CDMTCS Research Report Series

A Constructive Proof of Gleason's Theorem

Fred Richman
Department of Mathematics
Florida Atlantic University

Douglas Bridges
Department of Mathematics
University of Waikato

CDMTCS-038
May 1997

Centre for Discrete Mathematics and
Theoretical Computer Science

A CONSTRUCTIVE PROOF OF GLEASON'S THEOREM

Fred Richman
Florida Atlantic University
Boca Raton, FL 33431

Douglas Bridges
University of Waikato
Hamilton, New Zealand

1. Introduction

Gleason's theorem states that any totally additive measure on the closed subspaces, or projections, of a Hilbert space of dimension greater than two is given by a positive operator of trace class. In this paper we give a constructive proof of that theorem.

A **measure** μ on the projections of a real or complex Hilbert space assigns to each projection P a nonnegative real number $\mu(P)$ such that if $\sigma = \sum P_i$, where the P_i are mutually orthogonal, then $\mu(\sigma) = \sum \mu(P_i)$. Such a measure is determined by its values on the one-dimensional projections. Let W be the measure of the identity projection, and P_x the projection onto the 1-dimensional space spanned by the unit vector x. Then the measure μ is determined by the real-valued function $f(x) = \mu(P_x)$ on the unit sphere, a function which has the property that

$$\sum_{e \in E} f(e) = W$$

for each orthonormal basis E. Gleason calls such a function f a **frame function of weight** W. If T is a positive operator of trace class, then $f(x) = \langle Tx, x\rangle$ is a frame function. Gleason's theorem is that every frame function arises in this way.

The original reference for Gleason's theorem is [4], which can also be found in Hooker [6]. Cooke, Keane and Moran [3] gave a proof that is elementary in the sense that it does not appeal to the theory of representations of the orthogonal group, which the original proof does. However, some of the reasoning in [3] seems hopelessly nonconstructive, so we follow the general outline of [4] until we come to the end of the 3-dimensional real case, at which point we modify some arguments in [3] rather than attempt a constructive development of the necessary representation theory.

Any Hermitian form B on a finite-dimensional inner product space gives rise to a frame function $f(x) = B(x, x)$ whose weight is equal to the trace of the matrix of B. The essence of Gleason's theorem is the following converse.

Theorem 1. *If f is a bounded real-valued function on the unit sphere of an inner product space of dimension at least 3, and f is a frame function on each 3-dimensional subspace, then $f(x) = B(x, x)$ for some bounded Hermitian form B. That is, f is a quadratic form.*

Theorem 1 is the part of Gleason's theorem that requires the overwhelming bulk of the work to prove. All but the last section of this paper is devoted to it.

To finish the proof of Gleason's theorem we must construct, when f is a nonnegative frame function, a positive operator T of trace class so that $B(x,x) = \langle x, Tx\rangle$. Classically, the existence of an operator T such that $B(x,x) = \langle x, Tx\rangle$ follows immediately from the Riesz representation theorem if the space is complete. But there is a constructive problem: we may not be able to compute the norm of the linear functional $B(\cdot, y)$ for each y, which norm would be $\|Ty\|$ if we could construct the operator T. However, if $B(x,x)$ is a nonnegative frame function, then B is approximable arbitrarily closely by a form that vanishes on the orthogonal complement of some finite-dimensional subspace, and the operator corresponding to such a form approximates a positive operator T of trace class.

If f is a bounded function on the unit sphere of a normed linear space, then we will also use the letter f to denote the function defined on nonzero vectors v by

$$f(v) = \|v\|^2 f\left(\|v\|^{-1} v\right),$$

and its unique extension to the whole space. Note that any nonnegative frame function is bounded.

The proof of Theorem 1 breaks up into several parts.

1. If a bounded function on the unit sphere of a space of dimension at least two is a quadratic form on each 2-dimensional subspace, then it is a quadratic form. (Theorem 7)

2. If a nonnegative frame function on a 2-dimensional space is a quadratic form on each 2-dimensional completely real subspace, then it is a quadratic form. (Theorem 10)

3. Every nonnegative frame function on $\mathbf{R}^3$ is a quadratic form.

   (a) Every such frame function is uniformly continuous. (Corollary 18)

   (b) Every uniformly continuous such frame function is a quadratic form. (Theorem 23)

Part 1 is Lemma 3.4 of Gleason's paper, attributed to Jordan and von Neumann [7]. It is given short shrift in [3] where it is pretty much dismissed as straightforward. We define a Hermitian form B, in the obvious way, by

$$B(x,y) = \frac{f(x+y) - f(x-y)}{4} + i\frac{f(x-iy) - f(x+iy)}{4},$$

with the second term missing in the real case. This is the hypothesized form on any 2-dimensional subspace; the question is whether it is globally a Hermitian form.

We first prove continuity of B by showing how to get from x to x' by traveling short distances on 2-dimensional subspaces. A constructive problem here is that you cannot put an arbitrary vector in a 2-dimensional space (the vector might be too small to tell in which direction it is pointing, if any).

Part 2 is Lemma 3.3 of Gleason's paper. Here again we first prove that f is uniformly continuous so that we can use approximation techniques, whereas Gleason appeals to the highly nonconstructive Bolzano-Weierstrass theorem to construct a point where f achieves its maximum. As in the proof of Part 1, the argument is complicated by the fact that two vectors whose inner product is real need not demonstrably be contained in 2-dimensional completely real subspace.

To show 3(a), we have to circumvent the computation of two infima that occur in Gleason's treatment. The first computation is overcome by a sort of logical trick—the *negative least upper bound principle*. The second is more serious because Gleason extracts the modulus of continuity from it. We get around this by an argument which enables the calculation of the modulus of continuity without considering the infimum. For 3(b) we follow [3], using approximation techniques made possible by the fact that the frame function is known to be uniformly continuous.

It was claimed in [5] that there can be no constructive proof of Gleason's theorem in $\mathbf{R}^3$. The argument is essentially that the principal axes theorem does not admit a constructive proof, a well-known fact (see, for example, [2] page 21, and [8]). This is a tangential issue that does not touch the heart of Gleason's theorem. You can show that the frame function is a quadratic form, and you can construct bases for which the matrix of this form is arbitrarily close to a diagonal matrix, although you may not be able to construct a basis for which the matrix is diagonal. In fact, Gleason did not state his theorem in terms of diagonal matrices, but in terms of operators and bilinear forms, which *can* be constructed. The formulation given in [5] is taken from [3].

An attempt was made in [1] to formulate Gleason's theorem in $\mathbf{R}^3$ so that it admits a constructive proof. The author's best candidate was along the right lines: find diagonal matrices that approximate the frame function. However the formulation was flawed by the tacit assumption that the entries in the diagonal form are known in advance. As the infimum and supremum of the frame function are among these entries, this is a big assumption: the frame function is not uniformly continuous *a priori*—which seems to be pretty much what you need to compute the extrema—and a lot of work goes into proving that it is.

