A PYTHAGOREAN INEQUALITY

RUSSELL M. REID

(Communicated by Palle E. T. Jorgensen)

Abstract. Let \{v_1, v_2, v_3, \ldots \} be a sequence of elements of a Hilbert space, and suppose that (one or both of) the inequalities \( d^2 \sum a_i^2 \leq \| \sum a_i v_i \|^2 \leq D^2 \sum a_i^2 \) hold for every finite sequence of scalars \( \{a_i\} \). If an element \( v_0 \) is adjoined to \( \{v_i\} \), then the resulting set satisfies (one or both of) \( d^2 \sum a_i^2 \leq \| \sum a_i v_i \|^2 \leq D^2 \sum a_i^2 \), where, denoting the norm of \( v_0 \) by \( r \) and its distance from the closed linear span of the \( v_i \) by \( \delta \),

\[
d_0^2 = d^2 + \frac{1}{2} \left( r^2 - d^2 - \sqrt{(r^2 + d^2)^2 - 4d^2\delta^2} \right)
\]

and

\[
D_0^2 = D^2 + \frac{1}{2} \left( r^2 - D^2 + \sqrt{(r^2 + D^2)^2 - 4D^2\delta^2} \right).
\]

Both bounds are best possible. If \( v_0 \) is in the span of the original set, the expressions above simplify to \( d_0 = 0 \) and \( D_0 = D^2 + r^2 \). If the original set is a single unit vector \( v_1 \), so \( d = D = 1 \), and if \( v_0 \perp v_1 \) is a unit vector so \( \delta = 1 \), then the above is \( (a^2 + b^2) \leq \|av_0 + bv_1\|^2 \leq (a^2 + b^2) \), the Pythagorean Theorem.

Several consequences are deduced. If \( v_i \) are unit vectors, \( \sum a_i^2 = 1 \), and \( \delta_i \) is the distance from \( v_i \) to the span of its predecessors (so that the volume of the parallelepiped spanned by the \( v_i \) is \( V_n = \delta_1 \delta_2 \cdots \delta_n \)), the above result is used to show that \( \| \sum_{i=0}^n a_i v_i \| \geq V_n/2^{n/2} \).

1. An Estimate of the Norm of a Linear Combination

This work derives an estimate of the norm of a linear combination in terms of coordinate norms, the resulting inequalities being valid in finite or infinite dimensions. For simplicity of statement and proof two separate theorems are stated. Despite the obvious symmetry, Theorem 1 is likely the more useful, for there are fewer tools which provide nonzero lower bounds for the norm of a sum.

Theorem 1. Let \( \{v_i\}, \ i = 1, 2, 3, \ldots \), be a sequence of elements of a Hilbert space \( H \), and suppose that the inequality

\[
d^2 \sum a_i^2 \leq \| \sum a_i v_i \|^2
\]

Received by the editors June 15, 1993.
1991 Mathematics Subject Classification. Primary 42A70, 15A42.
Key words and phrases. Norm inequality, Gram matrix, nonharmonic Fourier series.
holds for every finite equation of scalars \( \{a_i\} \). If an element \( v_0 \) is adjoined to \( \{v_i\} \), then the resulting set satisfies, for every finite sequence of scalars,

\[
d_0^2 \sum a_i^2 \leq \left\| \sum a_i v_i \right\|^2
\]

where, denoting the norm of \( v_0 \) by \( r \) and its distance from the closed linear span of the \( v_i \) by \( \delta \),

\[
d_0^2 = d^2 + \frac{1}{2} \left( r^2 - d^2 - \sqrt{(r^2 + d^2)^2 - 4d^2\delta^2} \right)
\]

If the bound \( d \) is sharp, so is \( d_0 \).

**Theorem 2.** Let \( \{v_i\}, i = 1, 2, 3, \ldots \), be a sequence of elements of a Hilbert space \( H \), and suppose that the inequality

\[
\left\| \sum a_i v_i \right\|^2 \leq D^2 \sum a_i^2
\]

holds for every finite sequence of scalars \( \{a_i\} \). If an element \( v_0 \) is adjoined to \( \{v_i\} \), then the resulting set satisfies, for every finite sequence of scalars,

\[
\left\| \sum a_i v_i \right\|^2 \leq D_0^2 \sum a_i^2
\]

where, denoting the norm of \( v_0 \) by \( r \) and its distance from the closed linear span of the \( v_i \) by \( \delta \),

\[
D_0^2 = D^2 + \frac{1}{2} \left( r^2 - D^2 + \sqrt{(r^2 + D^2)^2 - 4D^2\delta^2} \right)
\]

If the bound \( D \) is sharp, so is \( D_0 \).

