The complex plank problem

Keith M. Ball*

Abstract

It is shown that if $(v_j)_1^n$ is a sequence of norm 1 vectors in a complex Hilbert space and $(t_j)_1^n$, a sequence of non-negative numbers satisfying

$$\sum t_j^2 = 1,$$

then there is a unit vector z for which

$$|\langle v_j, z\rangle| \geq t_j$$

for every j. The result is a strong, complex analogue of the author's real plank theorem.

Introduction

The plank theorem, proved by the author in [1], states (slightly more than) the following:

Theorem 1 *If $(\phi_j)_1^\infty$ is a sequence of norm 1 linear functionals on a (real) Banach space X and $(t_j)_1^\infty$ is a sequence of non-negative numbers whose sum is less than 1, then there is a unit vector x in X for which*

$$|\phi_j(x)| > t_j$$

for every j.

*Supported in part by NSF DMS-9257020 and British Council Alliance PN 98.100.
2000 Mathematics Subject Classification: Primary 52C99; Secondary 51M16

The plank theorem can be viewed as a generalisation of the Hahn-Banach Theorem, as a sharp quantitative version of the uniform boundedness principle, or as a geometric "pigeon-hole" principle. For Hilbert space, the statement was proved much earlier by Bang [4] whose article answered the notorious plank problem of Tarski. The general case was conjectured by Bang.

The condition that the t_j add up to 1 is obviously sharp, if X is the space ℓ_1 and the ϕ_j are standard basis vectors of the dual ℓ_∞. For other spaces, for example Hilbert space, one might expect to be able to improve upon this condition. If the ϕ_j are orthonormal, then it is possible to "beat" any sequence (t_j) for which

$$\sum t_j^2 = 1.$$

So, in particular, with n vectors, one might hope to beat the sequence $(1/\sqrt{n})_{i=1}^n$. However, nothing remotely this strong is possible.

Consider $2n$ unit vectors in the plane, distributed uniformly around the circle, and call them $\phi_1, \ldots, \phi_{2n}$. Clearly, for any unit vector x in the plane, there is a j for which

$$|\langle \phi_j, x \rangle| \leq \sin(\pi/2n)$$

and this is barely an improvement upon the result given by Theorem 1. (The example was described using $2n$ vectors, but since they appear in opposite pairs, there are really only n.)

The purpose of this article is to show that the situation changes dramatically, if we replace real scalars with complex ones. The main theorem of the article is the following.

Theorem 2 *Let $(v_j)_1^n$ be a sequence of norm 1 vectors in a complex Hilbert space and $(t_j)_1^n$, a sequence of non-negative numbers satisfying*

$$\sum t_j^2 = 1.$$

Then there is a unit vector z for which

$$|\langle v_j, z \rangle| \geq t_j$$

for every j.

In between the publication of [1] and the present, there were a number of other plank-type theorems discovered; in particular, Nazarov [6] found an elegant and powerful approach to coefficient problems in harmonic analysis which is very much in the same spirit as Theorem 1 and the work of Bang. A delightful analogue of Nazarov's Theorem, in a non-commutative setting, was found by Lust-Piquard [5]. A survey of some of this material appears in [3]. Nazarov's main theorem is as follows.

Theorem 3 *Let f_j be unit functions in L_1 which satisfy*

$$\left\| \sum_j a_j f_j \right\| \leq M \left(\sum_j a_j^2 \right)^{1/2}$$

for some M, and all sequences (a_j). Let (t_j) be a sequence of positive numbers with

$$\sum_j t_j^2 = 1.$$

Then there is a function g in L_∞ with norm at most $15M^2$ and

$$|\langle f_j, g \rangle| \geq t_j$$

for every j.

On the face of it, the result of this article, Theorem 2, may appear to have more in common with Nazarov's Theorem than with Theorem 1 (since the former imposes only quadratic control on the t_j). In fact, the arguments used here have something in common with those of the author's previous article but rather less connection with Nazarov's. However, the important point about Theorem 2 is that it is a genuinely complex result: among other things, the proof involves the maximum modulus principle for holomorphic functions.