2. Reduction to two dimensions

A **bilinear form** on a real or complex vector space is a scalar-valued function $B(x, y)$ that is linear in x and conjugate linear in y. It is **Hermitian** if $B(y, x) = \overline{B(x, y)}$. A Hermitian form B is **positive** if $B(x, x) \geq 0$ for all x. An **inner product** is a scalar-valued, positive, bilinear form $\langle x, y \rangle$ such that $\langle x, x \rangle = 0$ only if $x = 0$. A bilinear form on an inner product space is **bounded** if it is bounded on the unit sphere $\{x : \langle x, x \rangle = 1\}$. A complete inner product space is a **Hilbert space**. Any finite-dimensional inner product space is complete. A **quadratic form** is a function of the form $B(x, x)$ where B is a Hermitian form.

The principal axes theorem concerns diagonalizing Hermitian forms on a finite dimensional Hilbert space. This cannot quite be done if the eigenvalues are not separated, because of the sensitivity of the eigenvectors to the data; but we can make the off-diagonal terms as small as we want.

We first consider the crucial 2-dimensional real case.

Lemma 2. *Let $f(x, y) = ax^2 + 2bxy + dy^2$ be a real quadratic form on a 2-dimensional Hilbert space. If a is within $\varepsilon > 0$ of the supremum of f on the unit circle, then*

$$b^2 < \varepsilon^2 + \varepsilon(a - d).$$

Proof. Clearly we may assume that $b \neq 0$, so the matrix of f has distinct eigenvalues

$$r_+ = \frac{a + d + \sqrt{(a-d)^2 + 4b^2}}{2} \qquad r_- = \frac{a + d - \sqrt{(a-d)^2 + 4b^2}}{2}$$

with orthogonal nonzero eigenvectors $(b, r_+ - a)$ and $(b, r_- - a)$. Thus the matrix is diagonalizable, and the maximum value of f on the unit circle is r_+. So $a + \varepsilon > r_+$, from which the result follows. ■

Note that if $b \neq 0$ or $a \neq d$, then the matrix can be diagonalized. The only difficulty occurs when we cannot distinguish the matrix from a scalar matrix, in which case it is almost diagonal.

Theorem 3. *Let B be a Hermitian form on a finite-dimensional Hilbert space, and $\varepsilon > 0$. Then B admits a matrix each of whose off-diagonal elements has modulus at most ε.*

Proof. We induct on the dimension of the space. Choose a unit vector v such that $B(v, v)$ is within η (to be determined) of the supremum M of $B(u, u)$ over the unit sphere, and let M' be the supremum of $|B(u, u)|$ over the unit sphere. We will show that $|B(u,v)| < \varepsilon$ for each unit vector u orthogonal to v. The orthogonal complement of v has smaller dimension than the space, so we will be done by induction.

Consider the real quadratic form $f(s,t) = B(su+tv, su+tv)$ for s and t real variables. This is equal to

$$s^2 B(u,u) + 2st\,\mathrm{Re}(B(u,v)) + t^2 B(v,v)$$

and its supremum on the unit circle is at most M. So, by Lemma 2,

$$\mathrm{Re}(B(u,v))^2 < \eta^2 + 2M'\eta.$$

In the complex case, replacing u by iu, we get the same inequality for $\mathrm{Im}(B(u,v))^2$, so

$$|B(u,v)|^2 < 2\eta^2 + 4M'\eta.$$

For small enough η, independent of u and v, the right side is smaller than ε^2. ■

Theorem 4. *Let f be a quadratic form on a finite-dimensional Hilbert space. If $|f| \le M$ on the unit ball, then*

$$f(x) - f(y) \le M\|x-y\|\|x+y\|$$

for all x and y.

Proof. We first show that the inequality holds when f is diagonalizable—that is, $f(x) = \sum a_i |x_i|^2$. In this case,

$$\begin{aligned} f(x) - f(y) &= \sum a_i(|x_i|^2 - |y_i|^2) \\ &\le \sum |a_i|\,|x_i - y_i||x_i + y_i|, \end{aligned}$$

where the inequality comes from the triangle inequality

$$|x_i|^2 \le |y_i^2| + |x_i^2 - y_i^2|.$$

The same inequality holds for $f(y) - f(x)$, so

$$(f(x) - f(y))^2 \le M^2\left(\sum |x_i - y_i||x_i + y_i|\right)^2.$$

But

$$(\|x-y\|\|x+y\|)^2 = \sum |x_i - y_i|^2 \sum |x_i + y_i|^2 \ge \left(\sum |x_i - y_i||x_i + y_i|\right)^2$$

by the Cauchy-Schwarz inequality.

So the inequality holds for nonnegative diagonal matrices. It follows from Theorem 3 that the inequality holds in general. ∎

If $x_1, \ldots, x_n$ is an orthonormal basis for a subspace F of an inner product space, then

$$Py = \sum_{i=1}^{n} \langle y, x_i \rangle x_i$$

defines the projection onto F, and $I - P$ is the projection onto $F^\perp$. We have $\|(I - P)y\| \leq \|y - z\|$ for all z in F. When $x \neq 0$, then x and y are independent if and only if $y \neq \langle y, x \rangle / \|x\|^2$ —that is, $(1 - P)y \neq 0$, where P is the projection onto the linear span of x (so $I - P$ is the projection onto its orthogonal complement).

Lemma 5. *Let P be the projection of an inner product space of dimension at least $n + 1$ onto an n-dimensional subspace F. Then*

$$\{y : (1 - P)y \neq 0\}$$

is open and dense. In particular, if x is a nonzero element of an inner product space of dimension at least 2, then the set of elements y such that $\{x, y\}$ is linearly independent is open and dense.

Proof. To say the space has dimension at least $n + 1$ means that there exist independent elements $x_0, x_1, \ldots, x_n$, which we may assume are orthonormal. We first construct a nonzero element u of $F^\perp$. Let $\bar{P}$ denote $I - P$, the projection onto $F^\perp$. If $\bar{P} > 0$, then we can take $u = x_i$, so we may assume that $\left\|\bar{P}x_i\right\|$ is as small as we wish for $i = 1, \ldots, n$. Hence $Px_1, \ldots, Px_n$ are independent elements of F and therefore form a basis, so we can write $Px_0 = \sum_{i=1}^{n} a_i P x_i$. Set $u = x_0 - \sum_{i=1}^{n} a_i x_i$, which is nonzero as the x_i are linearly independent.

Clearly the set in question is open. To show that it is dense, let y be arbitrary and u a small nonzero element in $F^\perp$. Then either $\bar{P}y \neq 0$ or $\bar{P}(y + u) \neq 0$, and both y and $y + u$ are near y.

The final statement of the lemma follows from the fact that if F is the one-dimensional subspace generated by x, and $\bar{P}y \neq 0$, then x and y are linearly independent. ∎

Lemma 6. *Let f be a function on an inner product space such that $|f(x)| \leq M\|x\|^2$ for all x.*

1. *If*

$$|f(x) - f(y)| > M\|x - y\|\|x + y\|$$

then $x \neq 0$ and $y \neq 0$.

2. *If the space has dimension at least 2, and the inequality*

$$|f(x) - f(y)| \leq M\|x - y\|\|x + y\|$$

holds whenever x and y are in a 2-dimensional subspace, then it holds for all x and y.

