The Pythagorean Theorem: I. The finite case

Richard V. Kadison†

Mathematics Department, University of Pennsylvania, Philadelphia, PA 19104-6395

Contributed by Richard V. Kadison, December 17, 2001

The Pythagorean Theorem and variants of it are studied. The variations evolve to a formulation in terms of noncommutative, conditional expectations on von Neumann algebras that displays the theorem as the basic result of noncommutative, metric, Euclidean Geometry. The emphasis in the present article is finite dimensionality, both ''discrete'' and ''continuous.''

1. Introduction and Theme

Most of us carry away from our earliest contact with elementary mathematics memories of two basic formulae from Euclidean Geometry: πr^2, the "area" of a circle with radius r, and $a^2 + b^2 = c^2$, the formula relating the lengths, a and b, of the two sides of a right triangle to the length, c, of the hypotenuse of that triangle. That last formula, the Pythagorean Theorem, is the most basic result of "metric" Euclidean Geometry.

In this article, we study that theorem and variants of it. Our study falls into two large parts: the case of "discrete dimensionality" and the case of "continuous dimensionality." Each of these parts, in turn, falls into two parts: finite dimensionality and infinite dimensionality. The primary focus of this article is discrete dimensionality in the finite case, although we discuss the continuous case in the last section (where the meaning of the discrete-continuous division will become clearer). At the same time, in that section, we formulate the Pythagorean Theorem in terms of (noncommutative) conditional expectations and note its "semicommutative" nature. In this context (noncommutative, finite-continuous-dimensional, metric Euclidean Geometry), we prove a fully noncommutative version of the theorem. The next article in this series deals with discrete dimensionality in the infinite case. Arguments become more involved in that case.

Elementary, mostly finite-dimensional, variants of the Pythagorean Theorem are examined in the next section, some of them new. A converse, which we refer to as the Carpenter's Theorem, is introduced. The proof of this converse is carried out by operator–matrix methods in the third section. In this same section, we view the Pythagorean Theorem in terms of traces, in terms of indices, and in terms of stochastic matrices.

The fourth section contains a discussion of the finite-continuous case. The Carpenter's Theorem is left open in that case as a subject for later elucidation.

2. Elementary Variations

To begin with, the Pythagorean Theorem refers to "plane geometry." Are there three-dimensional, n-dimensional, or even infinite-dimensional analogues of that theorem? Of course there are, and they are familiar—but first we must recast the theorem mildly. If we replace the two sides of the triangle by "orthogonal" axes and the hypotenuse by a vector x of length c, the "orthogonal projections" of that vector on the axes have lengths a and b satisfying $a^2 + b^2 = c^2$, by virtue of the Pythagorean Theorem. This is our first variation.

By choosing vectors e_1 and e_2 of length 1 (*unit vectors*) along the positive (orthogonal) axes, the projections of x on these axes allow us to "expand" x in terms of the *orthonormal basis* $\{e_1, e_2\}$ (for the plane). That is, we express x as the linear combination $c_1e_1 + c_2e_2$ of e_1 and e_2. In this case, $|c_1| = a$, $|c_2| = b$, and the length $\|x\|$ of x is c, where $a^2 + b^2 = c^2$. This is our second variation of the Pythagorean Theorem.

In this form, we can take the leap (our third variation) to *Hilbert space* $\mathcal{H}$ of any dimension. With $\{e_a\}_{a\in\mathbb{A}}$ an orthonormal basis for $\mathcal{H}$, and x in $\mathcal{H}$, there is an expansion, $x = \Sigma_{a\in\mathbb{A}}\, c_ae_a$, where the equality refers to convergence of finite subsums to x in the "metric" of the Hilbert space. The *inner product* of vectors x and y in $\mathcal{H}$ is denoted by $\langle x, y\rangle$, and the *length* (or *norm*) $\|x\|$ of x is $\langle x, x\rangle^{1/2}$. Convergence of $\Sigma_{a\in\mathbb{A}}\, c_ae_a$ is over the "net" of finite subsets of $\mathbb{A}$ (directed by inclusion). The *Parseval equality* tells us that $\|x\|^2 = \Sigma_{a\in\mathbb{A}}\, |c_a|^2$, which is a direct extension of the Pythagorean Theorem, to (Hilbert) space of any dimension.

In the context of "infinite-dimensional" Hilbert space, there is more to be said. Given a potential set of coefficients $\{c_a\}_{a\in\mathbb{A}}$, there is a (unique) vector x in $\mathcal{H}$ with expansion $\Sigma_{a\in\mathbb{A}}\, c_ae_a$ if and only if $\Sigma_{a\in\mathbb{A}}\, |c_a|^2$ converges (in which case $\Sigma_{a\in\mathbb{A}}\, |c_a|^2$ converges to $\|x\|^2$). Some aspect of this added information is present in the Pythagorean Theorem when that theorem is suitably formulated (our fourth variation): the positive numbers a and b are the lengths of the sides of a right triangle with hypotenuse of length c if and only if $a^2 + b^2 = c^2$. Carpenters use this aspect to check that their work is "true." We shall refer to this "converse" to the usual statement of the Pythagorean Theorem as *the Carpenter's Theorem*.

The "expansion" formulation of the Pythagorean Theorem involves projecting a vector onto orthogonal axes. We can reverse that and formulate the theorem (our fifth variation) in terms of the projections of vectors of equal length along the axes onto the line determined by a vector.

In this case, the lengths of the projections of the axis vectors of length c onto the line have lengths a and b such that $a^2 + b^2 = c^2$, again as a result of the Pythagorean Theorem. It is not an essential restriction in this formulation to insist that c be 1. We are, then, projecting orthonormal basis vectors onto the line. Can something of this nature be said for orthonormal bases in higher-dimensional spaces? Our sixth variation follows.

Proposition 1. *If $\{e_a\}_{a\in\mathbb{A}}$ is an orthonormal basis for the Hilbert space $\mathcal{H}$, then the sum of the squares of the lengths of the orthogonal projections of each e_a on every one-dimensional subspace of $\mathcal{H}$ is 1. If a real non-negative t_a is specified for each a and $\Sigma_{a\in\mathbb{A}}\, t_a^2 = 1$, then $\Sigma_{a\in\mathbb{A}}\, t_ae_a$ is a unit vector x in $\mathcal{H}$ that generates a one-dimensional subspace of $\mathcal{H}$ on which each e_a has projection of length t_a.*

†E-mail: kadison@math.upenn.edu.

 www.pnas.org/cgi/doi/10.1073/pnas.032677199

Proof: If x is a unit vector and $\mathcal{V}$ is the one-dimensional subspace of $\mathcal{H}$ spanned by x, then the orthogonal projection of e_a on $\mathcal{V}$ is $\langle e_a, x\rangle x$ and $\|\langle e_a, x\rangle x\|^2 = |\langle e_a, x\rangle|^2\|x\|^2 = |\langle e_a, x\rangle|^2$. From Parseval's equality,

$$1 = \|x\|^2 = \sum_{a \in \mathbb{A}} |\langle x, e_a\rangle|^2 = \sum_{a \in \mathbb{A}} |\langle e_a, x\rangle|^2.$$

Because $\langle e_{a'}, \Sigma_{a\in\mathbb{A}}\, t_a e_a\rangle = t_{a'}$, each e_a has a projection on the one-dimensional space generated by $\Sigma_{a\in\mathbb{A}}\, t_a e_a$ of length t_a. ■

