POINTWISE BOUNDS FOR ORTHONORMAL BASIS ELEMENTS IN HILBERT SPACES

EMMANUEL KOWALSKI

1. Introduction

The following question is somewhat implicit in parts of [K1]: let Y be a non-empty finite set, $\nu : Y \to]0,1]$ a *probability density* on Y, i.e., a map such that

$$\sum_{y\in Y} \nu(y) = 1;$$

let $L^2(Y,\nu)$ denote the finite-dimensional Hilbert space of complex-valued functions on Y with the inner product

$$\langle f, g\rangle = \sum_{y\in Y} \nu(y) f(y)\overline{g(y)}.$$

Then, how small can we make the L^∞ norms

$$\|\varphi\|_\infty = \max_{y\in Y} |\varphi(y)|$$

of the elements of a given orthonormal basis of $L^2(Y,\nu)$?

More formally, we are looking for the quantity

$$M_\infty(Y,\nu) = \min_{\mathcal{B}} \max_{\varphi\in\mathcal{B}} \|\varphi\|_\infty,$$

where $\mathcal{B}$ runs over all orthonormal basis of $L^2(Y,\nu)$. Note that this is indeed a minimum, since the set of all $\mathcal{B}$'s is compact (being an orbit of the unitary group of the inner product). Denoting

$$\nu_- = \min_{y\in Y} \nu(y),$$

we will show in Section 2 that

$$M_\infty(Y,\nu) = \frac{1}{\sqrt{|Y|\nu_-}}.$$

Once the question is phrased, it is also natural to consider infinite-dimensional Hilbert spaces, such as $L^2([0,1],\nu)$, where ν is a positive integrable function with total mass 1, and the inner product is

$$\langle f, g\rangle = \int_0^1 f(t)\overline{g(t)}\nu(t)dt.$$

Of course in this context, not all orthonormal basis consist of bounded functions, but some are, and finding the "most" efficient is again a fairly natural questions. We show in Section 4 that the situation is very different then and that, in most reasonable cases, a space of the type $L^2([0,1],\nu dt)$ has an orthonormal basis where all elements are of (constant) modulus 1.

2. Best bounds for finite-dimensional spaces

In this section, we assume that $Y \neq \emptyset$ is finite, and ν is as in the introduction, and we will compute the value of $M_\infty(Y,\nu)$ in terms of ν.

For this, we start with a lower bound which is "local" in the sense of using only the normalization condition of the norm, not the global feature of an orthonormal basis.

Lemma 1. *For any Y and ν, we have $\|f\|_\infty \geqslant 1$ for any function f with L^2-norm equal to 1. In particular, we have $M_\infty(Y,\nu) \geqslant 1$.*

Proof. This is simply because, for any element $f \in L^2(Y,\nu)$ with norm $\|f\| = 1$, we have

$$1 = \sum_{y\in Y} \nu(y)|f(y)|^2 \leqslant \|f\|_\infty^2 \sum_y \nu(y) = \|f\|_\infty^2.$$

Or, phrased differently, the norm of the identity mapping $L^\infty(Y) \to L^2(Y,\nu)$ is equal 1 (as is well known), hence

$$\|f\| \leqslant \|f\|_\infty$$

for any function f on Y. $\square$

Here is on the other hand an upper bound for $M_\infty(Y,\nu)$ coming from an explicit choice of basis.

Proposition 2. *For any non-empty Y and ν, we have*

$$M_\infty(Y,\nu) \leqslant \frac{1}{\sqrt{|Y|\nu_-}},$$

where

$$\nu_- = \min_{y\in Y} \nu(y).$$

In fact, for any fixed $y_0 \in Y$, there exists an orthonormal basis $\mathcal{B}$ of $L^2(Y,\nu)$ such that for all $\varphi \in \mathcal{B}$ we have

$$|\varphi(y)| = \frac{1}{\sqrt{|Y|\nu(y)}} \leqslant \frac{1}{\sqrt{|Y|\nu_-}}, \qquad \text{for all } y \in Y, \tag{1}$$