Proof. Of course if $x = 0$, then the inequality in (1) cannot hold; but we want to show that $x \neq 0$—that is, $\|x\| > 0$ —not just that x cannot be zero. As

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x)| + |f(y)| \leq M\|x\|^2 + M\|y\|^2 \\ &\leq M\|x\|^2 + M(\|x\| + \|y - x\|)(\|x\| + \|x + y\|) \\ &= M\|x\|(2\|x\| + \|x - y\| + \|x + y\|) + M\|x - y\|\|x + y\|, \end{aligned}$$

we see that if $|f(x) - f(y)| > M\|x - y\|\|x + y\|$, then $x \neq 0$. By symmetry, $y \neq 0$ also.

For (2), the problem is that we may not be able to construct a two-dimensional subspace containing x and y. By (1), we may assume that $x \neq 0$ and $y \neq 0$. From Lemma 5, if $u \neq 0$, then the set of z such that u and z are linearly independent is dense and open. Therefore we can find z arbitrarily close to x such that x and z are independent, and also y and z are independent. So

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f(z)| + |f(z) - f(y)| \\ &\leq M\|x - z\|\|x + z\| + M\|y - z\|\|y + z\|. \end{aligned}$$

The latter expression converges to $M\|y - x\|\|y + x\|$ as z goes to x. ■

Now we follow Gleason's proof of his Lemma 3.4 to show that f is a quadratic form.

Theorem 7. *Let f be a bounded function on the unit sphere of an inner product space of dimension at least 2. If the restriction of f to any 2-dimensional subspace is a quadratic form, then f is a quadratic form.*

Proof. Extend f in the standard way to the whole space so that $|f(x)| \leq M\|x\|^2$ for all x. From Lemma 6 we know that f is uniformly continuous on bounded subsets. Define a B by the polarization identity

$$B(x, y) = \frac{f(x + y) - f(x - y)}{4} + i\frac{f(x - iy) - f(x + iy)}{4},$$

where the second term is missing in the real case. This is a Hermitian form on any 2-dimensional subspace. The question is whether it is globally a Hermitian form.

It is obviously bounded. There are now three equations to check:

- $B(\lambda x, y) = \lambda B(x, y)$,
- $B(y, x) = \overline{B(x, y)}$, and
- $B(x + y, z) = B(x, z) + B(y, z)$.

Our hypotheses ensure that the first two hold whenever x and y are linearly independent. But any two points x, y are arbitrarily close to linearly independent vectors, so these equations hold by the continuity of B. The second equation also follows directly from the definition of B.

The third equation follows from the parallelogram equality

$$2f(x) + 2f(y) = f(x + y) - f(x - y),$$

which holds when x and y are linearly independent, and therefore, by continuity, for all x and y. Following Gleason, we write

$$\begin{aligned} 8 \operatorname{Re} B(x, z) + 8 \operatorname{Re} B(y, z) &= 2f(x + z) - 2f(x - z) + 2f(y + z) - 2f(y - z) \\ &= f(x + y + 2z) + f(x - y) - f(x + y - 2z) - f(x - y) \\ &= 4 \operatorname{Re} B(x + y, 2z) = 8 \operatorname{Re} B(x + y, z). \end{aligned}$$

In the complex case, replacing x by ix and y by iy gives

$$\operatorname{Im} B(x, z) + \operatorname{Im} B(y, z) = \operatorname{Im} B(x + y, z),$$

so B is bilinear. ■

3. Completely real subspaces

A **completely real subspace** is a real subspace K such that $\langle x, y\rangle$ is real for all x and y in K. The basic problem addressed in this section is to show that if a nonnegative frame function on a 2-dimensional complex Hilbert space is regular on each completely real 2-dimensional subspace, then it is continuous. From this we can derive Gleason's Lemma 3.3 that it is regular on the 2-dimensional complex space.

We first need to observe Gleason's Lemma 3.2, which says that if f is a nonnegative regular frame function of weight W on a finite-dimensional real Hilbert space, then for any unit vectors x and y,

$$|f(x) - f(y)| \leq 2W\|x - y\|.$$

To prove this, just calculate $B(x + y, x - y) = f(x) - f(y)$ and use the fact that B is bounded by W. This gives us a uniform Lipschitz condition on every completely real t2-dimensional subspace.

Lemma 8. *Let f be a nonnegative frame function on a 2-dimensional complex Hilbert space such that f is regular on each completely real 2-dimensional subspace. Then, for each $C > 0$ there exists K such that if*

- $0 < \|x\| \leq C$,
- $\|x - y\| < 1$,
- $\mathrm{Re}\langle x, y\rangle \neq 0$,
- $\mathrm{Im}\langle x, y\rangle \neq 0$, *and*
- $\langle x^\perp, y\rangle \neq 0$, *where $x^\perp$ is any nonzero vector orthogonal to x,*

then $|f(x) - f(y)| \leq K\sqrt{\|x - y\|}$.

Proof. We may write $x = (r, 0) \in \mathbf{C}^2$ with $0 < r \leq C$, and $y - x = (\lambda, \mu)$ with $\lambda = a + bi$, where $b \neq 0$, $\mu \neq 0$ and $r + a \neq 0$. If $b > 0$, we consider the sequence of points

$$x = (r, 0), \quad (r + a, -\sqrt{b}), \quad (r + \lambda, (r + a)i\sqrt{b}), \quad (r + \lambda, 0), \quad (r + \lambda, \mu) = y.$$

The stated conditions guarantee that each two adjacent points of this sequence generate a completely real subspace that is 2-dimensional. If

$$\varepsilon = \sqrt{\|x - y\|} = \sqrt{a^2 + b^2 + |\mu|^2},$$

then the distance between adjacent points can be bounded by a constant times ε, because $a, b, |\mu| \leq \varepsilon$.

For $b < 0$, replace $-\sqrt{b}$ and $\sqrt{b}$ by $-\sqrt{-b}$. ∎

Lemma 9. *Let f be a nonnegative frame function on a 2-dimensional complex Hilbert space such that f is regular on each completely real 2-dimensional subspace. Then f is uniformly continuous on bounded subsets.*

Proof. It suffices to show that f is uniformly continuous on the ball $S = \{x : \|x\| \leq C\}$. Let K be as in Lemma 8. We have to get rid of the last three inequalities in the hypothesis of that lemma, and the restriction that x be nonzero. For fixed nonzero x, each of the three inequalities defines a dense open subset of Hilbert space.

Therefore, given nonzero x and y, we can find z arbitrarily close to x such that x and z, and y and z, satisfy the inequalities. Hence

$$|f(x) - f(y)| \leq K\sqrt{\|x - z\|} + K\sqrt{\|y - z\|},$$

where the right-hand side approaches $K\sqrt{\|x - y\|}$ as z approaches y. So the conclusion holds for all nonzero x and y in S. Thus

$$|f(x) - f(y)| \leq K\sqrt{\|x - y\|} + W\|x - y\|\|x + y\|,$$

where W is the weight of f. The point of the second term is that, by part (1) of Lemma 6, it suffices to prove this inequality when x and y are both nonzero, which we have already done. So this inequality holds for all x and y in S, whence f is uniformly continuous on S. ■

Now we can give an approximation version of Gleason's proof of his Lemma 3.3 (which opens with an appeal to the Bolzano-Weierstrass theorem). We will show that f can be approximated by (diagonal) quadratic forms uniformly on bounded subsets. Hence f is a quadratic form.

