Largest m-Cube in an n-Cube: Partial Solution

Kay R. Pechenick DeVicci Shultz[1, *]

With a new preface by

Greg Huber[2]

[1] *490 Cooper Landing Road, Apt. 211, Cherry Hill, New Jersey 08002, USA*
[2] *Kavli Institute for Theoretical Physics, Kohn Hall, University of California, Santa Barbara, California 93106, USA*

Abstract

We consider the problem of finding the largest m-dimensional cube (or m-cube) which fits into an n-dimensional cube of unit side. Let the side of this maximal m-cube be denoted by f(m, n). We solve the problem completely for all cases where m divides n, as well as for m = 2, with all odd values of n. The solution is essentially unique for these cases, and for m divides n, f(m, n) = √(n/m), while for m = 2, n odd, f(m, n) = √[(4n-3)/8]. We also show that f(3, 4) = $\sqrt{x_0}$, where x_0 is the (unique) real root of $4x^4-28x^3-7x^2+16x+16=0$ such that $1 < x_0 < 4/3$; thus $\sqrt{x_0} \approx 1.007435$. We derive some inequalities involving f(m, n) in general. We describe two general methods for attempting to solve the general problem, and we discuss some unanswered questions, including the relationship to Prince Rupert's problem.

These results were all obtained by 1996; only the Table of Contents, the Preface and the Appendices were added in 2013, and this abstract was modified slightly.

---

[*] Current official name: Kay R. Shultz

Table of Contents

Preface

The paper that you hold in your hands (or perhaps all you're holding is a retinal pattern) has been a long time in coming. On the surface, it is the revision of a 1996 manuscript containing a long and detailed proof. This itself is the incarnation of work that goes back some 15 years earlier. In fact, since as far back as I can remember, Kay Pechenick (as she was then known, before she became Kay Pechenick DeVicci, before she became Kay Pechenick DeVicci Shultz) has been fascinated by the problem of determining the largest cube that would fit inside a tesseract. The question was first posed by Martin Gardner in his November 1966 *Scientific American* column, but the roots of the problem go back centuries. My tongue was only lightly planted in cheek when I wrote the following talk abstract for a 2013 conference in honor of Gregory Galperin on his 60th birthday:

> **The De Vicci Cube, and other mysteries from the fourth dimension (and beyond)**
>
> There is a secret cube that lives in four dimensions called the De Vicci Cube. "Surely you mean a 'hypercube' or a 'tesseract'?" you might be forgiven to ask. No, I mean a regular three-dimensional cube. But this is no ordinary cube. Leonardo Da Vinci didn't know of its existence, nor did the Freemasons, and nor does Dan Brown (for, if he did, there would surely be another bestseller and movie out). It was first foreshadowed by the Babylonians long ago. Then, in the seventeenth century, a certain Prince Rupert of the Rhine wagered on the outcome an optimization problem. He got his answer, and he won his bet, but he didn't solve the problem. The problem wasn't solved until Pieter Nieuwland examined it over 100 years later. The solution was
>
> 1.06066017177982128660126654315...
>
> which, *à la manière égyptienne*,
>
> = 1 + 1/17 + 1/545 + 1/561804 + ...
>
> Nieuwland was getting close to the Cube, but soon after his discovery, he died under very mysterious circumstances. And still the De Vicci Cube remained a secret, lurking in the sea of undiscovered things. What is the De Vicci Cube? We know that De Vicci's Cube has linear dimension
>
> 1.00743475688427937609825359524...
>
> = 1 + 1/135 + 1/36564 + …
>
> This was calculated by at least six people (all dead now) before K. De Vicci discovered it. The Cube has acquired De Vicci's name because she

> supplied an actual proof. Erdös knew about the Cube, but soon after learning of it, he died too. There is much more to say. More than this, you must attend the talk, as I have written too much already.

The history reads like a screenplay: It's got Babylonians, ne'er-do-well princes and kings, untimely death, Latin manuscripts, and the 4th dimension -- all in one package. Add to this the mystery of a little-known mathematical genius who lives in Cherry Hill, New Jersey... what more could one want? But unlike a Hollywood movie, the DeVicci Cube story is actually true.

Some historical notes and corrections are in order. The phrase "DeVicci Cube" derives from Steven R. Finch's coinage "the DeVicci Constant" (= 1.0074347…) in his book *Mathematical Constants*. Also, a study of Martin Gardner's extensive correspondence reveals that only three others had calculated the DeVicci Constant before DeVicci. They are Eugen Bosch (1966), G. de Josselin de Jong (1971), and Hermann Baer (1974). In addition, although his solution for the 3-cube in a 4-cube is incorrect, Andrew L. Clarke (in 1967) had already formulated the correct inequalities and other results for the m-cube in an n-cube generalized Rupert problem.

I first got involved with the problem when, after first hearing from Kay about her work, I was able to calculate the DeVicci Constant to six decimal places. Later I found a proof for the statement f(m, m+1) > 1 (in the notation of the present paper), and more recently, Terry Ligocki and I have used techniques from computational physics to deeply probe the full f(m,n) problem. It is true that both Paul Erdös (and Raphael Robinson) looked at the f(m,n) problem, but neither of them made any inroads into the problem.

In one of his earlier collections of columns, *Mathematical Carnival*, Gardner claimed that he had received seven solutions to his four-dimensional Rupert problem, but that it was difficult for him to evaluate them as none of them agreed with each other. Gardner re-reviewed the situation in 2001, in his *The Colossal Book of Mathematics*, where he devotes a page (page 172) to his generalization of the Rupert problem. He corrects himself there, but then goes on to mention two people, Kay R. Pechenick DeVicci and Kay R. Pechenick, not realizing that these are one and the same person.

Finally, many people are familiar with the generalized Prince Rupert problem from the one-page description in the book by Croft, Falconer and Guy *Unsolved Problems in Geometry*, which is both fortunate in terms of publicity, but unfortunate in terms of scholarship. Their entry on this problem was selectively paraphrased from personal letters sent to Richard Guy, they credit neither Pechenick DeVicci Shultz nor myself, and they get important details wrong -- a flawed reference.

The present manuscript is also not perfect: first and foremost, its length is daunting. But it is a primary reference, and it should be viewed as a source for future work on the Rupert problem.

Introduction

In a discussion of hypercubes of all dimensions [1], Martin Gardner posed the problem of finding the largest square which fits into a unit cube (i.e., a cube of unit side), as well as the largest cube in a unit tesseract. He gave the solution to the largest square in a cube, but he stated that the problem of the largest cube in a tesseract was still unsolved. Seven people had sent him solutions, all different, and since none of these proofs had been published or checked by mathematicians, he regarded the problem as still unsolved. Even in a later edition of his book [2], Gardner wrote that nothing had been published on the problem.

Soon after I started thinking about these problems, I started wondering about the largest square in a tesseract. This led me to generalize the problem to finding the largest m-cube (that is, an m-dimensional cube) in an n-cube, for all m and n for which the problem makes sense: $1 \leq m \leq n$.

I found the solution for all (m,n) where m divides n, and also for all n when m = 2. I have a candidate for (m,n) = (3,4), but I have not yet been able to prove that it is the largest cube in a tesseract.* Letting f(m,n) denote the side of the largest m-cube in an n-cube of unit side, I have a lower bound for f(3,4). I have proved several inequalities involving f(m,n), and I have done additional work on m = 3, n = 4, and for n = m + 1 in

* Yes, I have. See Appendix A, pA1.

general.

Method of Solution

Instead of thinking about the largest m-cube in an n-cube, we will consider a unit m-cube in an arbitrary orientation in Euclidean n-space, and try to find the smallest n-cube that contains the m-cube, where the n-cube is in a "standard orientation": i.e., all of its edges are parallel to Cartesian coordinate axes.

We may assume that one of the vertices of the m-cube is at the point $\vec{P}_0$, and that the m orthonormal vectors in n-space, $\vec{v}_i$, i = 1, 2, ..., m, are such that the 2^m vertices of the m-cube are at $\vec{P}_0$, $\vec{P}_0+\vec{v}_1$, $\vec{P}_0+\vec{v}_2$, . . , $\vec{P}_0+\vec{v}_m$, $\vec{P}_0+\vec{v}_1+\vec{v}_2$, . . . $\vec{P}_0+\vec{v}_1+\vec{v}_2+\ldots+\vec{v}_m$. We need to determine how far the m-cube extends in each of the n coordinate directions. Since the m-cube is the convex hull of its vertices, we need only consider the vertices. We let v_{ij} represent the jth component of $\vec{v}_i$, j = 1 through n. A little thought leads to the conclusion that R_j, the range in the jth direction in n-space, is given by

$$R_j = \sum_{i=1}^{m} |v_{ij}|$$

Then the smallest n-cube in a standard orientation which contains the m-cube has a side equal to

$$R = \max R_j, \quad j = 1, 2, \ldots, n.$$

Thus we want to minimize R, where

$$R = \max \left(\sum_{i=1}^{m} |v_{ij}| \right), \quad j = 1, 2, \ldots, n, \text{ where}$$

$$\sum_{k=1}^{n} v_{ik} v_{jk} = \delta_{ij}$$

(where $\delta_{ij} = 1$ for $i = j$, and 0 otherwise.)

Then f(m,n) is the reciprocal of the minimum value of R. We know that R really has a minimum and not just an infimum, because R is a continuous function of the v_{ij}, and the v_{ij} satisfying the above equation form a compact set in an mn-dimensional Euclidean space.

If all of the R_i are equal, then we call the m-cube "snug-fitting" -- it fits snugly into the n-cube in the sense that it cannot undergo any translation and remain within the n-cube.

In some of our calculations, we will find it convenient to use a different normalization, such that the n-cube does have unit side, instead of the m-cube; thus, all the $\vec{v}_i$ will still have equal lengths but they will (usually) not have unit length.

We will describe embeddings of m-cubes in n-cubes by m x n matrices such that each row of a matrix represents the components of one of the $\vec{v}_i$. Thus the matrix element in the ith row and the jth column will be v_{ij}. The rows of the matrix will be mutually perpendicular and have equal lengths as vectors. We want to minimize the maximum of the "column-sums" (the sum of the absolute values of the numbers in each column) while keeping the row vectors mutually perpendicular and leaving their lengths unchanged.

These v_{ij} matrices, which we shall also call V matrices, have the property that the lengths of the row vectors, the maximum range, and the orthogonality of the rows are unaffected by certain trivial transformations: the interchange of any two rows or columns, and the multiplication of the elements of any row or column by -1.

When does $f(m,n) = \sqrt{(n/m)}$?

For m = 1, an m-cube is a line segment. The longest line segment in an n-cube of unit side is the main diagonal, so $f(1,n) = \sqrt{n}$. If an m-cube is inscribed in an n-cube, then the longest line segment in the m-cube cannot be longer than the longest line segment in the n-cube, so that $f(m,n) \le \sqrt{(n/m)}$. When does equality occur, besides $m = 1$?

Assume $m \ge 2$. For $f(m,n) = \sqrt{(n/m)}$, each main diagonal of the m-cube must be a main diagonal of the n-cube. Thus,

$$\pm \vec{v}_1 \pm \vec{v}_2 \pm \dots \pm \vec{v}_m = (\pm 1, \pm 1, \dots, \pm 1)$$

for all combinations of + and - signs on the left-hand side. Now, change the sign in front of one of the $\vec{v}_i$, and subtract and divide by 2. Then

$$\vec{v}_i = \text{(combination of 1's, -1's, and 0's)}$$

Since all of the $\vec{v}_i$ have length $\sqrt{(n/m)}$, there are (n/m) ± 1's among the components of each of the $\vec{v}_i$. Thus n/m must be an integer. Each row of the matrix of $\vec{v}_i$ components has (n/m) ± 1's and the rest zeros.

The V matrix has m rows and n columns, so the entire matrix contains n ± 1's. There cannot be a column of all zeros, so each column contains exactly one ± 1. By making trivial transformations, we can transform the matrix into n/m copies of the $m \times m$ identity matrix placed side by side. Thus equality occurs if and only if m divides n, and the maximal m-cubes are unique up to trivial transformations. These maximal m-cubes are snug-fitting, and they have the property that all of the vertices of the m-cube are also vertices of the n-cube.

The V matrix for the largest square in a tesseract, for example, may be written as

$$\begin{bmatrix} 1\,0\,1\,0 \\ 0\,1\,0\,1 \end{bmatrix}$$

or as

$$\begin{bmatrix} 1\,1\,0\,0 \\ 0\,0\,1\,1 \end{bmatrix}$$

A projection of it is shown in Figure 1.

