Math. Proc. Camb. Phil. Soc. (1985), **98**, 117
Printed in Great Britain

An elementary proof of Gleason's theorem

By ROGER COOKE,† MICHAEL KEANE‡ and WILLIAM MORAN§

† *Department of Philosophy, Delft University of Technology.*
‡ *Department of Mathematics, Delft University of Technology.*
§ *Department of Mathematics, University of Adelaide.*

(*Received* 3 *October* 1984)

Abstract

Gleason's theorem characterizes the totally additive measures on the closed subspaces of a separable real or complex Hilbert space of dimension greater than two. This paper presents an elementary proof of Gleason's theorem which is accessible to undergraduates having completed a first course in real analysis.

Introduction

Let H be a separable Hilbert space over the real or complex field. A (normalized) *state* on H is a function assigning to H the value 1, assigning to each closed subspace of H a number in the unit interval, and satisfying the following additivity property: If any given subspace is written as an orthogonal sum of a finite or countable number of subspaces, then the value of the state on the given subspace is equal to the sum of the values of the state on the summands. States should be thought of as 'quantum mechanical probability measures'; they play an essential role in the quantum mechanical formalism. For an exposition of these ideas we refer to Mackey [**5**].

Examples of normalized states are obtained by considering positive self-adjoint trace class operators with trace 1 on H. Such operators correspond to preparation procedures in quantum mechanics. If A is such an operator, then it is easy to see that we can define a state by associating to each one dimensional subspace generated by a unit vector $x \in H$ the inner product $\langle Ax, x\rangle$ and extending to subspaces of dimension greater than one by countable additivity. States of this type are called *regular states*.

In his course on the mathematical foundations of quantum mechanics (see [**5**]) Mackey proposed the following problem: determine the set of states on an arbitrary real or complex Hilbert space. This problem was solved by Gleason in [**2**], and the principal result, known as Gleason's theorem, states that every state on a real or complex Hilbert space of dimension greater than two is regular. Gleason's proof uses the representation theory of $O(3)$, and relies on an intricate continuity argument. Because of the role which Gleason's theorem plays in the foundations of quantum mechanics, there have been several attempts to simplify its proof. Using elementary methods, Bell proved a special case of the theorem, namely, that there exist no states on the closed subspaces of a Hilbert space of dimension greater than two taking only the values zero and one [**1**]. Kochen and Specker proved a similar result for states restricted to a finite number of closed subspaces [**4**]. Piron produced an elementary proof of Gleason's theorem for the special case that the state is *extreme* (i.e. assigns the value 1 to some one dimensional subspace) [**6**].

In this article we give an elementary proof of Gleason's theorem in full generality. Although this proof is longer than Gleason's proof, we believe that it contributes to the intuitive understanding of the underlying reasons for the validity of the theorem. The structure of the argument is as follows. In § 1 we show that it is enough to handle the case $H = \mathbb{R}^3$. This was part of Gleason's original argument, and is well understood; the essential difficulty of the proof is the treatment of the case $H = \mathbb{R}^3$. For this purpose it is convenient to study a certain class of real-valued functions on the unit sphere of $\mathbb{R}^3$, called frame functions. §§ 2 and 3 are devoted to an exposition of the properties of frame functions and the statement of the theorem in the case of $\mathbb{R}^3$ in terms of frame functions. § 3 also contains two 'warmup theorems' whose contents were essentially known to 19th century mathematicians. Coupled with a basic lemma in § 4 (essentially due to Gleason and Piron), they yield a new proof for the extreme case, which is given in § 5. In § 6 we show that a weak form of continuity in the general case follows from the result of § 5, and in § 7 we treat the general case. The proofs in §§ 2–7 are accessible to undergraduates who have completed a first course in real analysis.

1. *Reduction to* $H = \mathbb{R}^3$

Let H be a real or complex separable Hilbert space, and let L be the set of closed subspaces of H. If $A \in L$, and $B \in L$, then we write $A \perp B$ if A and B are orthogonal. For $A_i \in L$, $i \in I$, $\overline{\bigvee_{i\in I} A_i}$ denotes the smallest closed subspace containing A_i for all $i \in I$. If x is a vector in H, then $\bar{x}$ denotes the one dimensional subspace generated by x.

