On the discrepancy between norms on tensor products of normed spaces

Stanislaw Szarek

Case Western Reserve U. & Sorbonne U.

Collaborators: G. Aubrun, L. Lami, C. Palazuelos, A. Winter

Arxiv: 1809.10616

http://www.cwru.edu/artsci/math/szarek/

Abstract

The projective and injective norms are extreme ones among natural tensor products of normed spaces. An obvious question is: *How much do they differ?* This question was considered by Grothendieck and Pisier (in the 1950s and 1980s), but - surprisingly - no quantitative analysis of the finite-dimensional case was ever made. As explained in the talk of G. Aubrun, this last question comes up naturally in the context of generalized probabilistic theories (GPTs) and XOR games, where it can be restated as: *How powerful are global strategies compared to local ones?*

We will show that the discrepancy between the projective and injective norms on a tensor product of two finite-dimensional normed spaces E and F is always lower-bounded by the power of the (smaller) dimension, with the exponent depending on the generality of the setup (e.g., $E = F$ or $\dim E = \dim F$). Some of the results are essentially optimal, but other can be likely improved. The methods involve a wide range of techniques from geometry of Banach spaces and random matrices.

Joint work with G. Aubrun, L. Lami, C. Palazuelos, A. Winter.

Outline

• projective and injective tensor norms: definitions, notation

• historical background; the infinite dimensional case; qualitative vs. quantitative

• a selection of discrepancy results and examples of tools from geometric functional analysis

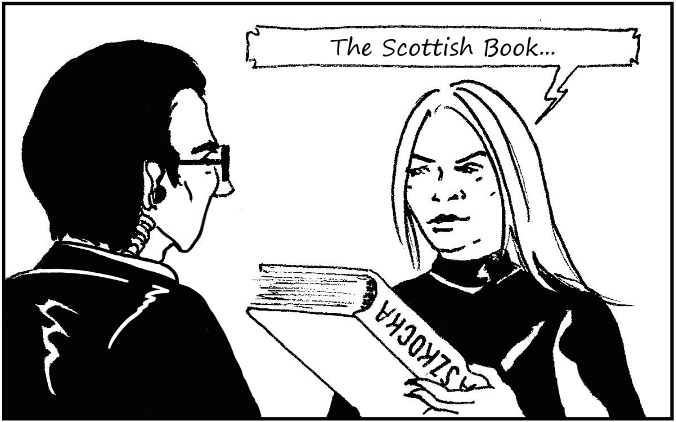
Buzzwords : Dvoretzky-Milman's theorem; p-summing norms; Chevet-Gordon's inequality; Grothendieck's inequality; K-convexity & the MM^*-estimate

Commercial break: *Alice and Bob Meet Banach*

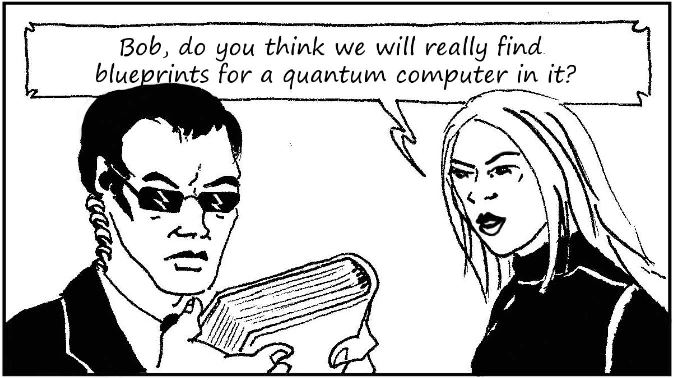
G. Aubrun and S. Szarek, *Alice and Bob Meet Banach. The interface between Asymptotic Geometric Analysis and Quantum Information Theory.* Mathematical Surveys and Monographs, American Mathematical Society, October 2017

And here is a comic strip (created by A. Garnier) that comes from the book, samples of which are available via the authors' web pages.

