

PISIER'S K -CONVEXITY INEQUALITY: AN EXPOSITION

NILAVA METYA

We aim to show the bound $K(X) \lesssim \log(1 + d_{BM}(X, \ell_2^m))$ on the K -convexity constant $K(X)$ of an m -dimensional Banach space $(X, \|\cdot\|_X)$.

1. PRELIMINARIES

Rademacher Projection. Consider the cube $E^n = \{\pm 1\}^n$ with the uniform measure μ . Then $L_2(E^n, \mu)$ is a Hilbert space whose orthonormal basis is the set $\{w_A \mid A \subseteq [n]\}$ of Walsh functions as described next. The Rademacher functions $r_i : E^n \rightarrow \{\pm 1\}$ are given by $r_i(\boldsymbol{\varepsilon}) = r_i(\varepsilon_1, \dots, \varepsilon_n) := \varepsilon_i$. For any subset $A \subseteq [n]$ define the Walsh function $w_A : E^n \rightarrow \{\pm 1\}$ as $w_A(\boldsymbol{\varepsilon}) := \prod_{i \in A} r_i(\boldsymbol{\varepsilon})$ with the agreement that $w_\emptyset \equiv 1$. Any function $f : E^n \rightarrow X$ (with $(X, \|\cdot\|_X)$ a Banach space) can be written as $f(\mathbf{x}) = \sum_{A \subseteq [n]} \hat{f}(A) w_A(\mathbf{x}) \forall \mathbf{x} \in E^n$

where the Fourier coefficients $\hat{f}(A) \in X$ are determined by $\hat{f}(A) = \sum_{\boldsymbol{\varepsilon} \in E^n} w_A(\boldsymbol{\varepsilon}) f(\boldsymbol{\varepsilon}) \mu(\boldsymbol{\varepsilon}) = \frac{1}{2^n} \sum_{\boldsymbol{\varepsilon} \in E^n} w_A(\boldsymbol{\varepsilon}) f(\boldsymbol{\varepsilon})$. The Rademacher projection is the operator $Rad_n : L_2(E^n; X) \rightarrow L_2(E^n; X)$

defined by $Rad_n(f) := \left(\mathbf{x} \mapsto \sum_{i \in [n]} \hat{f}(\{i\}) w_{\{i\}}(\mathbf{x}) = \sum_{i \in [n]} \hat{f}(\{i\}) x_i \right)$. Note that $L_2(E^n; X)$ is

the space of all X -valued functions on E^n with the norm $\|f\|_{L_2(E^n; X)} = \sqrt{\frac{1}{2^n} \sum_{\mathbf{x} \in E^n} \|f(\mathbf{x})\|_X^2}$. The K -convexity constant of X is

$$K(X) = \sup_n \|Rad_n\|_{L_2(E^n; X) \rightarrow L_2(E^n; X)}$$

and we say X is K -convex if $K(X) < \infty$.

Convolutions. We will consider E^n as a group, for Fourier analysis, with coordinate-wise multiplication. That is, if $\mathbf{x} = (x_1, \dots, x_n), \mathbf{y} = (y_1, \dots, y_n) \in E^n$ then $\mathbf{xy} = (x_1 y_1, \dots, x_n y_n)$. The identity is simply $\mathbf{1}$, the all-ones vector, and the inverse of $\mathbf{x} \in E^n$ is itself. Convolution is a powerful tool when it comes to Fourier analysis. Fix a Banach space $(X, \|\cdot\|)$. The convolution of $f : E^n \rightarrow X$ and $h : E^n \rightarrow \mathbb{R}$ is given by

$$f * h(\mathbf{x}) := \mathbb{E}_{\boldsymbol{\varepsilon}} [f(\mathbf{x}\boldsymbol{\varepsilon}) h(\boldsymbol{\varepsilon})] = \frac{1}{2^n} \sum_{\boldsymbol{\varepsilon} \in E^n} f(\mathbf{x}\boldsymbol{\varepsilon}) h(\boldsymbol{\varepsilon}).$$