Definition. A function $p\colon L \to [0, 1]$ is called a *state* if for all sequences $\{A_i\}_{i=1}^{\alpha}$, $A_i \in L$, $i = 1 \ldots$; with $A_i \perp A_j$, for $i \neq j$:

$$p\left(\overline{\bigvee_{i=1}^{\alpha} A_i}\right) = \sum_{i=1}^{\alpha} p(A_i).$$

Definition. p is called *regular* if there exists a self-adjoint trace class operator A on H such that for all unit vectors $x \in H$

$$p(\bar{x}) = \langle Ax, x\rangle.$$

LEMMA. *The following statements are equivalent:*

(i) *p is regular.*

(ii) *There is a symmetric continuous bilinear form B on H such that*

$$p(\bar{x}) = B(x, x).$$

Moreover, both A and B are uniquely determined in this way by p.

Proof. See Halmos [**3**] (§§ 2 and 3). ∎

LEMMA. *If the restriction of p to every two-dimensional subspace of H is regular, then p is regular* (*the restriction need not be normalized*).

Proof. For each two-dimensional subspace E of H we can find a symmetric continuous bilinear form B_E such that $B_E(x, x) = p(\bar{x})$ $(x \in E, \|x\| = 1)$. For $\|x\| = \|y\| = 1$, choose a two-dimensional subspace $E(x, y)$ containing x and y and define

$$B(x,y) = B_E(x, y).$$

http://dx.doi.org/10.1017/S0305004100063313

It is straightforward to check that B can be uniquely extended to a symmetric continuous bilinear form on H, and that $p(\bar{x}) = B(x,x)$ ($\|x\| = 1$). ∎

We shall call a closed real-linear subspace of H *completely real* if the inner product on this real linear subspace takes only real values.

LEMMA. *If p is a state on a two-dimensional complex Hilbert space H, and if p is regular on every completely real subspace, then p is regular.*

Proof. We first show that there is a one-dimensional subspace $\bar{x}$ such that $p(\bar{x})$ is maximal. Put

$$M = \sup_{x \in H} p(\bar{x}).$$

Choose a sequence $x_n \in H$ such that $\lim_{n\to\infty} p(\bar{x}_n) = M$. By passing to a subsequence, assume $\lim_{n\to\infty} x_n = x$. Clearly there exist θ_n such that $\langle e^{i\theta_n} x_n, x\rangle$ is real and non-negative, and passing again to a subsequence, we may assume that $\lim_{n\to\infty} \theta_n = \theta$. By continuity of the scalar product, the limit $\langle e^{i\theta} x, x\rangle = e^{i\theta}\|x\|^2$ is also real, and hence $e^{i\theta} = 1$. Thus $\lim_{n\to\infty} e^{i\theta_n} x_n = x$, and for each n the vectors x and $e^{i\theta_n} x_n$ are in the same completely real subspace. By uniform equicontinuity of regular states it follows that $p(\bar{x}) = M$.

Now for any $y \in H$ there exists θ such that $\langle x, e^{i\theta} y\rangle$ is real; hence $p(\overline{e^{i\theta} y})$ $(= p(\bar{y}))$ is equal to

$$M(\langle x, e^{i\theta} y\rangle)^2 + (1-M)(1-(\langle x, e^{i\theta} y\rangle)^2) = M|\langle x, y\rangle|^2 + (1-M)(1-|\langle x,y\rangle|^2),$$

and p is therefore regular. ∎

THEOREM. *If every state on $\mathbb{R}^3$ is regular, then every state on a real or complex separable Hilbert space H of dimension greater than two is regular.*

Proof. Every state on H necessarily induces a continuous symmetric bilinear form on every completely real three-dimensional subspace, and every completely real two-dimensional subspace can be embedded in a completely real three-dimensional subspace. It follows that the restriction of a state on H to any two-dimensional completely real subspace is regular, and from the above lemmata it follows that every state on H is regular. ∎

2. *Frame functions*

In this section, we define frame functions, collect some of their properties, and give examples. Denote by S the unit sphere of a fixed three-dimensional real Hilbert space. If s and s' are elements (i.e. vectors) of S, the angle between s and s' is designated by $\theta(s,s')$. If $\theta(s,s') = \pi/2$, we write $s \perp s'$.

Definition. A *frame* is an ordered triple (p,q,r) of elements of S such that $p \perp q$, $p \perp r$ and $q \perp r$.

Given a frame (p,q,r), each point in S (and in the vector space) can be uniquely expressed as $xp + yq + zr$, with $x, y, z \in \mathbb{R}$. We call (x,y,z) the frame coordinates of the point with respect to the given frame.