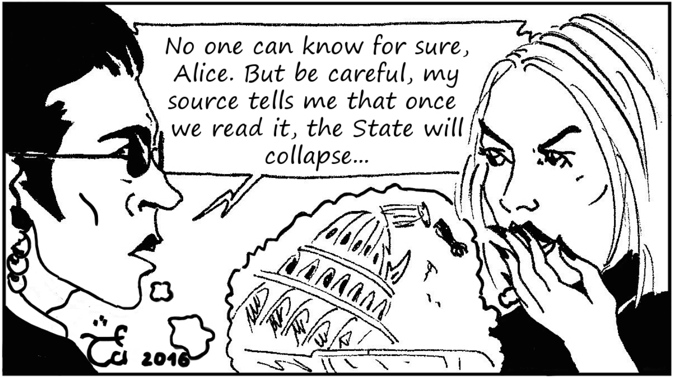
Alice and Bob Meet Banach (1)



Alice and Bob Meet Banach (2)



Alice and Bob Meet Banach (3)



Definitions and notation : the projective norm

If X, Y are real Banach spaces, we will consider norms on $X \otimes Y$ (the algebraic tensor product) verifying

$$\|x \otimes y\| = \|x\| \cdot \|y\|. \tag{1}$$

By the triangle inequality, every such norm must satisfy

$$\|z\| \leqslant \min\left\{\sum_i \|x_i\| \cdot \|y_i\| : \ z = \sum_i x_i \otimes y_i\right\} \tag{2}$$

and replacing "$\leqslant$" by "$:=$" in (2) we get the definition of the projective tensor norm $\|z\|_{X \otimes_\pi Y}$, the largest norm on $X \otimes Y$ verifying (1), also denoted sometimes by $\|z\|_{X \widehat{\otimes} Y}$. We will usually write simply $\|z\|_\pi$.

Definitions and notation : duality and the injective norm

For the smallest "reasonable" norm on $X \otimes Y$ it is most convenient to appeal to duality: if $x^* \in X^*, y^* \in Y^*$, we want $x^* \otimes y^*$ to induce a functional on $X \otimes Y$ whose norm is $\|x^*\| \cdot \|y^*\|$, which implies

$$\|z\| \geqslant \max\left\{(x^* \otimes y^*)(z) : \|x^*\| \leqslant 1,\ \|y^*\| \leqslant 1\right\}. \tag{3}$$

Again, replacing "$\geqslant$" by "$:=$" in (3) we get the definition of injective tensor norm $\|z\|_{X\otimes_\varepsilon Y}$ (or simply $\|z\|_\varepsilon$), denoted sometimes by $\|z\|_{X\check{\otimes} Y}$. Finally, observe that $\|z\|_\varepsilon$ is also the norm of z as a bilinear form on $X^* \times Y^*$.

An equivalent way to relate these two notions (at least in the finite dimensional case) is

$$X \otimes_\varepsilon Y = (X^* \otimes_\pi Y^*)^*.$$

If the spaces are infinite dimensional, completions are required and there are reflexivity issues, but we will largely ignore this side of the story and – unless explicitly stated otherwise – will assume that $\dim X, \dim Y < \infty$.

An equivalent language: tensor products of convex sets

In geometric functional analysis, we often identify norms on a finite dimensional vector space V with symmetric convex bodies:

$$X = (V, \|\cdot\|) \quad \to \quad B_X := \{x : \|x\| \leqslant 1\} \quad = \quad \text{the unit ball of } X \quad (4)$$

$$V \supset K \to \|x\|_K := \inf\{t \geqslant 0 : x \in tK\} = \text{ the Minkowski functional of } K$$

In this setting we define the projective tensor product as

$$K \otimes_\pi L := \mathrm{conv}\{x \otimes y : x \in K, y \in L\}$$

and the previous definitions can be restated as

$$B_{X \otimes_\pi Y} := B_X \otimes_\pi B_Y \quad \text{and} \quad B_{X \otimes_\varepsilon Y} := (B_{X^*} \otimes_\pi B_{Y^*})^\circ\,,$$

where $K^\circ := \{x \in V^* : \forall y \in K \ \langle y, x\rangle \leqslant 1\}$ is the polar of K.

Since the operations (4) are order reversing, the largest tensor norm corresponds to the smallest tensor product of sets and *vice versa*.