The question answered by this article was raised by D. Symes, [7], in connection with the Hopenwasser conjecture concerning the Jacobsen radical of a CSL algebra; but at present, the implications in this direction are still unclear. Similar issues turn up in robust control theory. But perhaps the most intriguing aspect of this theorem is the intrinsic information it provides concerning the geometry of the sphere in $\mathbf{C}^n$.

Some years after the publication of Theorem 1 the author used Bang's Theorem (or more precisely, its proof) to find efficient lattice packings of spheres in high dimensional Euclidean spaces: [2]. An obvious question is whether the complex theorem can be used to improve the sphere-packing bound. It turns out, perhaps surprisingly, that the estimate given by the complex result is asymptotically identical to the earlier one.

For the sake of clarity, most of the ensuing discussion will deal with the special case of Theorem 2 in which

$$t_j = 1/\sqrt{n}$$

for each j. The final section of the article describes the small modifications needed to handle the general case. Section 1 contains a reformulation of the problem which provides the starting point for the proof, as well as a discussion which is intended to motivate the rest of the argument. The main argument is found in Section 2.

1 Preliminaries

Any sequence (v_j) of n (unit) vectors in a complex Hilbert space can be regarded, for the present purposes, as belonging to $\mathbf{C}^n$. Let A denote the $n \times n$ matrix with these vectors as rows. The aim is to find a unit vector z satisfying

$$|(Az)_j| \geq 1/\sqrt{n}$$

for each j. Equivalently, we look for a vector z of norm $\sqrt{n}$ for which

$$|(Az)_j| \geq 1$$

for each j.

Let $\mathbf{1}$ denote the vector $(1, 1, \ldots, 1)$ of length n. Any z of norm $\sqrt{n}$ can be written as $U\mathbf{1}$ for some (indeed many) unitary matrices U. One way to look for z is thus to look for a unitary matrix U satisfying

$$|(AU\mathbf{1})_j| \geq 1$$

for each j. If w_j is chosen to be $1/(AU\mathbf{1})_j$ for each j and W is the diagonal matrix with diagonal entries $w_1, \ldots, w_n$ then $|w_j| \leq 1$ for each j and

$$WAU\mathbf{1} = \mathbf{1}. \tag{1}$$

The last statement says that the vector $\mathbf{1}$ is an eigenvector of the matrix WAU, with eigenvalue 1. An analogy with the argument used to prove Theorem 1 suggests that one might hope that W and U can be chosen in such a way that the matrix WAU is positive semi-definite, Hermitian. If so, then WAU is the positive square-root of the matrix WAA^*W^* and equation (1) is equivalent to the equation

$$WAA^*W^*\mathbf{1} = \mathbf{1}.$$

The crucial advantage of using this equation is that the matrix U has disappeared.

Since the rows of A are unit vectors, the matrix $H = AA^*$ is a positive semi-definite Hermitian matrix whose diagonal entries are all equal to 1: a Gram matrix. Theorem 2 (with $t_j = 1/\sqrt{n}$) is thus a consequence of the following.

Theorem 4 *Let $H = (h_{jk})$ be an $n \times n$ (complex) Gram matrix. Then there are complex numbers $w_1, \ldots, w_n$ of absolute value at most 1 for which*

$$w_j \sum_k h_{jk} \bar{w}_k = 1$$

for every j.

The proof of this theorem will be the subject of the next section. But in order to motivate the proof somewhat, it is worth comparing Theorem 4 with the lemma of Bang [4] used in the real case.

Lemma 5 *Let $H = (h_{jk})$ be an $n \times n$ real Gram matrix and $r_1, \ldots, r_n$ a sequence of positive numbers. Then there are signs $\epsilon_1, \ldots, \epsilon_n$ for which*

$$\epsilon_j r_j \sum_k h_{jk} r_k \epsilon_k \geq r_j^2 \tag{2}$$

for every j.

Bang's Lemma is proved by choosing the signs so as to maximise the expression

$$\sum_{jk} \epsilon_j r_j h_{jk} r_k \epsilon_k.$$

A similar argument gives the following complex analogue.