Definition. A *frame function* is a function $f\colon S \to \mathbb{R}$ such that the sum

$$f(p) + f(q) + f(r)$$

has the same value for each frame (p,q,r). This value, called the *weight* of f, will be denoted by $w(f)$.

The following obvious properties of frame functions will be useful in the sequel.

(P_1) *The set of frame functions is a vector space, and*

$$w(\alpha f) = \alpha w(f),$$
$$w(f+g) = w(f)+w(g)\ (\alpha \in \mathbb{R}, f,\ g \text{ frame functions}).$$

(P_2) *If f is a frame function, $f(-s) = f(s)\quad (s \in S)$.*

(P_3) *If f is a frame function, and if $s, t, s't' \in S$ all lie on the same great circle and $s \perp t$, $s' \perp t'$, then*

$$f(s)+f(t) = f(s')+f(t').$$

To illustrate the use of P_3 we prove:

(P_4) *Let f be a frame function with* $\sup f(s) = M < \infty$ *and* $\inf f(s) = m > -\infty$. *Let $\xi > 0$ and let $s \in S$ with $f(s) > M-\xi$. Then there is $t \in S$ with $s \perp t$ and $f(t) < m+\xi$.*

Proof. Given s with $f(s) > M-\xi$, choose $\delta > 0$ such that $f(s) > M-\xi+\delta$, and t' such that $f(t') < m+\delta$. Then s and t' determine a great circle on S, and if t, s' are chosen on this great circle with $s \perp t$, $s' \perp t'$, P_3 yields:

$$f(t) = f(s')+f(t')-f(s) < M+m+\delta-(M-\xi+\delta)$$
$$= m+\xi. \blacksquare$$

Next we give examples of frame functions. Obviously, constants are frame functions. If we fix a vector $p_0 \in S$, then for any frame (p,q,r) the frame coordinates of p_0 with respect to (p,q,r) are given by:

$$(\cos\theta(p_0,p),\ \cos\theta(p_0,q),\ \cos\theta(p_0,r)),$$

and the sum of the squares of these three numbers is one since $p_0 \in S$. Hence

$$f(s) = \cos^2\theta(p_0,s)$$

is a frame function, with $w(f) = 1$. Next, fix a frame (p_0,q_0,r_0) and a triple (α,β,γ) of real numbers. Let (x_0,y_0,z_0) denote the frame coordinates of a point $s \in S$ with respect to (p_0,q_0,r_0). By the above and by P_1,

$$f(s) = \alpha x_0^2+\beta y_0^2+\gamma z_0^2 \tag{$*$}$$

is a frame function, with $w(f) = \alpha+\beta+\gamma$. Now recall that if Q is any quadratic form on our Hilbert space, then there exists a frame (p_0,q_0,r_0) and a triple (α,β,γ) of real numbers such that the restriction of Q to S is given by $(*)$. Hence we have proved the following result:

PROPOSITION. *Let A be a linear operator from the given three dimensional Hilbert space to itself, and let*

$$Q(s) = \langle s, As\rangle$$

be the quadratic form associated with A. Then the restriction of Q to S is a frame function whose weight is the trace of A.

Note that p_0, q_0, r_0 are eigenvectors of $\frac{1}{2}(A+A^T)$ with respective eigenvalues α, β, γ.

Our last example shows that frame functions can be wildly discontinuous. Let $\psi\colon \mathbb{R} \to \mathbb{R}$ be any map such that $\psi(x+y) = \psi(x)+\psi(y)$ for all $x, y \in \mathbb{R}$. Then if f is a frame function, so is $\psi(f)$, and ψ can be chosen to have arbitrary values on a basis of $\mathbb{R}$ over $\mathbb{Q}$. Of course Cauchy's classical theorem on functional equations tells us that if ψ is bounded on an interval, then $\psi(x) = cx$ for some constant c. This example shows that the restriction to bounded frame functions in the following theorem is essential.

3. *Statement of Gleason's theorem* [2]

We now state the result to be proved.

GLEASON'S THEOREM. *Let f be a bounded frame function. Define*

$$M = \sup f(s)$$
$$m = \inf f(s)$$
$$\alpha = w(f) - M - m.$$

Then there exists a frame (p, q, r) such that if the frame coordinates with respect to (p, q, r) of $s \in S$ are (x, y, z),

$$f(s) = Mx^2 + \alpha y^2 + mz^2$$

for all $s \in S$.