**Proof.** We prove only Theorem 1; Theorem 2 is similar.

Let \( v_0 \) be given, with its norm denoted by \( r \) and its distance from the closed linear span of the original set \( \{v_i\} \) denoted by \( \delta \).

Let \( \{a_0, a_1, \ldots, a_n\} \) is a finite sequence of scalars. If \( a_0 = 0 \), the theorem follows from the observation that \( d_0 \) never exceeds \( d \): letting \( \delta = r \sin \theta \), observe that the radical becomes \( \sqrt{(r^2 + d^2)^2 + 4d^2r^2\cos^2\theta} \), and therefore the term in parentheses in (3) is negative.

Therefore, suppose that \( a_0 \neq 0 \). It will be convenient to use \( (\cdot, \cdot) \) to represent the \( l_2 \) inner product and \( (\cdot, \cdot) \) to represent the Hilbert space inner product. Inequality (1) is

\[
d^2 \leq \frac{(G a, a)}{(a, a)}
\]

where \( a \) represents the sequence of coefficients \( \{a_1, a_2, \ldots, a_n\} \) and \( G \) is the Gram matrix whose \( (i, j) \)th entry is \( \langle v_i, v_j \rangle \).

Consider a linear combination \( a_0 v_0 + a_1 v_1 + \cdots + a_n v_n \) with \( a_0 \) nonzero. Let \( V \) be the span and \( G \) be the Gram matrix of the finite set \( \{v_1, v_2, \ldots, v_n\} \), and let \( G_0 \) be the Gram matrix of \( \{v_0, v_1, \ldots, v_n\} \). \( G \) and \( G_0 \) are Hermitian and nonnegative definite; for simplicity of exposition suppose that the Hilbert space \( H \) is over the field of real scalars so that its inner product is symmetric.

Inequality (1) asserts that the smallest eigenvalue of \( G \) is no smaller than \( d^2 \); it is desired to show that no eigenvalue of \( G_0 \) is smaller than \( d_0^2 \).
Eigenvectors of $G$ form an orthonormal basis for $\mathbb{R}^n$; suppose the eigenvalues of $G$ are $d^2 \leq \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots \leq \lambda_n$, with corresponding eigenvectors $\alpha_1, \alpha_2, \ldots, \alpha_n$. Let $e_1, e_2, \ldots, e_n$ be the corresponding orthonormal elements of $H$; for example, if $\alpha_1 = (c_1, c_2, \ldots, c_n)$, then

$$e_1 = (c_1 v_1 + c_2 v_2 + \cdots + c_n v_n)/\sqrt{\lambda_1}.$$ 

The argument that follows is not materially affected by the possibility of repeated eigenvalues; note here that if some $\lambda_j = \lambda_{j+1}$, the corresponding $\alpha_j$ and $\alpha_{j+1}$ are not unique and neither are $e_j$ and $e_{j+1}$. The reader who ignores the occasional remark concerning repeated eigenvalues will have lost little in what follows. Three claims:

(1) If $v_0$ is orthogonal to $e_2, e_3, \ldots, e_n$, so that $v_0 = (r \cos \theta)e_1 + (r \sin \theta)w$, for $w$ a unit vector in $V^\perp$, then the smallest eigenvalue of $G_0$ is

$$\lambda_1 + \frac{1}{2} \left( r^2 - \lambda_1 - \sqrt{(r^2 + \lambda_1)^2 - 4\lambda_1 r^2 \sin^2 \theta} \right)$$

$$= \lambda_1 + \frac{1}{2} \left( r^2 - \lambda_1 - \sqrt{(r^2 - \lambda_1)^2 + 4\lambda_1 r^2 \cos^2 \theta} \right).$$

(2) The above is the worst case: if $v_0 = (r \cos \theta)v + (r \sin \theta)w$ for unit vectors $v \in V$ and $w \in V^\perp$, the smallest eigenvalue of $G_0$ is not smaller than the expression (8) above.