Lemma 6 *Let $M = (m_{jk})$ be an $n \times n$ (complex) Hermitian matrix. Suppose that real numbers $\theta_1, \ldots, \theta_n$ are chosen to maximise the sum*

$$\sum_{jk} e^{i\theta_j} m_{jk} e^{-i\theta_k}.$$

Then for every j

$$e^{i\theta_j} \sum_k m_{jk} e^{-i\theta_k}$$

is real and at least as large as m_{jj}.

Proof Let

$$F(\theta_1, \ldots, \theta_n) = \sum_{jk} e^{i\theta_j} m_{jk} e^{-i\theta_k}.$$

The partial derivative of F with respect to θ_j is

$$2\Re \left(i e^{i\theta_j} \sum_k m_{jk} e^{-i\theta_k} \right).$$

This is zero for each j and so

$$e^{i\theta_j}\sum_k m_{jk}e^{-i\theta_k}$$

is real for each j.

The second partial derivative of F with respect to θ_j is

$$-2\Re\left(e^{i\theta_j}\sum_k m_{jk}e^{-i\theta_k}\right)+2m_{jj}.$$

Since this is non-positive at a maximum

$$e^{i\theta_j}\sum_k m_{jk}e^{-i\theta_k}\geq m_{jj}$$

for each j. □

Lemma 6 looks as though it might provide a way to tackle Theorem 4. Given a Gram matrix $H=(h_{jk})$, and positive numbers $r_1,\ldots,r_n$, one may apply Lemma 6 to the matrix (m_{jk}) given by

$$m_{jk}=r_j h_{jk} r_k$$

so as to find a sequence $\theta_1,\ldots,\theta_n$, so that for each j,

$$s_j=e^{i\theta_j}r_j\sum_k h_{jk}r_k e^{-i\theta_k}$$

is real and no less than r_j^2.

Now regard the vector $(s_1,\ldots,s_n)$ as a function of $(r_1,\ldots,r_n)$. Clearly, if any r_j is zero, then the corresponding s_j is also zero. So, by renormalising the vectors it is possible to regard the function as a map from the simplex to itself, which preserves faces. If the map is continuous, then Brouwer's fixed point theorem guarantees a collection of r_j for which

$$s_j=1$$

for every j. The second part of Lemma 6 now ensures that these r_j all have absolute value at most 1.

The problem with this approach is that the vector s is constructed by maximising over the θ_j and there is thus no reason why it should be a continuous function of r. For $n = 2$ it is trivially continuous, and oddly enough it is continuous for $n = 3$. But for $n \geq 4$ there seems to be no obvious way to make a continuous selection from among possible maxima.

2 The proof of Theorem 4

The purpose of this section is to show that if $H = (h_{jk})$ is an $n \times n$ (complex) Gram matrix, then there are complex numbers $w_1, \ldots, w_n$ of absolute value at most 1 so that

$$w_j \sum_k h_{jk} \bar{w}_k = 1 \tag{3}$$

for every j. From now on, we assume, as we may, that H is positive definite (not merely semi-definite).

The equations (3) can be regarded as a system of $2n$ quadratic equations in the real and imaginary parts of the w_j. The real and imaginary parts of the expressions

$$w_j \sum_k h_{jk} \bar{w}_k$$

are obviously continuous functions of the real and imaginary parts of the w_j, but these functions do not seem to be "pinned down" in any ways that allow the use of a fixed point theorem. However, the existence of a solution can be shown directly using a variational argument.

Lemma 7 *Suppose $M = (m_{jk})$ is a positive definite Hermitian matrix and the complex numbers $u_1, \ldots, u_n$ are chosen so as to minimise the double sum*

$$\sum_{jk} u_j m_{jk} \bar{u}_k,$$

subject to the condition

$$\prod |u_j| = 1.$$

Then for some positive λ*,*

$$u_j \sum_k m_{jk} \bar{u}_k = \lambda^2$$

for every j*.*

The proof of this lemma is an immediate application of Lagrange multipliers. Clearly, once the u_j have been found using the lemma, one can replace them by $w_j = u_j/\lambda$ so as to get 1 on the right hand side, instead of λ^2.