In particular, the proposition of § 2 provides all bounded frame functions. We remark that the above representation implies that $m \leqslant \alpha \leqslant M$; if $m < \alpha < M$, then the frame (p, q, r) is unique up to change of sign; if $m \leqslant \alpha < M$ then p is unique up to change of sign, and similarly for $m < \alpha \leqslant M$.

In order to clarify the idea behind our proof of the above result, we now state and prove a theorem which might be called an 'abelianized' version of Gleason's theorem. Its content was essentially known to 19th century mathematicians.

'WARMUP' THEOREM I. *Let $f\colon [0,1] \to \mathbb{R}$ be a bounded function such that for all $a, b, c \in [0,1]$ with $a+b+c = 1$, $f(a)+f(b)+f(c)$ has the same value $\tilde{w} = w(f)$. Then $f(a) = (\tilde{w} - 3f(0))\,a + f(0)$ for all $a \in [0,1]$.*

Proof. By subtracting a constant, we may assume $f(0) = 0$. Now take $c = 0$, $b = 1-a$ to obtain

$$f(a) = \tilde{w} - f(1-a),$$

and then set $c = 1-(a+b)$ to obtain

$$f(a)+f(b) = \tilde{w} - f(1-(a+b)) = f(a+b)$$

for all $a, b, a+b \in [0,1]$. This implies immediately that

$$f(a) = \tilde{w}a$$

for all rational a, and for general $a \in [0,1]$ and $n \geqslant 1$ with $na \leqslant 1$ we have

$$f(na) = nf(a).$$

Hence as a tends to 0, $f(a)$ must tend to 0 because f is bounded, and thus

$$\lim_{a \to 0} f(a+b) = f(b)$$

for all $b \in [0,1)$. Thus $f(a) = \tilde{w}a$ for all $a \in [0,1]$. ∎

The above formulation was chosen in order to make the analogy with Gleason's theorem clear. Actually, we shall use the following modified version in our proof.

'WARMUP' THEOREM II. *Let C be a finite or countable subset of* $(0, 1)$. *Let* f: $[0,1]\backslash C \to \mathbb{R}$ *be a function such that*

(1) $f(0) = 0$.
(2) *If* $a, b \in [0,1]\backslash C$ *and* $a < b$, *then* $f(a) \leqslant f(b)$.
(3) *If* $a, b, c \in [0,1]\backslash C$ *and* $a+b+c = 1$, *then* $f(a)+f(b)+f(c) = 1$.

Then $f(a) = a$ *for all* $a \in [0,1]\backslash C$.

Proof. The set

$$\tilde{C} = \{rc: c \in C, r \text{ rational}\} \cup \{r(1-c): c \in C, r \text{ rational}\}$$

is at most countable, so that there exists a point $a_0 \in (0,1)$ with $a_0 \notin \tilde{C}$. Now if r is a rational number such that $ra_0 \in [0,1]$, then neither ra_0 nor $1-ra_0$ belong to C, since $a_0 \notin \tilde{C}$. As in the proof above, we conclude that

$$f(ra_0)+f(r'a_0) = f((r+r')\,a_0)$$

for rational r, r' with ra_0, $r'a_0$, $(r+r')\,a_0 \in [0,1]$, and hence

$$f(ra_0) = rf(a_0)$$

for rational r with $ra_0 \in [0,1]$. It now follows from (2) that $f(a) = a$ for all $a \in [0,1]\backslash C$. ∎

4. *The basic lemma*

In this paragraph, we prove a basic lemma to be used in the following two sections. We fix a vector $p \in S$, to be thought of as the north pole, and use the following notation.

$$N = \{s \in S: \theta(p,s) \leqslant \pi/2\} = \text{'northern hemisphere'},$$
$$E = \{s \in S: s \perp p\} = \text{'equator'}.$$

For each $s \in N$, set

$$l(s) = \cos^2\theta(p,s) = \text{'latitude' of } s,$$

and define for $0 \leqslant l \leqslant 1$:

$$L_l = \{s \in N: l(s) = l\} = \text{'}l\text{th parallel'}.$$

Thus $L_1 = \{p\}$ and $L_0 = E$.

For $s \in N\backslash\{p\}$, there is a unique vector $s^\perp \in N$ such that $s \perp s^\perp$ and $l(s)+l(s^\perp) = 1$ ($s^\perp$ is the 'coldest' vector orthogonal to s). The great half circle D_s defined by

$$D_s = \{t \in N: t \perp s^\perp\} \quad (s \in N\backslash\{p\})$$

will be called the *descent* through s; it is the great circle through s which has s as its northernmost point. (For $e \in E$, $D_e = E$). We can now state the basic lemma:

BASIC LEMMA. *Let f be a frame function such that*

(1) $f(p) = \sup_{s \in S} f(s)$, *and*
(2) $f(e)$ *has the same value value for all* $e \in E$.