Equivalently, from the Pythagorean Theorem, we can specify the distances s_a from e_a to the one-dimensional space subject to the condition that $\Sigma_{a\in\mathbb{A}}\, 1 - s_a^2 = 1$. Of course, the question of orthogonal projections of basis elements may be asked when the projections are made onto a subspace of $\mathcal{H}$ of dimension other than 1. What is the situation if, for example, $\mathcal{V}$ is an m-dimensional subspace of $\mathcal{H}$? In this case, choosing an orthonormal basis $\{f_1, \ldots, f_m\}$ for $\mathcal{V}$, we have that the projection of e_a on $\mathcal{V}$ is $\Sigma_{j=1}^m \langle e_a, f_j\rangle f_j$ of length whose square is $\Sigma_{j=1}^m |\langle e_a, f_j\rangle|^2$. Now, $\Sigma_{a\in\mathbb{A}} \Sigma_{j=1}^m |\langle e_a, f_j\rangle|^2$ converges, because all terms are real and non-negative, and

$$\sum_{j=1}^{m} \sum_{a \in \mathbb{A}} |\langle e_a, f_j\rangle|^2 = \sum_{j=1}^{m} \|f_j\|^2 = \sum_{j=1}^{m} 1 = m,$$

from Parseval's equality. We have proved our seventh variation.

Proposition 2. *The sums of the squares of the lengths of the projections of the elements of an orthonormal basis for a Hilbert space $\mathcal{H}$ onto an m-dimensional subspace of $\mathcal{H}$ is m.*

Our eighth variation is an interesting, although small, alteration of *Proposition 2*. We emphasize it as our definitive (geometric) formulation of the finite-dimensional Pythagorean Theorem because it puts in evidence a property that will play an important role in our extension of the Carpenter's Theorem to infinite dimensions.

Proposition 3. *If a is the sum of the squares of the lengths of the projections of r elements of an orthonormal basis $\{e_1, \ldots, e_n\}$ for an n-dimensional Hilbert space $\mathcal{H}$ onto an m-dimensional subspace $\mathcal{H}_0$, and b is the sum of the squares of the projections of the remaining $n - r$ basis elements on the orthogonal complement $\mathcal{H}_0'$, then*

$$a - b = m - n + r.$$

Proof: If a_j is the square of the length of the projection of e_j on $\mathcal{H}_0$, then $1 - a_j$ is the square of the length of its projection on $\mathcal{H}_0'$. Thus $a = a_1 + \cdots + a_r$, $b = 1 - a_{r+1} + \cdots + 1 - a_n$, and $m = a_1 + \cdots + a_n$ from *Proposition 2*. It follows that

$$a - b = a_1 + \cdots + a_n - n + r = m - n + r.$$

Another proof, one that does not make use of *Proposition 2*, which is not available, of course, in the infinite-dimensional case, follows. Let $\{e_1, \ldots, e_n\}$ be an orthonormal basis for $\mathcal{H}$. Let $\{f_1, \cdots, f_m\}$ and $\{f_{m+1}, \cdots, f_n\}$ be orthonormal bases for $\mathcal{H}_0$ and $\mathcal{H}_0'$, respectively. The projection y of e_j on $\mathcal{H}_0$ is $\Sigma_{k=1}^m \langle e_j, f_k\rangle f_k$, and $\|y\|^2 = \Sigma_{k=1}^m |\langle e_j, f_k\rangle|^2$. Thus $a = \Sigma_{j=1}^r \Sigma_{k=1}^m |\langle e_j, f_k\rangle|^2$. The projection of e_j on $\mathcal{H}_0'$ is $\Sigma_{k=m+1}^n \langle e_j, f_k\rangle f_k$, and the square of its length is $\Sigma_{k=m+1}^n |\langle e_j, f_k\rangle|^2$, which is $1 - \Sigma_{k=1}^m |\langle e_j, f_k\rangle|^2$ because $1 = \|e_j\|^2 = \Sigma_{k=1}^n |\langle e_j, f_k\rangle|^2$, from Parseval's equality. Thus

$$b = \sum_{j=r+1}^{n} \left(1 - \sum_{k=1}^{m} |\langle e_j, f_k\rangle|^2\right) = n - r - \sum_{j=r+1}^{n} \sum_{k=1}^{m} |\langle e_j, f_k\rangle|^2$$

and

$$a - b = \sum_{j=1}^{r} \sum_{k=1}^{m} |\langle e_j, f_k\rangle|^2 + \sum_{j=r+1}^{n} \sum_{k=1}^{m} |\langle e_j, f_k\rangle|^2 - n + r$$

$$= \sum_{j=1}^{n} \sum_{k=1}^{m} |\langle e_j, f_k\rangle|^2 - n + r$$

$$= \sum_{k=1}^{m} \sum_{j=1}^{n} |\langle e_j, f_k\rangle|^2 - n + r = m - n + r. \quad ■$$

We note, especially, that the difference $a - b$ is an integer however we split the basis for projection onto $\mathcal{H}_0$ and $\mathcal{H}_0'$. If we move a basis element from those projected onto $\mathcal{H}_0'$ to those projected onto $\mathcal{H}_0$, we increase the difference by 1; if we move a basis element in the opposite sense, we decrease the difference by 1, clearly not affecting the integrality of the difference. In the next section, we introduce matrix methods and give another proof.

Once again, we can ask whether the lengths of the projections of the basis elements can be specified subject to the condition that the sum of their squares is m (the Carpenter's Theorem for this case). That is, given such a specification, is there an m-dimensional subspace of $\mathcal{H}$ on which the projections of the basis elements have those lengths? Equivalently, from the Pythagorean Theorem, can we find an m-dimensional subspace of $\mathcal{H}$ from which the basis elements have specified distances not greater than 1, subject to the condition that subtracting their squares from 1 produces numbers that sum to m? The affirmative answer to these questions provides our ninth and tenth variations. Their proof requires more involved arguments.

3. Operator–Matrix Methods

We assume, first, that $\mathcal{H}$ has finite dimension n, and that $\{e_1, \dots, e_n\}$ is an orthonormal basis for $\mathcal{H}$. Let $\mathcal{H}_0$ be an m-dimensional subspace of $\mathcal{H}$ and E the orthogonal projection of $\mathcal{H}$ onto $\mathcal{H}_0$. If (a_{jk}) is the matrix of E relative to $\{e_j\}$, then $a_{jk} = \langle Ee_k, e_j\rangle$ for all k and j. Since $E = E^* = E^2$,

$$a_{jj} = \langle Ee_j, e_j\rangle = \langle E^2 e_j, e_j\rangle = \langle Ee_j, Ee_j\rangle = \|Ee_j\|^2,$$

and $\Sigma_{j=1}^n a_{jj} = \Sigma_{j=1}^n \|Ee_j\|^2$. It follows that the sum of the squares of the lengths of the projections of the basis elements $e_1, \dots, e_n$ onto $\mathcal{H}_0$ is the trace of E. Of course, there is a unitary transformation U of $\mathcal{H}$ onto itself that maps $\mathcal{H}_0$ onto the m-dimensional space generated by $\{e_1, \dots, e_m\}$. The projection E_m that has that space as its range has matrix (b_{jk}) relative to $\{e_j\}$ such that $b_{11}, \dots, b_{mm}$ are 1 and $b_{m+1\,m+1}, \dots, b_{nn}$ are 0. Since $UEU^{-1} = E_m$, E and E_m have the same trace m. This proves again our sixth variation, $\mathrm{tr}(E) = \Sigma_{j=1}^n \|Ee_j\|^2 = m$, where "tr" is the functional that assigns to a matrix its usual (non-normalized) trace, the sum of its diagonal entries. It tells us, too, that another (our eleventh) variation of the Pythagorean Theorem is the assertion:

The trace of a projection with m-dimensional range is m.

