A Combinatorial Version of Vaaler's Theorem

Keith Ball

Department of Mathematics
University College London
Gower Street, London, WC1E 6BT, UK
kmb@math.ucl.ac.uk

and

Maria Prodromou

Department of Mathematics
University College London
Gower Street, London, WC1E 6BT, UK
maria@math.ucl.ac.uk

Abstract

In 1979 Vaaler proved that every d-dimensional central section of the cube $[-1,1]^n$ has volume at least 2^d. We prove the following sharp combinatorial analogue. Let K be a d-dimensional subspace of $\mathbb{R}^n$. Then, there is a probability measure P on the section $[-1,1]^n \cap K$, so that the quadratic form

$$\int_{[-1,1]^n \cap K} v \otimes v \, dP(v)$$

dominates the identity on K (in the sense that the difference is positive semi-definite).

1 Introduction

In [Va] Vaaler proved that for every d and n, every d-dimensional central section of the cube $[-1,1]^n$ has volume at least 2^d. His result provided a sharp version of Siegel's Lemma in the geometry of numbers and was used by Bombieri and Vaaler himself ([BV]) for applications in Diophantine approximation. Vaaler's theorem is obviously sharp since the sections by d-dimensional coordinate subspaces are cubes of volume 2^d.

If $(\epsilon_i)_1^d$ are IID choices of sign and $x = (x_i)$ is a vector in $\mathbb{R}^d$ then

$$\mathrm{E}\left(\sum x_i \epsilon_i\right)^2 = \sum x_i^2.$$

Thus, if P is the uniform probability measure on the corners of the cube $[-1,1]^d$ then the quadratic form

$$\int_{[-1,1]^d} v \otimes v \, dP(v)$$

is the identity on $\mathbb{R}^d$.

In this paper, we prove the following sharp combinatorial version of Vaaler's Theorem.

Theorem 1. *Let K be a d-dimensional subspace of $\mathbb{R}^n$. Then, there is a probability measure P on $[-1,1]^n \cap K$, with*

$$\int_{[-1,1]^n \cap K} v \otimes v \, dP \geq I_K \tag{1}$$

where the dominance is in the sense of positive definite operators.

Thus, each section of the cube not only has large volume but it is also "fat in all directions" in the same way as a cube.

Observe that if we start with the uniform probability on the corners of the n-dimensional cube and project it orthogonally onto the subspace K, we will obtain a probability measure that yields the identity (in the above sense). However, for most subspaces, the support of this projected measure will extend far outside the section $[-1,1]^n \cap K$ so it will not be a suitable choice in the theorem.

Coordinate subspaces show that Theorem 1 is sharp in the sense that we cannot guarantee to beat a larger multiple of the identity. What is more surprising is that lower-dimensional cubes do not provide the only extreme cases. For example, the section of the 3-dimensional cube perpendicular to its main diagonal is a regular hexagon whose corners are points like $(1,-1,0)$ which are at distance $\sqrt{2}$ from the origin. If we take the traces of the operators appearing in equation (1) we obtain

$$\int_{[-1,1]^n \cap K} |v|^2 \geq \dim(K).$$

So the probability measure guaranteed by Theorem 1 must be supported on the corners of the hexagon and we cannot beat any multiple of the identity larger than 1. A similar argument works for the diagonal section of the cube in any odd dimension. The existence of a large family of subspaces for which the inequality is sharp makes it highly unlikely that we could write down the measure we wish to find in any reasonably explicit way. Our argument will build the probability as the end result of a sequence of linked optimisation problems.

Theorem 1 can be reformulated in a variety of ways. It is a consequence of the Pietsch Factorisation Theorem (see eg. [P]) that the so-called 2-summing norm $\pi_2(T)$ of a map $T : X \to \ell^2$ from a Banach space into Hilbert space is equal to the least C for which there is a probability measure P on the unit ball in X^* for which

$$|Tx|^2 \leq C^2 \int |\phi(x)|^2 \, dP(\phi)$$

for every $x \in X$. Thus Theorem 1 can be rewritten as the following "lifting" theorem for 2-summing norms

Theorem 2. *Let K be a d-dimensional subspace of $\mathbb{R}^n$ and $T : \mathbb{R}^n \to K$ the orthogonal projection onto K. Let $\tilde{T}$ be the map induced by T on the quotient space $\ell_1^n / \ker T$, as in the commuting diagram below. Then, $\tilde{T}$ is 2-absolutely summing, and $\pi_2(\tilde{T}) \leq 1$.*

$$\begin{tikzcd} \ell_1^n \arrow[r, "id"] \arrow[d, "q"'] & \ell_2^n \arrow[d, "T"] \\ \ell_1^n / \ker(T) \arrow[r, "\tilde{T}"] & K \end{tikzcd} \tag{2}$$