$$\varphi(y_0) = \frac{1}{\sqrt{|Y|\nu(y_0)}}. \tag{2}$$

Proof. Start with the uniform density case $\nu = \nu_u$, given by

$$\nu_u(y) = 1/|Y|$$

for all $y \in Y$. Using an arbitrary bijection

$$\sigma\,:\,Y \to \mathbf{Z}/|Y|\mathbf{Z},$$

such that $y_0 \mapsto 0$, we derive an isometry between $L^2(Y,\nu_u)$ and the Hilbert space of functions on the finite abelian group $\mathbf{Z}/|Y|\mathbf{Z}$ with the normalized counting measure (or pedantically the probability Haar measure). This allows us to transport to $L^2(Y,\nu_u)$ the basis of characters

$$\psi_a\,:\,\begin{cases}\mathbf{Z}/|Y|\mathbf{Z} \to \mathbf{C}\\ x \mapsto e(ax/|Y|)\end{cases}$$

for $a \in \mathbf{Z}/|Y|\mathbf{Z}$ (as is traditional, we write $e(z) = \exp(2i\pi z)$ for $z \in \mathbf{C}$). All those functions have L^∞ norm equal to 1, and in fact, are of constant modulus, and are equal to 1 at the origin 0, so the basis functions $\varphi_a(y) = \psi_a(\sigma(y))$, satisfy

$$|\varphi_a(y)| = 1, \qquad \varphi_a(y_0) = 1,$$

for all a and $y \in Y$.

Now to treat the general case to the uniform density, we simply consider the obvious isometry

$$\begin{cases} L^2(Y,\nu) \to L^2(Y,\nu_u) \\ f \mapsto f\sqrt{|Y|\nu} \end{cases}$$

If we take an orthonormal basis $\mathcal{B}$ of $L^2(Y,\nu_u)$ with each element of constant absolute value 1, and evaluating to 1 at y_0, as given above, we find that

$$\mathcal{B}^* = \{\varphi/\sqrt{|Y|\nu}\}_{\varphi\in\mathcal{B}}$$

is an orthonormal basis of $L^2(Y,\nu)$ such that (1) and (2) hold. ☐

Note that this upper bound already shows that

$$M_\infty(Y,\nu_u) \leqslant 1,$$

and therefore there is equality by the first lemma. However, it is not immediate that the basis above is optimal for a general non-uniform choice of ν: the isometry we used only implies a priori that

$$M_\infty(Y,\nu_u) \geqslant \frac{1}{\sqrt{|Y|\nu_+}}$$

where $\nu_+ = \max \nu(y)$. Note that (by positivity) the relation

$$1 = \|1\|^2 = \sum_{y\in Y} \nu(y)$$

also implies that

$$\nu_- \leqslant \frac{1}{|Y|} \leqslant \nu_+$$

with equality for one of these if and only if there is equality for both, and if and only if $\nu = \nu_u$. So if ν is not the standard uniform density, the upper bound of the Proposition is strictly larger than the trivial local lower bound.

It turns out however that the upper bound is the right one.

Proposition 3. *Let Y and ν be as above. We have*

$$M_\infty(Y,\nu) = \frac{1}{\sqrt{|Y|\nu_-}},$$

Here is an amusing corollary:

Corollary 4. *Let Y and ν be as above. If there exists an orthonormal basis $\mathcal{B}$ such that $|\varphi(y)|$ is constant for all $\varphi \in \mathcal{B}$, then $\nu = \nu_-$.*

Indeed, the constant modulus must necessarily be one, hence

$$1 = M_\infty(Y, \nu) = \frac{1}{\sqrt{|Y|\nu_-}}$$

so we have $\nu_- = |Y|^{-1}$, which shows that density must be the uniform density.