Note that the expectation is taken with respect to the uniform measure on the boolean cube. The fact that $\mathbf{x}\varepsilon \stackrel{D}{=} \varepsilon$, for a fixed $\mathbf{x} \in E^n$, allows us to write $f * h(\mathbf{x}) = \mathbb{E}_\varepsilon [f(\varepsilon)h(\mathbf{x}\varepsilon)]$. The intuition that Rad_n is the ‘linear’ part is captured by the fact that $\text{Rad}_n f = f * \ell$ where $\ell := \sum_{i=1}^n w_{\{i\}}$. Indeed, $f * \ell(\mathbf{x}) = \sum_{i \in [n]} \underbrace{\sum_{\varepsilon \in E^n} \frac{f(\varepsilon)\varepsilon_i}{2^n}}_{\hat{f}(\{i\})} x_i = \sum_{i \in [n]} \hat{f}(\{i\})x_i = \text{Rad}_n(f)(\mathbf{x})$.

An interesting way that Fourier analysis interacts with convolutions is that it is the natural way for preserving multiplicative structure, in the sense that $f * g(\mathbf{x}) = \sum_{A \subseteq [n]} \hat{f}(A)\hat{g}(A)w_A(\mathbf{x})$.

In other words, $\widehat{f * g}(A) = \hat{f}(A)\hat{g}(A)$. This is due to the orthonormality of the Walsch characters. We will also use the following fact about ‘sub-multiplicativity of norms’ under convolutions, which has been proved in Appendix A.1.

Lemma 1. *Let $f : E^n \rightarrow X, h : E^n \rightarrow \mathbb{R}$. Then*

$$\|f * h\|_{L_2(E^n; X)} \leq \|f\|_{L_2(E^n; X)} \|h\|_{L_1(E^n)}.$$

A highly nonlinear approximation to ℓ . Recall the ‘almost moving delta mass’ with

real parameter t : $\alpha(\mathbf{x}) = \alpha(t, \mathbf{x}) := \prod_{i=1}^n (1 + tx_i) = \sum_{A \subseteq [n]} t^{|A|} w_A(\mathbf{x})$. The large powers of t

works in our favour to help ignore the terms larger than 1. However, we will play around a bit with α to get precisely what we need. Consider the function $\varphi_r(\theta) := \frac{2r-1}{r} \frac{\sin(r\theta)}{\sin^2 \theta}$ on $[0, 2\pi]$. Consider $\Gamma_r := \{\frac{k \cdot 2\pi}{4r} \mid k \in [0, 4r-1] \cap \mathbb{Z}\}$ and $\Delta_r = \Gamma_r \setminus \{0, \pi\}$ each equipped with the uniform measure, where r is an odd integer parameter to be decided later. As well as the function $g_r(\mathbf{x}) = \mathbb{E}_{\theta \sim \Delta_r} [\varphi_r(\theta) \alpha(\frac{\sin \theta}{2}, \mathbf{x})]$. Expanding we get

$$g_r(\mathbf{x}) = 2\mathbb{E}_{\theta \sim \Delta_r} \left[\varphi_r(\theta) \alpha\left(\frac{\sin \theta}{2}, \mathbf{x}\right) \right] = 2 \sum_{A \subseteq [n]} \mathbb{E}_{\theta \sim \Delta_r} \left[\varphi_r(\theta) \left(\frac{\sin \theta}{2}\right)^{|A|} \right] w_A(\mathbf{x}).$$

So $\hat{g}_r(A) = 2\mathbb{E}_{\theta \sim \Delta_r} \left[\varphi_r(\theta) \left(\frac{\sin \theta}{2}\right)^{|A|} \right]$.

The following facts, proven in Appendix A.2, help showing that g linearly approximates g_1 :

Proposition 2. (a) $\mathbb{E}_{\theta \sim \Delta_r} [\varphi_r(\theta) \sin^k \theta] = \delta_{k,1}$ for $0 \leq k \leq r$.

(b) $\mathbb{E}_{\theta \sim \Delta_r} [|\varphi_r(\theta)|] \leq 4r$.