It is a simple (and pretty well-known) fact that if $(x_i)_1^d$ is a sequence of unit vectors in $\mathbb{R}^d$ then there is a unit vector v in $\mathbb{R}^d$ for which

$$|\langle v, x_i \rangle| \leq \frac{1}{\sqrt{d}}, \quad \text{for all} \quad i = 1, \ldots, n.$$

It follows from Theorem 1 that this fact can be generalised:

Theorem 3. *Let $(x_i)_1^n \subset \mathbb{R}^d$ be a sequence of vectors that satisfy $\sum_{i=1}^n |x_i|^2 = d$. Then there exists a unit vector $v \in \mathbb{R}^d$, such that*

$$|\langle v, x_i\rangle| \;\; \leq \;\; \frac{1}{\sqrt{d}}, \quad \textit{for all} \quad i = 1, \ldots, n. \tag{3}$$

In this article Theorem 1 and Theorem 3 will both be obtained from the following.

Theorem 4. *Let $(u_i)_{i=1}^n$ be a sequence of vectors in $\mathbb{R}^d$ that satisfies $\sum_{i=1}^n u_i \otimes u_i = I_d$ and Q a positive semi-definite quadratic form on $\mathbb{R}^d$. Then there is a vector $w \in \mathbb{R}^d$, such that $|\langle w, u_i\rangle| \leq 1$, for all $i = 1, \ldots, n$ and*

$$w^T Q w \geq \mathrm{tr}(Q).$$

In Section 2 we shall give the main argument: the proof of Theorem 4. The most intriguing feature of the proof is that it is to some extent constructive. In view of the importance in Diophantine approximation of finding vectors with small inner product with a given sequence, it is natural to ask whether there is a lattice version of Theorem 3 that can be proved without the averaging technique implicit in the proof of the Dirichlet-Minkowski box principle. In Section 3 we deduce Theorem 1 and in Section 4, Theorem 3.

2 Proof of Theorem 4

Let C be the set

$$\{x \in \mathbb{R}^d : |\langle x, u_i\rangle| \leq 1, \; \forall i = 1, \ldots, n\}.$$

Our aim is to find a point $w = (w_1, \ldots, w_d) \in C$, that satisfies $w^T Q w \geq \mathrm{tr}(Q)$, for the positive semi-definite quadratic form Q.

Assume that Q is diagonal with respect to the standard basis of $\mathbb{R}^d$, with eigenvalues $s_1 \geq s_2 \geq \cdots \geq s_d \geq 0$. We wish to find a point $w \in C$ that satisfies

$$\sum_{i=1}^d s_i w_i^2 \geq \sum_{i=1}^d s_i. \tag{4}$$

Note that $v^T(u_i \otimes u_i)v$ is equal to the number $\langle v, u_i\rangle^2$ so for any vector $v = (v_i)$

$$\sum_{i=1}^d v_i^2 = v^T \left(\sum_{i=1}^n u_i \otimes u_i\right) v = \sum_{i=1}^n \langle v, u_i\rangle^2. \tag{5}$$

For each m between 1 and d let C_m be the section of C by the subspace of $\mathbb{R}^d$ spanned by the first m basis vectors.

$$C_m = C \cap \{x \in \mathbb{R}^d : x = (x_1, \ldots, x_m, 0, \ldots, 0)\}, \quad m = 1, \ldots, d.$$

We shall construct inductively a sequence of points $w(1)$, $w(2)$ and so on, with $w(m)$ being a corner of C_m. Each $w(m)$ will be "large" as measured by a certain quadratic form (a different one for each m). The last of these quadratic forms will be just

$$w \mapsto \sum_{i=1}^{d} s_i w_i^2$$

so that the last point in the sequence, $w(d)$ will satisfy the conclusion of the theorem.