(Please insert Figure 1 here.) Figure 1

The largest 4-cube in a 12-cube may be written as

$$\begin{bmatrix} 1\,1\,1\,0\,0\,0\,0\,0\,0\,0\,0\,0 \\ 0\,0\,0\,1\,1\,1\,0\,0\,0\,0\,0\,0 \\ 0\,0\,0\,0\,0\,0\,1\,1\,1\,0\,0\,0 \\ 0\,0\,0\,0\,0\,0\,0\,0\,0\,1\,1\,1 \end{bmatrix}$$

Looking for the Largest Cube in a Tesseract

To try to find the largest cube in a tesseract, I computed formulas for the components of three orthonormal vectors $\vec{v}_i$, in 4-space, which were functions of 6 angles (similar to the 3 Euler angles used to specify the orientation of a rigid body in 3-space). For each orientation (with the angles varying in steps of 5°), the computer computed R, and we printed the cases in which R < 1. The results suggested that the smallest value of

R occurs when the cube is snug-fitting ($R_1 = R_2 = R_3 = R_4$) and the V matrix has a triangle of zeros:

$$\begin{bmatrix} A & B & 0 & 0 \\ F & G & H & 0 \\ K & L & M & P \end{bmatrix}$$

We shall prove that, of all matrices with orthonormal rows and zeros as above, the minimum value of R occurs when all "column-sums" are equal, and we find this value of R.

The above matrix can be parametrized in terms of 3 angles (where $s\alpha \equiv \sin\alpha$, $c\alpha \equiv \cos\alpha$, etc.) :

$$\begin{bmatrix} -s\alpha & -c\alpha & 0 & 0 \\ -c\alpha\, s\beta & s\alpha\, s\beta & -c\beta & 0 \\ -c\alpha\, c\beta\, s\gamma & s\alpha\, c\beta\, s\gamma & s\beta\, s\gamma & -c\gamma \end{bmatrix}$$

We want to find the values of α, β, and γ which minimize R.

From the above matrix,

$R_1 = |s\alpha| + |c\alpha\, s\beta| + |c\alpha\, c\beta\, s\gamma|$, etc.

If we let α' be the angle (in the first quadrant) such that $\sin\alpha' = |\sin\alpha|$, $\cos\alpha' = |\cos\alpha|$, and similarly for β and γ, then we may write

$R_1 = s\alpha' + c\alpha'\, s\beta' + c\alpha' c\beta' s\gamma'$

$R_2 = c\alpha' + s\alpha's\beta' + s\alpha'c\beta's\gamma'$

$R_3 = c\beta' + s\beta's\gamma'$

$R_4 = c\gamma'$

with $0 \le \alpha',\beta',\gamma' \le 90°$.

Since $R = \max R_i$, $i = 1$ to 4, is a continuous function of α', β', and γ', on a compact set, there must be a minimum value. We can prove that this minimum value is strictly less than 1, by showing that we can choose the angles so that R is less than 1.

First, choose α' so that $0 < \alpha' < 90°$. Then $\sin\alpha'$ and $\cos\alpha'$ are < 1. Next, choose β' such that $s\alpha'+c\alpha's\beta'$ and $c\alpha'+s\alpha's\beta'$ and $c\beta'$ are < 1. Finally, choose γ' so that R_1, R_2, R_3, and R_4 are all < 1.

We will now prove that the minimum of R occurs at values of α', β', and γ' which make all of the R_i equal.

First we show that the minimum does not occur on the boundary (where at least one of α', β', γ' is 0 or 90°). If $\alpha' = 0$, then $c\alpha' = 1$ so that $R_2 \ge 1$. Similarly, if $\beta' = 0$, then $R_3 \ge 1$, and if $\gamma' = 0$, then $R_4 \ge 1$.

If $\alpha' = 90°$, then $c\alpha' = 0$ so we have zeros at least in the following places:

X 0 0 0
0 X X 0
0 X X X

Then in the upper left corner we must have a 1, so $R \geq 1$. If $\beta' = 90°$, then $\cos \beta' = 0$, so we have

X X 0 0
X X 0 0
0 0 X X

The upper left 2 x 2 matrix now corresponds to the largest square in a square, so that $\max(R_1, R_2) \geq 1$, so $R \geq 1$. Finally, if $\gamma' = 90°$ then we have $\cos \gamma' = 0$, so the matrix looks like this:

X X 0 0
X X X 0
X X X 0

and now we have the largest cube in a cube, so $\max(R_1, R_2, R_3) \geq 1$, so $R \geq 1$. Thus the minimum does not occur on the boundary, so we can now assume that $0 < \alpha', \beta', \gamma' < 90°$.

Suppose $R = R_1 > R_2, R_3, R_4$. (We say that the only "maximal column" is the first column.) Then we can decrease R_1 slightly by decreasing γ'. This decreases R.

Similarly, if $R = R_2$ and the only maximal column is column 2, then we decrease γ', thus decreasing R. This also works if the maximal column is column 3, or even if there are 2 or 3 maximal columns among the first three columns.

If instead R_4 is largest, we increase γ' slightly, which decreases R. Next, suppose R_1 and R_4 are largest. Note that $\partial^2 R_1/\partial\alpha'^2 < 0$, while R_4 is independent of α'. Thus, if $\partial R_1/\partial\alpha' > 0$, we decrease α' slightly, decreasing R_1 and leaving R_4 unchanged. If $\partial R_1/\partial\alpha' < 0$, we increase α' to decrease R_1. If $\partial R_1/\partial\alpha'$ happens to be zero, then since the second derivative is negative we can decrease R_1 by changing α' slightly in either direction. This reduces this case to the case in which only R_4 is equal to R. (We could have changed β' instead.)

If R_2 and R_4 are largest, a similar argument applies. If R_3 and R_4 are largest, a similar argument shows that we can decrease R_3 by changing β' slightly.

If R_1, R_3, and R_4 or R_2, R_3, and R_4 are largest, then we change α' slightly, reducing it to the case in which only R_3 and R_4 are largest.

The last case to consider is the one where R_1, R_2, and R_4 are largest. Since $R_1 = R_2$,

$$(s\alpha')(1-s\beta'-c\beta'\,s\gamma') = (c\alpha')(1 - s\beta' - c\beta'\,s\gamma').$$

Thus either $\alpha' = 45°$ or $s\beta' + c\beta' s\gamma' = 1$, or both.

If $\alpha' = 45°$, then we can change β' slightly so that R_1 and R_2 both decrease, while R_4 is unchanged.

If $s\beta' + c\beta' s\gamma' = 1$, then $R_1 = \sin \alpha' + \cos \alpha'$, but then $R_1 \geq 1$, so this is not a maximal cube. (Also, then R_1 cannot be equal to $\cos \gamma'$.)

We have thus found that the minimum R occurs when α', β', and γ' are chosen such that $R_1 = R_2 = R_3 = R_4 = R$. This will give us the largest cube in a tesseract with the triangle of 3 zeros in its V matrix. The 3 equations in α', β', and γ', with $\cos \gamma' = R$, can be reduced to an equation for R alone (after doing a lot of algebra):

$$4R^4 - 6\sqrt{2}\, R^3 + 11\, R^2 - 6\sqrt{2}\, R + 2 = 0.$$

We can get rid of the square roots, which occur because $\alpha' = 45°$, and we obtain

$$16\, R^8 + 16\, R^6 - 7\, R^4 - 28\, R^2 + 4 = 0.$$

Finally, letting $x = 1/R^2$, we get

$$4x^4 - 28x^3 - 7x^2 + 16x + 16 = 0.$$

There may be extraneous roots, but we know that at least one root is genuine, because of our previous work, and corresponds to the largest cube in a tesseract with the triangle of zeros in its V matrix.

Using the quartic equation formula, we find that there is only one root in the appropriate range (that is, between 1 and 4/3); we call this root x_0. The side of this cube in a unit tesseract is $\sqrt{x_0}$, where

$x_0 = 7/4 + r/2 - d/2$ where

$d = \sqrt{(105/4 - y + 90/r)}$, $r = \sqrt{(14 + y)}$, $y = a + b - 7/12$,

$a = [\ 184841/1728 + (5/2)\sqrt{1290}\ ]\text{^}(1/3)$

and $b = [\ 184841/1728 - (5/2)\sqrt{1290}\ ]\text{^}(1/3)$

(where ^ denotes exponentiation)

These formulas yield $\sqrt{x_0} \approx 1.00743475688$.

Thus we now have a rigorous proof that $f(3,4) \geq \sqrt{x_0}$ where x_0 is the root of $4x^4 - 28x^3 - 7x^2 + 16x + 16 = 0$ which is between 1 and 4/3.

Other cases where n = m + 1

Similar arguments apply when m = 1, 2, 4, and 5; the number of angles is equal to m . The matrices have the forms

$[\ -\sin\alpha \quad -\cos\alpha\ ]$ (m=1) which becomes $[\ -\sqrt{2}/2,\ -\sqrt{2}/2\ ]$

$$\begin{bmatrix} -\sin\alpha & -\cos\alpha & 0 \\ -c\alpha\, s\beta & s\alpha\, s\beta & -c\beta \end{bmatrix}$$

(for m = 2) which becomes

$$\begin{bmatrix} -\sqrt{2}/2 & -\sqrt{2}/2 & 0 \\ -\sqrt{2}/6 & \sqrt{2}/6 & -(2/3)\sqrt{2} \end{bmatrix}$$

or, with a different normalization:

$$\begin{bmatrix} 3/4 & 3/4 & 0 \\ 1/4 & -1/4 & 1 \end{bmatrix}$$

and for m = 4, the triangle of zeros looks like this:

$$\begin{bmatrix} X & X & 0 & 0 & 0 \\ X & X & X & 0 & 0 \\ X & X & X & X & 0 \\ X & X & X & X & X \end{bmatrix}$$

while for m = 5, we have

$$\begin{bmatrix} X & X & 0 & 0 & 0 & 0 \\ X & X & X & 0 & 0 & 0 \\ X & X & X & X & 0 & 0 \\ X & X & X & X & X & 0 \\ X & X & X & X & X & X \end{bmatrix}$$

(In the next section we will give a procedure for computing the sines and cosines in the matrix elements, for all m.)

Although I have not given a rigorous inductive proof, it seems clear that f(m,m+1) is always strictly greater than 1. For m = 2, we obtain Martin Gardner's square in a cube. However, we have not yet proved that it is optimal.

By considering these matrices for $n = m+1$, with the triangles of zeros and the parametrizations in terms of sines and cosines of m angles, we can derive a lower bound for $f(m, m+1)$ which is larger than 1. We shall not describe the proof here, but the result is that, for $m > 1$,

f(m, m+1) > {1 - (1 - √2/2) ^ [3^(m-1)] / (32) ^ ([3^(m-1) - 1]/2)} ^ (-1)

where ^ denotes exponentiation.

In deriving this lower bound, I actually used a formula from special relativity! I needed a "nice" function of two variables to use in place of the "max" function. Both variables were less than 1, and I wanted the function to be less than 1, and at least as large as both of them. So I used the velocity addition formula from special relativity (with c, the velocity of light, set equal to 1) !

Plane (2-dimensional) rotations

For any V matrix, we can obtain a new one by a rotation involving 2 columns, where, say, column i becomes (column i) $\cos\theta$ + (column j) $\sin\theta$, and column j becomes (column j) $\cos\theta$ - (column i) $\sin\theta$. In fact, the matrices in the previous section can be obtained by starting with a matrix such as

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad \text{(for } m = 3\text{)}$$

or

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \qquad \text{(for } m = 4\text{)}$$

and performing m rotations through different angles, in which the ith rotation involves columns i and m+1.

Using small rotations to decrease R_i (or "column-sums")

In this section we shall prove two useful theorems, although we will discuss specific cases to make the proofs less abstract and easier to understand.

Suppose that $(A, B, C)^T$ and $(D, E, F)^T$ are the ith and jth columns, respectively, of a V matrix. Let a = sign of A, b = sign of B, etc. Assume that the ith column contains no zeros. Let us make a rotation: col. i $\rightarrow$ (col. i) cos θ + (col. j) sin θ, col. j $\rightarrow$ (col. j) cos θ - (col. i) sin θ. Then, $A \rightarrow A\cos\theta + D\sin\theta$, so that $|A| \rightarrow |A\cos\theta + D\sin\theta| = a(A\cos\theta + D\sin\theta)$, for θ sufficiently near zero. We find that $d|A|/d\theta$ at $\theta = 0$ is aD, while $d^2|A|/d\theta^2$ at $\theta = 0$ is $-|A|$. Thus $dR_i/d\theta$ at $\theta = 0$ is aD +

$bE + cF$, while $d^2R_i/d\theta^2$ at $\theta = 0$ is $- |A| - |B| - |C|$, which is always negative.