Corollary 3. $\hat{g}_r(A) = \delta_{k,1}$ if $0 \leq |A| \leq r$ and $|\hat{g}_r(A)| \leq \frac{4r}{2^r}$ if $|A| > r$.

Proof. The equality case directly follows from Proposition 2(a).

Now say $|A| > r$. Then $|\hat{g}_r(A)| \leq \frac{2}{2^{|A|}} \mathbb{E} [|\varphi_r(\theta)|] \leq \frac{2}{2^{r+1}} \cdot 4r = \frac{4r}{2^r}$ by Proposition 2(b). ■

Corollary 4. $\|g_r\|_{L_1(E^n)} = \mathbb{E}_{\boldsymbol{\varepsilon}} [|g_r(\boldsymbol{\varepsilon})|] \leq 8r$.

Proof.

$$\begin{aligned}
\frac{1}{2} \mathbb{E}_{\boldsymbol{\varepsilon}} [|g_r(\boldsymbol{\varepsilon})|] &\leq \mathbb{E}_{\boldsymbol{\varepsilon}} \mathbb{E}_{\theta \sim \Delta_r} \left[\left| \varphi_r(\theta) \prod_{i=1}^n \left(1 + \frac{\sin \theta}{2} \varepsilon_i \right) \right| \right] \\
&= \mathbb{E}_{\theta \sim \Delta_r} \mathbb{E}_{\boldsymbol{\varepsilon}} \left[|\varphi_r(\theta)| \prod_{i=1}^n \left(1 + \frac{\sin \theta}{2} \varepsilon_i \right) \right] && \because 1 + \frac{\sin \theta}{2} \varepsilon_i > 0 \\
&= \mathbb{E}_{\theta \sim \Delta_r} \left[|\varphi_r(\theta)| \prod_{i=1}^n \mathbb{E}_{\boldsymbol{\varepsilon}} \left(1 + \frac{\sin \theta}{2} \varepsilon_i \right) \right] \\
&= \mathbb{E}_{\theta \sim \Delta_r} [|\varphi_r(\theta)|] \stackrel{\text{Proposition 2(b)}}{\leq} 4r.
\end{aligned}$$

■

Banach-Mazur distance.

Definition 5. If E, F are isomorphic Banach spaces, their Banach-Mazur distance is

$$d(E, F) = d_{BM}(E, F) := \inf \left\{ \|T\| \cdot \|T^{-1}\| \mid T : E \xrightarrow{\sim} F \right\}.$$

For finite dimensional Banach spaces X , we want to look at their distances¹ from a Hilbert space H of the same dimension which will be used for bounding norms of convolutions. We will be specially interested in $H = \ell_2^m$ where $m := \dim X$.

Let $f : E^n \rightarrow X, h : E^n \rightarrow \mathbb{R}$. Consider the Fourier representation $h = \sum_A \hat{h}(A) w_A$. Fix any Hilbert space H isomorphic to X , and let $D := d_{BM}(X, H)$. For any $\varepsilon > 0, \exists T = T_\varepsilon : X \xrightarrow{\sim} H$ such that $\|T\| \|T^{-1}\| \leq (1 + \varepsilon)D$. Note that $f * h = \sum_A \hat{f}(A) \hat{h}(A) w_A$ with $\hat{f}(A) \in X, \hat{h}(A) \in \mathbb{R}$ and $w_A : E^n \rightarrow \{\pm 1\}$ are the Walsh functions, $\forall A \subseteq [n]$. So $T(f * h) = \sum_A T(\hat{f}_A) \hat{h}(A) w_A$. We thus have

$$\begin{aligned}
\|f * h\|_{L_2(E^n; X)} &= \|T^{-1} \circ T \circ (f * h)\|_{L_2(E^n; X)} \\
&\leq \|T^{-1}\| \|T(f * h)\|_{L_2(E^n; H)} \\
&= \|T^{-1}\| \sqrt{\sum_A \left\| \hat{h}(A) T(\hat{f}(A)) \right\|_H^2} \\
&= \|T^{-1}\| \sqrt{\sum_A |\hat{h}(A)| \left\| T(\hat{f}(A)) \right\|_H^2}
\end{aligned}$$