Theorem 10. *Let f be a bounded frame function on a 2-dimensional complex Hilbert space which is regular on each 2-dimensional completely real subspace. Then f is regular.*

Proof. First we reduce to the nonnegative case. If $f(z) \leq M\|z\|^2$ for all z, then the equation

$$f(z) = M\|z\|^2 - \left(M\|z\|^2 - f(z)\right)$$

shows that f is the difference of two nonnegative frame functions, each of which is regular on each completely real 2-dimensional subspace.

Because f is uniformly continuous on bounded subsets, we can find a unit vector y such that $f(y)$ is close to $\sup f$. Let z be a unit vector orthogonal to y. Gleason's calculation is this. Let λ and μ be nonzero complex numbers, and

$$z' = \frac{\mu}{|\mu|}\frac{|\lambda|}{\lambda}z,$$

which is also a unit vector orthogonal to y. So

$$\begin{aligned} f(\lambda y + \mu z) &= f((|\lambda|/\lambda)(\lambda y + \mu z)) \\ &= f(|\lambda|y + |\mu|z') = f(y)|\lambda|^2 + 2b|\lambda\mu| + (W - f(y))|\mu|^2 \end{aligned}$$

The last expression is not a (complex) quadratic form, because the middle term is not $2b\,\mathrm{Re}\,\lambda\bar{\mu}$. However, Lemma 2 shows that b is small; so f is approximated by the quadratic form obtained by omitting the middle term.

We assumed that λ and μ were nonzero. The alternative is that one of them is very small, in which case the approximation

$$f(\lambda y + \mu z) \approx f(y)|\lambda|^2 + (W - f(y))|\mu|^2$$

is obviously good. ■

4. Frame functions in $\mathbf{R}^3$

Gleason's Theorem 2.8 is essentially that every nonnegative frame function in $\mathbf{R}^3$ is uniformly continuous. The proof uses the existence of a point that approximates the infimum of certain a positive function. We cannot assume that such a point exists, but we can show that it cannot fail to exist. To be precise, here is **the negative least upper bound principle** (stated as a greatest lower bound principle).

Lemma 11. *Let S be a nonempty set of real numbers that is bounded below. Let ε be a positive real number. Then the following statement cannot be false:*

There exists $x \in S$ such that $y \geq x - \varepsilon$ for all $y \in S$.

Proof. Suppose the statement is false, and let a be a lower bound for S. If there were $x \in S$ with $x \leq a + \varepsilon$, then the statement would be true, which it is not. So $x \geq a + \varepsilon/2$ for each $x \in S$, that is, $a + \varepsilon/2$ is a lower bound for S. Iterating this argument we see that $a + n\varepsilon/2$ is a lower bound for S for each positive integer n. So S is empty, contrary to hypothesis. ■

Note that only the fact that S cannot be empty was used, not that S contained an element, which is what we mean by "nonempty".

What good is such an eccentric principle? There are two places in Gleason's proof of his Theorem 2.8 where he uses the fact that you can find a point in a set that approximates the infimum of that set. The conclusion of Theorem 2.8, or rather our revised version of it, is that for all x, y, if $\|x - y\| < \delta$, then $|f(x) - f(y)| \leq \varepsilon$. The condition $|f(x) - f(y)| \leq \varepsilon$ a negative one, equivalent to its double negation. So if we can derive it from the existence of a point that approximates the infimum of a set, then we can derive it from the double negation of that existence, which our principle says is true.

If we simply follow Gleason's proof, we cannot write down δ in advance, which is essential for the above analysis. In Gleason's proof, δ depends on another infimum. We have to calculate δ by an entirely different method.

Let N denote the punctured, open, northern hemisphere: the set of all points with latitude in the interval $(0, \pi/2)$.

For points r and s on the sphere, with $\|r - s\| < 2$, the **open disk between r and s** consists of those points x on the sphere such that

$$\left| x - \frac{r+s}{2} \right| < \left| \frac{r-s}{2} \right|,$$

the spherical disk with r and s as antipodal points.

For z a point of N other than the pole p, let G_z denote the set of all points $x \in N$ such that there exists y with the property that y is on the East-West great circle through x, and z is on the East-West great circle through y. That is, you can get from x to z in two East-West steps. The following lemma is essentially Gleason's Lemma 2.5, which states that G_z has a nonempty interior, although his proof contains a form of the additional information that we need.

Lemma 12. *If z is a point of N, then G_z contains the open disk between z and the pole p.*

Proof. Look at the projection from 0 on the plane tangent to p. Great circles become straight lines. Then z is on the East-West great circle through y exactly when $y \cdot (y - z) = 0$. If y is far enough along the East-West circle through x, then $y \cdot (y - z) > 0$. So G_z contains all x such that $x \cdot (x - z) < 0$, which is the inside of the circle with z and the pole 0 at opposite ends of a diameter. ■

Corollary 13. *Let r and s be points in N, with r due south of s. If x is in the open disk between r and s, then $x \in G_r$ and $s \in G_x$.*

Proof. Clearly $x \in G_r$. To show that $s \in G_x$, let the north pole be the origin in 3-dimensional space. We want to show that

$$\left\| s - \frac{x}{2} \right\|^2 \leq \left\| \frac{x}{2} \right\|^2$$

—that is, that

$$s \cdot (x - s) \geq 0.$$

The plane perpendicular to s at s goes through the south pole, so the circle it cuts out on the sphere contains the circle between r and s. Hence x is on the opposite side of this plane from the origin, and therefore the displayed inequality holds. ■

Let X be a subset of the sphere. We say that the **oscillation of f on X is at most α** if $|f(x) - f(x')| \leq \alpha$ for all x, x' in X.

Lemma 14. *Let p be the north pole and $0 < r < 1$. Set $r' = r^2/12$. Let f be a frame function such that the oscillation of f on $\{x : \|x - p\| < r\}$ is at most α. Then the oscillation of f on $\{x : \|x - e\| < r'\}$ is at most 2α for each point e on the equator.*

Proof. The key observation is that $|f(x) - f(y)|$ is invariant under $90°$ rotation of the great circle joining x and y. Let v be the point due south of e with $\|v - e\| = r/2$. For each point $u \neq v$, let $\rho(u)$ be the point that is $90°$ further along from v on the great circle C_u joining v to u. So, in particular, $p = \rho(e)$.

If $\|x - e\| < r'$, let q be the point on C_x such that $\|q - v\| = r/2$. We want to show that $\|q - e\| \leq 2r'$. (This is not a very tight bound.) First fix some notation. Let $\theta(t)$ denote the angular distance corresponding to the Euclidean distance t, and $\theta(u_1, u_2)$ the angular distance $\theta(\|u_1 - u_2\|)$ between the points u_1 and u_2. So $t = 2\sin\theta(t)/2$.