Proof. We consider the linear identity isomorphism

$$L^2(Y, \nu) \xrightarrow{i} L^\infty(Y)$$

and proceed to compute its norm. Let $\mathcal{B}$ be an arbitrary orthonormal basis of $L^2(Y, \nu)$. Then for any function f on Y, we have

$$f = \sum_{\varphi \in \mathcal{B}} c_\varphi \varphi,$$

where $c_\varphi = \langle f, \varphi \rangle$, hence by Cauchy's inequality we obtain

$$\begin{aligned} \|f\|_\infty &\leqslant \sum_{\varphi \in \mathcal{B}} |c_\varphi| \|\varphi\|_\infty \\ &\leqslant \Big(\sum_{\varphi \in \mathcal{B}} \|\varphi\|_\infty^2\Big)^{1/2} \Big(\sum_{\varphi} |c_\varphi|^2\Big)^{1/2} = \Big(\sum_{\varphi \in \mathcal{B}} \|\varphi\|_\infty^2\Big)^{1/2} \|f\|. \end{aligned}$$

It follows that

$$\|i\| \leqslant \Big(\sum_{\varphi \in \mathcal{B}} \|\varphi\|_\infty^2\Big)^{1/2}. \tag{3}$$

Moreover, for certain orthonormal basis, there is equality. Namely, assume $\mathcal{B}^*$ is an orthonormal basis for which there exists $y \in Y$ and $\alpha \in \mathbf{R}$ such that

$$\varphi^*(y) = \|\varphi^*\|_\infty e(\alpha) \tag{4}$$

for all $\varphi^* \in \mathcal{B}^*$ (i.e., evaluation at a single point computes all L^∞-norms, up to a phase which does not depend on φ^*; we say that such a basis is "peaked"). Then defining f by means of the expansion

$$f = \sum_{\varphi^* \in \mathcal{B}^*} \|\varphi^*\|_\infty \varphi^*,$$

we have for all $x \in Y$ the upper bound

$$|f(x)| \leqslant |f(y)| = e(-\alpha) f(y) = \sum_{\varphi^* \in \mathcal{B}^*} \|\varphi^*\|_\infty^2,$$

i.e., there is equality in the previous computation. This means that

$$\|i\| = \Big(\sum_{\varphi^* \in \mathcal{B}^*} \|\varphi^*\|_\infty^2\Big)^{1/2}$$

for any peaked orthonormal basis. In particular, the right-hand side is independent of the choice of such a basis $\mathcal{B}^*$.

Now observe that the second part of Proposition 2 exactly states that there exists a peaked orthonormal basis $\mathcal{B}^*$ where

$$\|\varphi^*\|_\infty = \frac{1}{\sqrt{|Y|\nu_-}},$$

for $\varphi \in \mathcal{B}^*$, the peak point being any $y \in Y$ such that $\nu(y) = \nu_-$ (and α being 0).

We therefore derive, for an arbitrary $\mathcal{B}$ again, that

$$\sum_{\varphi \in \mathcal{B}} \|\varphi\|_\infty^2 \geqslant \|i\|^2 = \sum_{\varphi \in \mathcal{B}^*} \|\varphi^*\|_\infty^2 = |Y| \times \frac{1}{|Y|\nu_-} = \frac{1}{\nu_-},$$

which in particular implies that there exists one element $\varphi \in \mathcal{B}$ at least such that

$$\|\varphi\|_\infty^2 \geqslant \frac{1}{|Y|\nu_-}.$$

This clearly implies the result. $\square$

Remark 5. In fact, from the proof we can remark that a peaked orthonormal basis $\mathcal{B}$ is *characterized* by the formula

$$\frac{1}{\nu_-} = \|i\|^2 = \sum_{\varphi \in \mathcal{B}} \|\varphi\|_\infty^2$$

which seems quite a bit "softer" than the definition.

Indeed, if (3) is sharp coming from $\mathcal{B}$, there must (in particular) exist some $f \neq 0$ and some $y \in Y$ such that

$$\|f\|_\infty = |f(y)| = \Bigl|\sum_{\varphi} c_\varphi \varphi(y)\Bigr| = \sum_{\varphi} |c_\varphi||\varphi(y)| = \sum_{\varphi} |c_\varphi|\|\varphi\|_\infty.$$

Moreover, $c_\varphi \neq 0$ for all φ, since otherwise Cauchy's inequality leads to a better estimate for $\|f\|_\infty$ than the one desired, so the last equality requires that

$$|\varphi(y)| = \|\varphi\|_\infty$$

for each element of the orthonormal basis, and the last but one requires that the phase of $\varphi(y)/|\varphi\|_\infty$ is independent of φ.