¹warning: d_{BM} is not an honest metric

$$\begin{aligned}
&\leq \|T^{-1}\| \left(\max_{A \subseteq [n]} |\hat{h}(A)| \right) \sqrt{\sum_A \|T(\hat{f}(A))\|_H^2} \\
&= \left(\max_{A \subseteq [n]} |\hat{h}(A)| \right) \|T^{-1}\| \|T(f)\|_{L_2(E_2^n; H)} \\
&\leq \left(\max_{A \subseteq [n]} |\hat{h}(A)| \right) \|T^{-1}\| \|T\| \|f\|_{L_2(E_2^n; X)} \\
&\leq (1 + \varepsilon) D \left(\max_{A \subseteq [n]} |\hat{h}(A)| \right) \|f\|_{L_2(E_2^n; X)}.
\end{aligned}$$

Since this is true $\forall \varepsilon > 0$, we conclude that

Lemma 6. *Let $f : E^n \rightarrow X, h : E^n \rightarrow \mathbb{R}$, where X is m -dimensional, and consider the Fourier representation $h = \sum_A \hat{h}(A) w_A$. Then*

$$\|f * h\|_{L_2(E^n; X)} \leq d(X, \ell_2^m) \left(\max_{A \subseteq [n]} |\hat{h}(A)| \right) \|f\|_{L_2(E_2^n; X)}.$$

2. THE FINAL PROOF

We have established $g_r = \sum_A \hat{g}(A) w_A = \sum_{i \in [n]} w_{\{i\}} + \sum_{|A| > r} \hat{g}(A) w_A$ with $|\hat{g}(A)| \leq \frac{4r}{2^r}$ (small) for $|A| > r$. Letting $\psi_r := \sum_{|A| > r} \hat{g}(A) w_A$, we have $g_r = \ell + \psi_r$. We note that $\hat{\psi}_r(A) = \begin{cases} 0 & \text{if } |A| \leq r \\ \hat{g}(A) & \text{if } r < |A| \leq n \end{cases}$. Thus, g_r and ℓ are ‘close’ functions. Now let’s start bounding $\|f * \ell\|_{L_2(E^n; X)}$. Let $D := d(X, \ell_2^{\dim X})$.

Firstly note that

$$\begin{aligned}
\|f * \ell\|_{L_2(E^n; X)} &= \|f * (g_r - \psi_r)\|_{L_2(E^n; X)} \\
&\leq \|f * g_r\|_{L_2(E^n; X)} + \|f * \psi_r\|_{L_2(E^n; X)} \\
&\leq \|f\|_{L_2(E^n; X)} \|g_r\|_{L_1(E^n)} + D \left(\max_{A \subseteq [n]} |\hat{\psi}_r(A)| \right) \|f\|_{L_2(E_2^n; X)} \\
&\quad \text{[Lemma 1]} \qquad \qquad \qquad \text{[Lemma 6]} \\
&= \left(\|g_r\|_{L_1(E^n)} + D \left(\max_{A \subseteq [n]} |\hat{\psi}_r(A)| \right) \right) \|f\|_{L_2(E_2^n; X)} \\
&\leq \left(\underset{\text{[Corollary 3]}}{8r} + D \cdot \underset{\text{[Corollary 3]}}{4r \cdot 2^{-r}} \right) \|f\|_{L_2(E_2^n; X)} \\
&= 8r (1 + D \cdot 2^{-(r+1)}) \|f\|_{L_2(E_2^n; X)}.
\end{aligned}$$

To complete the proof, choose r to be an odd number such that $2^{r-1} \leq D+1 \leq 2^r$ so that $8r(1 + D \cdot 2^{-(1+r)}) \leq 8(\lg(1+D) + 1) \left(1 + \frac{D}{2(D+1)}\right) \lesssim \lg(1+D)$ thus proving that $\|Rad_n\|_{L_2(E^n;X) \rightarrow L_2(E^n;X)} \lesssim \lg(1+D) \forall n$. This means $K(X) \lesssim \lg(1 + d(X, \ell_2^{\dim X}))$.