Considering operators rather than tensors

Since $X^* \otimes Y$ is canonically isomorphic to $\mathcal{L}(X, Y)$, it is also possible to avoid talking about tensors and rephrase all questions in terms of operators. In that setting, if $z = \sum_i |y_i\rangle\langle x_i^*|$, then

$$\|z\|_\varepsilon = \|z : X \to Y\|,$$

the operator norm, while $\|z\|_\pi = \min \sum_i \|y_i\| \cdot \|x_i^*\|$ (the minimum over all representations) is the nuclear norm. Moreover, appealing to duality we have

$$\|z\|_\pi = \max_{\|w : Y \to X\| \leqslant 1} \operatorname{tr} wz.$$

This allows to analyze both concepts in terms of operator norms, which are arguably conceptually simpler. In particular

$$\rho(X, Y) := \max_{z \in X \otimes Y,\, z \neq 0} \frac{\|z\|_\pi}{\|z\|_\varepsilon} = \max_{\|w : Y \to X^*\| \leqslant 1,\, \|z : X^* \to Y\| \leqslant 1} \operatorname{tr} wz.$$

Grothendieck and Pisier

Tensor products of normed spaces were studied in detail by Grothendieck in 1950s. In particular, he proposed and studied 14 "natural tensor norms" and posed a number of open questions, one of which was whether the norms $\|\cdot\|_{X\otimes_\pi Y}$ and $\|\cdot\|_{X\otimes_\varepsilon Y}$ can be equivalent when when $\dim X = \dim Y = \infty$.

It was a surprise when in 1980s Pisier answered this question in the positive, even more so because he showed earlier that if $\dim X \to \infty$ and $\dim Y \to \infty$, then

$$\rho(X, Y) \quad \to \quad \infty.$$

Also surprisingly, no quantitative analysis of the finite-dimensional case was made until very recently.

Some special cases

If $\mathcal{H}, \mathcal{K}$ are Hilbert (inner product) spaces, the situation is very simple: $\|\cdot\|_\varepsilon$ is the operator (spectral) norm, while $\|\cdot\|_\pi$ is the trace class norm and so

$$\rho(\mathcal{H}, \mathcal{K}) = \min\{\dim \mathcal{H}, \dim \mathcal{K}\}.$$

(This in particular saturates the easy general upper bound for $\rho(X, Y)$.)
For a general lower bound, a naive attempt is to appeal now to the John's theorem, which says that if $\dim X = n = \dim \mathcal{H}$, where $\mathcal{H}$ is a Hilbert space, then $d(X, \mathcal{H}) \leqslant n^{1/2}$, where

$$d(E, F) = \min\{\|v : E \to F\| \cdot \|v^{-1} : F \to E\|\}$$

is the Banach-Mazur distance. This allows to obtain some nontrivial information; for example using v, v^{-1} certifying $d(X, \mathcal{H}) \leqslant n^{1/2}$ as w, z in

$$\rho(X, \mathcal{H}) = \max_{\|w:\mathcal{H}\to X^*\|\leqslant 1,\ \|z:X^*\to\mathcal{H}\|\leqslant 1} \operatorname{tr} wz$$

we obtain $\rho(X, \mathcal{H}) \geqslant n^{1/2}$. The same circle of ideas allows to handle the case of different dimensions: $\rho(X, \mathcal{H}) \geqslant \min\{\dim X, \dim \mathcal{H}\}^{1/2}$.

Some special cases, cont'd

The same argument proves a cute equality $\rho(X, X^*) = \dim X$, but it doesn't help in the general case: by a 1981 result of Gluskin $\max\{d(E, F) : \dim E = \dim F = n\} = \Theta(n)$ and no nontrivial lower bound can be directly inferred.

Here are other interesting special cases that can be handled. If (say) $\dim X \geqslant n$, then

$$\rho(X, \ell_1^n) \geqslant (n/2)^{1/2} \text{ and } \rho(X, \ell_\infty^n) \geqslant (n/2)^{1/2}.$$

The first inequality follows by relating $\rho(X, \ell_1^n)$ to the so-called p-summing norms of the identity on X; these concepts were fashionable in 1970s and 1980s. The second one is then a consequence of (generally true) $\rho(X, Y) = \rho(X^*, Y^*)$. No substantial improvement is possible since $\rho(\mathcal{H}, \ell_1^n) = n^{1/2}$ (easy), but we do not know whether $(n/2)^{1/2}$ can be replaced by $n^{1/2}$ in general.