This shows that we can always decrease R_i by making a small rotation. This theorem, which we shall call Theorem 1, can often be used to decrease R in particular situations to show that a particular V matrix is not optimal. For instance, if a V matrix with no zeros is not snug-fitting, then one can make small rotations to decrease R, so the m-cube cannot be optimal.

Even if the ith column contains zeros, we may still decrease its column-sum if the jth column contains a zero in every row in which the ith column contains a zero (unless, of course, the ith column contains all zeros).

Theorem 1 can be used, in most places, instead of our arguments about increasing or decreasing the angles α', β', and γ', to show that we can decrease R for our cube in the tesseract. One has to be careful sometimes to perform the small rotations in a particular order, since rotating columns i and j may change a zero in column j to a nonzero number, which may make prevent some other rotation from decreasing R.

Another useful theorem about two columns is especially useful where $R_i = R_j$, and shows that sometimes we can decrease R_i and R_j simultaneously. Let the ith and jth columns be $(A, B, C)^T$ and $(D, E, F)^T$ as before. Suppose that A, B, C, and D are positive and E and F are negative. Then $dR_i/d\theta$ at $\theta = 0$ is $aD + bE + cF$, while $dR_j/d\theta$ at $\theta = 0$ is $-dA - eB - fC$. So $dR_i/d\theta$ at $\theta = 0$ is $D + E + F = D - |E| - |F|$ while $dR_j/d\theta$ at $\theta = 0$ is $- A + B + C$. Assuming that neither derivative is zero, the two derivatives have the same sign if $A - B - C$ and $D - |E| - |F|$ have opposite signs, and then R_i and R_j can both be reduced by the same small rotation. We shall call this Theorem 2.

Both of these theorems will be used in the next section, in order to find the largest square in an odd-dimensional cube.

The Largest Square in an Odd-Dimensional Cube

To find the largest square in an n-dimensional cube for all odd n, we shall find it convenient to distinguish two possibilities:

(1) $f(2,3) > 1$

(2) $f(2,3) = 1$

We may already know that $f(2,3) > 1$, but we will pretend we don't know that, to show that we can prove it and derive the largest square in any odd-dimensional cube, without being clever enough to guess it first.

Consider possibility (1). We normalize the V matrix so that the largest column-sum is 1. If $f(2,3) > 1$, then we can append a 2x2 identity matrix to the 2 x 3 V matrix to obtain a valid V matrix for $n = 5$, so that $[f(2,5)]^2 \geq [f(2,3)]^2 + 1$. We can continue this process indefinitely, to show that if $f(2,3) > 1$, then $f(2,5) > \sqrt{2}$, $f(2,7) > \sqrt{3}$, and so on.

We will first prove that, for maximality, all column-sums are equal. That is, the square is snug-fitting: it touches all hyperfaces of the n-dimensional cube.

First of all, the optimal V matrix cannot have any columns which have both entries equal to zero. For if it had an all-zero column, the matrix would correspond to the largest square in an (n-1)-cube, and we would be in possibility (2). Thus, each column of the V matrix either has both entries nonzero (we will call this type 1), or has a zero in the second row (type 2) or in the first row (type 3), but not both. The idea of the proof is to attempt to use Theorem 1 to show that, for each of the three types of columns, all maximal columns can be "shrunk" (have their R_i values decreased), by performing small rotations with other columns of the same type. So we have to determine when this will not work. If this procedure were always possible, we would know already that the maximal square in any odd-dimensional cube must be snug-fitting, if possibility (2) is correct. Actually, columns with no zeros can be made smaller by rotating them with any type of column. The only difficulty occurs if there are maximal columns with a zero, with no corresponding non-maximal columns to rotate them with. We may assume that all columns with both elements nonzero have been shrunk and are non-maximal. Thus, the matrix looks like this:

$$\begin{array}{ccc}
 & \text{all equal} & \\
 & \downarrow & \\
\left[\begin{array}{c} \text{Nonzero} \\ \text{part} \end{array}\right. & \begin{array}{|c|} \text{nonzero} \\ 0\,0\,.\,.\,.\,0 \end{array} & \left.\begin{array}{c} 0\,0\,.\,.\,.\,0 \\ \text{nonzero} \end{array}\right] \\
\text{(all non-} & & \uparrow \\
\text{maximal)} & & \text{all equal}
\end{array}$$

Because the two rows of the matrix are orthogonal, the shorter rows of the nonzero part are also orthogonal. Then we can reduce the nonzero elements of type 3 by some constant factor, and multiply the second row type 1 elements by a common factor to increase them slightly, so that $\vec{v}_2$ remains the same length. We can do a similar thing for the type 2 columns. This reduces R. This will work unless there are no type 1 columns; that is, every column has a zero. So we will show that such a V matrix cannot be maximal.

If the maximum column-sum is normalized to 1, then for n odd, since all elements are ≤ 1, then $[f(2,n)]^2$ is at most $(n-1)/2$. But then we have case (2), a contradiction. This completes the proof that any optimal matrix is snug-fitting.

We will also need to show that neither $\vec{v}_1+\vec{v}_2$ nor $\vec{v}_1-\vec{v}_2$ can have any zero components. Either one of these would represent a line segment in a unit (n-1)-cube, which is the diagonal of the inscribed square. To obtain the side of the square, we divide $\sqrt{(n-1)}$ by $\sqrt{2}$. Thus, f(2,n) cannot be greater than $\sqrt{[(n-1)/2]}$, which is possibility (2).

Therefore, the V matrix can be written as

$$\left[\begin{array}{cc|c|c|c} a_1 & a_2 \ldots & b_1 \ldots & 1 \ldots 1 & 0 \ldots . 0 \\ 1-a_1 & 1-a_2 \ldots & -(1-b_1) \ldots & 0 \ldots 0 & 1 \ldots . 1 \end{array}\right]$$

where none of the a_i or b_i are equal to 1/2.

Next, we assume, with no loss of generality, that one of the a_i is greater than 1/2. Then, by Theorem 2, each of the b_i must be greater than 1/2; otherwise we could reduce at least two column-sums. Applying this theorem again, we find that all of the a_i must also be greater than 1/2. So we can now write the matrix as

$$\left[\begin{array}{cc|c|c|c} 1/2 + c_1 & 1/2 + c_2 \ldots & 1/2 + d_1 \ldots & 1 \ldots 1 & 0 \ldots 0 \\ 1/2 - c_1 & 1/2 - c_2 \ldots & -(1/2-d_1) \ldots & 0 \ldots 0 & 1 \ldots 1 \end{array}\right]$$

All c_i and d_i are strictly between 0 and 1/2. In the above matrix we now have 4 possible types of columns, and the numbers of columns of each type will be called i+, i-, i1, and i0, from left to right.

Now we can use the method of Lagrange multipliers to find possible maximal V matrices. There are two constraints: the two rows of the matrix, $\vec{v}_1$ and $\vec{v}_2$, are orthogonal and have equal lengths. We want to maximize, say, $|\vec{v}_1|^2$. The constraints are

$$\sum_{i=1}^{i+}(1/2+c_i)(1/2-c_i) - \sum_{i=1}^{i-}(1/2+d_i)(1/2-d_i) = 0$$

$$\sum_{i=1}^{i+}(1/2+c_i)^2 + \sum_{i=1}^{i-}(1/2+d_i)^2 + i1 = \sum_{i=1}^{i+}(1/2-c_i)^2 + \sum_{i=1}^{i-}(1/2-d_i)^2 + i0$$

where we wish to maximize

$$\sum_{i=1}^{i+}(1/2+c_i)^2 + \sum_{i=1}^{i-}(1/2+d_i)^2 + i1.$$

These equations simplify to:

$$(1/4)[(i+) - (i-)] - \sum_{i=1}^{i+} c_i^2 + \sum_{i=1}^{i-} d_i^2 = 0$$

$$2\sum_{i=1}^{i+} c_i + 2\sum_{i=1}^{i-} d_i + i1 - i0 = 0$$

where we want to maximize

$$\sum_{i=1}^{i+} (c_i^2+c_i) + \sum_{i=1}^{i-}(d_i^2 + d_i) + (1/4)[(i+) + (i-)] + i1.$$

However, there is a theorem which will tell us that none of the "points" we find by using Lagrange multipliers will be maxima. This is the second derivative test for constrained relative maxima. The proof of this theorem is given as a homework exercise in [3]. This theorem is a generalization of the second derivative test used when maximizing an unconstrained function of several variables, to determine whether a particular critical point is a maximum, a minimum, or a saddle point. The unconstrained theorem tells us to look at the matrix of second partial derivatives of the function: if this matrix is negative definite, then the point is a local maximum, and if the point is a local maximum then the matrix is negative semidefinite.

In the constrained version of the theorem, the test is the same, except that in place of the function to be maximized, we use this function plus the sum of the Lagrange multipliers multiplied by the corresponding constraints.

In our particular optimization problem, the variables are the c_i and d_i, and there must be at least one c_i and at least one d_i. This is because, as we have already shown, we never obtain a maximal square in an odd-dimensional cube when every column of the V matrix contains a zero, and, since the rows are orthogonal, the existence of a c_i requires at least one

d_i, and vice versa.

The matrix of 2nd partial derivatives in our constrained problem is very simple: It is a diagonal matrix, and the diagonal elements consist of $2(1-\lambda_1)$ for every c_i and $2(1+\lambda_1)$ for every d_i, where λ_1 is the Lagrange multiplier corresponding to the first (orthogonality) constraint. Since $2(1+\lambda_1) + 2(1-\lambda_1)$ is positive, it follows that $2(1+\lambda_1)$ or $2(1-\lambda_1)$ or both must be positive, so we never have a maximum.

Therefore, the only possible maxima occur in cases for which Lagrange multipliers cannot be used. There are 2 constraints. If there are 3 or more variables, then Lagrange multipliers can be used provided that the matrix of first partial derivatives of the constraints with respect to the c_i and d_i has rank 2. This matrix is:

$$\left[\begin{array}{ccc|ccc} -2c_1 & -2c_2 & \ldots & 2d_1 & 2d_2 & \ldots \\ 2 & 2 & \ldots & 2 & 2 & \ldots \end{array}\right]$$

Since the c_i and d_i are all positive, there are always 2 linearly independent rows.

The only other possibility is that Lagrange multipliers cannot be used because there are at least as many constraints as variables. That is, $(i+) + (i-) \leq 2$. We have already seen that there must be at least one c_i

and one d_i . Thus optimality occurs when $i_+ = i_- = 1$.

For $n = 3$ (the square in the cube), since $|1/2+c_1|>|1/2-c_1|$ and $|1/2+d_1|>|1/2-d_1|$, we must have $i_1 = 0$ and $i_0 = 1$. We obtain the V matrix

$$\begin{bmatrix} 3/4 & 3/4 & 0 \\ 1/4 & -1/4 & 1 \end{bmatrix}$$

so we now know that $f(2,3) = \sqrt{(9/8)}$, and the optimal square is unique up to trivial transformations. Now we know that possibility 1 is indeed correct, so we need not consider possibility 2.

For $n \geq 5$ and odd, the V matrix has the form

$$\left[\begin{array}{cc|ccc|ccc} 1/2+c_1 & 1/2+d_1 & 1 & \ldots & 1 & 0 & \ldots & 0 \\ 1/2-c_1 & -(1/2-d_1) & 0 & \ldots & 0 & 1 & \ldots & 1 \end{array}\right]$$

Since $|\vec{v}_1| = |\vec{v}_2|$, the first row must have more zeros than ones. Thus $i_0 \geq (n-1)/2$ and $i_1 \leq (n-3)/2$.

Now, $f(2,n) \leq f(2, n+1) = \sqrt{[(n+1)/2]}$. From the ones in the second row we have $i_0 < [f(2,n)]^2$. Combining these, i_0 must be less than $(n+1)/2$.

The only possibility is that $i_0 = (n-1)/2$, so that $i_1 = (n-3)/2$. So the nonzero elements in the V matrix are the same as for $m = 2$, $n = 3$.

We have thus shown that for all odd n,

$f(2,n) = \sqrt{[(n-1)/2 + 1/8]} = \sqrt{[(4n-3)/8]}$

and the optimal square is both snug-fitting and unique up to trivial transformations.