Then if $s \in N\backslash\{p\}$ *and if* $s' \in D_s$,

$$f(s) \geqslant f(s').$$

Proof. Set $f(p) = M$. Property P_4 implies that

$$f(e) = m = \inf_{s \in S} f(s) \quad (e \in E).$$

http://dx.doi.org/10.1017/S0305004100063313

Let $s \in N\backslash\{p\}$ and $s' \in D_s$. Choose $t, t' \in D_s$ with $s \perp t$, and $s' \perp t'$. By property P_3,

$$f(s)+f(t) = f(s')+f(t').$$

and using the fact that $t \in E$ we obtain

$$f(s)-f(s') = f(t')-f(t) = f(t')-m \geqslant 0. \blacksquare$$

Later on we shall need an

APPROXIMATE VERSION OF THE BASIC LEMMA. *Let f be a frame function and $\xi > 0$ such that*

(1) $f(p) > \sup_{s\in S} f(s) - \xi$, *and*
(2) $f(e)$ *has the same value for all* $e \in E$.

Then if $s \in N\backslash\{p\}$ and if $s' \in D_s$, $f(s) > f(s') - \xi$.

Proof. As above, property P_4 yields

$$f(e) < m+\xi \quad (e \in E),$$

and with exactly the same choices of t and t':

$$f(s)-f(s') = f(t')-f(t) > f(t')-m-\xi \geqslant -\xi. \blacksquare$$

5. *Simple frame functions*

In this paragraph, we show that Gleason's theorem is true under two additional hypotheses on frame functions. We begin with a geometric lemma due to Piron [**6**].

GEOMETRIC LEMMA. *Let $s, t \in N\backslash(p)$ such that $l(s) > l(t)$. Then there exist $n \geqslant 1$ and $s_0, \ldots, s_n \in N\backslash(p)$ such that $s_0 = s$, $s_n = t$, and for each $1 \leqslant i \leqslant n$:*

$$s_i \in D_{s_{i-1}}.$$

Proof. To facilitate the calculations, we transfer this problem to the plane tangent to S at p by projecting each point of N onto this plane from the origin (center of the sphere S). Points of the same latitude on S are projected onto circles centred at p, and the descent through s becomes the straight line through s tangent to the latitude circle at s (see Fig. 1). In the simplest case, s and t lie on a ray from the origin. In this case

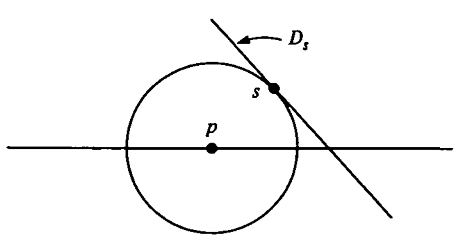

Fig. 1.

we may choose $n = 2$ and pick s_1 as in Fig. 2. Now fix $s_0 = s = (x, 0)$ (in $\mathbb{R}^2$ coordinates) and $n \geqslant 1$. Choose $s_1 \ldots s_n$ successively such that $s_i \in D_{s_{i-1}}$ and such that the angle

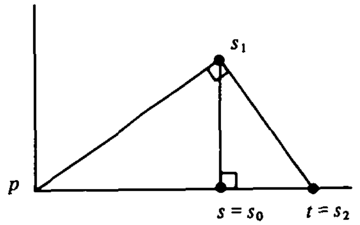

Fig. 2.

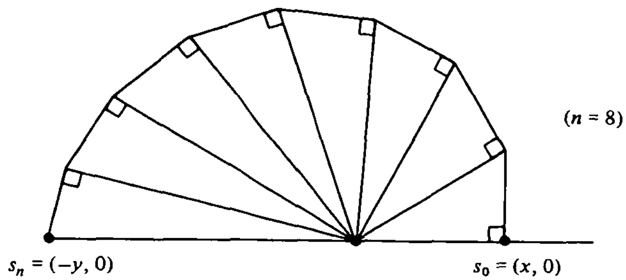

Fig. 3.

between s_i and s_{i+1} in the plane is π/n (see Fig. 3). Then s_n has coordinates $(-y, 0)$ and we wish to show that $y - x \to 0$ as $n \to \infty$. Let d_k be the distance of s_k from the origin. Then $d_0 = x$ and $d_n = y$. For each i we have

$$d_{i+1}/d_i = 1/\cos(\pi/n),$$

and hence

$$1 \leqslant y/x = d_n/d_0 = \prod_{i=1}^{n} \frac{d_i}{d_{i-1}} = \frac{1}{(\cos \pi/n)^n} \leqslant \frac{1}{(1-\pi^2/2n^2)^n},$$

which approaches 1 as n tends to infinity. The lemma is proved.∎

We now come to the main result of this section.