The problem with this construction of the w_j is that it forces us to use a *minimiser* of the double sum, because the sum will not be bounded above, on the (real) variety of points satisfying

$$\prod |u_j| = 1.$$

But Lemma 6 of the last section shows that we should be trying to *maximise* the double sum over sequences with specified absolute values, if we are to find w_j whose absolute values are small.

To deal with this problem, we apply Lemma 7, not to the Gram matrix H, but to its inverse H^{-1}, (roughly in the hope that a minimum for H^{-1} will be converted into a maximum for H). Observe that if Lemma 7 yields u_j so that

$$u_j \sum_k (H^{-1})_{jk} \bar{u}_k = 1$$

for each j, then the numbers

$$w_j = \bar{u}_j^{-1}$$

will satisfy

$$w_j \sum_k (H)_{jk} \bar{w}_k = 1, \tag{4}$$

for each j. Moreover, as long as the u_j have been chosen using Lemma 7, it is automatic that whenever c_j are complex numbers for which $\prod |c_j| = 1$, then

$$\sum_{jk} c_j u_j (H^{-1})_{jk} \bar{u}_k \bar{c}_k \geq n. \tag{5}$$

The problem is to show that $|w_j| \leq 1$ for each j.

To clarify the problem, define a new matrix M by

$$m_{jk} = w_j h_{jk} \bar{w}_k$$

and observe that its inverse will be given by

$$(M^{-1})_{jk} = u_j (H^{-1})_{jk} \bar{u}_k.$$

Then for each j

$$|w_j|^2 = m_{jj}$$

and in view of equations (4) and (5) it suffices to prove the following lemma.

Lemma 8 *Suppose that M is a positive definite Hermitian matrix satisfying*

1. $M\mathbf{1} = \mathbf{1}$
2. *Whenever $c = (c_j)$ is a complex vector for which $\prod |c_j| = 1$, then*

$$cM^{-1}\bar{c} \geq n.$$

Then $m_{jj} \leq 1$ for each j.

The second condition states that the sum

$$\sum_{jk} c_j (M^{-1})_{jk} \bar{c}_k$$

achieves its minimum value over the set of vectors satisfying $\prod |c_j| = 1$, at the vector $\mathbf{1}$. Clearly, it would be convenient to use only the fact that there is a *local* minimum at $\mathbf{1}$, since this property has a simple analytic characterisation. Unfortunately, the lemma is false (even in dimension $n = 3$) if global minimum is replaced by local. For this reason the proof of Lemma 8 requires a "jump" away from $\mathbf{1}$.

Proof of Lemma 8 Observe that if $c = bM$ then

$$cM^{-1}\bar{c} = bM\bar{b}.$$

The second condition of the Lemma can thus be reexpressed in terms of M as follows: whenever $b = (b_j)$ is a complex vector for which $\prod |(bM)_j| = 1$, then

$$bM\bar{b} \geq n.$$

Equivalently, for any vector b with

$$bM\bar{b} = n$$

$\prod |(bM)_j| \leq 1$.

Suppose, for a contradiction, that some diagonal entry of M is too large; $m_{11} = r > 1$, say. Let $z_2, z_3, \ldots, z_n$ be the remaining entries in the first row of M, so that M looks like

$$\begin{pmatrix} r & z_2 & z_3 & \ldots & z_n \\ \bar{z}_2 & m_{22} & \ldots & & \\ \bar{z}_3 & m_{32} & \ldots & & \\ \vdots & & & & \\ \bar{z}_n & & & & \end{pmatrix}$$

We shall examine the effect of M on vectors of the form

$$b = (z, 1, 1, \ldots, 1)$$

for different complex numbers z. Notice that since $M\mathbf{1} = \mathbf{1}$

$$\sum_2^n z_k = 1 - r$$

and hence

$$\sum_2^n \bar{z}_k = 1 - r.$$

Similarly, for each $k \geq 2$,

$$\sum_2^n m_{jk} = 1 - z_k.$$

Hence

$$\prod |(bM)_k| = |1 - r + rz|. \prod_2^n |1 - z_k + z_k z|. \tag{6}$$

The quantity $bM\bar{b}$ can also be expressed in terms of the z_j,

$$\begin{aligned} bM\bar{b} &= r|z|^2 + 2\Re\left(z \sum_2^n z_j\right) + \sum_2^n (1 - z_j) \\ &= r|z|^2 + 2(1 - r)\Re z + r - 2 + n \end{aligned}$$