Thus, for example, when n = 5, the optimal V matrix may be written as

$$\begin{bmatrix} 3/4 & 3/4 & 0 & 0 & 1 \\ 1/4 & -1/4 & 1 & 1 & 0 \end{bmatrix}$$

Inequalities involving f(m,n)

It is geometrically evident that $f(m,n) \leq f(m,n+1)$. This inequality can also be derived by adding a column of zeros to a V matrix. Similarly, it is obvious that $f(m,n) \leq f(m-1,n)$, which may be derived by deleting a row from a V matrix.

Because we can put an m-cube in an n-cube and then a k-cube in the m-cube, we see that $f(k,m) \times f(m,n) \leq f(k,n)$.

However, we have also shown that $f(m,m+1) > 1$. Using this fact, we can strengthen our other inequalities to $f(m,n) < f(m,n+1)$ and $f(m,n) < f(m-1,n)$.

If we start with a V matrix with the maximum column-sum equal to one, then we can adjoin an m x m identity matrix to one side, which shows that $[f(m,n+m)]^2 \geq [f(m,n)]^2 + 1$.

From a diagram such as Figure 2, we can show that $f(km,kn) \geq f(m,n)$, where k is a positive integer.

(Please insert Figure 2 here.) Figure 2

More generally, we can show that

$f(m_1+m_2+...+m_k, n_1+...+n_k) \geq \min[f(m_1,n_1), ..., f(m_k,n_k)]$

The inequality $f(k,m)\ f(m,n) \leq f(k,n)$ can be rewritten as

$f(m,n) \leq f(k,n) / f(k,m)$.

Setting k = 1, we obtain

$f(m,n) \leq \sqrt{(n/m)}$,

as we found earlier (with equality if and only if m divides n).

Setting k = 2, we find that

$f(m,n) \leq f(2,n) / f(2,m)$.

We now know that if j is even, then $f(2,j) = \sqrt{(j/2)}$, whereas if j is odd, then $f(2,j) = \sqrt{[(4j-3)/8]}$. If m is even and n is odd, we obtain

$f(m,n) \leq f(2,n)/f(2,m) = \sqrt{[(4n-3)/(4m)]}$,

which is a better bound than $\sqrt{(n/m)}$ except when m = 2. Thus, for example, if $n = m + 1$, we have $f(m, m+1) \leq \sqrt{[(4m+1)/4m]}$. So, for $m = 4$, $n = 5$, the new bound gives $f(4,5) \leq \sqrt{(17/16)}$, whereas the earlier bound gave $f(4,5) < \sqrt{(5/4)}$.

When does $f(m,n) = f(2,n)/f(2,m)$?

For m even and n odd, we shall investigate when equality occurs in

$$f(m,n) \leq \sqrt{[(4n - 3) / (4m)]} .$$

For $m = 2$, it is trivially true that $f(2,n) = f(2,n) / f(2,2)$, so we will only consider $m \geq 4$. When is a largest square in an m-cube also one of the largest squares in the n-cube, with the m-cube in the n-cube? We shall use normalization appropriate for a unit n-cube. Then,

$$\sum_{i=1}^{m/2} \vec{v}_i = \vec{w}_1, \qquad \sum_{i=m/2+1}^{m} \vec{v}_i = \vec{w}_2$$

where $\vec{w}_1$ = something like (3/4, 3/4, 0, 0, 1) and $\vec{w}_2$ = something like (1/4, -1/4, 1, 1, 0).

(That is, for $i = 1, 2$, $j = 1$ to n, $w_{ij} = 0, \pm 1, \pm 1/4, \pm 3/4$, with $\pm 3/4$ and $\pm 1/4$ appearing exactly twice, etc.) so that $|w_{1i}| + |w_{2i}| = 1$ for all $i = 1$ to n. Thus,

$$\left|\sum_{i=1}^{m/2} v_{ij}\right| + \left|\sum_{i=m/2+1}^{m} v_{ij}\right| = 1 \text{ for all } j = 1, 2, \ldots, n.$$

But $\left|\sum_{i=1}^{m/2} v_{ij}\right| \leq \sum_{i=1}^{m/2} |v_{ij}|$

and $\left|\sum_{i=m/2+1}^{m} v_{ij}\right| \leq \sum_{i=m/2+1}^{m} |v_{ij}|$

and $\sum_{i=1}^{m} |v_{ij}| \leq 1$, so

$$\sum_{i=1}^{m} |v_{ij}| = 1$$

Now, $\sum_{i=1}^{m/2} v_{ij} = w_{1j}$ and $\sum_{i=m/2+1}^{m} v_{ij} = w_{2j}$

If $w_{1j} \geq 0$ for some j, then $v_{ij} \geq 0$ for all $i \leq m/2$

If $w_{1j} \leq 0$ for some j, then $v_{ij} \leq 0$ for all $i \leq m/2$.

Similarly for w_{2j} and $i > m/2$.

However, the $\vec{v_i}$ are all all orthogonal, so that if some $v_{ij} \neq 0$ for some $i \leq m/2$, then all other $v_{ij} = 0$ for the same j and other $i \leq m/2$. Similarly for $i > m/2$.

Therefore, for every $i \leq m/2$, if one of the v_{ij} is equal to w_{1j}, then for all other j, $v_{ij} = 0$. Similarly for $i > m/2$ and w_{2j}.

But the $\vec{v_i}$ all have equal squared length. The squared lengths cannot be equal unless for each $i \leq m/2$ there is a $v_{ij} = \pm 3/4$ and for each $i > m/2$ there is a $v_{ij} = \pm 1/4$, or vice versa. There are only two $v_{ij} = \pm 3/4$ and only two $v_{ij} = \pm 1/4$, so m can only be 4. The only possibility is that for each $i \leq m/2$, the v_{ij} consist of one $\pm 3/4$ and $(n-3)/4$ ± 1's, and the rest zeros, and for each $i > m/2$ the v_{ij} consist of one $\pm 1/4$ and $(n-1)/4$

± 1's, and the rest zeros (or vice versa).

However, $(n-3)/4$ and $(n-1)/4$ cannot both be integers, a contradiction. This proves that for m even and ≥ 4, and n odd, $f(m,n)$ is strictly less than $\sqrt{[(4n-3)/(4m)]}$.

General Infinitesimal Rotations

We now consider a general method of finding the largest m-cube(s) in an n-cube, by applying an arbitrary infinitesimal rotation to the $\vec{v}_i$. This technique reduces the problem to a system of polynomial equations and inequalities. We will illustrate this method for $m = 3$, $n = 4$, but it can be applied to any m and n.

In 4-dimensional space, an infinitesimal rotation matrix has the form

$$1 + E = \begin{bmatrix} 1 & e_{12} & e_{13} & e_{14} \\ -e_{12} & 1 & e_{23} & e_{24} \\ -e_{13} & -e_{23} & 1 & e_{34} \\ -e_{14} & -e_{24} & -e_{34} & 1 \end{bmatrix} \qquad (1)$$

Let the matrix $1 + E$ in (1) be multiplied by the transpose of the V matrix:

$$V = \begin{bmatrix} A\,B\,C\,D \\ F\,G\,H\,J \\ K\,L\,M\,P \end{bmatrix}$$

$$V^T = \begin{bmatrix} A\,F\,K \\ B\,G\,L \\ C\,H\,M \\ D\,J\,P \end{bmatrix}$$

Then

$$[\,1 + E\,]\,[V^T] = \begin{bmatrix} A+Be_{12}+Ce_{13}+De_{14}, \ldots & \\ -Ae_{12}+B+Ce_{23}+De_{24}, \ldots & \\ \cdot & \\ \cdot & \end{bmatrix}$$

(a 4x3 matrix).

If we take absolute values and assume for now that the matrix V contains no zeros, we find that the infinitesimal rotation converts R_1 to R_1', R_2 to R_2', etc., where

$R_1'-R_1 = a(Be_{12}+Ce_{13}+De_{14}) + f(\ldots$

$R_2'-R_2 = b(\ \ldots\ldots \qquad) + \ldots$ (etc.)

$R_3'-R_3 = c(\ \ldots\ldots \qquad) + \ldots$

$R_4'-R_4 = d(\qquad\qquad) + \ldots$

(where a = sign A, b = sign B, etc.; that is, $a = 1$ if A is positive and $a = -1$ if A is negative, etc.). This may be written as

$$\begin{bmatrix} R_1'-R_1 \\ R_2'-R_2 \\ R_3'-R_3 \\ R_4'-R_4 \end{bmatrix} = \begin{bmatrix} T \end{bmatrix} \begin{bmatrix} e_{12} \\ e_{13} \\ e_{14} \\ e_{23} \\ e_{24} \\ e_{34} \end{bmatrix}$$

where T is the following 4 x 6 matrix:

$$\begin{bmatrix} aB+fG+kL & aC+fH+kM & aD+fJ+kP & 0 & 0 & 0 \\ -bA-gF-lK & 0 & 0 & bC+gH+lM & bD+gJ+lP & 0 \\ 0 & -cA-hF-mK & 0 & -cB-hG-mL & 0 & cD+hJ+mP \\ 0 & 0 & -dA-jF-pK & 0 & -dB-jG-pL & -dC-jH-pM \end{bmatrix}$$

Now, suppose that that the original V matrix has only one maximal column: the first. Then we can decrease R if there is some choice of the e_{ij}'s for which $R_1' - R_1 < 0$. This is possible provided that T_{11}, T_{12}, and T_{13} are not all zero. In general, if V has any number of maximal columns, we can decrease R if the matrix Tmod consisting of the corresponding rows of T has maximum rank. So, to find all possible cube orientations which might be locally optimal, we need to determine when Tmod has less than maximum rank.

The procedure is slightly more complicated if V contains zeros in any of its maximal columns. For instance, if $A = 0$, then we have to deal with absolute values such as $| Be_{12}+Ce_{13}+De_{14} |$ in the expression for $R_1'-R_1$. We can proceed as follows: if $A = 0$, then we consider the inequality

$$Be_{12} + Ce_{13} + De_{14} > 0$$

and replace a by 1 in the equation for $R_1'-R_1$.

Including this equation with the others corresponds to adding another row to Tmod:

$$[\, B\ C\ D\ 0\ 0\ 0 \,]$$

Then, if we cannot make Tmod applied to (or multiplied by) $\begin{bmatrix} e_{12} \\ e_{13} \\ e_{14} \\ e_{23} \\ e_{24} \\ e_{34} \end{bmatrix}$ equal to anything we like, it must be that Tmod has less than maximum rank.

We may also consider the effects of $| Be_{12} + Ce_{13} + De_{14} | = 0$ and $|Be_{12} + Ce_{13} + De_{14}| < 0$. But these lead to matrices of the same rank and need not be treated separately.

We have been assuming that Tmod has no more rows than it has columns. If this condition is not satisfied, then, instead of using a Tmod matrix, we need to solve a system of equations; for instance, we could be considering the case in which V is snug-fitting ($R = R_1 = R_2 = R_3 = R_4$) and the matrix elements A, G, and M are zero, in addition to the orthonormality of the rows of V. This leads to 9 polynomial equations in 9 unknowns. Not surprisingly, we have not been able to solve these equations and find out whether we can obtain a larger cube in a tesseract than the one we have already discussed. However, there are algorithms for doing this, which are described in [4].

Now consider a cube in a tesseract with $R = R_1 > R_2, R_3, R_4$, with no zeros in the first column of V. For possible optimality, Tmod must have rank zero, where

$$\text{Tmod} = [\ aB+fG+kL,\ aC+fH+kM,\ aD+fJ+kP,\ 0,\ 0,\ 0\ ].$$

This implies that the first three elements of Tmod are zero. Thus,

$$a(B,C,D) + f(G,H,J) + k(L,M,P) = (0,0,0)$$

This means that $\pm\vec{v}_1\pm\vec{v}_2\pm\vec{v}_3$ contains 3 zero components for some combination of signs. However, it is impossible for $\pm\vec{v}_1\pm\vec{v}_2\pm\vec{v}_3$ to contain even one or more zero components for any combination of signs, by a generalization of Phil DeVicci's Theorem. For if it does, then a main diagonal of the cube fits into a unit cube parallel to one of the hyperfaces of the unit tesseract. Therefore, the cube has a side less than or equal to 1 and is not optimal. (Philip DeVicci proved that for $m = 2$, $n = 3$, $\vec{v}_1+\vec{v}_2$ could not have any zero components.)

Next, let us consider $R = R_1 = R_2 > R_3, R_4$ with no zeros in the first 2 columns of V. Then

$$Tmod = \begin{bmatrix} aB+fG+kL, & aC+fH+kM, & aD+fJ+kP, & 0 & 0 & 0 \\ -bA-gF-lK, & 0 & 0 & bC+gH+lM, & bD+gJ+lP, & 0 \end{bmatrix}$$

has rank 1 or less, so some linear combination of the two rows of Tmod is zero. This is obviously impossible by GPT (generalization of Phil's theorem). In fact, the same thing happens when there are 3 maximal columns containing no zeros. An optimal V matrix with no zeros must be snug-fitting, as we proved previously.