(3) The expression (8) above is an increasing function of $\lambda_1$ and of $\sin \theta$ so that establishing the above two statements establishes the theorem: computing with a finite subset of the original $\{v_i\}$ produces values for $r \sin \theta$ and $\lambda_1$ which are overestimates of $\delta$ and $d^2$.

The proof of (2) actually includes the result (1), but the exposition is simpler if the two are separated.

Suppose then that $v_0 = (r \cos \theta)e_1 + (r \sin \theta)w$, for a unit vector $w$ in $V^\perp$. The Gram matrix $G_0$ is

$$G_0 = \begin{bmatrix}
\langle v_0, v_0 \rangle & \langle v_0, v_1 \rangle & \langle v_0, v_2 \rangle & \cdots & \langle v_0, v_n \rangle \\
\langle v_1, v_0 \rangle & \langle v_1, v_1 \rangle & \langle v_1, v_2 \rangle & \cdots & \langle v_1, v_n \rangle \\
\langle v_2, v_0 \rangle & \langle v_2, v_1 \rangle & \langle v_2, v_2 \rangle & \cdots & \langle v_2, v_n \rangle \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\langle v_n, v_0 \rangle & \langle v_n, v_1 \rangle & \langle v_n, v_2 \rangle & \cdots & \langle v_n, v_n \rangle 
\end{bmatrix}.$$ 

Notice that $n - 1$ of the eigenvectors of $G_0$ are the eigenvectors $\alpha_2, \alpha_3, \ldots, \alpha_n$ of $G$ with a zero adjoined to the first position, with the same eigenvalues; repeated eigenvalues are again repeated with the same multiplicity. For example, if $\alpha_k = (b_1, b_2, \ldots, b_n)$, then

$$G = \begin{bmatrix}
\langle v_0, v_0 \rangle & \langle v_0, v_1 \rangle & \langle v_0, v_2 \rangle & \cdots & \langle v_0, v_n \rangle \\
\langle v_1, v_0 \rangle & \langle v_1, v_1 \rangle & \langle v_1, v_2 \rangle & \cdots & \langle v_1, v_n \rangle \\
\langle v_2, v_0 \rangle & \langle v_2, v_1 \rangle & \langle v_2, v_2 \rangle & \cdots & \langle v_2, v_n \rangle \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\langle v_n, v_0 \rangle & \langle v_n, v_1 \rangle & \langle v_n, v_2 \rangle & \cdots & \langle v_n, v_n \rangle 
\end{bmatrix} \begin{bmatrix}
0 \\
b_1 \\
b_2 \\
\vdots \\
b_n 
\end{bmatrix} = \begin{bmatrix}
0 \\
\lambda_k b_1 \\
\lambda_k b_2 \\
\vdots \\
\lambda_k b_n 
\end{bmatrix}.$$
where the zero remains in the first position because \( v_0 \) is orthogonal to \( e_k = b_1 v_1 + \cdots + b_n v_n \).

Now suppose that \( x_1 = (c_1, c_2, \ldots, c_n) \). The two remaining eigenvectors are in the form \( v = (x, c_1, c_2, \ldots, c_n) \). To determine \( x \), compute the product \( G_0 v \) directly, observing that, because \( \sum_j c_j \langle v_i, v_j \rangle \) is the \( i \)th row of \( G_{x_1} \),

\[
\langle v_i, v_0 \rangle = \left\langle v_i, \frac{r \cos \theta}{\sqrt{\lambda_1}} \sum_j c_j v_j \right\rangle = \frac{r \cos \theta}{\sqrt{\lambda_1}} \sum_j c_j \langle v_i, v_j \rangle = r \cos \theta \sqrt{\lambda_1} c_i.
\]