Clearly $\theta(x, e) \leq \theta(r')$ and $\theta(q, v) = \theta(r/2)$. Moreover,

$$\theta(r/2) - \theta(r) \leq \theta(x, v) \leq \theta(r/2) + \theta(r'),$$

so

$$\theta(r') \geq |\theta(q, v) - \theta(x, v)| = \theta(q, x).$$

Hence $\theta(q, e) \leq \theta(q, x) + \theta(x, e) \leq 2\theta(r')$, and therefore

$$\|q - e\| \leq 2r'.$$

Now we claim that

$$\|\rho(q) - p\| \leq r/2. \tag{*}$$

This gives the desired result, because if $\rho_x(v)$ denotes the point that is $90°$ north of v on C_x, then both $\rho(x)$ and $\rho_x(v)$ are within $r/2$ of $\rho(q)$ and therefore in $\{x : \|x - p\| < r\}$. So

$$|f(v) - f(x)| = |f(\rho_x(v)) - f(\rho(x))| \leq \alpha.$$

To verify (*), let β be the angle between C_x and C_e. Then $\|\rho(q) - p\| \leq \beta$ because β is the angular distance between C_x and C_e at their point of greatest separation. Drop a perpendicular from q, or e, to a point t on the diameter through v. Then

$$\|q - t\| = \|e - t\| \geq r/2\sqrt{2} \geq r/3.$$

Consider the triangle qet and the similar triangle $\rho(q)p0$. As $\|q - e\| \leq 2r'$, we have $\|\rho(q) - p\| \leq 6r'/r = r/2$. ∎

For the purpose of iteration, let h denote the function such that $h(r) = r^2/12$. Note that h and all of its iterates are strictly increasing functions.

Lemma 15. *Let $r' = h^2(r)$. If the oscillation of f on $\{x : \|x - p\| < r\}$ is at most α, then the oscillation of f on $\{x : \|x - t\| < r'\}$ is at most 4α for each point t on the sphere.*

Proof. We can go from the north pole to any point in the metric complement of the two poles in two steps of arc-length $\pi/2$. So the conclusion is true on a dense subset of the sphere and therefore on the sphere itself. ■

If f is a frame function of weight W, and p is a point on the sphere, then the **symmetrization** of f with respect to p is the function

$$g(x) = f(x) + f(\sigma x),$$

where σ is clockwise rotation by $\pi/2$ about p. This is a frame function of weight $2W$ which is constant on the equator (taking p as the north pole). Note that $g(p) = 2f(p)$.

The argument in the next lemma is taken from Gleason [4, Theorem 2.8].

Lemma 16. *Let p be the north pole, f a nonnegative frame function, and g its symmetrization with respect to p. If r and s are points of N, such that s is on the East-West great circle through r, then $g(r) \leq g(s) + 2f(p)$. Hence if $r \in G_s$, then $g(r) \leq g(s) + 4f(p)$.*

Proof. Let W be the weight of f. The East-West great circle C through r meets the equator at a point q. As $g(q) = W - f(p)$, we have

$$2W \geq g(r) + g(q) = g(r) + W - f(p).$$

So, for any $r \in N$, not just the one referred to in the hypothesis,

$$g(r) \leq W + f(p).$$

Now take $s \neq r$ on $C \cap N$, and let t be the point orthogonal to s in $C \cap N$. Then

$$g(r) + W - f(p) = g(r) + g(q) = g(s) + g(t) \leq g(s) + W + f(p),$$

so $g(r) \leq g(s) + 2f(p)$ for any $r \in N$ and any $s \neq r$ in $C \cap N$. Repeating this, we get

$$g(r) \leq g(s) + 4f(p) \quad \text{if } r \in G_s.$$

■

Theorem 17. *Let f be a nonnegative frame function on $\mathbf{R}^3$ of weight W, and η a positive number less than 1. Let x and y be points on the sphere, such that*

$$\|x - y\| < h^4\left(\frac{\eta}{4W+4}\right).$$

If there exists a point p such that $f(p) \leq f(z) + \eta/2$ for all points z on the sphere, then $|f(x) - f(y)| \leq 200\eta$.

Proof. Either $f(p) \leq \eta$ or $f(p) > \eta/2$. In the latter case, we may consider $f'(z) = f(z) - (f(p) - \eta/2)$. As $f(p) \leq f(z) + \eta/2$, the function f' is nonnegative. Clearly $f'(p) = \eta/2 < \eta$, and f' is a frame function of weight $W - 3(f(p) - \eta/2) \leq W$, so the hypothesis also holds for f' because h^4 is increasing. As $f(x) - f(y) = f'(x) - f'(y)$, we may assume, by passing to f' if necessary, that $f(p) \leq \eta$.

By Lemma 16,

$$g(r) \leq g(s) + 4\eta \quad \text{if } r \in G_s.$$

Now we part company with Gleason and find points a and b on some meridian in N such that $|g(a) - g(b)| < 4\eta$ and $\|a - b\| > \eta/(2W+2)$. To do so, choose a positive integer n so that $2W/\eta < n < 2W/\eta + 2$, and divide a longitude line in N into n equal segments with endpoints $p = x_0, \ldots, x_n$. Then $x_i \in G_{x_{i+1}}$ if $0 < i < n$, so

$$g(x_i) \leq g(x_{i+1}) + 4\eta$$

for all $i < n$, the case $i = 0$ being trivial. Either $g(x_{i+1}) \leq g(x_i) + 4\eta$ for some i, in which case we clearly have our a and b, or else $g(x_{i+1}) \geq g(x_i) + \eta$ for all i. But the latter would show that

$$g(x_n) \geq g(x_0) + n\eta > g(x_0) + 2W,$$

which is absurd. To see that $\|a - b\| > \eta/(2W+2)$, note that

$$\frac{1}{n} > \frac{\eta}{2W + 2\eta} > \frac{\eta}{2W+2},$$

so it suffices to show that $n\|a - b\| > 1$. The left-hand side represents the length of the polygonal path from x_0 to x_n which is at least $\|x_0 - x_n\| - \sqrt{2}$.

Now suppose that x is within the circle on N between a and b. We have

$$g(a) \leq g(x) + 4\eta \text{ and } g(x) \leq g(b) + 4\eta,$$

so

$$g(a) - 4\eta \leq g(x) \leq g(b) + 4\eta.$$

Thus the oscillation of g within that circle is at most $g(b) - g(a) + 8\eta \leq 12\eta$. This circle contains the ball of radius $\eta/(4W+4)$ around the point halfway between a and b.

So we get a ball of radius $h^2(\eta/(4W+4))$ around p where the oscillation of g is at most 48η. As $g(p) \leq 2\eta$ this says that $g(x) \leq 50\eta$ in that ball. So the oscillation of f is at most 50η in that ball, whence every point is the center of a ball of radius $h^4(\eta/(W+1))$ in which the oscillation of f is at most 200η. ■

Corollary 18. *If f is a bounded frame function in $\mathbf{R}^3$, then f is uniformly continuous on the unit sphere.*

Proof. Clearly it suffices to prove this for nonnegative frame functions. Suppose we are in the context of the theorem. We know that if there exists a point p with the property described, then $|f(x) - f(y)| \leq 200\eta$. So if $|f(x) - f(y)| > 800\eta$, then there cannot exist a point p with that property. But that would contradict Lemma 11, the negative least upper bound principle. ■

5. From uniformly continuous to regular

Rather than developing the representation theory of the orthogonal group from a constructive point of view, we follow the elementary treatment of Cooke, Keane and Moran [3] for this final step towards Gleason's Theorem.

First we prove an approximate form of their Warm-up Theorem I, for uniformly continuous functions.