From these same considerations, we see that prescribing the squares of the lengths of the projections of the basis elements on an m-dimensional subspace of $\mathcal{H}$ amounts to prescribing the diagonal of the matrix, relative to that basis, of the projection with that subspace as range. Our Carpenter's Theorem question, in this case, becomes:

Is an ordered n-tuple, $\langle a_1, \dots, a_n\rangle$ *of numbers in* $[0, 1]$ *with sum m the diagonal of some idempotent self-adjoint* $n \times n$ *matrix*?

This has an affirmative answer. (Together with the ninth variation, it provides an extension, our twelfth variation, of the fourth variation.) For its proof, we make use of a variant of a combinatorial–geometric lemma used in ref. 1.

Definition 4: With $(a_1, \dots, a_n)$ $(= \tilde{a})$ a point in $\mathbb{R}^n$ and Π the group of permutations of $\{1, \dots, n\}$, we let $\mathcal{K}_{\tilde{a}}$ be the (closed) convex hull of $\{(a_{\pi(1)}, \dots, a_{\pi(n)})\ (= \pi(\tilde{a})): \pi \in \Pi\}$ $(= \Pi(\tilde{a}))$. We refer to $\mathcal{K}_{\tilde{a}}$ as *the permutation polytope generated by* $\tilde{a}$.

Lemma 5. *If* $a_1 \geq a_2 \geq \cdots \geq a_n$, $b_1 \geq b_2 \geq \cdots \geq b_n$, *and* $a_1 + \cdots + a_n = b_1 + \cdots + b_n$, *then the following are equivalent*:

(*i*) $(b_1, \dots, b_n)$ $(= \tilde{b}) \in \mathcal{K}_{\tilde{a}}$;

(*ii*) $b_1 \leq a_1$, $b_1 + b_2 \leq a_1 + a_2, \dots, b_1 + \cdots + b_{n-1} \leq a_1 + \cdots + a_{n-1}$;

(*iii*) *There are points* $(a_1^{(1)}, \dots, a_n^{(1)})$ $(= \tilde{a}_1), \dots, (a_1^{(n)}, \dots, a_n^{(n)})$ $(= \tilde{a}_n)$ *in* $\mathcal{K}_{\tilde{a}}$ *such that* $\tilde{a}_1 = \tilde{a}$, $\tilde{a}_n = \tilde{b}$, *and* $\tilde{a}_{k+1} = t\tilde{a}_k + (1-t)\tau(\tilde{a}_k)$ *for each k in* $\{1, \dots, n-1\}$, *some transposition* τ *in* Π, *depending on k*, *and some t in* $[0, 1]$, *depending on k*.

Proof: (*i*)→(*ii*). From the assumption that $a_1 \geq \cdots \geq a_n$, we conclude that $a_1 + \cdots + a_j \geq a_{\pi(1)} + \cdots + a_{\pi(j)}$, for each j in $\{1 \dots, n\}$ and π in Π. Thus for each convex combination $\tilde{b}$ of points in $\Pi(\tilde{a})$ and j in $\{1, \dots, n\}$, $b_1 + \cdots + b_j \leq a_1 + \cdots + a_j$.

(*iii*)→(*i*). As $\pi(\tilde{d}) \in \Pi(\tilde{a})$ when $\tilde{d} \in \Pi(\tilde{a})$, $\pi(\tilde{c}) \in \mathcal{K}_{\tilde{a}}$ when $\tilde{c} \in \mathcal{K}_{\tilde{a}}$. Thus $\tilde{a}_1 = \tilde{a} \in \mathcal{K}_{\tilde{a}}$, $\tilde{a}_2 = t\tilde{a}_1 + (1-t)\tau(\tilde{a}_1) \in \mathcal{K}_{\tilde{a}}$, $\dots$, $\tilde{b} = \tilde{a}_n = t'\tilde{a}_{n-1} + (1-t')\tau'(\tilde{a}_{n-1}) \in \mathcal{K}_{\tilde{a}}$.

(*ii*)→(*iii*). If $b_1 < a_j$ for all j in $\{2, \dots, n\}$, then $b_j \leq b_1 < a_j$ for all such j, and $b_1 + \cdots + b_n < a_1 + \cdots + a_n$, contrary to assumption. Let m be the smallest number in $\{2, \dots, n\}$ such that $a_m \leq b_1$. Since $a_m \leq b_1 \leq a_1$, there is a t in $[0, 1]$ such that $b_1 = ta_1 + (1-t)a_m$. Let τ be the transposition that interchanges 1 and m. Let $\tilde{a}_1$ be $\tilde{a}$ and $\tilde{a}_2$ be $t\tilde{a}_1 + (1-t)\tau(\tilde{a}_1)$. Then

$$\begin{aligned}(a_1^{(2)}, \dots, a_n^{(2)}) &= (ta_1 + (1-t)a_m, a_2, \dots, a_{m-1}, ta_m + (1-t)a_1, a_{m+1}, \dots, a_n)\\ &= (b_1, a_2, \dots, a_{m-1}, a_1 + a_m - b_1, a_{m+1}, \dots, a_n).\end{aligned}$$

As $b_{m-1} \leq b_{m-2} \leq \cdots \leq b_1 < a_{m-1} \leq \cdots \leq a_2$, by choice of m,

$$b_1 \leq a_1^{(2)} (=b_1), b_1 + b_2 \leq a_1^{(2)} + a_2^{(2)} = b_1 + a_2, \dots, b_1 + \cdots + b_j \leq a_1^{(2)} + \cdots + a_j^{(2)}\ (j < m).$$

If $m \leq j \leq n-1$, then $a_1^{(2)} + \cdots + a_j^{(2)} = a_1 + \cdots + a_j \geq b_1 + \cdots + b_j$.

Suppose now that we have constructed $\tilde{a}_1, \dots, \tilde{a}_j$ such that $\tilde{a}_{k+1} = t\tilde{a}_k + (1-t)\tau(\tilde{a}_k)$ for each k in $\{1, \dots, j-1\}$ ($t \in [0, 1]$ and τ is a transposition in Π depending on k), such that $b_1 = a_1^{(k)}, \dots, b_{k-1} = a_{k-1}^{(k)}$ for each k in $\{2, \dots, j\}$ and

$$b_1 \leq a_1^{(k)}, b_1 + b_2 \leq a_1^{(k)} + a_2^{(k)}, \dots, b_1 + \cdots + b_{n-1} \leq a_1^{(k)} + \cdots + a_{n-1}^{(k)}$$

for each k in $\{1, \dots, j\}$. Then

$$b_1 + \cdots + b_j \leq a_1^{(j)} + \cdots + a_{j-1}^{(j)} + a_j^{(j)} = b_1 + \cdots + b_{j-1} + a_j^{(j)}.$$

Hence $b_j \leq a_j^{(j)}$. In addition, for k in $\{1, \cdots, j-1\}$,

$$a_1^{(k+1)} + \cdots + a_n^{(k+1)} = a_1^{(k)} + \cdots + a_n^{(k)} = \cdots = a_1 + \cdots + a_n = b_1 + \cdots + b_n.$$

Thus $a_n^{(j)} \leq b_n \leq b_j$, because $b_1 + \cdots + b_{n-1} \leq a_1^{(j)} + \cdots + a_{n-1}^{(j)}$. Let m be the smallest number in $\{j+1, \dots, n\}$ such that $a_m^{(j)} \leq b_j$. Then

$$(*)\qquad b_{j+1} \le b_j \le a_{j+1}^{(j)}, \dots, b_{m-1} \le b_j \le a_{m-1}^{(j)}.$$