THEOREM. *Let f be a frame function such that for some point $p \in S$*

(1) $f(p) = M := \sup_{s \in S} f(s)$,
(2) $f(e)$ *takes the constant value m for all $e \in E$.*

Then $f(s) = m + (M-m)\cos^2 \theta(s,p)$, for all $s \in S$.

Proof. By property P_4, $m = \inf_{s \in S} f(s)$, so that if $M = m$ the theorem is true. If $M \neq m$, then we may assume that $m = 0$ and $M = 1$ (replace f by $(1/(M-m))(f-m)$). Let $s, t \in N \backslash (p)$ with $l(s) > l(t)$. Then by the geometric lemma and the basic lemma of the preceding section, we have

$$f(s) \geqslant f(t).$$

For each $l \in [0, 1]$, define:

$$\bar{f}(l) = \sup\{f(s) \colon s \in N,\, l(s) = l\},$$
$$\underline{f}(l) = \inf\{f(s) \colon s \in N,\, l(s) = l\}.$$

Then $\bar{f}(1) = \underline{f}(1) = 1$, $\bar{f}(0) = \underline{f}(0) = 0$, and if $l, l' \in [0, 1]$ with $l < l'$, it follows from the above that

$$\bar{f}(l) \leqslant \underline{f}(l').$$

http://dx.doi.org/10.1017/S0305004100063313

Hence the set $C := \{l : \bar{f}(l) > \underline{f}(l)\}$ is at most countable, as

$$\sum_{l \in C} (\bar{f}(l) - \underline{f}(l)) \leqslant 1.$$

For $l \in [0,1] \backslash C$, define

$$f(l) = \bar{f}(l) = \underline{f}(l).$$

If $l, l', l'' \in [0,1]$ with $l + l' + l'' = 1$, then there exists a frame (q, q', q'') with $l(q) = l$, $l(q') = l'$, $l(q'') = l''$. That is, the function f satisfies the hypotheses of Warmup Theorem II, and we conclude that

$$f(l) = l \quad \text{for} \quad l \in [0,1] \backslash C.$$

But this implies that $C = \varnothing$, so that for each $s \in N$,

$$f(s) = f(l(s)) = l(s) = \cos^2 \theta(s, p).$$

The theorem now follows from property P_2.

6. *Extremal values*

In this section we use the results of the preceding section to show that bounded frame functions attain their extremal values. Let f be a bounded frame function,

$$M = \sup_{s \in S} f(s),$$

and choose a sequence $p_n \in S$, $n \geqslant 1$, such that $\lim_{n\to\infty} f(p_n) = M$. Since S is compact, we may assume by passing to a subsequence that p_n converges, and we set

$$p = \lim_{n \to \infty} p_n.$$

Assume also $p_n \in N$ for all n. Our goal is to show that $f(p) = M$.

STEP 1. *Changing coordinates*

For each n, we would like to look at p_n as the north pole, instead of p. We do this as follows. Choose and fix a point $e_0 \in E$ and let C_0 denote the great circle segment from p to e_0. Let $\rho_n : S \to S$ be the rigid motion of S which takes p to p_n and some point, say c_n, on C_0 to p. Obviously

$$\lim_{n \to \infty} c_n = p.$$

Now define the sequence g_n of frame functions by setting

$$g_n(s) = f(\rho_n s) \quad (s \in S).$$

We note the following properties:

(1) $\lim_{n\to\infty} g_n(p) = M$.
(2) $M = \sup_{s \in S} g_n(s)$ and $m = \inf_{s \in S} f(s) = \inf_{s \in S} g_n(s)$ for each $n \geqslant 1$.
(3) $g_n(c_n) = f(p)$ for each $n \geqslant 1$.