Lemma 19. *Let F be a uniformly continuous real-valued function on $[0,1]$ that vanishes at 0, and let W be a real number. If $|F(a)+F(b)+F(c)-W| \leq \eta$ whenever $a+b+c=1$, then $|F(t) - Wt| \leq 3\eta$ for all $t \in [0,1]$.*

Proof. We will show that

$$|F(t) - Wt| \leq 3\eta$$

for each dyadic rational number $t \in [0,1]$. This argument does not appeal to continuity, and uses induction on n. As F is uniformly continuous, it then follows that since the inequality holds on a dense subset of $[0,1]$, it holds on all of $[0,1]$.

We assume that the displayed inequality holds for dyadic numbers t with denominator at most 2^n, and we want to show that it holds for denominator 2^{n+1}. Note that $F(0) = 0$, $|F(1) - W| \leq \eta$, and $|F(1/2) - W/2)| \leq \eta/2$, so we may assume that $n \geq 1$. Let k be an odd number between 0 and 2^{n+1}. For $k < 2^n$ consider

$$k2^{-n-1} + k2^{-n-1} + (1 - k2^{-n}) = 1.$$

We have

$$|2F(k2^{-n-1}) + F(1 - k2^{-n}) - W| \le \eta,$$

whence, by induction,

$$|2F(k2^{-n-1}) - Wk2^{-n}| \le 4\eta.$$

So for $k2^{-n-1}$ with $k < 2^n$ odd, we obtain the desired inequality with 2η instead of 3η.

For $k > 2^n$ consider

$$k2^{-n-1} + (2^{n+1} - k)2^{-n-1} = 1.$$

We have

$$\left|F(k2^{-n-1} + F((2^{n+1} - k)2^{-n-1}) - W\right| \le \eta.$$

As $2^{n+1} - k < 2^n$, using the 2η bound from the previous case we have

$$|F(k2^{-n-1}) - Wk2^{-n-1}| \le 3\eta. \qquad \blacksquare$$

The next thing to prove is that, given f and ε, we can find a point p so that the symmetrization g of f with respect to p can be approximated within ε by a regular function.

First we have some spherical trigonometry. Let θ denote latitude and φ longitude. The equation of the East-West great circle through the point $(\theta_0, 0)$ is

$$\tan\theta = \tan\theta_0 \cos\varphi,$$

so along that circle the derivative of θ with respect to φ is

$$\frac{d\theta}{d\varphi} = -\sin\theta_0 \frac{\cos^2\theta}{\cos\theta_0} \sin\varphi.$$

If we restrict to latitudes such that

$$\cos\theta \le 2\cos\theta_0,$$

then

$$-\frac{d\theta}{d\varphi} = \sin\theta_0 \frac{\cos^2\theta}{\cos\theta_0} \sin\varphi \le 2\sin 2\theta_0 \sin\varphi \le 2\varphi,$$

so

$$\theta \ge \theta_0 - \varphi^2.$$

Now suppose we take n East-West steps of length a/n, measured in longitude. At each step, the latitude will go down by at most $(a/n)^2$, so, in total, the latitude will go down by at most a^2/n.

Lemma 20. *For each $\delta > 0$ and weight W, there is $\eta > 0$ such that if f is a nonnegative frame function of weight W, and p is a point on the unit sphere such that $f(p) < \eta$, then the oscillation on any latitude of the symmetrization of f with respect to p is at most δ.*

Proof. Consider p to be the north pole of the unit sphere. We can compute a modulus of continuity for f, and hence for any symmetrization g, from W. We can also compute $\eta_0 > 0$ and $\theta_0 \in (0, \pi/2)$ from W so that if $\eta < \eta_0$, then the oscillation of g on the polar cap of latitude $2\theta_0 - \pi/2$ is less than δ. Choose $\xi \in (0, \theta_0)$ such that if $\|x - y\| < \xi$, then $|g(x) - g(y)| < \delta/2$, and so that $\cos(\theta_0 - \xi) \leq 2\cos\theta_0$. Note that $\cos(\theta - \xi') \leq 2\cos\theta$ whenever $0 \leq \theta \leq \theta_0$ and $0 \leq \xi' \leq \xi$. Choose a positive integer $n > \pi^2/\xi$, and $\eta < \eta_0$ so that $2n\eta < \delta/2$.

Let x and x' be points in the northern hemisphere with the same latitude θ. If $\theta > 2\theta_0 - \pi/2$, then we are in the polar cap, so $|g(x) - g(x')| < \delta$. We may therefore assume that $\theta < \theta_0$. Let a be the difference in longitudes of x and x'. Take n steps along East-West great circles to go from x' to a point y due south of x. The difference in latitudes between x and y is at most a^2/n, while, from Lemma 16,

$$g(y) \geq g(x') - 2n\eta > g(x') - \delta/2.$$

As $a^2/n \leq \pi^2/n < \xi$, by continuity we have $g(y) \leq g(x) + \delta/2$. So

$$g(x) \geq g(y) - \delta/2 > g(x') - \delta$$

for arbitrary points x and x' on the same latitude. ∎

Combining these two lemmas, we show that each nonnegative frame function has a symmetrization that is arbitrarily close to a quadratic form.

Theorem 21. *For each $\delta > 0$ and weight W, there is $\eta > 0$ such that if*

- *f is a nonnegative frame function of weight W,*
- *p is a point on the unit sphere where $f(p) < \eta$, and*
- *g is the symmetrization of f with respect to p,*

then, for each t on the unit sphere, $|Q(t) - g(t)| \leq \delta$ where Q is the quadratic form $Q(t) = W(1 - (t \cdot p)^2)$.

Proof. Think of p as the north pole of the unit sphere. Note that g is a nonnegative frame function of weight $2W$, and that $g(e) = W - f(p)$ for all points e on the equator. From Lemma 20 there is $\eta > 0$ such that the oscillation of g on latitudes is at most $\delta/9$.

For t on the sphere, let $\ell(t) = (t \cdot p)^2$. Then $Q(t) = W(1 - \ell(t))$. If t_1, t_2, t_3, are orthogonal vectors on the closed northern hemisphere, then $\ell(t_1) + \ell(t_2) + \ell(t_3) = 1$. Conversely, if $a_1 + a_2 + a_3 = 1$, for $a_i \in [0,1]$, then there exist orthogonal vectors t_1, t_2, t_3 on the closed northern hemisphere such that $a_i = \ell(t_i)$.

For each φ define F_φ on $[0,1]$ by setting $F_\varphi(\ell(t)) = W - g(t)$ where t is a point in the closed northern hemisphere with longitude φ. Then $|F_\varphi(a_i) - (W - g(t_i))| \leq \delta/9$, the bound on the oscillation of g on latitudes, whence

$$|F_\varphi(a_1) + F_\varphi(a_2) + F_\varphi(a_3) - W| < \delta/3.$$

It follows from Lemma 19 that $|F_\varphi(\ell(t)) - W\ell(t)| \leq \delta$. So

$$|Q(t) - g(t)| = |W - g(t) - W\ell(t)| \leq \delta$$

for any t in the closed northern hemisphere and therefore for any t on the sphere. ■

Corollary 22. *Let f be a bounded frame function of weight W, and $m = \inf f$ or $m = \sup f$. Let Q be the quadratic form given by*

$$Q(t) = W - m + (3m - W)(t \cdot p)^2 = 2m(t \cdot p)^2 + (W - m)(1 - (t \cdot p)^2).$$

For each $\delta > 0$, there is $\eta > 0$ such that if

- *p is a point on the unit sphere such that $|f(p) - m| < \eta$, and*
- *g is the symmetrization of f with respect to p,*

then $|Q(t) - g(t)| \leq \delta$ for all t on the unit sphere.