The key will be to "dehomogenise" the hypothesis on M by examining only those numbers z for which $bM\bar{b} = n$: ie. for which

$$r|z|^2 + 2(1 - r)\Re z + r - 2 = 0.$$

These are the numbers that lie on the circle with radius $1/r$ and centre $1 - 1/r$. To obtain a contradiction it suffices to show that there is a z on this circle where the product (6) is more than 1.

The first factor of the product, $|1 - r + rz| = r|z - (1 - 1/r)|$, is also constant, with value 1, on the circle in question. So it suffices to show that there is a point z on the circle where

$$\prod_2^n |1 - z_j + z_j z| > 1.$$

As a function on the *circle*, this product may have a local maximum at 1, where the value is 1. But the product is the absolute value of a polynomial,

$$p(z) = \prod_2^n (1 - z_j + z_j z).$$

So, by the maximum modulus principle, it suffices to find a point in the corresponding *disc*, where the product is greater than 1. But

$$p'(1) = 1 - r < 0$$

so as a function on the *disc*, $|p(z)|$ cannot have a local maximum at 1. □

It is perhaps worth mentioning that the variational construction of the w_j, used in the proof, could be described without reference to polar decomposition or to the Gram matrix associated with A. However, the proof that the construction works, automatically brings in the Gram matrix.

3 The general case

Let A be an $n \times n$ complex matrix whose rows are unit vectors and $t = (t_j)$ be a sequence of positive numbers with

$$\sum_j t_j^2 = 1.$$

The aim is to find a complex vector z of length at most 1, satisfying

$$|(Az)_j| \geq t_j$$

for each j.

In this case z will be of the form Ut for some unitary matrix U, and by the argument above it will suffice to prove the following.

Theorem 9 *Let $H = (h_{jk})$ be an $n \times n$ (complex) Gram matrix and t_j a sequence of positive numbers. Then there are complex numbers $w_1, \ldots, w_n$ with $|w_j| \leq t_j$ for each j and*

$$w_j \sum_k h_{jk} \bar{w}_k = t_j^2$$

for every j.

To find the w_j, minimise the expression

$$\sum_{jk} u_j (H^{-1})_{jk} \bar{u}_k$$

subject to the condition

$$\prod_j |u_j|^{t_j^2} = 1.$$

The rest of the argument is the same as that above, except that one has to look for a large value of

$$\prod_2^n |t_j^2 - z_j + z_j z|^{t_j^2}.$$

This is not necessarily the absolute value of a holomorphic function, but its logarithm is subharmonic, and hence satisfies a maximum principle.

References

[1] K. M. Ball, The plank problem for symmetric bodies, Invent. Math. 104 (1991), 535-543.

[2] K. M. Ball, A lower bound for the optimal density of lattice packings, Duke Math. J. 68 (1992), 217-221.

[3] K. M. Ball, in Handbook of Banach Spaces, W. B. Johnson and J. Lindenstrauss eds. Elsevier (2000), to appear.

[4] T. Bang, A solution of the "Plank problem", Proc. Amer. Math. Soc. **2** (1951), 990-993.

[5] F. Lust-Piquard, On the coefficient problem: a version of the Kahane-Katznelson-DeLeeuw Theorem for spaces of matrices, J. Funct. Anal. **149** (1997), 352-376.

[6] F. Nazarov, The Bang solution of the coefficient problem, Algebra i Analiz **9** (1997), 272-287.

[7] D. Symes, Personal communication.

K. M. Ball
Department of Mathematics,
University College London,
Gower Street,
LONDON WC1E 6BT
ENGLAND.
e-mail: kmb@math.ucl.ac.uk