Next, we consider the case in which only the first column is maximal, with $A = 0$ and no other zeros in this column. Thus we have

$$\text{Tmod} = \begin{bmatrix} B+fG+kL & C+fH+kM & D+fJ+kP & 0 & 0 & 0 \\ B & C & D & 0 & 0 & 0 \end{bmatrix}$$

or, more simply,

$$\begin{bmatrix} fG+kL & fH+kM & fJ+kP & 0 & 0 & 0 \\ B & C & D & 0 & 0 & 0 \end{bmatrix}$$

must have rank 1 or less. This implies that (B,C,D) and f(G,H,J)+k(L,M,P) are parallel.

Recall that in this case

$$V = \begin{bmatrix} 0 & B & C & D \\ F & G & H & J \\ K & L & M & P \end{bmatrix}$$

Since $\vec{v}_1 \perp \vec{v}_2$, (B,C,D) $\perp$ (G,H,J) and (B,C,D) $\perp$ (L,M,P). Also, f and k are each 1 or -1. So (B,C,D) $\perp$ f(G,H,J) + k(L,M,P). But the only way that two vectors can be both parallel and perpendicular is if one or both are zero. If (B,C,D) were zero, then $|\vec{v}_1|$ would be zero, which is impossible. Thus, (G,H,J) ± (L,M,P) must be zero, which implies that $\vec{v}_2 \pm \vec{v}_3$ must contain three zeros. But $\vec{v}_2 \pm \vec{v}_3$ cannot contain two or more zeros by GPT. For if s is the side of the cube in the unit tesseract, then $\vec{v}_2 \pm \vec{v}_3$ has length $s\sqrt{2}$. This must fit parallel to one of the coordinate planes inside the tesseract, and cannot be longer than the diagonal of a square of side 1, so $s\sqrt{2} \leq \sqrt{2}$.

Thus s ≤ 1 and the cube cannot be optimal.

Similar proofs can be used to show that A = 0 is also impossible if the first 2 or the first 3 columns are the only maximal ones and do not contain any additional zeros. Again we find that only the snug-fitting case can be optimal.

Similar proofs also apply to 2 zeros either in the same row or the same column of V. That is, if A and B are the only zero elements in the first 2 columns, and only these columns are maximal, then V cannot be optimal, nor can it be optimal if there is a third maximal column containing no zeros; also, if A and F are the only zero elements in the first column, and only this column is maximal, then V cannot be optimal, nor can it be optimal if there are one or two other maximal columns containing no zeros.

However, if the 2 zeros are in neither the same row nor the same column, then the problem is more difficult. If A and G are zero and are the only zero elements in the first two columns, and the first two columns are the only maximal columns, then we can prove that V is not optimal, and because of the difficulty of the proof, we shall describe it here in detail. Tmod is shown below:

$$\begin{bmatrix} kL & fH+kM & fJ+kP & 0 & 0 & 0 \\ B & C & D & 0 & 0 & 0 \\ -lK & 0 & 0 & bC+lM & bD+lP & 0 \\ -F & 0 & 0 & H & J & 0 \end{bmatrix}$$

Assume that some linear combination of the rows of Tmod is zero, with coefficients a_1, a_2, a_3, a_4, respectively.

If only a_1 is nonzero, then $f(H,J)+k(M,P)=0$, so V cannot be optimal by GPT. Similarly if only a_3 is nonzero.

If only a_2 is nonzero, then B, C, and D are all zero, which is impossible. Similarly if only a_4 is nonzero.

If only a_1 and a_2 are nonzero, then

$a_1[f(G,H,J)+k(L,M,P)]+a_2(B,C,D) = 0.$

Thus $f(G,H,J)+k(L,M,P) = -(a_2/a_1)(B,C,D)$. (2)

But $f(G,H,J)+k(L,M,P) \perp (B,C,D)$.

Taking the dot product of both sides of (2) with (B,C,D) yields $B = C = D = 0$, which is impossible.

If only a_1 and a_3 are nonzero, then $f(H,J)+k(M,P) = 0$ (non-optimal by GPT).

If only a_1 and a_4 are nonzero, then H and J are 0 (non-optimal).

Only a_2 and a_3 nonzero -- C and D are 0 (non-optimal).

Only a_2 and a_4 nonzero -- C and D are 0 (non-optimal).

Only a_3 and a_4 nonzero -- this is similar to a_1 and a_2 nonzero.

If exactly 3 of the a_i are nonzero, similar situations occur. Thus we now assume that all four a_i are nonzero.

We have $a_1[f(H,J)+k(M,P)] + a_2(C,D) = 0$

and $a_3[b(C,D)+l(M,P] + a_4(H,J) = 0$.

Let $\vec{u} = (C,D)$, $\vec{v} = (H,J)$, and $\vec{w} = (M,P)$.

Then $a_1(f\vec{v}+k\vec{w}) + a_2\vec{u} = 0$, and $a_3(b\vec{u}+l\vec{w}) + a_4\vec{v} = 0$.

Neither $\vec{u}$ nor $\vec{v}$ is zero (either would make V non-optimal). Also, if $\vec{w} = 0$ then $\vec{u}$ and $\vec{v}$ are parallel. But $\vec{u}$ and $\vec{v}$ are also perpendicular (because A = G = 0) and nonzero, which is a contradiction. So $\vec{w}$ is also nonzero. Thus,

$$\vec{u} = C_1(f\vec{v}+k\vec{w})$$

$$\vec{v} = C_2(b\vec{u}+l\vec{w}),$$

with C_1 and C_2 nonzero. Substituting each of the last 2 equations into the other yields

$$(1 - C_1C_2fb)\vec{u} = C_1(flC_2+k)\vec{w} \tag{3}$$

$$(1 - C_1C_2fb)\vec{v} = C_2(bkC_1+l)\vec{w} \tag{4}$$

If $flC_2+k = 0$, then $C_1C_2fb = 1$. Since f, b, and l are ±1, we have $C_1C_2 = fb$, and $C_2 = -kfl$, so $C_1 = -kbl$. Thus $\vec{u} = \pm\vec{v}\pm\vec{w}$, so V is non-optimal by GPT.

Similarly, if $bkC_1+l = 0$, then $C_1C_2fb = 1$, so $C_1C_2 = fb$, and $C_1 = -lbk$, $C_2 = -lfk$, so $\vec{v} = \pm\vec{u}\pm\vec{w}$, so V is non-optimal. Thus the right-hand sides of (3) and (4) are nonzero. This shows that $\vec{u}$ and $\vec{w}$ are parallel, and so are $\vec{v}$ and $\vec{w}$. Since $\vec{u}$, $\vec{v}$, and $\vec{w}$ are all nonzero, it follows that $\vec{u}$ and $\vec{v}$ are parallel. Since $\vec{u}$ and $\vec{v}$ are also perpendicular to each other, we have a contradiction, which proves that a_1 through a_4 are all zero and thus Tmod cannot have less than rank 4 in this case, so V is non-optimal.

If A and G are zero and are the only zero elements in the first three columns, which are the only maximal columns, then it is trivial to show that a_5 must be zero. Then this case reduces to the previous case. Thus we have shown that if A = G = 0 and the first 2 columns (at least) are maximal and A and G are the only zeros in the maximal columns, then V must be snug-fitting to be optimal.

To illustrate an even more difficult case, consider the case of A = G = M = 0, with $R = R_1 = R_2 = R_3 > R_4$. Thus V looks like this:

$$V = \begin{bmatrix} 0 & B & C & D \\ F & 0 & H & J \\ K & L & 0 & P \end{bmatrix}$$

and

$$\text{Tmod} = \begin{bmatrix} kL & fH & fJ+kP & 0 & 0 & 0 \\ -lK & 0 & 0 & bC & bD+lP & 0 \\ 0 & -hF & 0 & -cB & 0 & cD+hJ \\ B & C & D & 0 & 0 & 0 \\ -F & 0 & 0 & H & J & 0 \\ 0 & -K & 0 & -L & 0 & P \end{bmatrix}$$

Then we will have 9 equations in 9 unknowns: 6 equations for orthonormality: $CH + DJ = 0$, $B^2+C^2+D^2 = 1$, etc.; also, $R_1 = R_2$, and $R_2 = R_3$ and finally det(Tmod) = 0. With no loss of generality, D, J, and P may be assumed positive (we shall show later that a maximal cube in a tesseract cannot have more than 3 zeros in its V matrix). Also, we may assume that $C > 0$, which makes $H < 0$; and assume $L > 0$, which makes $B < 0$, and assume $F > 0$, so that $K < 0$. Thus in Tmod, $k = -1$, $f = 1$, $b = -1$, etc.; there is only one set of signs to consider. But we still have a horrible set of 9 polynomial equations in 9 unknowns.

In general, we always get a system of polynomial equations with integer coefficients (including the setting of subdeterminants to zero), together with inequalities such as $C > 0$, $H < 0$, and $C - H > D + J + P$.

The snug-fitting case with exactly two zeros in the same row, and all other elements non-zero, can be shown to be non-optimal, but we omit the proof.

We have seen that an optimal V matrix with no zeros must be snug-fitting. We shall derive equations for one of these sub-cases. There are more than one sub-case because each element of V may be positive or negative. Many of these sub-cases can be eliminated as equivalent to others, because we may change all the signs in any row or column, or permute rows or columns. Other sub-cases can be eliminated by the constraint that the rows of V are orthogonal. Still other sub-cases can be shown to be non-optimal by Theorem 2 in the section on plane rotations, together with GPT. Here is an example:

$$\begin{bmatrix} A & B & C & D \\ F & G & H & J \\ K & L & M & P \end{bmatrix} = \begin{bmatrix} + & + & + & + \\ + & + & - & - \\ + & - & + & - \end{bmatrix}$$

We use the normalization here in which all column-sums are equal to 1. Thus a cube cannot be maximal unless $A^2+B^2+C^2+D^2$ (or $F^2+G^2+H^2+J^2$, etc. -- they are all equal) is greater than 1. [In fact, it must be at least $(1.007434...)^2$.]

By GPT, no matrix element may be $\pm 1/2$. Then, by Theorem 2 applied to columns 1 and 2, either $|K|$ and $|L|$ are $> 1/2$ and $|A| + |F|$ and $|B| + |G| < 1/2$, or $|K|$ and $|L|$ are $< 1/2$ and $|A| + |F|$ and $|B| + |G| > 1/2$. Next we apply the theorems to columns 3 and 4, obtaining an analogous result. We can eliminate two of the "sub-sub-cases" because they lead to one or more row vectors with a length less than 1. Then we can apply Theorem 2 to columns 1 and 3, and then to 2 and 4, and similar arguments will eliminate the entire sub-case.

Now let us consider Tmod, which, with no non-maximal columns and no zeros, is identical with the 4x6 matrix T.

For an optimal cube, T must have rank less than 4. This is equivalent to the property that all 4x4 subdeterminants are zero; but it is easier in this case to consider linear combinations of the rows of T, with coefficients a_1, a_2, a_3, and a_4, and set the linear combinations equal to zero.

From GPT, no element of T such as aB+fG+kL can be equal to zero.

In

$$V = \begin{bmatrix} A & B & C & D \\ F & G & H & J \\ K & L & M & P \end{bmatrix}$$

let us assume that J, M, and P are negative, with all other elements of V positive. We now set

$$a_1\ (B+G+L) + a_2\ (-A-F-K) = 0$$

$$a_1\ (C+H+M) + a_3\ (-A-F+K) = 0$$

$$a_1\ (D+J+P) + a_4\ (-A+F+K) = 0$$

$$a_2\ (C+H+M) + a_3\ (-B-G+L) = 0$$

$$a_2\ (D+J+P) + a_4\ (-B+G+L) = 0$$

$$a_3\ (D+J-P) + a_4\ (-C+H+M) = 0$$

However, the cube is snug-fitting, so that $A+F+K = B+G+L = C+H-M = D-J-P$. We may normalize V so that these sums and differences are all equal to 1. We find that $a_1 = a_2$, so the equations simplify. Note that if any one of a_1, a_2, a_3, and a_4 is zero, then they are all zero. So we may assume that they are all nonzero.