Because $a_m^{(j)} \le b_j \le a_j^{(j)}$, there is a t in $[0, 1]$ such that $b_j = ta_j^{(j)} + (1 - t)a_m^{(j)}$. Let τ be the transposition that interchanges j and m, and let $\tilde{a}_{j+1}$ be $t\tilde{a}_j + (1 - t)\tau(\tilde{a}_j)$. Then

$$(a_1^{(j+1)}, \dots, a_n^{(j+1)}) = (b_1, \dots, b_j, a_{j+1}^{(j)}, \dots, a_{m-1}^{(j)}, a_j^{(j)} + a_m^{(j)} - b_j, a_{m+1}^{(j)}, \dots, a_n^{(j)}).$$

If $j + 1 = n$, we are through. If not, we must show that $b_1 + \cdots + b_k \le a_1^{(j+1)} + \cdots + a_k^{(j+1)}$, for each k in $\{1, \dots, n - 1\}$, to carry the construction forward. If $1 \le k \le j$, then

$$b_1 + \cdots + b_k = a_1^{(j+1)} + \cdots + a_k^{(j+1)}.$$

If $j + 1 \le k \le m - 1$, then from $(*)$,

$$b_1 + \cdots + b_j + b_{j+1} + \cdots + b_k \le b_1 + \cdots + b_j + a_{j+1}^{(j)} + \cdots + a_k^{(j)} = a_1^{(j+1)} + \cdots + a_k^{(j+1)}.$$

Finally, if $m \le k \le n - 1$, then

$$b_1 + \cdots + b_k \le a_1^{(j)} + \cdots + a_k^{(j)} = a_1^{(j+1)} + \cdots + a_k^{(j+1)}. \quad \blacksquare$$

Theorem 6. *Let φ be the mapping that assigns to each self-adjoint $n \times n$ matrix (a_{jk}) the point $(a_{11}, \dots, a_{nn})$ $(= \tilde{a})$ in $\mathbb{R}^n$, $\mathcal{K}_m$ be the range of φ restricted to the set $\mathcal{P}_m$ of projections of rank m, where $m \in \{0, \dots, n\}$, and $\mathcal{K}$ be the range of φ restricted to the set $\mathcal{P}$ of projections. Then $\tilde{a} \in \mathcal{K}_m$ if and only if $0 \le a_{jj} \le 1$, for each j and $\Sigma_{j=1}^n a_{jj} = m$, and $\tilde{a} \in \mathcal{K}$ if and only if $0 \le a_{jj} \le 1$, for each j, and $\Sigma_{j=1}^n a_{jj} \in \{0, \dots, n\}$.*

Proof: Let (a_{jk}) $(= A)$ be a self-adjoint matrix and U be the unitary matrix with $\xi \sin\theta$, $\sin\theta$ at the j, j and k, k entries, respectively, $-\cos\theta$, $\xi\cos\theta$ at the j, k and k, j entries, respectively, 1 at all diagonal entries other than j, j and k, k, and 0 at all other entries, where ξ is a complex number of modulus 1 such that $\overline{\xi a_{jk}} = -\xi a_{jk}$. Then UAU^{-1} has $a_{jj}\sin^2\theta + a_{kk}\cos^2\theta$ at the j, j entry, $a_{jj}\cos^2\theta + a_{kk}\sin^2\theta$ at the k, k entry, and a_{hh} at the hh entry when $h \ne j, k$. Letting t be $\sin^2\theta$, τ be the transposition of $\{1, \dots, n\}$ that interchanges j and k, and $\tilde{a}_\tau$ be $(a_{\tau(1),\tau(1)}, \dots, a_{\tau(n),\tau(n)})$, we see that

$$\varphi(UAU^{-1}) = t\tilde{a} + (1 - t)\tilde{a}_\tau.$$

Because $VEV^{-1} \in \mathcal{P}_m$ for each unitary V, when $E \in \mathcal{P}_m$, we see that, when $\tilde{a} \in \mathcal{K}_m$, so is $t\tilde{a} + (1 - t)\tilde{a}_\tau$, for each t in $[0, 1]$ and each transposition τ of $\{1, \dots, n\}$.

As noted, $t\tilde{a} + (1 - t)\tilde{a}_\tau \in \mathcal{K}_m$ when $\tilde{a} \in \mathcal{K}_m$, for each t in $[0, 1]$ and each transposition τ of $\{1, \dots, n\}$. From *Lemma 5*, $\mathcal{K}_m$ contains the permutation polytope $\mathcal{K}_{\tilde{a}}$ of each $\tilde{a}$ in $\mathcal{K}_m$. Now the point $\tilde{a}$ whose first m coordinates are 1 and whose last $n - m$ coordinates are 0 is in $\mathcal{K}_m$. If $\tilde{b} = (b_1, \dots, b_n)$, $0 \le b_j \le 1$ for each j in $\{1, \dots, n\}$ and $\Sigma_{j=1}^n b_j = m$, then it follows that $b_1 \le 1, b_1 + b_2 \le 1 + 1, \dots, b_1 + \cdots + b_m \le m, b_1 + \cdots + b_{m+1} \le m + 0, \dots, b_1 + \cdots + b_{n-1} \le m$. Again, from our lemma, $\tilde{b} \in \mathcal{K}_{\tilde{a}} \subseteq \mathcal{K}_m$. Thus $\mathcal{K}_m$ is as described in the statement. In particular, $\mathcal{K}_m$ is convex.

Since $\mathcal{K} = \cup_{m=0}^n \mathcal{K}_m$, $\mathcal{K}$ is as described in the statement. $\blacksquare$

We present another proof of our twelfth variation (*Theorem 6*) and extend the information contained there slightly, to yield our thirteenth variation. Specifically, we prove the following result.

Theorem 7. *If $\langle a_1, \dots, a_n\rangle$ is an ordered n-tuple of numbers in $[0, 1]$ with sum a positive integer, then there is an idempotent self-adjoint $n \times n$ matrix with diagonal entries $a_1, \dots, a_n$ and all entries real.*

Proof: Our proof proceeds by induction on m, the sum of $a_1, \dots, a_n$. In the case where m is 1, we let E_1 be the projection matrix (acting on $\mathbb{C}^n$ in the standard manner) that has range spanned by the vector $(x =)$ $(a_1^{1/2}, \dots, a_n^{1/2})$. Let $\{e_j\}$ be the orthonormal basis for $\mathbb{C}^n$ where e_j is the n-tuple with 1 at the jth coordinate and 0 at all others. The matrix for E_1 relative to this basis has $\langle E_1 e_k, e_j\rangle$ as its j, kth entry. Since

$$\langle E_1 e_k, e_j\rangle = \langle\langle e_k, x\rangle x, e_j\rangle = \langle a_k^{1/2} x, e_j\rangle = a_k^{1/2} a_j^{1/2},$$

each entry of the matrix is a non-negative real number (positive, when no a_j is 0, and perforce none is 1 in this case, unless $n = 1$). The jth diagonal entry is $\langle E_1 e_j, e_j\rangle$ $(= a_j \ge 0)$, as desired.

