslant \Big\langle \sum_{i\in\sigma_2} t_i Ae_i, \sum_{j\in\sigma_1} q_j u_j\Big\rangle \leqslant \Big|\sum_{j\in\sigma_1} q_j u_j\Big|\Big|\sum_{i\in\sigma_2} t_j Ae_j\Big|$$
$$\leqslant 2\sqrt{|\sigma_2|}\Big|\sum_{i\in\sigma_2} t_j Ae_j\Big|.$$

The final step

- Let $A : \ell_2^n \to \ell_2^n$ with $|Ae_i| = 1$ for all $i = 1, \ldots, n$.

The final step

- Let $A : \ell_2^n \to \ell_2^n$ with $|Ae_i| = 1$ for all $i = 1, \dots, n$.
- We set $x_i = Ae_i$ and, by the previous step, we have $\sigma_2 \subset [n]$ such that $|\sigma_2| \geqslant c_2 n/\|A\|^2$ and for the operator $S : X = \langle x_i : i \in \sigma_2 \rangle \to \ell_1^n$ defined by

$$Sx_i = \frac{e_i}{\sqrt{|\sigma_2|}}, \qquad i \in \sigma_2.$$

we have that $\|S\| \leqslant 1/c_2$.

The final step

- Let $A : \ell_2^n \to \ell_2^n$ with $|Ae_i| = 1$ for all $i = 1, \ldots, n$.
- We set $x_i = Ae_i$ and, by the previous step, we have $\sigma_2 \subset [n]$ such that $|\sigma_2| \geqslant c_2 n/\|A\|^2$ and for the operator $S : X = \langle x_i : i \in \sigma_2 \rangle \to \ell_1^n$ defined by

$$Sx_i = \frac{e_i}{\sqrt{|\sigma_2|}}, \qquad i \in \sigma_2.$$

  we have that $\|S\| \leqslant 1/c_2$.
- The adjoint operator $S^* : \ell_\infty^n \to X$ is 2-summing and

$$\pi_2(S^*) \leqslant K_G \|S^*\| \leqslant \frac{K_G}{c_2}.$$

The final step

- Let $A : \ell_2^n \to \ell_2^n$ with $|Ae_i| = 1$ for all $i = 1, \dots, n$.
- We set $x_i = Ae_i$ and, by the previous step, we have $\sigma_2 \subset [n]$ such that $|\sigma_2| \geqslant c_2 n/\|A\|^2$ and for the operator $S : X = \langle x_i : i \in \sigma_2 \rangle \to \ell_1^n$ defined by

$$Sx_i = \frac{e_i}{\sqrt{|\sigma_2|}}, \qquad i \in \sigma_2.$$

  we have that $\|S\| \leqslant 1/c_2$.
- The adjoint operator $S^* : \ell_\infty^n \to X$ is 2-summing and

$$\pi_2(S^*) \leqslant K_G \|S^*\| \leqslant \frac{K_G}{c_2}.$$

- Then, S^* is factorized as $S^* = U \circ D$, where $U : \ell_2^n \to X$ with $\|U\| \leqslant \pi_2(S^*)$ and $D : \ell_\infty^n \to \ell_2^n$ is a diagonal operator with $De_i = \lambda_i e_i$ for some $\lambda_i \in \mathbb{R}$ with $\sum_{i=1}^n \lambda_i^2 \leqslant 1$.

The final step

- Let $A : \ell_2^n \to \ell_2^n$ with $|Ae_i| = 1$ for all $i = 1, \ldots, n$.
- We set $x_i = Ae_i$ and, by the previous step, we have $\sigma_2 \subset [n]$ such that $|\sigma_2| \geqslant c_2 n/\|A\|^2$ and for the operator $S : X = \langle x_i : i \in \sigma_2 \rangle \to \ell_1^n$ defined by

$$Sx_i = \frac{e_i}{\sqrt{|\sigma_2|}}, \qquad i \in \sigma_2.$$

  we have that $\|S\| \leqslant 1/c_2$.
- The adjoint operator $S^* : \ell_\infty^n \to X$ is 2-summing and