Also, $-A-F+K = -1+2K$, etc.

$$a_1 / a_3 = (A+F-K)/(C+H+M) = (1-2K)/(1+2M)$$

$$a_1 / a_4 = (1-2A)/(1-2D)$$

$$a_2 / a_3 \ (= a_1 / a_3) = (1-2L)/(1+2M)$$

$$a_2 / a_4 \ (= a_1 / a_4) = (1-2B)/(1-2D)$$

$$a_3 / a_4 = (1-2H)/(1+2J)$$

From these equations we see that $A = B$ and $K = L$. In fact, since $R_1 = R_2$, the first 2 columns of V are identical. Thus we have only 9 unknowns instead of 12. In addition, we get another equation: since

$$(a_1/a_3)(a_3/a_4)(a_4/a_1) = 1,$$

we obtain $(1+2M)(1+2J)(1-2A) = (1-2K)(1-2H)(1-2D)$. We end up with 9 equations in 9 unknowns.

Zeros in the V matrix for $m = 3$, $n = 4$

Here we show that there cannot be more than 3 zeros in the V matrix for a maximal cube in a tesseract, and we prove some other restrictions as well, which are necessary for optimality.

We previously found that if V contains either 0, 1, or 2 zeros in its maximal columns, and is optimal, then the cube must be snug-fitting.

If there are 3 zeros all in different rows and different columns, then the cube is either snug-fitting or there are 3 maximal columns -- the columns containing the zeros.

If there are 3 zeros in one column, the cube is not optimal. The same applies to 3 zeros in one row. If there are 3 zeros in a triangle:

X X 0 0
X X X 0
X X X X

where the X's may be zero or non-zero, the cube must be snug-fitting. This is our candidate for optimality. There cannot be any additional zeros.

There are 2 other cases to consider with 3 zeros. One is

0 X X X
0 X X X
X 0 X X

(case 1) and the other is

0 0 X X
X X 0 X
X X X X

(case 2).

In both cases above, the X's may be zero or non-zero. In case 1, V can be parametrized as follows:

$$\begin{bmatrix} 0 & \sin\beta & -\cos\beta\sin\gamma & -\cos\beta\cos\gamma \\ 0 & \cos\beta & \sin\beta\sin\gamma & \sin\beta\cos\gamma \\ \cos\alpha & 0 & -\sin\alpha\cos\gamma & \sin\alpha\sin\gamma \end{bmatrix}$$

We see that $R_2 = |\sin\beta| + |\cos\beta|$, which is ≥ 1 for all β , so that such a cube cannot be optimal.

Finally, in case 2, for $\vec{v}_1 \perp \vec{v}_2$, either v_{14} or v_{24} (or both) must be zero. If $v_{14} = 0$, we have a row of 3 zeros, so this is not optimal. If $v_{24} = 0$, we may parametrize V as follows:

$$\begin{bmatrix} 0 & 0 & \sin\alpha & \cos\alpha \\ \sin\beta & \cos\beta & 0 & 0 \\ \cos\gamma\cos\beta & -\cos\gamma\sin\beta & \sin\gamma\cos\alpha & -\sin\gamma\sin\alpha \end{bmatrix}$$

We introduce primed angles with $0 \leq \alpha',\beta',\gamma' \leq 90°$. Then

$$R_1 = \sin\beta' + \cos\gamma' \cos\beta'$$

$$R_2 = \cos\beta' + \cos\gamma' \sin\beta'$$

$$R_3 = \sin\alpha' + \sin\gamma' \cos\alpha'$$

$$R_4 = \cos\alpha' + \sin\gamma' \sin\alpha'$$

Assume first that $0 < \alpha',\beta',\gamma' < 90°$. If there is only one maximal column, we may increase or decrease γ to decrease R.

If columns 1 and 2, or 3 and 4, are the only maximal columns, we may change γ to decrease both maximal column-sums.

If columns 1 and 3, or 1 and 4, or 2 and 3, or 2 and 4 are the maximal columns, we may change α or β slightly so that one column-sum decreases and the other is unchanged.

If columns 1, 2, and 3, or 1, 2, and 4 are maximal, then we change α slightly. If columns 1, 3, and 4, or 2, 3, and 4 are maximal, we change β slightly.

If α' or β' are 0 or 90°, then $R \geq 1$.

If γ' is 0 or 90°, then one of the column-sums will be $\sin\alpha' + \cos\alpha'$ or $\sin\beta' + \cos\beta'$, so $R \geq 1$.

Thus an optimal cube with this form must be snug-fitting, with $0 < \alpha',\beta',\gamma' < 90°$. By doing some simple algebra we find that α', β', and γ' are all 45°, so that $R = 1/2 + \sqrt{2}/2 > 1$, so that even this is not optimal.

The above also shows that for $m = 3$, $n = 4$, V may not have 4 or more zeros. Thus the only possibilities are those shown below (where the X's represent non-zero matrix elements):

```
XXXX    0XXX    0XXX    0XXX    XX00
XXXX    XXXX    0XXX    X0XX    XXX0
XXXX    XXXX    XXXX    XXXX    XXXX
```

which must be snug-fitting to be optimal, and

```
0XXX
X0XX
XX0X
```

for which at least the first three columns must be maximal for possible optimality. Notice that we have eliminated almost all non-snug-fitting cases!

Lagrange Multipliers (again)

For general m, n, an alternative technique to the method of infinitesimal rotations is Lagrange multipliers. For instance, consider

$$\begin{bmatrix} A & B & C & D \\ F & G & H & J \\ K & L & M & P \end{bmatrix}$$

with columns 1, 2, and 3 maximal, and A = G = M = 0. Then with no loss of generality we may assume that D, J, P, C, L, and F are positive, and that B, H, and K are negative.

We want to maximize $A^2+B^2+C^2+D^2$, with $A^2+B^2+C^2+D^2-F^2-G^2-H^2-J^2 = 0$, $A^2+B^2+C^2+D^2-K^2-L^2-M^2-P^2 = 0$, $AF+BG+CH+DJ= 0$, etc.; also, $|A|+|F|+|K| = 1$, etc. so that $F-K = 1$, $L-B = 1$, and $C-H < 1$, and $D+J+P < 1$. However, we may eliminate A, G, and M from the problem entirely, since they are all zero. Thus we want to maximize $B^2+C^2+D^2$, etc. Let the Lagrange multipliers be λ_1, λ_2, etc. We want to apply the constrained 2nd derivative test in [3], and the diagonal elements of the matrix of 2nd partial derivatives include $2(1+\lambda_1+\lambda_2)$ (3 times, for B, C, and D), $-2\lambda_1$ (3 times), and $-2\lambda_2$ (3 times). These add up to 2, which is positive, so at least one diagonal element is positive. So none of the "points" which we get from Lagrange multipliers are maxima. The only possibility is that the matrix of first partial derivatives of the (equality) constraints with respect to B, C, etc., has less than maximum rank; that is, since there are now 8 equality constraints and 9 variables, all 8 x 8 subdeterminants can be set to 0 simultaneously. So again we get a system of polynomial equations with integer coefficients, together with a set of strict inequalities of a similar form.

The matrix of first partial derivatives is

$$\begin{bmatrix} 2B & 2C & 2D & -2F & -2H & -2J & 0 & 0 & 0 \\ 2B & 2C & 2D & 0 & 0 & 0 & -2K & -2L & -2P \\ 0 & H & J & 0 & C & D & 0 & 0 & 0 \\ L & 0 & P & 0 & 0 & 0 & 0 & B & D \\ 0 & 0 & 0 & K & 0 & P & F & 0 & J \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Instead, we may use the other normalization in which $|A|+|F|+|K| = |B|+|G|+|L|$, with $A^2+B^2+C^2+D^2 = 1$, etc., and minimize $|A|+|F|+|K|$. This also leads to the same sorts of polynomial equations and inequalities with integer coefficients.

Further Comments and Unsolved Problems

We have seen that the problem of the largest m-cube in an n-cube can be reduced to a finite number of systems of simultaneous polynomial equations and inequalities with integer coefficients, either by using Lagrange multipliers or by using our earlier method. In [4] it is shown how to solve such systems of equations, even with arbitrary coefficients, using an algorithm which eliminates the variables one by one. If there are only a finite number of complex solutions, then one could test each real solution to find which one maximizes the function to be maximized, of those which satisfy the inequalities. Then one could also prove that the maximum value is an algebraic number if the coefficients are integral, rational, or even arbitrary real algebraic numbers. However, this argument breaks down if there is a continuous infinity of complex solutions. It would be nice to be able to prove that, even in this case, the optimal value is algebraic [i.e., that $f(m,n)$ is algebraic for all m, n], and if it is, there remains the question of whether there is an algorithm which produces the exact value of $f(m,n)$ (by finding a polynomial equation which it satisfies, and describing it as, say, the 47th smallest real root of that equation).

It would not be surprising if we do sometimes get a continuous infinity of solutions, since if the V matrix has 2 or more non-maximal columns, these columns could be rotated into each other to produce a continuous infinity of solutions, all with the same value of R.

Many other geometry problems can be put into the form of optimization problems in which the function to be optimized is a polynomial function of several variables, with polynomial (equality and inequality) constraints, all with coefficients that are either integers or, in some cases, real algebraic numbers. These include the unsolved problem of finding the largest regular dodecahedron in a regular icosahedron (and vice versa); it would be interesting to be able to know immediately that the ratio of the edges or volumes must be an algebraic number. Even the kissing sphere problem in any number of dimensions can be reduced to problems of this form. (These problems are discussed in [5, 6]. See also [7, pp. 52, 53], where the vertices of the coordinates of regular dodecahedra and icosahedra are given.)

At this time, we still do not know whether our candidate for the largest cube in a tesseract is optimal. Perhaps a computer-assisted proof will finally settle this question, unless there is a more elegant way to prove it. It would also be interesting to know if the largest m-cube in an n-cube is snug-fitting for all m and n, as it is for the cases for which we have solved the problem. Perhaps future work will find additional simple patterns in f(m,n), as we found for m = 2. For instance, one of our general inequalities was $[f(m,n+m)]^2 \geq [f(m,n)]^2 + 1$. It seems entirely possible at this time that the equal sign always applies. In my opinion, either a proof or a counterexample would be interesting.

Prince Rupert's Problem

Prince Rupert is discussed in [8, 9], and Prince Rupert's problem and a generalization of it are discussed in [9,10].

Prince Rupert (1619-1682) was born in Prague and educated in the Netherlands. He was also called Rupert of the Rhine or of the Palatinate, and he was a nephew of King Charles I of Great Britain and Ireland. He was a royalist cavalry commander in the English civil war. He also dabbled in scientific experiments and became a member of the Royal Society. Not only did he have a geometry problem named after him, but he also had an alloy named after him (still called Prince's metal), and he studied the properties of quickly cooled drops of glass, known as Rupert drops, in the laboratory which he had built for himself.

Prince Rupert proposed, but apparently did not solve, the following problem: "To perforate a cube in such a way that a second cube of the same size may pass through the hole" [9, p. 73]. Pieter Nieuwland (1764-1794) not only found that this was possible, but he also found the size of the largest cube which could pass through such a hole.

Many authors seem to assume, without proof, that this problem is equivalent to finding the largest square in a cube. The reader may enjoy investigating this question (the problems do seem to give the same numerical result for $m = 2$, $n = 3$) for an arbitrary number of dimensions [that is, is the problem of finding the largest m-cube in an (m+1)-cube equivalent to finding the largest (m+1)-cube which can pass through a "hypercubical" hole in another (m+1)-cube?]. For the most trivial case, that of passing a line segment through a hole in a square, the square breaks up into 2 pieces, and for the limiting size, the area of the square disappears completely.

References

1. M. Gardner, Mathematical Carnival, Random House, Inc., New York, pp. 48, 53, 54, 1977.

2. M. Gardner, Mathematical Carnival, Knopf, New York, pp. 48, 53, 54, 272, 1989.

3. W. H. Fleming, Functions of Several Variables, Addison-Wesley, Reading, Mass., ©1965; 2nd printing, 1966, p.134, problem 11. (There appears to be a typographical error in that the signs are reversed.)

4. D. Cox, J. Little, and D. O'Shea, Ideals, Varieties, and Algorithms, Springer-Verlag, New York, 1992.

5. H. T. Croft, K. J. Falconer, and R. K. Guy, Unsolved Problems in Geometry, Springer-Verlag, New York, p.52, problem B3 (maximal regular polyhedra inscribed in regular polyhedra) and p. 121, problem D12 (kissing numbers of convex sets), 1991.

6. H. T. Croft, On maximal polyhedra inscribed in a regular polyhedron, Proc. London Math. Soc. (3) 41, pp. 279-296, 1980.