Similarly,

\[
\sum_i c_i \langle v_0, v_i \rangle = \left\langle v_0, \sum_i c_i v_i \right\rangle = \left\langle v_0, \sqrt{\lambda_1} e_1 \right\rangle = r \cos \theta \sqrt{\lambda_1}
\]

so that

\[
\begin{bmatrix}
\langle v_0, v_0 \rangle & \langle v_0, v_1 \rangle & \langle v_0, v_2 \rangle & \cdots & \langle v_0, v_n \rangle \\
\langle v_1, v_0 \rangle & \langle v_1, v_1 \rangle & \langle v_1, v_2 \rangle & \cdots & \langle v_1, v_n \rangle \\
\langle v_2, v_0 \rangle & \langle v_2, v_1 \rangle & \langle v_2, v_2 \rangle & \cdots & \langle v_2, v_n \rangle \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\langle v_n, v_0 \rangle & \langle v_n, v_1 \rangle & \langle v_n, v_2 \rangle & \cdots & \langle v_n, v_n \rangle
\end{bmatrix}
= \begin{bmatrix}
x \\
\c_1 \\
\c_2 \\
\vdots \\
\c_n
\end{bmatrix}
\]

Eigenvectors correspond to the values of \( x \) which are solutions of the quadratic equation

\[
x (r \cos \theta \sqrt{\lambda_1} + \lambda_1) = r^2 x + r \cos \theta \sqrt{\lambda_1},
\]

and the eigenvalues are

\[
x r \cos \theta \sqrt{\lambda_1} + \lambda_1.
\]

Choosing the smaller value of \( x \), which produces an eigenvalue smaller than \( \lambda_1 \) which must therefore be the smallest eigenvalue of \( G_0 \), and replacing \( \cos^2 \theta \) with \( 1 - \sin^2 \theta \) gives (8). It is not difficult to show that both eigenvalues are the same just when \( r = \delta_0 \), which happens when \( v_0 \) is in \( V^\perp \), in which case it is easy to see that \( d_0 = d \).

The second claim above can be approached the same way. Suppose that \( v_0 = \sum_{i=1}^n s_i e_i + \delta_0 w \), for a unit vector \( w \) in \( V^\perp \), where \( r \cos \theta = \sqrt{\sum s_i^2} \) and \( r \sin \theta = \delta_0 \), the distance from \( v_0 \) to the span of the finite set \( \{v_1, v_2, \ldots, v_n\} \). The Gram matrix \( G_0 \) still has the form (9), but the expressions (11) and (12) must be generalized. Denoting the coordinates of the \( \alpha_k \) by \( c_{ki} \), so that \( \alpha_k = (c_{1k}, c_{2k}, \ldots, c_{nk}) \), and \( e_k = \frac{1}{\sqrt{\lambda_k}} \sum_{i=1}^n c_{ki} v_i \), the analog to (11) is, again using
the observation that \( \sum_{i=1}^{n} \langle v_i, v_k \rangle c_{ij} \) is the \( k \)th element of \( G \lambda_j = \lambda_j \alpha_j \),

\[
\langle v_0, v_k \rangle = \left\langle \sum_{j=1}^{n} s_j e_j + \delta_0 w, v_k \right\rangle = \sum_{j=1}^{n} s_j \langle e_j, v_k \rangle
\]

\[
= \sum_{j=1}^{n} s_j \left( \frac{1}{\sqrt{\lambda_j}} \sum_{i=1}^{n} c_{ij} v_i, v_k \right) = \sum_{j=1}^{n} \frac{s_j}{\sqrt{\lambda_j}} \sum_{i=1}^{n} \langle v_i, v_k \rangle c_{ij}
\]

\[
= \sum_{j=1}^{n} \frac{s_j}{\sqrt{\lambda_j}} \lambda_j c_{kj} = \sum_{j=1}^{n} s_j \sqrt{\lambda_j} c_{kj}.
\]