The quantum case

The case that is of relevance to quantum theory is when X, Y are spaces of Hermitian matrices endowed with the trace class norm. We have then

$$\rho(X, Y) = \Theta(\min\{\dim X, \dim Y\}^{3/4}).$$

Here is the idea behind the $O(\cdot)$ argument. For simplicity, consider $X = Y$ to be spaces of $k \times k$ matrices, so $n = \dim X = \dim Y = k^2$. We note first that $z \in X \otimes_\varepsilon Y$ can be thought of as a bilinear form on $X^* \times Y^*$ and that $X^* = Y^*$ is (the self-adjoint part of) the C^*-algebra $\mathcal{A}$ of $k \times k$ matrices with the usual operator norm. Thus we are in the realm of the Haagerup-Pisier non-commutative Grothendieck inequality, which says that for such bilinear form there are states φ, ψ on $\mathcal{A}$ such that

$$|z(a, b)| \leqslant 2\|z\|_\varepsilon \, \varphi(a^2)^{1/2} \, \psi(b^2)^{1/2} \quad \text{for all} \quad a, b \in \mathrm{Re}\, \mathcal{A}.$$

With this information, we need to upper-bound $\|z\|_\pi$.

The quantum case, conclusion

We need to upper-bound $\|z\|_\pi$, or the nuclear norm of $z : \mathcal{A} \to \mathcal{A}^*$, using

$$|z(a,b)| \leqslant 2\|z\|_\varepsilon\, \varphi(a^2)^{1/2}\, \psi(b^2)^{1/2} \quad \text{for all} \quad a,b \in \mathrm{Re}\,\mathcal{A}.$$

Here is a calculation which is not quite right, but supplies the gist of the trick. Let $\varphi = \sum_i \lambda_i |u_i\rangle\langle u_i|$ be the spectral decomposition. We will estimate the nuclear norm of $z : \mathcal{A} \to \mathcal{A}^*$ (say, with $\|z\|_\varepsilon \leqslant 1$) by writing

$$z(a) = \sum_{i,j} \mathrm{tr}(aE_{ji}) z(E_{ij}), \quad \text{or} \quad z = \sum_{i,j} |z(E_{ij})\rangle\langle E_{ji}|$$

where $E_{ij} = |u_i\rangle\langle u_j|$. For a single term, we have

$$\|z(E_{ij})\|_{\mathcal{A}^*} = \max_{\|b\|_{\mathcal{A}} \leqslant 1} |z(E_{ij}, b)| \leqslant 2\varphi(|E_{ij}|^2)^{1/2} \leqslant 2\lambda_i^{1/2}$$

(note that $\psi(b^2) \leqslant 1$ if $\|b\|_{\mathcal{A}} \leqslant 1$) and summing over i, j gives $2k \sum_i \lambda_i^{1/2} \leqslant 2k^{3/2} = 2n^{3/4}$ as a bound on $\|z\|_\pi$ ($4n^{3/4}$ if we don't cheat).

The general case, or cases

Modulo logarithmic factors (indicated by * in the Ω notation), we have:

- $X = Y, \dim X = n$: $\rho(X, X) = \Omega^*(n^{1/2})$ (almost optimal, see $X = \ell_1^n$)
- $\dim X = \dim Y = n$: $\rho(X, Y) = \Omega^*(n^{1/6})$
- $\dim X = n \leqslant \dim Y$: $\rho(X, Y) = \Omega^*(n^{1/8})$

The upper bounds are respectively $(2n)^{1/2}$, $n^{1/2}$, and again $n^{1/2}$. We know that the upper bound $(2n)^{1/2}$ is not sharp, but we do not know whether the factor 2 can be removed. It is conceivable that all these quantities are actually $\Omega^*(n^{1/2})$ or even $\Omega(n^{1/2})$.

The toolbox for the general case

This is again based on $\rho(X, Y) = \max_{\|w:Y\to X^*\|\leqslant 1,\, \|z:X^*\to Y\|\leqslant 1} \operatorname{tr} wz$ and an appropriate relaxation of the choices $w = v, z = v^{-1}$. First, we define the factorization constant of X through Y as

$$\mathrm{f}(X, Y) := \inf_{u,v} \left\{ \|u : X \to Y\| \cdot \|v : Y \to X\| \;:\; vu = \mathrm{Id}_X \right\},$$

which allows $\dim X \neq \dim Y$ and means that a subspace "well-isomorphic" to X is "well-complemented" in Y. Next, the weak factorization constant is

$$\mathrm{wf}(X, Y) := \inf_{u,v} \left\{ \mathbf{E} \left[\|u : X \to Y\| \cdot \|v : Y \to X\| \right] \;:\; \mathbf{E}\left[vu\right] = \mathrm{Id}_X \right\},$$

where u, v are now operator-valued random variables.
Clearly $\mathrm{wf}(X, Y) \leqslant \mathrm{f}(X, Y) \leqslant d(X, Y)$ and one easily checks that

$$\rho(X', Y) \leqslant \mathrm{wf}(X', X)\rho(X, Y) \quad \text{and} \quad \rho(X, Y) \geqslant \frac{\dim X}{\mathrm{wf}(X, Y^*)}.$$