We take the inductive step. Suppose our assertion has been established when $a_1, \dots, a_n$ has sum $m - 1$ (where m is an integer 2 or greater). Assume that $a_1 + \cdots + a_n = m$. Let k be the smallest integer j for which $a_1 + \cdots + a_j \ge m - 1$ and a be $m - 1 - \Sigma_{r=1}^{k-1} a_r$. By inductive hypothesis, there is a self-adjoint idempotent E_2 with matrix (a_{jr}) relative to the basis $\{e_j\}$, such that each a_{jr} is real, with diagonal $a_1, \dots, a_{k-1}, a, 0, \dots, 0$. Let F_2 be E_2 with the $k + 1, k + 1$ entry replaced by 1. Each a_{jr} with j or r greater than k is 0 (since $E_2 \ge 0$). Hence F_2 is a projection. Let $W_k(\theta)$ be the unitary operator whose matrix relative to the basis $\{e_j\}$ has $\sin\theta$ at the k, k and $k + 1, k + 1$ entries, $-\cos\theta$ and $\cos\theta$ at the $k, k + 1$ and $k + 1, k$ entries, respectively, 1 at all other diagonal entries, and 0 at all other off-diagonal entries. Let $p(k, \theta, F_2)$ be $W_k(\theta)F_2W_k(\theta)^*$. Relative to the basis $\{e_j\}$, the matrix of $p(k, \theta, F_2)$ has diagonal entries $a_1, \dots, a_{k-1}, a\sin^2\theta + \cos^2\theta, a\cos^2\theta + \sin^2\theta, 0, \dots, 0$. The j, r entry is a_{jr} when both j and r do not exceed $k - 1$ and 0 when either j or r is greater than $k + 1$. The entries in the kth row of the matrix for $p(k, \theta, F_2)$ are $a_{k1}\sin\theta, \dots, a_{kk-1}\sin\theta, a\sin^2\theta + \cos^2\theta, (a - 1)\sin\theta\cos\theta, 0, \dots, 0$. The entries in the $k + 1$st row are $a_{k1}\cos\theta, \dots, a_{kk-1}\cos\theta, (a - 1)\sin\theta\cos\theta, a\cos^2\theta + \sin^2\theta, 0, \dots, 0$. The entries in the kth column are $a_{1k}\sin\theta, \dots, a_{k-1k}\sin\theta, a\sin^2\theta + \cos^2\theta, (a - 1)\sin\theta\cos\theta, 0, \dots, 0$ and in the $k + 1$st column are $a_{1k}\cos\theta, \dots, a_{k-1k}\cos\theta, (a - 1)\sin\theta\cos\theta, a\cos^2\theta + \sin^2\theta, 0, \dots, 0$.

By choice of k, $m - 1 \le \Sigma_{r=1}^{k-1} a_r + a_k$, whence $a = m - 1 - \Sigma_{r=1}^{k-1} a_r \le a_k \le 1$. For an appropriate choice θ_2 of θ, $a\sin^2\theta_2 + \cos^2\theta_2 = a_k$, and

MATHEMATICS

$$a\cos^2\theta_2 + \sin^2\theta_2 = a + 1 - a_k = m - \sum_{r=1}^{k} a_r = \sum_{r=k+1}^{n} a_r.$$

Let $p(k, \theta_2, F_2)$ be F_3. Each entry in the matrix for F_3 is real.

The projection $p(k+1, \theta, F_3)$ has as its diagonal entries $a_1, \ldots, a_k$, $(\Sigma_{r=k+1}^{n} a_r)\sin^2\theta$, $(\Sigma_{r=k+1}^{n} a_r)\cos^2\theta, 0, \ldots, 0$. Again, for an appropriate choice θ_3 of θ, $(\Sigma_{r=k+1}^{n} a_r)\sin^2\theta_3 = a_{k+1}$. Thus the projection $p(k+1, \theta_3, F_3)$ $(= F_4)$ has as its diagonal entries $a_1, \ldots, a_{k+1}$, $\Sigma_{r=k+2}^{n} a_r, 0, \ldots, 0$. We continue with this construction, forming $p(k+2, \theta, F_4)$ next and so forth, until we consider $p(n-1, \theta, F_{n-k+1})$. Choosing θ_{n-k+1} appropriately, we let F_{n-k+2} be the self-adjoint idempotent matrix $p(n-1, \theta_{n-k+1}, F_{n-k+1})$. The diagonal entries of the matrix for F_{n-k+2} are $a_1, \ldots, a_{n-1}, a_n$, and all entries are real. ∎

Remark 8. When we constructed $p(k, \theta, F_2)$ in the preceding argument, if F_2 is replaced by A with the same matrix except that the $k+1, k+1$ entry is b rather than 1, then the matrix of $p(k, \theta, A)$ has diagonal entries $a_1, \ldots, a_{k-1}$, $a\sin^2\theta + b\cos^2\theta$, $a\cos^2\theta + b\sin^2\theta, 0, \ldots, 0$. The entries of the kth and $k+1$st rows and columns remain the same except that "$(a-1)$" becomes "$(a-b)$," and "$a\sin^2\theta + \cos^2\theta$" and "$a\cos^2\theta + \sin^2\theta$" become "$a\sin^2\theta + b\cos^2\theta$" and "$a\cos^2\theta + b\sin^2\theta$." All other entries remain the same. This general process of transforming a matrix by our unitary matrix so that two segments of the diagonal are altered by replacing their terminal and initial elements by convex combinations of the two in such a way that the sum of the original elements is the same as the sum of the replacements will be referred to as *splicing*.

We have applied this general construction once, when b is 1 and for the rest with 0 for b. With 1 for b, $(a-1)\sin\theta_2\cos\theta_2$ appears at the $k+1, k$ and $k, k+1$ entries of F_3. Because $a < 1$ and $\theta_2 \in (0, \pi/2)$, in general these entries are negative, even though E_1 has a matrix of non-negative real entries, and F_2 may have all its entries real and non-negative. At a lecture on this topic, Frank Hansen raised the possibility of constructing our projection with specified diagonal so that all its entries are real and non-negative.[‡] This is accomplished in the case of a one-dimensional projection by the construction given in *Proposition 1*. It would be interesting to know whether this is possible in general, and whether the construction can be altered to produce such a projection.

Remark 9. As noted at the end of Section 2, the question of whether there is an m-dimensional subspace of our n-dimensional Hilbert space from which the elements of a given orthonormal basis $\{e_1, \ldots, e_n\}$ have distances $r_1, \ldots, r_n$, respectively, is equivalent, by the Pythagorean Theorem, to the existence of such a subspace on which $e_1, \ldots, e_n$ have orthogonal projections of lengths $t_1, \ldots, t_n$, respectively, where $r_j^2 + t_j^2 = 1$. Because this latter question is answered affirmatively by *Theorem 6* if and only if $t_1^2 + \cdots + t_n^2 = m$, and $\Sigma_{j=1}^{n} r_j^2 + t_j^2 = n$, the former question is answered affirmatively if and only if $0 \le r_j \le 1$ and $r_1^2 + \cdots + r_n^2 = n - m$. This last variation, our fourteenth, is equivalent to the assertion that there is an "m-plane" through the origin tangent to each of the spheres $S_1, \ldots, S_n$ with centers at $e_1, \ldots, e_n$ and radii $r_1, \ldots, r_n$, respectively, if and only if $0 \le r_j \le 1$ and $r_1^2 + \cdots + r_n^2 = n - m$.