If any \( s_j \) are zero, the corresponding eigenvalues \( \lambda_j \) are undisturbed, with eigenvectors as before with a zero adjoined in the first position. Repeated eigenvalues are repeated with the same multiplicity. If any of the \( s_j \) corresponding to a repeated eigenvalue \( \lambda_j \) are nonzero, it is simplest to note that one could choose the original \( e_j \) in such a way that one of them is parallel to the projection of \( v_0 \) into the associated eigenspace and the others are perpendicular. By that device, one may suppose that each nonzero \( s_j \) corresponds to a distinct eigenvalue. The eigenvectors corresponding to nonzero \( s_j \) have the form, summing over the nonzero \( s_j \) and supposing that there are \( m \) such,

\[
v = \left[ \begin{array}{c}
\sum_{j=1}^{m} \beta_j c_{1j} \\
\sum_{j=1}^{m} \beta_j c_{2j} \\
\vdots \\
\sum_{j=1}^{m} \beta_j c_{nj}
\end{array} \right] = \left[ \begin{array}{c}
x \\
\beta_1 \alpha_1 \\
\vdots \\
\beta_m \alpha_m
\end{array} \right].
\]

Inserting this into \( G_0 \), using (14) above only in the left column gives

\[
\left[ \begin{array}{c}
\sum_{j=1}^{m} s_j \sqrt{\lambda_j} c_{1j} \\
\sum_{j=1}^{m} s_j \sqrt{\lambda_j} c_{2j} \\
\sum_{j=1}^{m} s_j \sqrt{\lambda_j} c_{3j} \\
\vdots \\
\sum_{j=1}^{m} s_j \sqrt{\lambda_j} c_{nj}
\end{array} \right] = G.
\]

To compute the first entry in the product \( G_0 v \), it is convenient to use the leftmost form in (15) for \( v \). Remembering that the entries of \( \alpha_k \) are \( c_{kj} \) and using a subscript \( k \) outside parentheses to indicate the \( k \)th entry of a vector,

\[
(G_0 v)_0 = r^2 x + \sum_{k=1}^{n} \langle v_0, v_k \rangle \sum_{j=1}^{m} \beta_j c_{kj} = r^2 x + \sum_{j=1}^{m} \beta_j \sum_{k=1}^{n} c_{kj} \langle v_0, v_k \rangle
\]

\[
= r^2 x + \sum_{j=1}^{m} \beta_j \left( \sum_{k=1}^{n} c_{kj} v_k \right) = r^2 x + \sum_{j=1}^{m} \beta_j \langle v_0, \sqrt{\lambda_j} e_j \rangle
\]

\[
= r^2 x + \sum_{j=1}^{m} \beta_j s_j \sqrt{\lambda_j}.
\]
The rightmost form in equation (15) is more convenient for calculating the other entries of the product:

$$(G_0v)_k = x \sum_{j=1}^{m} s_j \sqrt{\lambda_j} c_{kj} + (\beta_1 G\alpha_1 + \beta_2 G\alpha_2 + \cdots + \beta_m G\alpha_m)_k$$

(18)

$$= x \sum_{j=1}^{m} s_j \sqrt{\lambda_j} c_{kj} + \sum_{j=1}^{m} \beta_j \lambda_j c_{kj} = \sum_{j=1}^{m} (xs_j \sqrt{\lambda_j} + \beta_j \lambda_j) c_{kj}$$

so that

$$G_0v = \begin{bmatrix}
\sum_{j=1}^{m} \beta_j s_j \sqrt{\lambda_j} + xr^2 \\
\sum_{j=1}^{m} (xs_j \sqrt{\lambda_j} + \beta_j \lambda_j) c_{1j} \\
\sum_{j=1}^{m} (xs_j \sqrt{\lambda_j} + \beta_j \lambda_j) c_{2j} \\
\vdots \\
\sum_{j=1}^{m} (xs_j \sqrt{\lambda_j} + \beta_j \lambda_j) c_{nj}
\end{bmatrix}.$$ 

(19)

In order for $v$ to be an eigenvector (with eigenvalue $\lambda$) it is sufficient that, separately for each $j$,

$$\beta_j \lambda = xs_j \sqrt{\lambda_j} + \beta_j \lambda_j$$

(20)

and that

$$\sum_{j=1}^{m} \beta_j s_j \sqrt{\lambda_j} + xr^2 = \lambda x.$$ 

(21)

Dividing (21) by $x$, in the $j$th term substituting the value for $x$ provided by solving (20) for $x$, gives an equation whose $m + 1$ solutions are eigenvalues of $G_0$:

$$\frac{s_1^2 \lambda_1}{\lambda - \lambda_1} + \frac{s_2^2 \lambda_2}{\lambda - \lambda_2} + \cdots + \frac{s_m^2 \lambda_m}{\lambda - \lambda_m} = \lambda - r^2.$$ 

(22)

Considering the left and right sides of this equation to be functions of $\lambda$, one can see the nature of the solutions graphically by sketching each function and looking for their intersections. The left side is a weighted sum of terms of the form

$$f(x) = \frac{a}{x - a};$$

(23)

so a sketch of equation (22) is easy (see Figure 1).