STEP 2. *Symmetrization*

Denote by $\hat{p}: S \to S$ the right-hand rotation by 90° of S around the pole p. For each $n \geqslant 1$, set

$$h_n(s) = g_n(s) + g_n(\hat{p}s) \quad (s \in S).$$

The sequence h_n of frame functions ($P1$) has the following properties:

(1) $\sup_{s \in S} h_n \leqslant 2M \quad$ for all $\quad n \geqslant 1$.
(2) $\inf_{s \in S} h_n \geqslant 2m \quad$ for all $\quad n \geqslant 1$.

http://dx.doi.org/10.1017/S0305004100063313

(3) $\lim_{n\to\infty} h_n(p) = 2M$.
(4) Each h_n is constant on E (by $P3$).
(5) $h_n(c_n) \leqslant M+f(p)$ for all $n \geqslant 1$.

STEP 3. *Limit*

We consider each h_n as a point in the product space

$$[2m, 2M]^S.$$

Under the product topology, this space is compact, so that the sequence h_n has an accumulation point, which we denote by h. Then:

(1) $h(p) = 2M = \sup_{s\in S} h(s)$.
(2) h is constant on E.
(3) h is a frame function, since the frame functions form a closed subset of $[2m, 2M]^S$.

By the theorem of § 5, h is continuous (and has a special form, which does not interest us here).

STEP 4. $f(p) = M$

Choose $\epsilon > 0$, and choose $c \in C_0$ such that $h(c) > 2M-\epsilon$. Applying the approximate version of the basic lemma to h_n and noting that we can reach c from c_n in two steps (easiest case of the geometric lemma) for sufficiently large n, we obtain

$$h_n(c_n) > h_n(c) - 2\delta_n$$

with $\delta_n > 2M - h_n(p) \to 0$ as $n \to \infty$. Now choose a subsequence $n_j \to \infty$ such that

$$\lim_{j\to\infty} h_{n_j}(c) > 2M - \epsilon.$$

It then follows that (step 2, 5)

$$M+f(p) \geqslant \liminf_{j\to\infty} h_{n_j}(c_{n_j}) \geqslant \lim_{j\to\infty} (h_{n_j}(c) - 2\delta_n) > 2M-\epsilon,$$

so that $f(p) > M-\epsilon$. Hence we have proved:

THEOREM. *Bounded frame functions attain their extremal values.*

7. *The general case*

We now prove the theorem stated in § 3. Choose $p \in S$ such that $f(p) = M$, and $r \in S$, $r \perp p$, such that $f(r) = m$. This is possible because of P_4 and the theorem of § 5. Choose q orthogonal to p and r, and set $f(q) = \alpha$. We may assume that $m < \alpha < M$, since otherwise the result of § 4 applies to f or $-f$ and the proof is finished. As in § 5 we let $\hat{p}$, $\hat{q}$, $\hat{r}$ denote the 90° right-hand rotations about p, q, and r.

We shall now use the theorem of § 5 to obtain information concerning f. It is sufficient to know that f belongs to the space of quadratic frame functions. Taking p as the north pole, the function

$$f(s)+f(\hat{p}s)$$

takes the constant value $m+\alpha$ on the equator, and attains its supremum $2M$ at p. Letting

$$g(s) = M\cos^2\theta(s,p) + m\cos^2\theta(s,r) + \alpha\cos^2\theta(s,q),$$

http://dx.doi.org/10.1017/S0305004100063313

we have from §4

$$f(s)+f(\hat{p}s) = 2M\cos^2\theta(s,p)+(m+\alpha)(1-\cos^2\theta(s,p)) \tag{$*$}$$
$$= g(s)+g(\hat{p}s),$$
$$f(s)+f(\hat{r}s) = g(s)+g(\hat{r}s)$$

(the second equation follows by analogous reasoning, since $-f$ is a frame function taking its supremum $-m$ at r).

Now let (x,y,z) denote the (p,q,r)-frame coordinates of $s \in S$.

Claim. (a) If either $x = y$, $x = z$, or $y = z$, then $f(s) = g(s)$;
(b) If either $x = -y$, $x = -z$, $y = -z$, then $f(s) = g(s)$.