To get a more quantitative bound, we can use John's theorem which states that $d(X, \ell_2^{\dim X}) \leq \sqrt{\dim X}$, thus giving $K(X) \lesssim \log \dim X$.

APPENDIX A.

A.1. **Proof of Lemma 1.** We want to show $\|f * h\|_{L_2(E^n; X)} \leq \|f\|_{L_2(E^n; X)} \|h\|_{L_1(E^n)}$.

$$\begin{aligned}
\|f * h\|_{L_2(E^n; X)}^2 &= \int_{E^n} \|f * h(\mathbf{x})\|_X^2 d\mathbf{x} \\
&= \int_{E^n} \|\mathbb{E}_\epsilon [h(\epsilon)f(\mathbf{x}\epsilon)]\|_X^2 d\mathbf{x} \\
&= \int_{E^n} (\mathbb{E}_\epsilon [|h(\epsilon)| \|f(\mathbf{x}\epsilon)\|_X])^2 d\mathbf{x} \\
&= \int_{E^n} \left(\mathbb{E}_\epsilon \left[\sqrt{|h(\epsilon)|} \cdot \sqrt{|h(\epsilon)|} \|f(\mathbf{x}\epsilon)\|_X \right] \right)^2 d\mathbf{x} \\
&\stackrel{\text{Cauchy-Schwarz}}{\leq} \int_{E^n} \mathbb{E}_\epsilon [|h(\epsilon)|] \cdot \mathbb{E}_\epsilon [|h(\epsilon)| \|f(\mathbf{x}\epsilon)\|_X^2] d\mathbf{x} \\
&= \mathbb{E}_\epsilon [|h(\epsilon)|] \cdot \mathbb{E}_\epsilon \left[|h(\epsilon)| \int_{E^n} \|f(\mathbf{x}\epsilon)\|_X^2 d\mathbf{x} \right] \\
&= \mathbb{E}_\epsilon [|h(\epsilon)|] \cdot \mathbb{E}_\epsilon \left[|h(\epsilon)| \int_{E^n} \|f(\mathbf{x})\|_X^2 d\mathbf{x} \right] \\
&= (\mathbb{E}_\epsilon [|h(\epsilon)|])^2 \int_{E^n} \|f(\mathbf{x})\|_X^2 d\mathbf{x} \\
&= \|h\|_{L_1(E^n)}^2 \|f\|_{L_2(E^n; X)}^2.
\end{aligned}$$

■

A.2. **Proof of Proposition 2.**

(a) We first want to show $\mathbb{E}_{\theta \sim \Delta_r} [\varphi_r(\theta) \sin^k \theta] = \delta_{k,1}$ for $0 \leq k \leq r$.

- Say $k = 0$. $\varphi(\theta) = -\varphi(2\pi - \theta)$ and since the expectation is taken with respect to the uniform measure and for every $\theta \in \Delta_r$, we also have $2\pi - \theta \in \Delta_r$, the statement is true for $k = 0$.
- Say $k = 1$. We will show $\mathbb{E}_{\theta \sim \Delta_r} [\varphi_r(\theta) \sin \theta] = 1$ by inducting on odd r . For now, just consider the sum

$$\begin{aligned}
\sum_{\theta \in \Delta_r} \frac{\sin(r\theta)}{\sin \theta} &= \sum_{\theta \in \Delta_r} \frac{e^{ir\theta} - e^{-ir\theta}}{e^{i\theta} - e^{-i\theta}} = \sum_{\theta \in \Delta_r} \sum_{j=0}^{r-1} e^{ij\theta} \cdot e^{-i(r-1-j)\theta} \\
&= \sum_{\theta \in \Delta_r} \sum_{j=0}^{r-1} e^{i(2j-r+1)\theta} = \sum_{j=0}^{r-1} \sum_{\theta \in \Delta_r} e^{i(2j-r+1)\theta} \\
&= \sum_{j=0}^{r-1} \left(\sum_{\theta \in \Gamma_r} e^{i(2j-r+1)\theta} - 1 - \cos(\underbrace{(2j-r+1)\pi}_{\text{even}}) \right)
\end{aligned}$$