Figure 1. Eigenvalue intersections
It is straightforward to observe that, among sums having the form of the left side of equation (22) with fixed \( r \) and \( \delta \), and therefore a fixed value for \( \sqrt{r^2 - \delta^2} = \sum s_j^2 \), the one with the leftmost intersection with the line \( y = x - r^2 \) is the one with \( s_1 = \sqrt{r^2 - \delta^2} = r \cos \theta \) and all other \( s_i \) zero. One might perhaps remark here that in such a case, equation (22) is a quadratic equation for \( \lambda \), whose smaller root gives the result (8). For example, one may notice that the only intersection of two curves of the form (23) is at \( x = 0 \) and that (23) is an increasing function of \( a \) when \( x > 0 \), so that for all positive \( \lambda \),
\[
\frac{\lambda_1}{\lambda - \lambda_1} < \frac{\lambda_2}{\lambda - \lambda_2}.
\]

In the region \( 0 < \lambda < \lambda_1 \), all terms in equation (22) have the same sign; after normalizing one has the mean of a collection of functions, the smallest of which is the \( \lambda_1 \) term. The intercept of the line is \( (-r^2) \), and the intercept of the sum on the left is \( -\sum s_j^2 \); the two intercepts are identical if \( \delta_0 = 0 \), but otherwise the line is beneath the sum at \( x = 0 \); its intersection with any curve of the form of equation (22) is a minimum when that sum is minimum, which is when all \( s_j \) except the first are zero.

2. Applications

The two properties that characterize a Riesz basis for a Hilbert space are completeness and independence, where the required form of independence for a set \( \{f_1, f_2, f_3, \ldots \} \) is (see [1, 6]) that there exist a nonzero \( d \) and finite \( D \) such that, for every finite set of scalars \( \{a_i\} \),
\[
(24) \quad d^2 \sum a_i^2 \leq \left\| \sum a_if_i \right\|^2 \leq D^2 \sum a_i^2.
\]

Theorems 1 and 2 of this work combine to offer an improvement of a classical theorem of Levinson [3], which states that the completeness of a system \( \{e^{i\lambda_n t}\} \) in \( L^2[-A, A] \) is unaffected if some \( \lambda_n \) is replaced by another number. We add a restriction that the new exponential not be one of the originals, or else the resulting set is of course not independent.

**Theorem 3.** The completeness and independence of a system of complex exponentials \( \{e^{i\lambda_n t}\} \) in \( L^2[-A, A] \) are unaffected if some \( \lambda_n \) is replaced by any number distinct from the remaining \( \lambda_i \).

**Proof.** We show only independence; Levinson proved completeness (closure.)

If the given set of exponentials is independent in the sense of (24), then it must be minimal; no element of the set can be a linear combination of the others. If some \( \{e^{i\lambda_k t}\} \) is removed from the set, the resulting set is incomplete, or else the original set could not have been independent. The reduced set is independent; let \( d \) and \( D \) be its upper and lower bounds. A specified \( e^{i\lambda_k t} \) cannot be in the closed linear span of the others, since they are an incomplete set [4].

Let \( d_0 > 0 \) and \( D_0 < \infty \) be the values given by Theorem 1 and Theorem 2, and let \( \{f_1, f_2, f_3, \ldots \} \) be the new set of complex exponentials. Then the inequality (24) with the weakened bounds \( d_0 \) and \( D_0 \) obtains for any finite set of scalars \( \{a_0, a_1, \ldots, a_n\} \), showing that the new set is independent.
If the given set of complex exponentials is not independent, and still is not independent when some \( e^{ik_n t} \) is removed, it will also fail to be independent when a new element \( e^{ik_m t} \) is added. If the original set failed to be independent but becomes independent upon the removal of the term \( e^{ik_n t} \), that term can be approximated arbitrarily closely by linear combinations of the remaining elements; hence (see [4]) so can any other complex exponential \( e^{ik_m t} \), so that the modified set also fails to be independent. □  

We remark that if the original set of exponentials is independent but not complete, any exponential can be adjoined to it and the resulting set will still be independent.