Proof of claim. (a) Note that $\hat{r}(x,y,z) = (-y,x,z)$; $\hat{p}(x,y,z) = (x,-z,y)$. Applying these operations in succession, one verifies:

$$\hat{p}\hat{p}\hat{r}(x,x,z) = (-x,-x,-z),$$
$$\hat{p}\hat{r}\hat{r}(x,z,z) = (-x,-z,-z),$$
$$\hat{r}\hat{p}\hat{p}\hat{p}\hat{r}(x,y,x) = (-x,-y,-x).$$

Suppose $s = (x,x,z)$. Since $f(s) = f(-s)$, $g(s) = g(-s)$ (by property P_2), we conclude from (x):

$$f(s)+f(\hat{r}s) = g(s)+g(\hat{r}s),$$
$$f(\hat{r}s)+f(\hat{p}\hat{r}s) = g(\hat{r}s)+g(\hat{p}\hat{r}s),$$
$$f(\hat{p}\hat{r}s)+f(\hat{p}\hat{p}\hat{r}s) = g(\hat{p}\hat{r}s)+g(\hat{p}\hat{p}\hat{r}s);$$

subtracting the second equation from the sum of the first and third; we conclude that $f(s) = g(s)$. The other two cases under (a) are proved similarly.

(b) Suppose $s = (x,-x,z)$; then $\hat{r}(x,-x,z) = (x,x,z)$, which lies on the great circle $x = y$. From (a) we know that $f(\hat{r}s) = g(\hat{r}s)$, and from $(*)$ we conclude that also $f(s) = g(s)$. The other two cases in (b) are proved similarly, and the claim is proved.

Now define $h := g-f$. h is clearly a frame function, and the claim implies that $h(p) = h(q) = h(r) = 0$, so that the weight of h is zero. We also know that h is zero on the six great circles $x = \pm y$, $x = \pm z$, $y = \pm z$. The proof is completed by showing that h is identically zero. Assume that h is not identically zero; then by §5 we may put

$$M' := \sup h = h(p'),$$
$$m' := \inf h = h(r'),$$
$$\alpha' := h(q');\ q' \perp r', q' \perp p'.$$

The argument is broken into four steps.

(i) $M' = -m'$: Assume that $m' > -M'$. Then $\alpha' < 0$, and by P_3, α' is the maximal value of h on the great circle orthogonal to p'. However, the great circle $x = y$ must intersect the former great circle in at least two points, and at these two points h must take the value zero. Considering $-h$, we derive a contradiction from the assumption $m' < -M'$ by the same argument.

(ii) $\alpha' = 0$: This follows immediately from (i) and the fact that h has weight zero.

(iii) $h(x',x',z') = M(x'^2-z'^2)$, where (x',y',z') denote the (p',q',r')-frame coordinates. Using the previous two steps, this follows from the claim, upon substituting h for f and $M(x'^2-z'^2)$ for g.

http://dx.doi.org/10.1017/S0305004100063313

(iv) On the great circle $x' = y'$, h takes the value zero at exactly the following four points: (x', x', x'), $(x', x', -x')$, $(-x', -x', x')$ and $(-x', -x', -x')$.

The great circles $x = y$, $x = z$ and $y = z$ intersect in the two points: (x, x, x) and $(-x, -x, -x)$. As h is zero on these great circles, we see that the great circle $x' = y'$ must pass through the points (x, x, x) and $(-x, -x, -x)$, since otherwise there would be six points on $x' = y'$ at which h takes the value zero. The great circles $x = -y$ and $x = -z$ intersect at $(x, -x, -x)$ and $(-x, x, x)$. $x' = y'$ must also intersect these points, since otherwise it would intersect $x = -y$ and $x = -z$ at four points, making six points at which h would take the value zero on $x' = y'$. However, there is only one great circle, passing through the four points (x, x, x), $(-x, -x, -x)$, $(x, -x, -x)$ and $(-x, x, x)$, namely $y = z$. It follows that $y = z$ and $x' = y'$ describe the same great circle, and therefore h must take the value zero at all points of $x' = y'$. This contradicts step (iv) and the theorem is proved. ∎

REFERENCES

[1] J. S. BELL. On the problem of hidden variables in quantum mechanics. *Rev. Modern Phys.* **38** (1966), 447–452.

[2] A. M. GLEASON. Measures on the closed subspaces of a Hilbert Space. *J. Math. Mech.* **6** (1957), 885–893.

[3] P. R. HALMOS. *Introduction to Hilbert Space* (Chelsea, 1957).

[4] S. KOCHEN and E. P. SPECKER. The problem of hidden variables in quantum mechanics. *J. Math. Mech.* **17** (1967), 59–67.

[5] G. W. MACKEY. *Mathematical Foundations of Quantum Mechanics* (Addison-Wesley, 1963).

[6] C. PIRON. *Foundations of Quantum Physics* (Addison-Wesley, 1976).