$$= \sum_{j=0}^{r-1} \left(\sum_{\theta \in \Gamma_r} e^{i(2j-r+1)\theta} - 2 \right) = \sum_{j=0}^{r-1} \sum_{\theta \in \Gamma_r} e^{i(2j-r+1)\theta} - 2r.$$

Recall that for any integer x , $\sum_{\theta \in \Gamma_r} e^{ix\theta} = 4r \cdot \mathbf{1}[x \equiv 0 \pmod{4r}]$ because Γ_r divides 2π into $4r$ equally spaced angles, whence the above sum becomes $4r - 2r = 2r$. Thus, $\mathbb{E}_{\theta \sim \Delta_r} [\varphi_r(\theta) \sin \theta] = \frac{1}{4r-2} \cdot \frac{2r-1}{r} \sum_{\theta \in \Delta_r} \frac{\sin(r\theta)}{\sin \theta} = 1$.

- For $k \geq 2$ we consider $q = k - 2 \in [0, r - 2] \cap \mathbb{Z}$ and are interested in $\mathbb{E}_{\theta \sim \Delta_r} [\varphi_r(\theta) \sin^k \theta] = \frac{1}{2r} \sum_{\theta \in \Delta_r} \sin(r\theta) \sin^q \theta = \frac{1}{2r} \sum_{\theta \in \Gamma_r} \sin(r\theta) \sin^q \theta$. Let's instead look at $\sum_{\theta \in \Gamma_r} \sin(r\theta) \sin^q \theta = \sum_{\theta \in \Gamma_r} \left(\frac{e^{ir\theta} - e^{-ir\theta}}{2i} \right) \left(\frac{e^{i\theta} - e^{-i\theta}}{2i} \right)^q$. For each $\theta \in \Gamma_r$, a typical term in the expansion of the power looks like (upto constants in \mathbb{C}) $e^{i\theta(2j-q)}$ for $j = 0, 1, \dots, q$. Combining with the first bracket gives a typical term like $e^{i\theta(2j-q \pm r)}$. $2j - q \pm r$ is always in $[-2r, 2r]$ so all terms on expansion, after $\sum_{\theta \in \Gamma_r}$ are 0, unless $2j - q \pm r = 0$ for some $0 \leq j \leq q \leq r - 2$. If $2j - q - r = 0$ then $0 \geq 2j - 2q = r + q > 0$, a contradiction. If $2j - q + r = 0$ then $0 \leq j = q - r < 0$, a contradiction again. Hence all terms have their exponent nonzero modulo $4r$, hence zero.

(b) Next we come to showing $\mathbb{E}_{\theta \sim \Delta_r} [|\varphi_r(\theta)|] \leq 4r$. Recall that $\varphi_r(\theta) = \frac{2r-1}{r} \frac{\sin(r\theta)}{\sin^2 \theta}$ where θ takes the values $\frac{2\pi k}{4r}$ for $k \in [4r - 1] \setminus \{2r\}$. By the symmetry of θ in all four coordinates to account for the fact that we only want to look at $|\sin^2 \theta|$, we have

$$\mathbb{E}_{\theta \sim \Delta_r} [|\varphi_r(\theta)|] \leq \frac{4}{4r-2} \cdot \frac{2r-1}{r} \sum_{k=1}^r \frac{1}{\sin^2 \left(\frac{2\pi k}{4r} \right)} \leq \frac{2}{r} \sum_{k=1}^r \frac{r^2}{j^2} \leq 2r \cdot \frac{\pi^2}{6} \leq 4r$$

where we used the inequality $\sin t \geq \frac{2}{\pi}t$ for $0 \leq t \leq \frac{\pi}{2}$.