Proof. We prove the corollary for $m = \inf f$. The case $m = \sup f$ follows upon replacing f by $-f$.

Clearly $f' = f - m$ is a nonnegative frame function of weight $W' = W - 3m$. Choose η for δ and W'. Then $f'(p) < \eta$ and $g' = g - 2m$. From the theorem, $|Q'(t) - g'(t)| \leq \delta$, where $Q'(t) = W'(1 - (t \cdot p)^2)$. Clearly $Q' = Q - 2m$, so $|Q(t) - g(t)| = |Q'(t) - 2m - (g'(t) - 2m)| \leq \delta$. ■

The six great circles $x = \pm y$, $y = \pm z$ and $z = \pm x$ on the unit sphere divide the sphere into 24 triangles. The 14 vertices are

$$\left(\pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}\right), \quad (\pm 1, 0, 0), \quad (0, \pm 1, 0), \quad (0, 0, \pm 1).$$

The first eight are the vertices of an inscribed cube. The other eight are the projections on the sphere of the centers of the faces of that cube.

How big are the triangles? They all look like

$$(0, 0, 1), \quad \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(\frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right).$$

The maximum distance between two vertices is $2/\sqrt{3}$. As the triangle lies in the quarter hemisphere $(0, 0, 1), (1/\sqrt{2}, 1/\sqrt{2}, 0), (1/\sqrt{2}, -1/\sqrt{2}, 0)$, its diameter is $2/\sqrt{3}$. So any great circle through a point in the interior of that triangle must intersect the boundary at a distance at most $1/\sqrt{3} + \varepsilon$ from the point.

Theorem 23. *Let f be a bounded frame function on the unit sphere in $\mathbf{R}^3$, and $\delta > 0$. Then there is a quadratic form Q on $\mathbf{R}^3$ so that $|f(x) - Q(x)| \leq 3\delta$ for all x on the sphere.*

Proof. Choose $\eta > 0$ according to Corollary 22. Let $M = \sup f$ and $m = \inf f$. Choose p on the sphere so that $f(p) > M - \eta$ and r so that $f(r) < m + \eta$. Let p' and r' be the images of p and r under $90°$ rotation of the great circle joining them. Then $f(p) + f(p') = f(r) + f(r')$ so

$$f(r') - f(p') = f(p) - f(r) > M - m,$$

whence either $f(p') < m + \eta$ or $f(r') > M - \eta$. Thus we may assume that p and r are perpendicular.

Choose q perpendicular to p and r. Let $(x, y, z) = (t \cdot p, t \cdot q, t \cdot r)$ denote the coordinates of a point t with respect to the frame (p, q, r). Let Q be the quadratic form

$$Q(t) = Mx^2 + (W - M - m)y^2 + mz^2$$

where W is the weight of f. We want to show that Q approximates f within 3δ.

Let $f'(t) = f(t) - Q(t)$. We want to show that $|f'| \leq 3\delta$. Let $\widehat{p}$ and $\widehat{r}$ denote the $90°$ clockwise rotations about p and r. Then

$$\widehat{p}(x, y, z) = (x, -z, y) \quad \text{and} \quad \widehat{r}(x, y, z) = (-y, x, z).$$

Note that

$$Q(t)+Q(\widehat{p}t)=2Mx^2+(W-M)(y^2+z^2)$$

and

$$Q(t)+Q(\widehat{r}t)=2mz^2+(W-m)(x^2+y^2),$$

so

$$|f'(t)+f'(\widehat{p}t)|=\left|f(t)+f(\widehat{p}t)-(2Mx^2+(W-M)(y^2+z^2))\right|\le\delta,$$

(the inequality coming from Corollary 22) and

$$|f'(t)+f'(\widehat{r}t)|=\left|f(t)+f(\widehat{r}t)-(2mz^2+(W-m)(x^2+y^2))\right|\le\delta.$$

Thus $|f'(t)+f'(\alpha t)|\le\delta$ if α is any one of the basic rotations $\widehat{p}$, $\widehat{r}$, $\widehat{p}^{-1}$, and $\widehat{r}^{-1}$. If α_n is a product of n basic rotations, then

$$f'(t)-(-1)^nf'(\alpha_nt)=f'(t)-(-1)^{n-1}f'(\alpha_{n-1}t)+(-1)^{n-1}(f'(\alpha_{n-1}t)+f'(\alpha\alpha_{n-1}t));$$

so, by induction,

$$|f'(t)-(-1)^nf'(\alpha_nt)|\le n\delta.$$

We can rotate each of the six great circles 180° with a product of three basic rotations:

$$\begin{aligned}
\widehat{p}\widehat{p}\widehat{r}(x,x,z) &= (-x,-x,-z),\\
\widehat{p}\widehat{r}\widehat{r}(x,z,z) &= (-x,-z,-z),\\
\widehat{r}\widehat{p}^{-1}\widehat{r}(x,y,x) &= (-x,-y,-x),\\
\widehat{p}\widehat{p}\widehat{r}^{-1}(x,-x,z) &= (-x,x,-z),\\
\widehat{r}\widehat{r}\widehat{p}(x,z,-z) &= (-x,-z,z),\\
\widehat{r}\widehat{p}\widehat{r}(x,y,-x) &= (-x,-y,x).
\end{aligned}$$

So

$$|2f'(t)|=|f'(t)+f'(-t)|=|f'(t)+f'(\alpha t)|\le3\delta$$

on each of those great circles. That is, $|f(t)-Q(t)|\le3\delta/2$ on those great circles. In particular,

$$\left|f(t)-M(x^2-z^2)+Wz^2\right|\le3\delta/2$$

on the great circle $y=z$.

Let $M'=\sup f'$ and $m'=\inf f'$, and let $\delta'>0$. Note that the weight W' of f' is zero. Choose $\eta'>0$ for δ' from Corollary 22 and choose a frame (p',q',r') such that $f'(p')>M'-\eta'$ and $f'(r')<m'+\eta'$. In this case, the approximating form is

$$Q'(t)=M'x'^2-(M'+m')y'^2+m'z'^2.$$

By the previous paragraph, for t on the great circle $y' = z'$, we have

$$\left|f'(t) - M'(x'^2 - z'^2)\right| \leq 3\delta'/2.$$

We will combine that with the inequality $|f'(t)| \leq 3\delta/2$ on the (p, q, r) great circles to show that M' is small.

There is t on the great circle $y' = z'$, (almost) within $1/\sqrt{3}$ of p' and on one of the six (unprimed) great circles. That is

$$(1 - x')^2 + 2z'^2 = (1 - x')^2 + y'^2 + z'^2 \leq 1/3.$$

Also $x'^2 + 2z'^2 = 1$. So $2(1 - x') \leq 1/3$ or $x' \geq 5/6$ so $2z'^2 \leq 11/36$ whence $x'^2 - z'^2 \geq 13/24 > 1/2$. Now $|M'(x'^2 - z'^2)| \leq 3(\delta + \delta')/2$ so $|M'| \leq 3(\delta + \delta')$. As $\delta' > 0$ was arbitrary, $|M'| \leq 3\delta$. Same for m', so $|f(t) - Q(t)| \leq 3\delta$ for any t. ■

To wrap up the proof of Theorem 1 we must observe that, in an inner product space of dimension at least 3, every completely real two-dimensional subspace is contained in a completely real three-dimensional subspace. This follows from Lemma 5 with F generated by an orthogonal basis for the completely real two-dimensional space.