Remark 6. The following geometric interpretation may well clarify this result (and make it obvious for those with strong geometric intuition): the norm of i is simply the "radius" of the smallest ball in the L^∞-norm containing the unit ball in $L^2(Y, \nu)$. Since the latter, seen in $\mathbf{R}^{2|Y|}$, is an ellipse with semiaxis lengths given by $1/\sqrt{\nu(y)}$, $y \in Y$ (each repeated twice for real and imaginary parts of evaluation at y), the equality $\|i\| = \nu_-^{-1/2}$ is quite obvious without further computation.

On the other hand, it is not (for the author) a priori clear for geometric reasons why there should exist an orthonormal basis with precisely this L^∞ norm, as we proved.

Remark 7. Is is natural to enquire further about how small $\|\varphi\|_\infty$ may be for elements of an orthonormal basis of a *subspace* of $L^2(Y, \nu)$. This is obviously more delicate because the lower bound $\|\varphi\|_\infty \geqslant 1$ may be the best possible even for a subspace of dimension > 1. For instance, it is easy to check that, for any density $\nu : Y = \{1, 2, 3\} \rightarrow]0, 1]$, one may always find a function f on Y such that the system $(1, f)$ is orthonormal.

3. Examples and various types of orthonormal basis

In this section, we consider interesting examples of Y and ν, and, motivated by the previous section, we introduce different conditions on orthonormal basis of $L^2(Y,\nu)$, and discuss their relations.

Let $\mathcal{B}$ be an orthonormal basis. We say that $\mathcal{B}$ is *optimal* if it satisfies

$$\max_{\varphi\in\mathcal{B}} \|\varphi\|_\infty = M_\infty(Y,\nu) = \frac{1}{\sqrt{|Y|\nu_-}}\ ; \tag{5}$$

and (because many examples will appear below), we say that $\mathcal{B}$ is *pointed* if there exists some $y\in Y$ such that

$$\frac{1}{\nu_-} = \|i\|^2 = \sum_{\varphi\in\mathcal{B}} |\varphi(y)|^2. \tag{6}$$

Recall $\mathcal{B}$ was said to be *peaked* if, as in the proof of Proposition 3, it "computes" $\|i\|$, i.e., we have

$$\frac{1}{\nu_-} = \|i\|^2 = \sum_{\varphi\in\mathcal{B}} \|\varphi\|_\infty^2, \tag{7}$$

or equivalently, if there exists some $y\in Y$ such that $|\varphi(y)| = \|\varphi\|_\infty$ for all $\varphi\in\mathcal{B}$, and $\varphi(y)/|\varphi\|_\infty = e(\alpha)$ with α independent of φ.

Now we note that if $\mathcal{B}$ is optimal, then it must also be peaked: indeed, (5) implies

$$\sum_{\varphi\in\mathcal{B}} \|\varphi\|_\infty^2 \leqslant \frac{1}{\nu_-},$$

which combined with the reverse inequality (3) leads to (7). Moreover, since there would be strict inequality, contradicting (3), unless $\|\varphi\|_\infty = M_\infty(Y,\nu)$ for all φ, it follows that a basis is optimal if and only if the L^∞ norm of all its elements is equal to $M_\infty(Y,\nu)$:

Proposition 8. *For any non-empty Y and ν, any orthonormal basis $\mathcal{B}$ satisfies*

$$\max_{\varphi\in\mathcal{B}} \|\varphi\|_\infty \geqslant \frac{1}{\sqrt{|Y|\nu_-}},$$

and if there is equality, there exists $y\in Y$ such that we have

$$\|\varphi\|_\infty = |\varphi(y)| = \frac{1}{\sqrt{|Y|\nu_-}}$$

for all $\varphi\in\mathcal{B}$.

An interesting family of examples of non-uniform density arises by taking $Y = G^\sharp$, the set of conjugacy classes of a finite group G, with

$$\nu(y) = \frac{|y|}{|G|}$$

for any conjugacy class $y\subset G$. This is not uniform if (and only if) G is not an abelian group.

The best known orthonormal basis of $L^2(Y,\nu)$ in this setting is the basis of *characters*

$$y\mapsto \operatorname{Tr}\rho(g)$$

where ρ runs over isomorphism classes of irreducible linear representations of G and g is any element of the conjugacy class y.