Let $w(1)$ be an extreme point of the line segment C_1. Since $w(1)$ belongs to the boundary of C, there will be at least one index i for which $|\langle w(1), u_i\rangle| = 1$. So equation (5) gives

$$w_1(1)^2 \geq 1.$$

The point $w(1)$ belongs to C_2 since the C_i are nested. Since the function $x \mapsto x_1^2$ is convex, it attains its maximum over the section C_2 at a vertex of C_2 which therefore also satisfies $w_1(2)^2 \geq 1$. This vertex is a point $w(2) = (w_1(2), w_2(2), 0, \ldots, 0)$ that lies on a $(d-2)$-dimensional face of C. So it belongs to at least two of the boundary hyperplanes of C and we will have $|\langle w(2), u_i\rangle| = 1$ for at least two indices i. Hence by equation (5) again

$$w_1(2)^2 + w_2(2)^2 \geq 2.$$

Thus, we have found a point $w(2)$ which satisfies

$$w_1(2)^2 \geq 1 \quad \text{and} \quad w_1(2)^2 + w_2(2)^2 \geq 2. \tag{6}$$

Since $(s_1 - s_2)$ and $(s_2 - s_3)$ are nonnegative, we can combine these inequalities to get

$$\begin{aligned} (s_1 - s_3)w_1(2)^2 + (s_2 - s_3)w_2(2)^2 &= (s_1 - s_2)w_1(2)^2 \\ &\quad +(s_2 - s_3)(w_1(2)^2 + w_2(2)^2) \\ &\geq (s_1 - s_2) + 2(s_2 - s_3) \\ &= (s_1 - s_3) + (s_2 - s_3). \end{aligned} \tag{7}$$

Now we maximise the function $x \mapsto (s_1 - s_3)x_1^2 + (s_2 - s_3)x_2^2$ over the section C_3. The maximum will occur at a corner $w(3)$ satisfying

$$\begin{aligned}(s_1 - s_3)w_1(3)^2 + (s_2 - s_3)w_2(3)^2 &\geq (s_1 - s_3) + (s_2 - s_3) \quad \text{and} \\ w_1(3)^2 + w_2(3)^2 + w_3(3)^2 &\geq 3. \end{aligned} \tag{8}$$

These inequalities can be combined to give

$$\begin{aligned}\sum_{i=1}^{3}(s_i - s_4)w_i(3)^2 &= \sum_{i=1}^{2}(s_i - s_3)w_i(3)^2 + (s_3 - s_4)\sum_{i=1}^{3} w_i(3)^2 \\ &\geq \sum_{i=1}^{2}(s_i - s_3) + 3(s_3 - s_4) = \sum_{i=1}^{3}(s_i - s_4). \end{aligned} \tag{9}$$

Continuing in this way, at the k-th step we maximise the function $x \mapsto \sum_{i=1}^{k-1}(s_i - s_k)x_i^2$ to get a corner $w(k)$ of C_k for which

$$\sum_{i=1}^{k}(s_i - s_{k+1})w_i(k)^2 \geq \sum_{i=1}^{k}(s_i - s_{k+1}). \tag{10}$$

For the final step, we may choose $s_{d+1} = 0$ and get a vertex $w(d)$ of $C_d = C$ for which

$$\sum_{i=1}^{d} s_i w_i(d)^2 = \sum_{i=1}^{d}(s_i - s_{d+1})w_i(d)^2 \geq \sum_{i=1}^{d}(s_i - s_{d+1}) = \sum_{i=1}^{d} s_i. \tag{11}$$

So we have found a point $w = w(d) \in C$ such that $w^T Q w \geq \mathrm{tr}(Q)$. This completes the proof. $\square$

3 Proof of Theorem 1

Our aim is to prove that there is a positive semi-definite symmetric operator H, and a sequence λ_i of nonnegative numbers with $\sum_{i=1}^{m} \lambda_i = 1$ such that the identity on K can be written as

$$I_K = \sum_{i=1}^{m} \lambda_i v_i \otimes v_i - H,$$

where each v_i belongs to $[-1,1]^n \cap K$ and we regard $v_i \otimes v_i$ as an operator on K. Our probability P will then assign mass λ_i to the point v_i (and m can be taken to be at most $\frac{d(d+1)}{2}$: the dimension of the space of symmetric matrices).

Assume this is false. Then we can separate I_K from the set $\overline{\mathrm{conv}}\{v \otimes v - H : v \in [-1,1]^n \cap K, H \geq 0\}$ by a hyperplane. So there is a linear functional f on the space of operators, such that

$$f(I_K) > f(v \otimes v) - f(H), \tag{12}$$

for all $v \in [-1,1]^n \cap K$ and all H. With respect to some orthonormal basis of K we may regard this functional as a $d \times d$ matrix $Q = (q_{ij})$, where for any $d \times d$ matrix $A = (a_{ij})$,

$$f(A) = \sum_{i,j=1}^{d} q_{ij} a_{ij} = \mathrm{tr}(Q^T A)$$

and without loss of generality we may assume that Q is symmetric.