6. Constructing the operator

Let $\mathcal{H}$ be an inner product space. We want a definition of a measure on projections, and of a nonnegative frame function, that does not require bases of $\mathcal{H}$ to formulate or to use. This allows a treatment of Gleason's theorem that is not restricted to spaces with bases, or to separable spaces, and does not rely on any countable axiom of choice.

A **measure** μ assigns to each projection P a nonnegative real number $\mu(P)$ so that

1. If $\sigma = \sum P_i$, where the P_i are mutually orthogonal, then $\mu(\sigma) = \sum \mu(P_i)$.
2. For each P, and $\varepsilon > 0$, there exists a finite-dimensional $P' \leq P$ such that $\mu(P) < \mu(P') + \varepsilon$.

The extra Condition 2 follows from Condition 1 if each summand of $\mathcal{H}$ has a basis. This latter condition can be proved classically by a simple application of Zorn's lemma, and constructively for separable $\mathcal{H}$ using countable choice (Gleason's theorem is normally stated for separable spaces).

The definition of a frame function on a finite-dimensional space needs no modification. If f is a nonnegative frame function that comes from a measure μ as defined above, then

1. f is a nonnegative frame function on each finite-dimensional subspace of $\mathcal{H}$, and

2. There is a number W, the weight of f, such that for each finite-dimensional subspace K and each $\varepsilon > 0$, there is a finite-dimensional subspace $F \supset K$ with the property that the weight of f on F is within ε of W.

We take this as our definition of a **nonnegative frame function of weight W on $\mathcal{H}$**, and call a finite-dimensional subspace F of $\mathcal{H}$ satisfying Condition (2) an **ε-subspace for** f. It is easy to see that if F is an ε-subspace for f, then the weight of f on each finite-dimensional subspace of $F^\perp$ is at most ε. In particular, $f(x) \leq \varepsilon$ if x is a unit vector in $F^\perp$.

If $\mathcal{H}$ has dimension at least 3, then Theorem 1 says that each nonnegative frame function on $\mathcal{H}$ is given by a positive form. In particular, nonnegative frame functions are uniformly continuous on bounded subsets of $\mathcal{H}$. Note that a nonnegative frame function is bounded by its weight.

If T is a bounded operator on an inner product space, then $B_T(x, y) = \langle Tx, y\rangle$ defines a bounded bilinear form that determines T. We carry over terminology from B_T to T. We say that a bilinear form B is **finite dimensional** if there is a finite-dimensional subspace F such that $B(x, y) = 0$ for $y \in F^\perp$. So an operator T is finite dimensional if there is a finite-dimensional subspace F such that $TF \subset F$ and $TF^\perp = 0$. We say that a bilinear form B is **compact** if it can be approximated by finite-dimensional forms.

For any bilinear form B, let $B_P(x, y)$ denote the bilinear form $B(Px, Py)$, and write $\|B\| \leq \varepsilon$ if $|B(x, y)| \leq \varepsilon$ for all unit vectors x and y. A bilinear form B is compact if and only if for each $\varepsilon > 0$ there exists a projection P onto a finite-dimensional subspace such that

$$\|B - B_P\| \leq \varepsilon.$$

The space of finite-dimensional operators on a Hilbert space is a metric space, the completion of which is the set of compact operators.

If B is a bilinear form, then $f(x) = B(x, x)$ is a frame function on each finite-dimensional subspace. We say that a positive form B is of **trace class** if f is a frame function, in which case the **trace** of B is the weight of f. Constructively, not every positive form can be written as B_T for some operator T, but the compact ones can if the space is complete.

Theorem 24. *If a positive bilinear form on an inner product space is of trace class, then it is compact.*

Proof. Let B be a positive bilinear form with trace t, and $\varepsilon > 0$. Let P be the projection onto an ε-subspace for the frame function $B(x,x)$, and $\bar{P} = 1 - P$. Then

$$|B(x,y) - B(Px,Py)| = |B(\bar{P}x,\bar{P}y) + B(Px,\bar{P}y) + B(\bar{P}x,Py)|$$

for unit vectors x and y. By the Schwarz inequality for positive forms, this is at most $\varepsilon + 2\sqrt{t\varepsilon}$. ∎

The next theorem is the only place where the inner product space must be assumed complete.

Theorem 25. *If B is a compact bilinear form on a Hilbert space, then there exists a compact operator T such that $B(x,y) = \langle Tx, y\rangle$ for all x and y.*

Proof. For each projection P onto a finite-dimensional subspace, there exists a unique linear transformation T_P of that subspace such that

$$B_P(x,y) = \langle T_P x, y\rangle$$

for all x and y. If $\|B - B_P\| \leq \varepsilon$, and $\|B - B_{P'}\| \leq \varepsilon'$, then $\|B_P - B_{P'}\| \leq \varepsilon + \varepsilon'$, so

$$|\langle T_P x - T_{P'} x, y\rangle| = |B_P(x,y) - B_{P'}(x,y)| \leq \varepsilon + \varepsilon'$$

for x and y in the unit ball. So $\|T_P(x) - T_{P'}(x)\| \leq \varepsilon + \varepsilon'$ for x in the unit ball. As the space is complete, this defines a compact operator T. ∎

That completes the proof of Gleason's theorem. Theorem 1 provides a positive form $B(x,y)$ such that $f(x) = B(x,x)$. The form B is of trace class because f is a nonnegative frame function. Theorem 24 says that B is compact, and Theorem 25 says that $B(x,y) = \langle Tx, y\rangle$ for a compact operator T, which is necessarily positive and of trace class.

References

[1] Billinge, Helen, *A constructive formulation of Gleason's theorem*, J. Philosophical Logic, 1997.

[2] Bridges, Douglas, *A constructive look at positive linear functionals on $\mathcal{L}(H)$*, Pacific J. Math. **95**(1981), 11–25.

[3] Cooke, Roger, Michael Keane and William Moran, *An elementary proof of Gleason's theorem*, Math. Proc. Camb. Phil. Soc. **98**(1985), 117–128.

[4] Gleason, Andrew M., *Measures on the closed subspaces of a Hilbert space*, J. Math. Mech. **6**(1957), 885–893.

[5] Hellman, Geoffrey, *Gleason's theorem is not constructively provable*, J. Philosophical Logic **22**(1993), 193–203.

[6] Hooker, Clifford Alan (ed.) *The logico-algebraic approach to quantum mechanics*, D. Reidel, 1975.

[7] Jordan, P. and John von Neumann, *On inner products in linear metric space*, Annals of Math. **36**(1935), 719–793.

[8] Scedrov Andre, *Diagonalization of continuous matrices as a representation of intuitionistic reals*, Ann. Pure Appl. Logic **30**(1986), 201–206.