Since we have

$$|\operatorname{Tr}\rho(g)| \leqslant \dim\rho = \operatorname{Tr}\rho(1)$$

for any $y \in Y$ ($\rho(g)$ is always unitary with respect to some inner product on the space of ρ, hence its eigenvalues are roots of unity, and moreover $\rho(1)$ is the identity), we see that this orthonormal basis is a peaked basis, and indeed it is well-known that

$$\sum_{\rho} \dim(\rho)^2 = |G| = \frac{1}{\nu_-}$$

(the conjugacy class of the identity gives the minimal value $\nu_- = 1/|G|$).

If G is not abelian, there are distinct character degrees, and therefore the basis of characters is not optimal (showing in particular that a peaked basis is not necessarily optimal). However, it is a fact that it is fairly close to being optimal in some important cases (see Chapter 5 of [K1]).

Here is the next general family of examples. They are related to the infinite-dimensional setting of the next section. Let $\nu : [0,1] \to [0,+\infty[$ be a measurable function such that

$$\int_0^1 \nu(t)dt = 1,$$

and consider the Hilbert space $L^2([0,1],\nu dt)$. A particular orthonormal basis of this space is given by the *orthogonal polynomials* $(p_n)_{n\geqslant 0}$ associated to ν, which determined uniquely by the properties that p_n is of degree n, has real coefficients, and positive leading term. Those polynomials are obtained by Gram-Schmidt orthonormalisation from the (algebraic) basis (x^n) of the dense subspace of polynomials (see for instance [Sz] for the general theory, noting that the basic interval there is $[-1,1]$ instead of $[0,1]$). Fix $n \geqslant 1$. One of the properties of p_n is that it has n distinct real zeros in $]0,1[$, say

$$x_1 < x_2 < \ldots < x_n,$$

and that there exist unique numbers $\lambda_i > 0$, $1 \leqslant i \leqslant n$, such that

$$\int_0^1 p(t)\nu(t)dt = \sum_{i=1}^{n} \lambda_i p(x_i)$$

for arbitrary polynomials $p \in \mathbf{C}[X]$ of degree $\deg(p) \leqslant 2n-1$. It follows in particular that if $Y = \{x_1,\ldots,x_n\}$, then

$$\mathcal{B} = (p_0,\ldots,p_{n-1})$$

(restricted to Y) is an orthonormal basis of the finite-dimensional space $L^2(Y,\lambda)$. It turns out, moreover, that λ_i may be expressed as follows (see, e.g., [Sz, (3.4.8)]):

$$\frac{1}{\lambda_j} = \sum_{i=0}^{n-1} p_i(x_j)^2.$$

In particular, it follows that $1/\lambda_-$ can be written

$$\frac{1}{\lambda_-} = \sum_{\varphi\in\mathcal{B}} |\varphi(y)|^2$$

for some $y \in Y$, in other words, the basis $\mathcal{B}$ of $L^2(Y,\lambda)$ is *pointed*. However, simple examples show that it is not peaked in general (e.g., for $\nu = 1$, where the p_n are L^2-normalized Legendre polynomials).

4. Infinite-dimensional examples

In this section we consider the space $L^2(\nu) = L^2([0,1], \nu dt)$ where ν is a non-negative measurable function on $[0,1]$ with total mass 1, and we wish to investigate how small the L^∞-norms of elements of an orthonormal basis may be. Thus we denote

$$N_\infty(\nu) = \min_{\mathcal{B}} \max_{\varphi \in \mathcal{B}} \|\varphi\|_\infty$$

where $\mathcal{B}$ runs over all orthonormal basis of $L^2(\nu)$, with obvious conventions when φ is not (essentially) bounded.

The proof of the previous section used the finite-dimensionality in two ways: to argue that $\nu_- > 0$, and in the proof of Proposition 3, to compute the norm of the identity map $L^2(Y,\nu) \to L^\infty(Y)$, which is not a continuous operator in the present case.

On the other hand, notice that Lemma 1 remains valid, in the sense that $\|\varphi\|_\infty \geqslant 1$ for any L^2-normalized $\varphi \in L^2(\nu)$; in fact, it remains true that the embedding $L^\infty \to L^2(\nu)$ is continuous, with norm 1. So $N_\infty(\nu) \geqslant 1$.