A set of exponentials which is independent on \( L_2[-\pi, \pi] \) in the sense of (24) is an interpolating sequence; for any sequence \( \{c_i\} \in l_2 \) there is a \( \phi \) in \( L_2 \) whose Fourier Transform satisfies \( \hat{\phi}(\lambda_n) = c_n \). The least-norm interpolating function is the transform of a \( \phi \in L_2 \) which is in the span of the exponentials; if \( \phi = \sum a_n e^{ik_n t} \), \( c = \{c_i\} \), \( a = \{a_i\} \), and \( G \) is the Gram matrix of the exponentials, then \( Ga = c \). The square of the norm of \( \phi \) is, taking \( l_2 \) inner products,

\[
a^T G a = a^T c = (G^{-1} c)^T c = c^T G^{-1} c,
\]

so that \( 1/d \) is the norm of the mapping from \( l_2 \) to \( L_2 \) and hence, to the Paley-Wiener space \( P \). The expression for \( d_0 \) in Theorem 1 gives the increase in the norm of the interpolating function due to an additional interpolating point.

**Theorem 4.** Let \( v_0, v_1, \ldots, v_n \) be unit vectors in a Hilbert space, let \( \theta_j \) denote the angle between \( v_j \) and the subspace spanned by its predecessors \( \{v_0, v_1, \ldots, v_{j-1}\} \), and let \( \delta_j \) be the distance from \( v_j \) to that subspace. Let \( \{a_i\} \) be a set of scalars normalized so that \( \sum a_i^2 = 1 \). Then

\[
(1 - \cos \theta_1)(1 - \cos \theta_2) \cdots (1 - \cos \theta_n) \leq \left\| \sum_{i=0}^{n} a_i v_i \right\|^2
\]

and

\[
\left( \frac{\delta_1^2}{2} \right) \left( \frac{\delta_2^2}{2} \right) \cdots \left( \frac{\delta_n^2}{2} \right) \leq \left\| \sum_{i=0}^{n} a_i v_i \right\|^2,
\]

with similar upper bounds.

**Proof.** A vector \( v_0 \) trivially satisfies the hypotheses of the theorem as a singleton set, with \( d_0^2 = D_0^2 = 1 \). Suppose that \( v_0, v_1, \ldots, v_n \) are unit vectors. Using (8) with \( r = 1 \) and \( d_0 = 1 \) gives a new lower bound \( d_1^2 = 1 - \cos \theta_1 \). Using (8) again with \( r = 1 \) and \( d_1^2 = 1 - \cos \theta_1 \), denoting by \( \theta_2 \) the angle between \( v_2 \) and the span of the set of its predecessors \( \{v_0, v_1\} \), one obtains

\[
(1 - \cos \theta_1) + \frac{1}{2} \left[ \cos \theta_1 - \sqrt{\cos^2 \theta_1 + 4(1 - \cos \theta_1) \cos^2 \theta_2} \right].
\]

A simpler but less sharp bound is obtained by overestimating the square root; it is easy to prove that for \( a \) and \( x \) between 0 and 1,

\[
\sqrt{a^2 + 4(1 - a)x^2} \leq a + 2(1 - a)x.
\]
Using this in the bound (27) gives a weaker, but still positive, lower bound 
\((1 - \cos \theta_1)(1 - \cos \theta_2)\). Repeating and normalizing so \(\sum \alpha_i^2 = 1\) to simplify the appearance gives (25).

Replacing \(\cos \theta\) with \(\sqrt{1 - \delta_i^2}\) and overestimating the square root with \(\sqrt{1 - \delta_i^2} < 1 - \delta_i^2/2\) gives (26). The product of the \(\delta_i^2\) is the determinant of the Gram matrix and hence, the square of the volume of the parallelootope whose edges are the \(v_i\) (see [2, Eq. 8.8.13]), giving the form of (26) given in the abstract. ♦

REFERENCES