The toolbox for the general case, cont'd

If $X = Y$, we select u, v "at random." At first, we choose a representation of X of $\mathbb{R}^n$ and let $u = n^{-1/2}G$, $v = n^{-1/2}G^\dagger$, where G is a GUE matrix. The tool which allows to estimate $\mathbf{E}\|G : X^* \to Y\|$ is the Chevet-Gordon inequality, which upper-bounds it by $n^{-1/2}$ times

$$\|\mathrm{Id} : \ell_2^n \to X^*\| \cdot \mathbf{E}\|g\|_Y + \|\mathrm{Id} : \ell_2^n \to Y^*\| \cdot \mathbf{E}\|g\|_X,$$

where g is the standard Gaussian vector on $\mathbb{R}^n$. If $X = Y$, the two terms coincide and – bounding similarly $\mathbf{E}\|G^\dagger : X^* \to Y\|$ – we need to control

$$\|\mathrm{Id} : \ell_2^n \to X^*\| \cdot \|\mathrm{Id} : X^* \to \ell_2^n\| \cdot \mathbf{E}\|n^{-1/2}g\|_X \cdot \mathbf{E}\|n^{-1/2}g\|_{X^*}.$$

For an appropriate representation of X of $\mathbb{R}^n$, the first two factors give $d(X^*, \ell_2^n) = d(X, \ell_2^n) \leqslant n^{1/2}$. The last two factors are essentially the same as spherical means, which can be controlled by the MM^*-estimate.

Mean (half-)width of $K \subset \mathbb{R}^n$ and the MM^*-estimate

If $|u| = 1$ and $w(K, u) := \sup_{x \in K} \langle u, x \rangle = \|u\|_{K^\circ}$, then $w(K, u) + w(K, -u)$ is the width of K in the direction of u. The average over u is the mean half-width of K.

The MM^*-estimate says that, for some well-balanced linear image $\tilde{K}$ of a centrally symmetric convex body $K \subset \mathbb{R}^n$ we can achieve $w(\tilde{K})w(\tilde{K}^\circ) = O(\log n)$.

Some additional tweaking is needed since we need to reconcile two requirements for the representation of X of $\mathbb{R}^n$, the one witnessing $d(X, \ell_2^n)$ and the other consistent with the MM^*-estimate, but ultimately gathering all bounds we get

$$\rho(X, X) = \Omega\left(\frac{\dim X}{d(X, \ell_2^n)\log^3 n}\right) \geqslant \Omega\left(\frac{n^{1/2}}{\log^3 n}\right),$$

as needed.

The case $X \neq Y$

There are too many "balancing requirements" to be simultaneously achievable, so instead the argument is based on the following trichotomy.

Let X be a normed space of dimension n. Then for every $1 \leqslant A \leqslant n^{1/2}$ at least one of the following holds

1. *X contains a subspace E of dimension $d = \Omega(n^{1/2})$ such that $d(E, \ell_\infty^d) = O(A\sqrt{\log n})$.*
2. *X^* contains a subspace F of dimension $d = \Omega(n^{1/2})$ such that $d(F, \ell_\infty^d) = O(A\sqrt{\log n})$.*
3. *X contains a subspace H of dimension $d = \Omega(A^2/\log n)$ such that $d(H, \ell_2^d) \leqslant 4$ and, additionally, H is $O(\log n)$-complemented in X.*

Since subspaces λ-isomorphic to ℓ_∞^d are automatically λ-complemented, each of the conditions above leads to an upper bound on $\mathrm{wf}(\ell_p^d, X)$ for the appropriate $p \in \{1, 2, \infty\}$. Given that $\rho(\ell_p^d, \ell_{p'}^{d'})$ are known, every combination of these conditions for X and Y leads to a lower bound on $\rho(X, Y)$, and the final step is optimizing over A.

THANK YOU!