Remark 10. Another proof of the formula of *Proposition 3* was promised earlier. With the notation established in that proposition and its proof, let F be the projection of $\mathcal{H}$ onto $\mathcal{H}_0$ and E the projection with range spanned by $\{e_1, \ldots, e_r\}$. Then $a = \Sigma_{j=1}^{r} \langle Fe_j, e_j\rangle = \mathrm{tr}(EFE)$ and $b = \Sigma_{r+1}^{n} \langle (I-F)e_j, e_j\rangle = \mathrm{tr}((I-E)(I-F)(I-E))$. Thus

$$\begin{aligned} a - b &= \mathrm{tr}(EFE) - \mathrm{tr}((I-E)(I-F)(I-E)) = \mathrm{tr}(EF) - \mathrm{tr}((I-E)(I-F)) \\ &= \mathrm{tr}(EF) + \mathrm{tr}((I-E)F) - \mathrm{tr}(I-E) = \mathrm{tr}(F) - \mathrm{tr}(I-E) \\ &= m - n + r. \end{aligned}$$

Formulated in matrix terms, this equality takes on the following form: If a is the sum of any r elements of the diagonal of an $n \times n$ matrix of a projection of rank m, and b is the sum of the result of subtracting each of the remaining $n - r$ diagonal elements from 1, then $a - b = m - n + r$. In these same matrix terms, a $(= \mathrm{tr}((FE)^*FE))$ is the trace of the principal upper $r \times r$ block of the matrix for F (relative to $\{e_j\}$) and also the sum of the squares of the absolute values of the entries in the matrix for FE, that is, the sum of those squares for the entries in the first r columns of the matrix for F. This sum of squares is the square of the Hilbert–Schmidt norm of FE (and of EF). We write "$\|FE\|_2^2$" for that sum. (More will be said about this in the infinite-dimensional case.) In this notation, our formula is $\|FE\|_2^2 - \|(I-F)(I-E)\|_2^2 = m - n + r$. Surprisingly (at first sight), $\|EFE\|_2^2 - \|(I-E)(I-F)(I-E)\|_2^2$ is also $m - n + r$. To prove this, note that $\|EFE\|_2^2$ is the sum of the squares of the absolute values of the matrix entries of the principal upper $r \times r$ block of the matrix for F, and $\|(I-E)(I-F)(I-E)\|_2^2$ is the same sum for the principal lower $(n-r) \times (n-r)$ block of $I - F$; their difference is $\|FE\|_2^2 - \|(I-E)(I-F)\|_2^2$ (at the same time $\|(I-E)(I-F)\|_2^2$ is $\|(I-F)(I-E)\|_2^2$). In addition,

$$\|EFE\|_2^2 - \|(I-E)(I-F)(I-E)\|_2^2 = \mathrm{tr}(E - (I-F)) = m - n + r,$$

by a straightforward trace computation of the type we used in proving the formula $\mathrm{tr}(EFE) - \mathrm{tr}((I-E)(I-F)(I-E)) = m - n + r$.

With the notation established in this remark, if we assume that the ranges of E and F and of their complements have intersections (0), then we may view $a - b$ $(= m - n + r)$ as the index of $E(I-F)$. To see this, note that the null space of $E(I-F)$ is $F(\mathcal{H}) \vee ((I-F)(\mathcal{H}) \wedge (I-E)(\mathcal{H}))$, which is $F(\mathcal{H})$, by assumption. (See ref. 3, proposition 2.5.14.) The null space of $(I-F)E$ $(= [E(I-F)]^*)$ is $(I-E)(\mathcal{H}) \vee (E(\mathcal{H}) \wedge F(\mathcal{H}))$, which is $(I-E)(\mathcal{H})$, by assumption. Thus the index of the operator $E(I-F)$ is $m - (n-r)$ $(= m - n + r)$.

[‡]Hansen, F., May 3, 2000, Copenhagen.

Remark 11. Another approach to proving our formula "$a - b = m - n + r$" results from stochastic–matrix methods. We describe the stochastic matrices, introducing some terminology and establishing some basic facts that will be useful to us. Although our present interest is the finite discrete case, these stochastic–matrix considerations will reappear in the infinite case.

For later use, we develop the basics in the infinite case as well as the finite. We deal with matrices having complex entries and the property that each row and each column sums to r. If r is 1 and all entries are non-negative real numbers, the matrices are the well-studied *doubly stochastic matrices* (the entries representing stationary transition probabilities from one state of a discrete Markov process to another). If an "r-sum" matrix has n rows and m columns (with n and m finite and r non-zero), then $n = m$ for summing each row, and then all the sums yields nr as the sum of all entries while summing each column, and then all those sums yields mr as the sum of all matrix entries.

We say that the submatrix A_0 of a matrix A whose rows are indexed by a set $\mathbb{A}$ and whose columns are indexed by a set $\mathbb{B}$ consisting of those entries in the rows corresponding to a given subset $\mathbb{A}_0$ of $\mathbb{A}$ and, at the same time, in the columns corresponding to a subset $\mathbb{B}_0$ of $\mathbb{B}$ is a *block* (in A, *the* $\mathbb{A}_0$, $\mathbb{B}_0$ *block*). The *complementary block* A_0' to A_0 is the $\mathbb{A}_0'$, $\mathbb{B}_0'$ block in A, where $\mathbb{A}_0' = \mathbb{A}\backslash\mathbb{A}_0$ and $\mathbb{B}_0' = \mathbb{B}\backslash\mathbb{B}_0$. The *weight* $\mathrm{w}(A_0)$ of the block A_0 is the sum of its entries. In the case where A_0 and hence A have an infinite number of entries, this sum is taken over the net of finite subsums, directed by inclusion, provided that net converges. If the entries of A are non-negative real numbers, r is positive, and A is infinite, then $\mathrm{w}(A)$ is ∞, for each row sums to r and there are an infinite number of rows. Of course, $\mathrm{w}(A_0)$ is finite when A_0 is a finite block. In this case, the sum of the entries in the (finite number of) rows and columns corresponding to A_0 is finite, whence $\mathrm{w}(A_0')$, the sum of the remaining entries in A, is ∞ (still under the assumption that A is an infinite matrix). Despite these observations, there are infinite blocks A_0, with infinite complements A_0', such that $\mathrm{w}(A_0)$ and $\mathrm{w}(A_0')$ are both finite. The article on the infinite discrete case to follow this article will contain a description of a method for generating such blocks.

The differences of the weights of complementary blocks of doubly stochastic matrices are intimately related to the Pythagorean Theorem. To describe that relation, we note first that each pair of orthonormal bases $\{e_j\}_{j\in\mathbb{Z}_0}$ and $\{f_j\}_{j\in\mathbb{Z}_0}$ of a Hilbert space $\mathcal{H}$, where $\mathbb{Z}_0 = \mathbb{Z}_+ \cup \mathbb{Z}_-$, $\mathbb{Z}_+$ are the positive integers, and $\mathbb{Z}_-$ are their negatives, gives rise to a doubly stochastic matrix. If $a_{jk} = |\langle e_j, f_k\rangle|^2$, then $\Sigma_{k\in\mathbb{Z}_0} a_{jk} = \|e_j\|^2 = 1$ for each j in $\mathbb{Z}_0$, from Parseval's equality, because $e_j = \Sigma_{k\in\mathbb{Z}_0}\langle e_j, f_k\rangle f_k$. Symmetrically, $\Sigma_{j\in\mathbb{Z}_0} a_{jk} = \|f_k\|^2 = 1$ for each k in $\mathbb{Z}_0$. Thus (a_{jk}) is a doubly stochastic infinite matrix. If U is the unitary operator on $\mathcal{H}$ such that $Uf_j = e_j$ for each j in $\mathbb{Z}_0$, then $\langle Uf_j, f_k\rangle = \langle e_j, f_k\rangle = u_{kj}$, the k, j entry of the matrix for U corresponding to the basis $\{f_j\}$. Thus $|u_{kj}|^2 = a_{jk}$.