$$\pi_2(S^*) \leqslant K_G \|S^*\| \leqslant \frac{K_G}{c_2}.$$

- Then, S^* is factorized as $S^* = U \circ D$, where $U : \ell_2^n \to X$ with $\|U\| \leqslant \pi_2(S^*)$ and $D : \ell_\infty^n \to \ell_2^n$ is a diagonal operator with $De_i = \lambda_i e_i$ for some $\lambda_i \in \mathbb{R}$ with $\sum_{i=1}^n \lambda_i^2 \leqslant 1$.
- It follows that S can be written as $S = D^* \circ U^*$, with $D^* e_i = \lambda_i e_i$. Note that

$$U^*(x_j) = \frac{1}{\lambda_j \sqrt{|\sigma_2|}} e_j, \qquad j \in \sigma_2.$$

The final step

- We define
$$\sigma = \{j \in \sigma_2 : |\lambda_j| \leqslant \sqrt{2/|\sigma_2|}\}.$$

The final step

- We define

$$\sigma = \{j \in \sigma_2 : |\lambda_j| \leqslant \sqrt{2/|\sigma_2|}\}.$$

- From $\sum_{i=1}^{n} \lambda_i^2 \leqslant 1$ we see that

$$1 \geqslant \sum_{j \in \sigma_2 \setminus \sigma} \lambda_j^2 \geqslant 2|\sigma_2 \setminus \sigma|/|\sigma_2|.$$

The final step

- We define
$$\sigma = \{j \in \sigma_2 : |\lambda_j| \leqslant \sqrt{2/|\sigma_2|}\}.$$
- From $\sum_{i=1}^{n} \lambda_i^2 \leqslant 1$ we see that
$$1 \geqslant \sum_{j \in \sigma_2 \setminus \sigma} \lambda_j^2 \geqslant 2|\sigma_2 \setminus \sigma|/|\sigma_2|.$$
- Therefore,
$$|\sigma| \geqslant \frac{|\sigma_2|}{2} \geqslant c_3 n/\|A\|^2.$$

The final step

- We define
$$\sigma = \{j \in \sigma_2 : |\lambda_j| \leqslant \sqrt{2/|\sigma_2|}\}.$$
- From $\sum_{i=1}^n \lambda_i^2 \leqslant 1$ we see that
$$1 \geqslant \sum_{j\in\sigma_2\setminus\sigma} \lambda_j^2 \geqslant 2|\sigma_2 \setminus \sigma|/|\sigma_2|.$$
- Therefore,
$$|\sigma| \geqslant \frac{|\sigma_2|}{2} \geqslant c_3 n/\|A\|^2.$$
- Finally, for any choice of real coefficients $(t_i)_{i\in\sigma}$ we have
$$\frac{K_G}{c_2}\Big|\sum_{j\in\sigma} t_j x_j\Big| \geqslant \Big|U^*\Big(\sum_{j\in\sigma} t_j x_j\Big)\Big| = \Big(\sum_{j\in\sigma}\Big|\frac{t_j}{\lambda_j\sqrt{|\sigma_2|}}\Big|^2\Big)^{1/2}$$
$$\geqslant \frac{1}{\sqrt{2}}\Big(\sum_{j\in\sigma} t_j^2\Big)^{1/2}.$$

The final step

- We define
$$\sigma = \{j \in \sigma_2 : |\lambda_j| \leqslant \sqrt{2/|\sigma_2|}\}.$$
- From $\sum_{i=1}^n \lambda_i^2 \leqslant 1$ we see that
$$1 \geqslant \sum_{j \in \sigma_2 \setminus \sigma} \lambda_j^2 \geqslant 2|\sigma_2 \setminus \sigma|/|\sigma_2|.$$
- Therefore,
$$|\sigma| \geqslant \frac{|\sigma_2|}{2} \geqslant c_3 n/\|A\|^2.$$
- Finally, for any choice of real coefficients $(t_i)_{i\in\sigma}$ we have
$$\frac{K_G}{c_2}\Big|\sum_{j\in\sigma} t_j x_j\Big| \geqslant \Big|U^*\Big(\sum_{j\in\sigma} t_j x_j\Big)\Big| = \Big(\sum_{j\in\sigma}\Big|\frac{t_j}{\lambda_j\sqrt{|\sigma_2|}}\Big|^2\Big)^{1/2} \geqslant \frac{1}{\sqrt{2}}\Big(\sum_{j\in\sigma} t_j^2\Big)^{1/2}.$$
- This proves the theorem.