7. H. S. M. Coxeter, Regular Polytopes, Dover, New York, 1973.

8. Encyclopædia Britannica, 1964 edition (under "Rupert")

9. D. J. E. Schrek, Prince Rupert's Problem and its extension by Pieter Nieuwland, Scr. Math. 16, pp. 73-80, 1950.

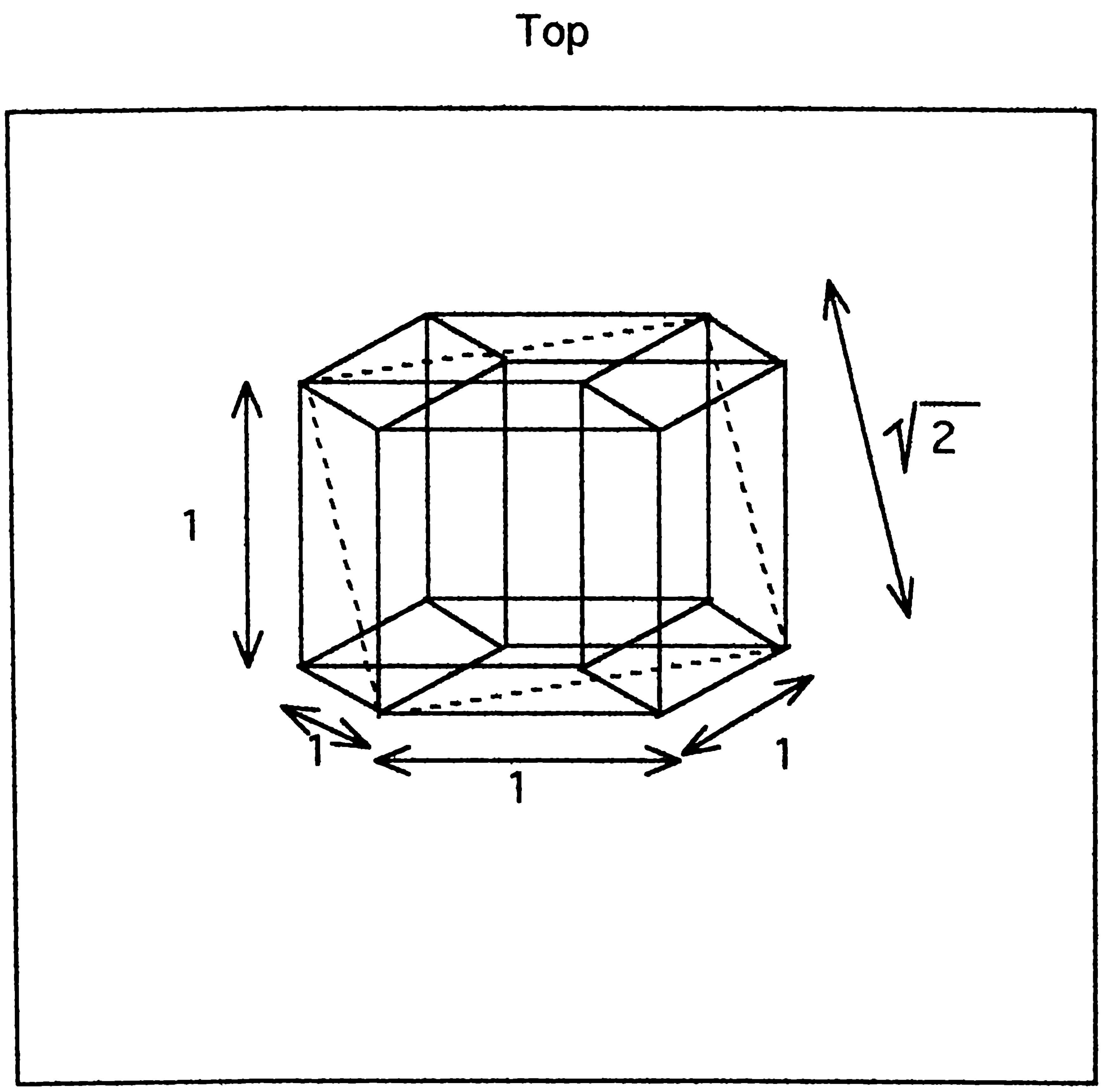
10. D. J. E. Schrek, same as Ref. 9 but pp. 261-267. (This is the conclusion of the previous article.)

Kay R. DeVicci

Partial Solution to the Largest m-Cube in an n-Cube

Figure 1

The largest square in a tesseract.



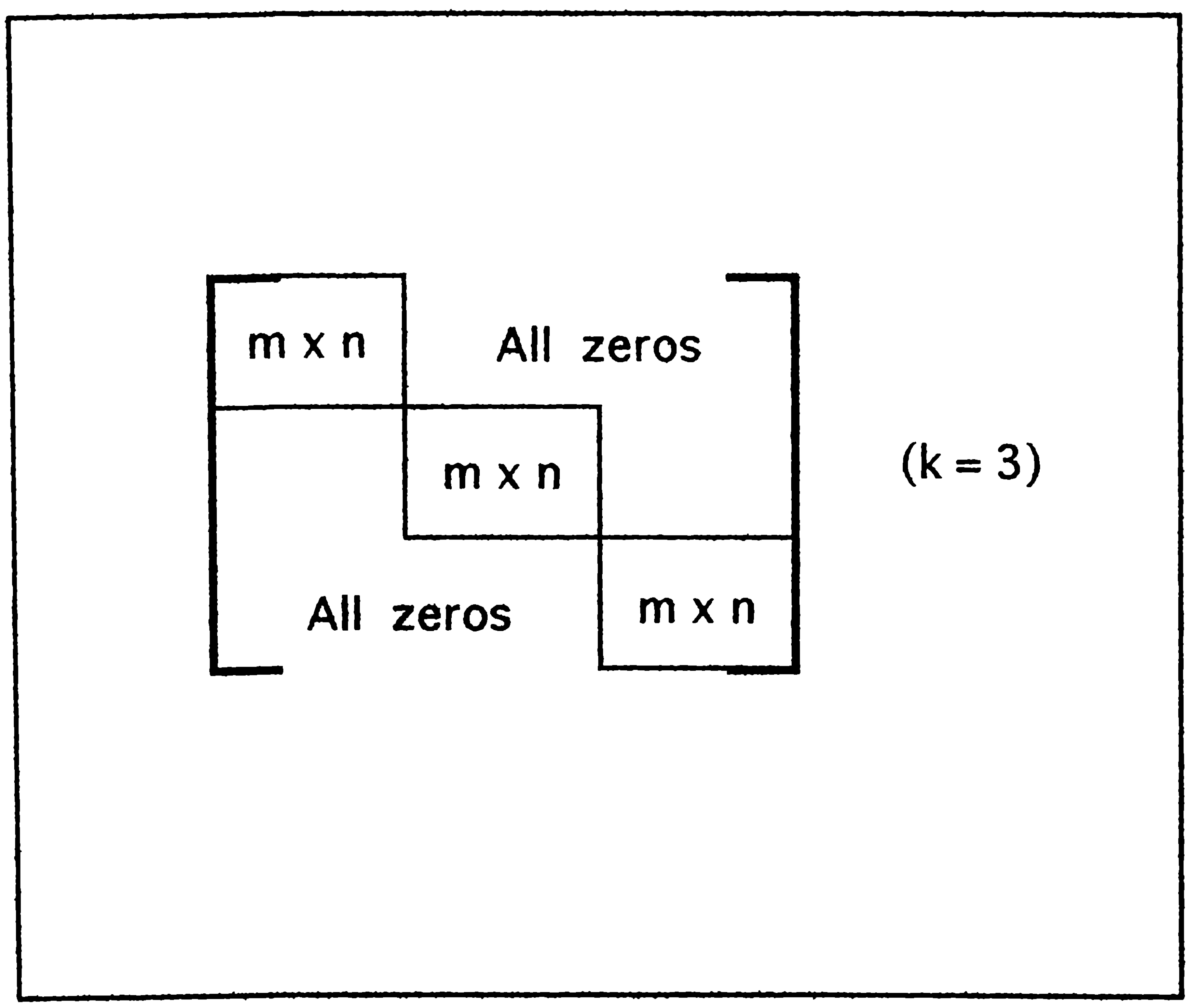
Kay R. DeVicci

Partial Solution to the Largest m-Cube in an n-Cube

Figure 2

(No figure title)

Top



Part 2 of "Partial Solution to the Largest m-Cube in an n-Cube"

Copyright © 1996 by
Kay R. DeVicci
131 Haines Drive
Moorestown, NJ 08057

Here we complete the work on the largest cube in a tesseract, proving that the previously found candidate is, in fact, the largest cube in a tesseract of unit side; this cube has a side of 1.0074 . . . , as shown previously.

We have to show that the remaining possibilities in the previous paper cannot be optimal. Fortunately, we can do this without having to solve large systems of simultaneous polynomial equations, so that a computer-assisted proof is not necessary.

Previously we found that an optimal cube in a tesseract had to be one (or more) of the following cases:

(1) No zeros in the V matrix
(2) Exactly one zero
(3) Exactly 2 zeros in the same row
(4) Exactly 2 zeros in the same column
(5) Exactly 2 zeros in neither the same row nor the same column
(6) Exactly 3 zeros in a diagonal
(7) Exactly 3 zeros in a triangle

Furthermore, in cases 1 through 5, and case 7, the cube must be snug-fitting. In case 6, the cube must either be snug-fitting or have 3 "column-sums" equal - the ones for the 3 columns containing a zero - and the remaining "column-sum" less than the others. (Recall that what we call a "column-sum" is actually a sum of absolute values.) In all of these cases except 7, we shall assume that the cube is optimal and derive a contradiction.

We shall find it convenient to normalize the V matrix so that the maximum column-sum is 1. Thus the 3 rows of V will all have squared length $\mathcal{L}^2$, where $\mathcal{L}$ is the side of the cube inscribed in a tesseract of unit side.

We will need the following theorem:

The identical column theorem: For a cube in a tesseract, if any 2 columns of V are identical, then the cube is not optimal.

Proof: Assume that V is normalized as above. Assume that, say, columns 1 and 2 are identical. That is, if V is

$$\begin{array}{cccc} A & B & C & D \\ F & G & H & J \\ K & L & M & P \end{array}$$

then $A = B$, $F = G$, and $K = L$. We have already shown that an optimal cube cannot have a column of all zeros in V. Thus the inner product of the first two columns is positive. Now, let us add another row to V, such that the resulting 4x4 matrix, which we shall call W, is the scalar $\mathcal{L}$ times an orthogonal matrix. Thus W can be written as

$$\begin{array}{cccc} A & A & C & D \\ F & F & H & J \\ K & K & M & P \\ w & x & y & z \end{array}$$

Since columns 1 and 2 of W must be orthogonal and of equal squared length, w and x must be nonzero and $w = -x$. However, column 3 of W must be orthogonal to both columns 1 and 2. This is possible only if $y = 0$. However, the squared length of each column is $\mathcal{L}^2$, which must be greater than 1 for optimality. Thus $C^2+H^2+M^2 = \mathcal{L}^2$. But $|C|+|H|+|M| \leq 1$, and $C^2 \leq |C|$, and similarly for H and M. Thus $\mathcal{L}^2 \leq 1$, so the cube is not optimal.

Case 1: All matrix elements of V are nonzero.

The V matrix is:

$$\begin{array}{cccc} A & B & C & D \\ F & G & H & J \\ K & L & M & P \end{array}$$

With no loss of generality, the first row and first column may all be assumed positive. (A, B, C, D, F, and K are all > 0.) First we consider the case in which at least 2 columns have the same pattern of signs. Thus we may assume that G and L are also positive. H and J may not both be positive, for rows 1 and 2 must be orthogonal. The same applies to M and P. Similarly, H and J cannot have the same pattern of signs as M and P. The only possibilities to consider, then, are: J, M, and P negative, all others positive; and just J and M negative, the rest positive. So we have:

$$\begin{array}{cccc} + & + & + & + \\ + & + & + & - \\ + & + & - & - \end{array} \qquad \begin{array}{cccc} + & + & + & + \\ + & + & + & - \\ + & + & - & + \end{array}$$

If in the second of these sub-cases, we multiply column 4 by (-1), and then interchange rows 1 and 2, we get the first sub-case. Thus the second is essentially equivalent to the first. This first case was considered in Part 1 of this paper (pages 43 to 45), where it was proved that, for optimality, columns 1 and 2 must be identical. The identical column theorem then shows that it cannot be optimal.