It turns out that this lower bound can not be improved in this setting!

Proposition 9. *Let ν be of the form $\nu(t) = \phi'(t)$ for a strictly increasing homeomorphism $\phi : [0,1] \to [0,1]$. Then we have $N_\infty(\nu) = 1$.*

The assumption is of course valid for all but the weirdest functions ν, since one can take

$$\phi(t) = \int_0^t \nu(u)du$$

for $0 \leqslant t \leqslant 1$, whenever the fundamental theorem of the calculus holds.

Proof. First consider $\nu = 1$, which corresponds to the uniform measure on $[0,1]$ (i.e., Lebesgue measure). Then Fourier series give the well-known orthonormal basis of additive characters $x \mapsto e(nx)$ where $n \in \mathbf{Z}$, and those (in perfect analogy with the uniform measure on finite sets) are uniformly bounded by 1. In particular, we have $N_\infty(1) = 1$.

Now the assumption on ν gives the general proof away: the map

$$\begin{cases} L^2([0,1]) \to L^2([0,1],\nu) \\ f \mapsto f \circ \phi \end{cases}$$

is an isometry by the change of variable formula $u = \phi(t)$, $u' = \phi'(t)dt = \nu(t)dt$:

$$\int_0^1 f(u)\overline{g(u)}du = \int_0^1 f(\phi(t))\overline{g(\phi(t))}\nu(t)dt,$$

and moreover it obviously satisfies $\|f \circ \phi\|_\infty = \|f\|_\infty$, so we can transfer the additive characters to obtain an orthonormal basis $\varphi_n(t) = e(n\phi(t))$ of $L^2(\nu)$, where all elements have constant modulus 1. □

Here are some standard examples of orthonormal basis, which shows that such optimal behavior does not necessarily occur naturally.

Example 10. The Haar system (H_n) is an orthonormal basis of $L^2([0,1])$ constructed as follows: let $H_1 = 1$ and for $n \geqslant 1$, write (uniquely) $n = 2^k + j$ for some $k \geqslant 0$ and $1 \leqslant j \leqslant 2^k$, and then let

$$H_n = \begin{cases} 2^{k/2}, & \text{if } 2j-2 < 2^{k+1}t < 2j-1, \\ -2^{k/2}, & \text{if } 2l-1 < 2^{k+1}t < 2j, \\ 0, & \text{otherwise} \end{cases}$$

One has $\|H_n\|_\infty = 2^{k/2} \approx \sqrt{n}$ for $n = 2^k + j$.

Example 11. We come back to orthogonal polynomials, as described in the previous section. Let (p_n) be the sequence of orthogonal polynomials associated to $L^2([0,1], \nu dt)$. In [Sz, Ch. 8] are found a number of upper bounds for $|p_n(t)|$, hence for $\|p_n\|_\infty$ (taking care to renormalize the interval from $[-1,1]$ of [Sz] to $[0,1]$).

If we take for instance $\nu = 1$, the orthogonal polynomials are, up to constants, the classical Legendre polynomials $(P_n)_{n\geqslant 0}$; more specifically, we have

$$p_n(t) = 2\sqrt{n+\frac{1}{2}}\; P_n(2t-1)$$

where

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n$$

is the Legendre polynomial as usually defined (orthogonal on $[-1,1]$ for Lebesgue measure). Since $|P_n|$ is maximal and equal to 1 at ± 1, we obtain

$$\|p_n\|_\infty = 2\sqrt{n+\frac{1}{2}}$$

for $n \geqslant 0$. This gives another example of orthonormal basis with unbounded L^∞-norms, while $N_\infty = 1$.

References

[K1] E. Kowalski: *The large sieve and its applications: arithmetic geometry, random walks, discrete groups*, Cambridge Univ. Tracts (to appear).

[Sz] G. Szegö: *Orthogonal polynomials*, Coll. Publ. XXIII, A.M.S (1939).

Université Bordeaux I - A2X, 351, cours de la Libération, 33405 Talence Cedex, France

E-mail address: `emmanuel.kowalski@math.u-bordeaux1.fr`