In the case of finite doubly stochastic matrices, we derive a formula relating the weights of complementary blocks (a "Pythagorean Theorem" for doubly stochastic matrices) that provides us with another proof of our formula, $a - b = m - n + r$.

Proposition 12. *If A is an $n \times n$ doubly stochastic matrix and A_0 is a block in A with p rows and q columns, then*

$$\mathrm{w}(A_0) - \mathrm{w}(A_0') = p - n + q.$$

Proof: The sum of the p rows of A corresponding to the p rows of A_0 is p, and the sum of the $n - q$ columns of A corresponding to the columns of A_0' is $n - q$. The difference of these sums, $p - n + q$, is $\mathrm{w}(A_0) - w(A_0')$. ∎

Given an orthonormal basis $\{e_1, \ldots, e_n\}$ for the n-dimensional Hilbert space $\mathcal{H}$ and an m-dimensional subspace $\mathcal{H}_0$ with orthogonal complement $\mathcal{H}_0'$, choose orthonormal bases $\{f_1, \ldots, f_m\}$ and $\{f_{m+1}, \ldots, f_n\}$ for $\mathcal{H}_0$ and $\mathcal{H}_0'$, respectively, and let a_{jk} be $|\langle e_j, f_k\rangle|^2$. As noted in the infinite-dimensional case, (a_{jk}) is a doubly stochastic matrix A, an $n \times n$ matrix, in this case. If A_0 is the $r \times m$ block whose entries are a_{jk} with j in $\{1, \ldots, r\}$ and k in $\{1, \ldots, m\}$, and F is the projection of $\mathcal{H}$ onto $\mathcal{H}_0$, then Fe_j is $\Sigma_{k=1}^m \langle e_j, f_k\rangle f_k$ and $(I - F)e_j$ is $\Sigma_{k=m+1}^n \langle e_j, f_k\rangle f_k$. Thus $\|Fe_j\|^2$ is the sum of the jth row of A_0, when $1 \le j \le r$, and $\|(I - F)e_{r+j}\|^2$ is the sum of the jth row of A_0', when $1 \le j \le n - r$. Thus these sums are a_j and $1 - a_{r+j}$, respectively, where a_p is the pth diagonal entry of the matrix for F relative to the basis $\{e_1, \ldots, e_n\}$. It follows that $\Sigma_{j=1}^r a_j - \Sigma_{j=r+1}^n 1 - a_j = \mathrm{w}(A_0) - \mathrm{w}(A_0') = m - n + r$ from *Proposition 12*. Again, with a the sum of the squares of the lengths of the projections of the r elements $e_1, \ldots, e_r$ of $\{e_1, \ldots, e_n\}$ onto $\mathcal{H}_0$ and b the sum of the squares of the lengths of the projections of the remaining $n - r$ basis elements onto $\mathcal{H}_0'$, $a - b = m - n + r$.

4. Finite Continuous Dimensionality

The Pythagorean and Carpenter's Theorems, in the form of *Proposition 2* and its operator–matrix variant (referring to the trace of a rank m projection), deal with projections on an n-dimensional Hilbert space $\mathcal{H}$ and the diagonals of their matrices with respect to a given orthonormal basis. Denoting by "$\mathcal{B}(\mathcal{H})$" the algebra of all (bounded) operators on $\mathcal{H}$ (also when $\mathcal{H}$ is infinite dimensional) and by "$\mathcal{A}$" the algebra of all operators in $\mathcal{B}(\mathcal{H})$ with diagonal matrices relative to the given basis, we have that $\mathcal{A}$ is a maximal abelian self-adjoint subalgebra of $\mathcal{B}(\mathcal{H})$ (a "masa," that is, if $TA = AT$ for each A in $\mathcal{A}$, then $T \in \mathcal{A}$, and $A^* \in \mathcal{A}$ when $A \in \mathcal{A}$). The masas are precisely the subalgebras of $\mathcal{B}(\mathcal{H})$ whose matrices are diagonal relative to some fixed orthonormal basis for $\mathcal{H}$. For our purposes, orthonormal bases and masas are interchangeable. The mapping Φ that assigns to T in $\mathcal{B}(\mathcal{H})$ the element $\Phi(T)$ in the masa $\mathcal{A}$ corresponding to the diagonal of the matrix for T (relative to the orthonormal basis associated with $\mathcal{A}$) has special properties. It is linear [from $\mathcal{B}(\mathcal{H})$ onto $\mathcal{A}$], maps positive operators to positive operators, and maps the identity operator I in $\mathcal{B}(\mathcal{H})$ to I. With A and B in $\mathcal{A}$, we have that $\Phi(ATB) = A\Phi(T)B$. A mapping such as Φ is said to be a *conditional expectation* [of $\mathcal{B}(\mathcal{H})$ onto $\mathcal{A}$]. For this Φ, $\mathrm{tr}(TA) = \mathrm{tr}(\Phi(T)A)$ for each A in $\mathcal{A}$. In particular, $\mathrm{tr}(T) = \mathrm{tr}(\Phi(T))$, for each T in $\mathcal{B}(\mathcal{H})$. Conversely, if $\mathrm{tr}(T) = \mathrm{tr}(\Phi'(T))$ when $T \in \mathcal{B}(\mathcal{H})$, for a conditional expectation Φ' of $\mathcal{B}(\mathcal{H})$ onto $\mathcal{A}$, then, again, $\mathrm{tr}(TA) = \mathrm{tr}(\Phi'(TA)) = \mathrm{tr}(\Phi'(T)A)$, for each A in $\mathcal{A}$. Thus $\mathrm{tr}([\Phi(T) - \Phi'(T)]A) = 0$, for each A in the algebra $\mathcal{A}$, and $\mathrm{tr}([\Phi(T) - \Phi'(T)][\Phi(T) - \Phi'(T)]^*) = 0$. It follows that $\Phi(T) - \Phi'(T) = 0$ for each T in $\mathcal{B}(\mathcal{H})$ and $\Phi = \Phi'$. When the conditional expectation Φ has the property that $\mathrm{tr}(T) = \mathrm{tr}(\Phi(T))$, for each T in $\mathcal{B}(\mathcal{H})$, we say that Φ *lifts the trace* [from $\mathcal{A}$ to $\mathcal{B}(\mathcal{H})$]. We have just proved that there is a unique conditional expectation of $\mathcal{B}(\mathcal{H})$ onto a masa that lifts the trace.

In these terms, with E a projection in $\mathcal{B}(\mathcal{H})$, $\mathrm{tr}(\Phi(E))$ is the sum of the squares of the lengths of the projections onto the range of E of the basis vectors corresponding to $\mathcal{A}$, where tr is the unique linear functional on $\mathcal{B}(\mathcal{H})$ such that $\mathrm{tr}(I) = n$ and $\mathrm{tr}(AB) = \mathrm{tr}(BA)$, for all A and B in $\mathcal{B}(\mathcal{H})$, and $\mathrm{tr}(E)$ is the rank m of E. Thus the equality

MATHEMATICS

$$(*) \qquad m = \mathrm{tr}(E) = \mathrm{tr}(\Phi(E))$$

is the Pythagorean Theorem as expressed in *Proposition 2*. In these same terms, the Carpenter's Theorem states that if $A \in \mathcal{A}$, $0 \leq A \leq I$, and $\mathrm{tr}(A) = m$, then there is a projection E in $\mathcal{B}(\mathcal{H})$ (necessarily of rank m), such that $\Phi(E) = A$.