So, writing $H = (h_{ij})$, condition (12) can be written

$$\begin{aligned} \mathrm{tr}(Q) &> \sum_{i,j=1}^{d} q_{ij}(v \otimes v)_{ij} - \sum_{i,j=1}^{d} q_{ij} h_{ij} = \sum_{i,j=1}^{d} q_{ij} v_i v_j - \sum_{i,j=1}^{d} q_{ij} h_{ij} \\ &= v^T Q v - \sum_{i,j=1}^{d} q_{ij} h_{ij}, \quad \text{for all } v, H. \end{aligned}$$

Since we can choose H to be any positive semi-definite matrix, this can only hold if Q is positive semi-definite. So, we have found a positive semi-definite symmetric Q, such that

$$\mathrm{tr}(Q) > v^T Q v, \tag{13}$$

for every $v \in [-1,1]^n \cap K$.

To obtain a contradiction, we will show that there is a point $v \in [-1,1]^n \cap K$, with $v^T Q v \geq \mathrm{tr}(Q)$. Let $e_1, e_2, \ldots, e_n$ be the standard basis of the ambient space $\mathbb{R}^n$. For each i let u_i be the orthogonal projection of e_i onto the subspace K. Clearly these points satisfy the hypothesis $\sum_{i=1}^n u_i \otimes u_i = I_K$ of Theorem 4. Therefore there is a vector $v \in K$ which satisfies $|\langle v, u_i \rangle| \leq 1$, for all i and $v^T Q v \geq \mathrm{tr}(Q)$. The conditions $|\langle v, u_i \rangle| \leq 1$, for all i imply that the point v belongs to the cube $[-1,1]^n$ (and hence to the section $[-1,1]^n \cap K$) since for each i, $|\langle v, e_i \rangle| = |\langle v, u_i \rangle| \leq 1$. This contradicts (13). $\square$

4 Proof of Theorem 3

Let $(x_i)_{i=1}^n \subset \mathbb{R}^d$ be a sequence of vectors for which $\sum_{i=1}^n \|x_i\|^2 = d$. Define the $d \times d$ matrix H to be $\sum_{i=1}^n x_i \otimes x_i$ and set $Q = H^{-1}$. Note that $\mathrm{tr}(H) = \sum_{i=1}^n \|x_i\|^2 = d$ and therefore $\mathrm{tr}(Q) \geq d$ (by the Cauchy-Schwarz inequality). Let $u_i = H^{-1/2} x_i$, for all $i = 1, \ldots, n$. The vectors u_i satisfy the hypothesis of Theorem 4 because

$$\sum_{i=1}^n u_i \otimes u_i = \sum_{i=1}^n H^{-\frac{1}{2}} x_i \otimes H^{-\frac{1}{2}} x_i = H^{-\frac{1}{2}} \left(\sum_{i=1}^n x_i \otimes x_i \right) H^{-\frac{1}{2}} = H^{-\frac{1}{2}} H H^{-\frac{1}{2}} = I_d.$$

Hence, there exists a vector $w \in \mathbb{R}^d$ such that $|\langle w, u_i \rangle| \leq 1$, for all $i = 1, \ldots, n$ and $w^T Q w \geq d$. The conditions $|\langle w, u_i \rangle| \leq 1$, can be re-written as follows:

$$|\langle w, u_i \rangle| = |\langle H^{-1/2} x_i, w \rangle| = |\langle x_i, H^{-1/2} w \rangle| \leq 1.$$

Now set $v = H^{-1/2} w \in \mathbb{R}^d$. Then $|\langle x_i, v \rangle| \leq 1$, for all $i = 1, \ldots, n$. Also,

$$\langle v, v \rangle = \langle H^{1/2} v, Q H^{1/2} v \rangle = \langle w, Qw \rangle \geq d.$$

So, $\|v\| \geq \sqrt{d}$. $\square$

References

[B] K. Ball, An elementary introduction to modern convex geometry. Flavors of geometry, 1–58, (MSRI. Publ.), **31**, Cambridge Univ. Press, Cambridge, 1997.

[BV] E. Bombieri and J. D. Vaaler, On Siegel's Lemma, *Inventiones Mathematicae*, **73** (1983), 11–32.

[P] G. Pisier, Factorization of linear operators and geometry of Banach spaces, CBMS Regional Conference Series in Mathematics, **60**. AMS, Providence, RI, 1986.

[Va] J. D. Vaaler, A geometric inequality with applications to linear forms, *Pacif. J. Math*, **83**, No 2, (1979), 543–553.