Trace considerations will play a role when we discuss the Pythagorean Theorem for the case of an infinite-dimensional projection in the next article, although there is no trace functional defined on *all* of $\mathcal{B}(\mathcal{H})$ when $\mathcal{H}$ is infinite dimensional. There are, however, subalgebras of $\mathcal{B}(\mathcal{H})$, the factors of type II_1, that serve as an infinite-dimensional generalization of $\mathcal{B}(\mathcal{H})$ when $\mathcal{H}$ is finite dimensional that are, in many ways, a more appropriate replacement than the infinite-dimensional $\mathcal{B}(\mathcal{H})$. For one thing, these factors have a (unique) trace functional defined on them. For another, they are simple algebras, whereas the infinite-dimensional $B(\mathcal{H})$ is not. They can be characterized as the simple algebras consisting of all operators commuting with a self-adjoint operator algebra and admitting a trace.

Examples of factors $\mathcal{M}$ of type II_1 are provided by (countably) infinite (discrete) groups G, each of whose conjugacy classes, other than that of the unit e of G, is infinite (*i.c.c groups*). Let $\mathcal{H}$ be $l_2(G)$, the Hilbert space of complex-valued functions φ on G such that $\Sigma_{g\in G} |\varphi(g)|^2 < \infty$, provided with the inner product: $\langle\varphi, \psi\rangle = \Sigma_{g\in G}\, \varphi(g)\overline{\psi(g)}$. If $R_h\varphi(g) = \varphi(gh)$, for each φ in $\mathcal{H}$ and g in G, then R_h is a unitary operator on $\mathcal{H}$ (*right translation by h*). The family $\{T{:}TR_h = R_hT,\ h \in G\}$ [those operators in $\mathcal{B}(\mathcal{H})$ commuting with all R_h], denoted by "$\mathcal{L}_G$" is a factor of type II_1 (the "left von Neumann group algebra" of G).

Let $\mathcal{M}$ be a factor of type II_1 and τ the unique "tracial state" on $\mathcal{M}$ [characterized as a linear functional on $\mathcal{M}$ such that $\tau(I) = 1$ and $\tau(AB) = \tau(BA)$, for each A and B in $\mathcal{M}$, but possessing many other properties]. Each spectral projection E for a self-adjoint A in $\mathcal{M}$ is a limit on vectors ("strong-operator" limit) of a sequence $p_n(A)$, where p_n is a polynomial function on the reals. Thus $TE = ET$ when $TA = AT$, and $E \in \mathcal{M}$. It follows that $\mathcal{M}$ is generated by (the "norm closure" of the linear span of) the projections in $\mathcal{M}$. Restricted to these projections, τ is a "dimension function"—$\tau(E)$ being the dimension of the range of E "relative to $\mathcal{M}$." In the case of $\mathcal{B}(\mathcal{H})$, where $\mathcal{H}$ has finite dimension n, we used "tr" in place of τ, and $\mathrm{tr}(I)$ is n as is appropriate, because there are minimal projections in $\mathcal{B}(\mathcal{H})$. There are no minimal projections in a factor of type II_1 and no "natural" projection to which to assign trace ("dimension") 1 other than I. The structural properties of factors $\mathcal{M}$ of type II_1 allow us to conclude that for each real number a in $[0, 1]$, there are projections E in $\mathcal{M}$ such that $\tau(E) = a$; that is, the range of the dimension function on $\mathcal{M}$ is the entire closed unit interval $[0, 1]$. Thus the factors of type II_1 provide us with a natural extension of the finite-dimensional $\mathcal{B}(\mathcal{H})$ to a central simple algebra in which each of the projections has finite "rank" and the dimensions of the projections form a "continuous" range of values.

An orthonormal basis relative to $\mathcal{M}$ is precisely what we arrived at in the case of $\mathcal{B}(\mathcal{H})$, with $\mathcal{H}$ finite dimensional, that is, a masa $\mathcal{A}$ in $\mathcal{M}$. In this case, there is a conditional expectation Φ of $\mathcal{M}$ onto $\mathcal{A}$ that lifts the trace, although it is more complicated to construct than passing to the diagonal of a matrix (see p. 403 of ref. 2). The paraphrased version of $(*)$,

$$(**) \qquad \tau(E) = \tau(\Phi(E)),$$

is the Pythagorean Theorem for the case of finite continuous dimensionality (that is, in a factor of type II_1). The Carpenter's Theorem for this case asserts that each A in $\mathcal{A}$ such that $0 \leq A \leq I$ is $\Phi(E)$ for some projection E in $\mathcal{M}$. This will be proved in a later article.

A factor $\mathcal{M}$ of type II_1 may be thought of and studied as a noncommutative algebra of (bounded) measurable functions on a (noncommutative) measure space, the projections in $\mathcal{M}$ serving as the "characteristic" (or "indicator") functions on the measure space. In the case of a classical measure space, the algebra of bounded measurable functions on the space is (isomorphic to) a masa in some $\mathcal{B}(\mathcal{H})$. Our Pythagorean Theorem describes a property ("lifting the trace") of a mapping (conditional expectation) from the projections in $\mathcal{M}$ to a masa $\mathcal{A}$ in $\mathcal{M}$. So that theorem describes a certain (trace, that is, integral) property of a mapping from a noncommutative, finite, continuous measure algebra ($\mathcal{M}$) to a commutative measure algebra ($\mathcal{A}$). In that sense, it is a semicommutative result in the metric Euclidean geometry of spaces with finite continuous dimensionality.

Let $\mathcal{N}$ be a von Neumann subalgebra of a factor $\mathcal{M}$ of type II_1 (that is, $\mathcal{N}$ is a self-adjoint subalgebra of $\mathcal{M}$ consisting of all operators that commute with some other self-adjoint algebra). By techniques akin to those used to prove classical Radon–Nikodým results (suitably modified to apply to the case of noncommutative measure spaces), it was shown (in 1950) that there is a (unique) conditional expectation Φ of $\mathcal{M}$ onto $\mathcal{N}$ that lifts the trace. Thus $\tau(E) = \tau(\Phi(E))$ for each projection E in $\mathcal{M}$. If $\mathcal{N}$ is noncommutative, for example, if it is a subfactor of $\mathcal{M}$, then the domain and range of Φ are noncommutative. In that case, the equality, $\tau(E) = \tau(\Phi(E))$, is a (fully) noncommutative version of the Pythagorean Theorem. Again, the Carpenter's Theorem would describe the range of Φ restricted to the projections in $\mathcal{M}$. There is even a version of *Proposition 3* that is valid in a factor $\mathcal{M}$ of type II_1. If $\mathcal{A}$ is a masa in $\mathcal{M}$, E is a projection in $\mathcal{A}$, and F is a projection in $\mathcal{M}$, then

$$\tau(EFE) - \tau((I - E)(I - F)(I - E)) = \tau(F) - \tau(I - E) = \tau(F) - 1 + \tau(E).$$

The computation of *Remark 10* applies to prove this.

The Pythagorean investigation can be extended to include C* algebras with faithful tracial states and their C* subalgebras. Under what conditions are there trace-lifting conditional expectations, and what are their ranges when restricted to the projections in the algebra?

1. Kadison, R. & Pedersen, G. K. (1985) *Math. Scand.* **57,** 249–266.
2. Kadison, R. & Ringrose, J. (1998) *Fundamentals of the Theory of Operator Algebras* (Am. Math. Soc., Providence, RI), Vol. 4.
3. Kadison, R. & Ringrose, J. (1997) *Fundamentals of the Theory of Operator Algebras* (Am. Math. Soc., Providence, RI), Vol. 1.