We still have the sub-case to consider in which no two columns have the same pattern of signs:

+ + + +
+ + - -
+ - + -

On pages 42 and 43 (Part 1) we showed how to rule out this case. Here we clarify the previous discussion. The idea is to show that, if this is snug-fitting (so that all "column-sums" are equal to 1), then two columns can be rotated into each other such that both column-sums decrease. Then this shows that this cube is the same size as a cube with only 2 maximal columns, and all such cubes were shown in Part 1 to be non-optimal. We use Theorem 2 (Part 1) to show that 2 column-sums can be decreased without affecting any of the others. It is convenient to put a circle or square around a + or - sign, where a circle means that the absolute value of the matrix element is < 1/2, while a square means > 1/2. If any matrix element were 1/2 or -1/2, then by GPT, $\mathcal{L} \leq 1$ and it would not be optimal. Then, using squares and circles, we can use rules such as: each row must contain at least one square, or $\mathcal{L}^2$ would be < 1. No column may contain more than 1 square, or the column-sum would be > 1. (Remember that a column-sum means a sum of absolute values, not a sum of the matrix elements themselves.)

a

In the above matrix of + and - signs, by comparing columns 1 and 2, we see that v_{31} and v_{32} have opposite signs, whereas v_{11} and v_{12}, as well as v_{21} and v_{22}, have the same sign. Thus row 3 is the "different" row, so that v_{31} and v_{33} must have either both circles or both squares. Similarly, comparing columns 3 and 4, row 3 again is different from the others, so v_{33} and v_{34} have either both circles or both squares. Two of these sub-sub-cases are then no good, because by applying the rules for circles and squares, we end up with 4 circles in at least one row. Then, by comparing columns 1 and 3, or 2 and 4, the entire sub-case is ruled out.

We have now proved that a matrix with no zeros cannot be an optimal square in a tesseract. To rule out more cases, we have to investigate when we can simultaneously reduce two column-sums when exactly one of the two columns contains exactly one zero. This is a little harder. When we consider $dRi/d\theta$ we find that if column i contains a zero, we have different one-sided derivatives for positive and negative θ. This is because when the zero gets rotated, the absolute value always increases. If we have, say,

$$\begin{matrix} 0 & D \\ B & E \\ C & F \end{matrix}$$

then if BE and CF have opposite signs, and $R_1 = R_2$, then the rule is that if the cube is optimal and B gets a circle, then E must also get a circle, while if C gets a circle then so does F. Otherwise either $\mathcal{L}$ would be ≤ 1 by GPT or it would be possible to decrease R_1 and R_2 simultaneously. If instead, $R_1 > R_2$ above, then, in order that R_1 cannot be decreased by rotating these 2 columns into each other, and still assuming that BE and CF have opposite signs, and again applying GPT, we must have $|D| + |E| > |F|$ and $|D| + |F| > |E|$. These rules will be proved in detail later.

Now we can rule out the 2 cases with 3 zeros in a diagonal. (Both snug and non-snug.) The V matrix is:

$$\begin{matrix} 0 & B & C & D \\ F & 0 & H & J \\ K & L & 0 & P \end{matrix}$$

We may assume D, J, and P to be all positive. We have $|F| + |K| = 1$, while from orthogonality $|F|\,|K| = JP$. We now find an expression for F^2+K^2. We have $F^2+K^2 = (|F|+|K|)^2 - 2|F|\,|K| = 1-2JP$. Similarly, $B^2+L^2 = 1-2DP$, and $C^2+H^2 = 1-2DJ$. Since each row of V has a squared length of $\mathcal{L}^2$, the sum of the squares of all of the matrix elements is $3\mathcal{L}^2$. Thus, $3\mathcal{L}^2 = F^2+K^2+B^2+L^2+C^2+H^2+D^2+J^2+P^2 = 3-2JP-2DJ-2DP+D^2+J^2+P^2$
$= 3 + D(D-J-P) + J(J-P-D)+P(P-D-J)$. We will soon prove that D, J, and P satisfy the triangle inequality, so that D-J-P, J-P-D, and P-D-J are all negative, so it will follow that $3\mathcal{L}^2 < 3$ so $\mathcal{L}^2 < 1$ (not optimal).

To prove the triangle inequality, first we consider the snug-fitting case. Suppose that $|F|<1/2$, so that F gets a circle around it. Then J also gets a circle, by our new theorem applied to columns 1 and 4 (and using the fact that FJ and KP have opposite signs). But row 2 cannot have all circles and zeros, so H must get a square. Thus C must be a circle. Applying the new theorem now to columns 3 and 4, since C and H have opposite signs, D gets a circle. Since row 1 cannot have all circles and zeros, B must be a square, so L must be a circle, and applying the new theorem again, P must be a circle. Since D, J, and P are all $< 1/2$ and their sum is 1, the triangle inequality is thus proved. (We also have to consider $|F| > 1/2$, but the proof is similar.)

Proving the triangle inequality for the non-snug case, where $R_1=R_2=R_3>R_4$, is slightly different since it uses the second part of our theorem. The idea is to assume that the cube is optimal, so that we cannot decrease R_1, R_2, or R_3 by rotating the corresponding column with the 4th column. We again obtain the same inequalities: $D+J>P$, etc.

This shows that for the snug case, if we assume that 2 column-sums cannot be simultaneously decreased without affecting the others, then it is not optimal. Thus for optimality, it is possible to find a cube with only 2 maximal columns, which has the same size. But in Part 1 such cubes were shown to be non-optimal. For the non-snug case, if we assume that no maximal column-sum can be decreased slightly, then it is not optimal, so again there is a cube of the same size with only 2 maximal columns.

Another easy case is 2 zeros in a diagonal, and snug, with all other elements of V nonzero. The V matrix is:

$$\begin{matrix} 0 & B & C & D \\ F & 0 & H & J \\ K & L & M & P \end{matrix}$$

We may assume that D, J, K, L, M, and P are positive:

No zeros, + snug-fitting

Subcase

$$\begin{array}{cccc} + & + & + & + \\ + & + & - & - \\ + & - & + & - \end{array}.$$

Compare col.'s 1 + 2.

$$\begin{array}{cccc} + & + & + & + \\ + & + & - & - \\ \boxplus & \boxminus & + & - \end{array}$$

No 2 squares in same column:

$$\begin{array}{cccc} \oplus & \oplus & + & + \\ \oplus & \oplus & - & - \\ \boxplus & \boxminus & + & - \end{array}$$

Compare col's 3 + 4:

$$\begin{array}{cccc} \oplus & \oplus & + & + \\ \oplus & \oplus & - & - \\ \boxplus & \boxminus & \boxplus & \boxminus \end{array}$$

No 2 squares in same column

$$\begin{array}{cccc} \oplus & \oplus & \oplus & \oplus \\ \oplus & \oplus & \ominus & \ominus \\ \boxplus & \boxminus & \boxplus & \boxminus \end{array}$$

Rows 1 + 2 are all circles— no good!

$$\begin{array}{cccc} \oplus & \oplus & + & + \\ \oplus & \oplus & - & - \\ \boxplus & \boxminus & \oplus & \ominus \end{array}.$$

Compare col's 1 + 3, and 2 + 4.

$$\begin{array}{cccc} \oplus & \oplus & + & + \\ \oplus & \oplus & \ominus & \ominus \\ \boxplus & \boxminus & \oplus & \ominus \end{array}$$

Row 2 is all circles— no good!

$$\begin{array}{cccc} + & + & + & + \\ + & + & - & - \\ \oplus & \ominus & + & - \end{array}$$

Compare col's 3 + 4:

$$\begin{array}{cccc} + & + & + & + \\ + & + & - & - \\ \oplus & \ominus & \oplus & \ominus \end{array}$$

4 circles in a row— no good!

$$\begin{array}{cccc} + & + & + & + \\ + & + & - & - \\ \oplus & \ominus & \boxplus & \boxminus \end{array}$$

No 2 squares in same column

$$\begin{array}{cccc} + & + & \oplus & \oplus \\ + & + & \ominus & \ominus \\ \oplus & \ominus & \boxplus & \boxminus \end{array}$$

Compare col's 1 + 3:

$$\begin{array}{cccc} + & + & \oplus & \oplus \\ \oplus & + & \ominus & \ominus \\ \oplus & \ominus & \boxplus & \boxminus \end{array}$$

Compare col's 2 + 4:

$$\begin{array}{cccc} + & + & \oplus & \oplus \\ \oplus & \oplus & \ominus & \ominus \\ \oplus & \ominus & \boxplus & \boxminus \end{array}$$

Row 2 is all circles— no good!

3 zeros in a diagonal

$$V = \begin{matrix} O & B & C & D \\ F & O & H & J \\ K & L & O & P \end{matrix}$$

Either snug ($R_1 = R_2 = R_3 = R_4$) or $R = R_1 = R_2 = R_3 > R_4$.

We may assume D, J, and $P > 0$. Then $|F| + |K| = 1$, and from orthogonality of the rows, $|F|\,|K| = JP$ (with either $F > 0$, $K < 0$ or vice versa).

$F^2 + K^2 = (|F| + |K|)^2 - 2|F|\,|K| = 1 - 2|F||K| = 1 - 2JP$.

Similarly, $B^2 + L^2 = 1 - 2DP$ and $C^2 + H^2 = 1 - 2DJ$.

Since each row of V has a squared length of $\mathcal{L}^2$ (do not confuse $\mathcal{L}$ with L), the sum of the squares of all matrix elements is $3\mathcal{L}^2$.

Thus $3\mathcal{L}^2 = F^2 + K^2 + B^2 + L^2 + C^2 + H^2 + D^2 + J^2 + P^2$.

So $3\mathcal{L}^2 = 3 - 2JP - 2DP - 2DJ + D^2 + J^2 + P^2$.

$3\mathcal{L}^2 = 3 + D(D - J - P) + J(J - P - D) + P(P - D - J)$.

We will soon prove that $D - J - P$, $J - P - D$, and $P - D - J$ are all negative, so that $3\mathcal{L}^2 < 3$ and $\mathcal{L} < 1$ (not optimal.)

Thus we want to prove the triangle inequality: $J + P > D$, $P + D > J$, $D + J > P$ both in the snug and non-snug cases. With no loss of generality, we can assume F, L, and C are positive, so by orthogonality, K, B, and H are negative.

Suppose $|F| < \frac{1}{2}$, so $|K| > \frac{1}{2}$.

$$\begin{matrix} 0 & - & + & + \\ + & 0 & - & + \\ \boxed{-} & + & 0 & + \end{matrix} \quad = \quad \begin{matrix} 0 & B & C & D \\ F & 0 & H & J \\ K & L & 0 & P \end{matrix}$$

Since, for optimality, columns 1 and 4 cannot be simultaneously decreased (this is for the snug case), $J < \frac{1}{2}$, and gets a circle.

$$\begin{matrix} 0 & - & + & + \\ + & 0 & - & + \\ \boxed{-} & + & 0 & + \end{matrix}$$

But row 2 cannot have all circles and zeros, so $|H| > \frac{1}{2}$, and gets a square.

$$\begin{matrix} 0 & - & + & + \\ + & 0 & \boxed{-} & + \\ \boxed{-} & + & 0 & + \end{matrix}$$

But $|C| + |H| = 1$, so C gets a circle. Since columns 3 and 4 cannot be simultaneously decreased, D gets a circle. Continuing in this manner, D, J, and P all have circles. D, J, and P are each $< \frac{1}{2}$, and their sum is 1, so the triangle inequality is proved.

(If we had assumed $|F| > \frac{1}{2}$ and $|K| < \frac{1}{2}$, the proof would have been similar.)

Non-snug case: Assume that the cube is optimal so that we cannot decrease R_1, R_2, or R_3 by rotating the corresponding column with column 4. This also leads to $D + J > P$, etc.

This shows that for the snug case, if we assume that 2 column-sums cannot be simultaneously decreased without affecting the others, then it is not optimal. Thus for optimality, it is possible to find a cube with only 2 maximal columns, which has the same size. But in part I, such cubes were shown to be non-optimal. For the non-snug case, if we assume that no maximal column-sum can be decreased slightly, then it is not optimal, so again there is a cube of the same size with only 2 maximal columns.

2 zeros in a diagonal
(and snug, with all other matrix elements nonzero).

$$\begin{matrix} 0 & \cdot & \cdot & + \\ \cdot & 0 & \cdot & + \\ + & + & + & + \end{matrix}$$

$$\begin{matrix} 0 & \cdot & + & + \\ \cdot & 0 & - & + \\ + & + & + & + \end{matrix} \qquad \begin{matrix} 0 & \cdot & - & + \\ \cdot & 0 & + & + \\ + & + & + & + \end{matrix}$$

(since rows 1 & 2 are $\perp$)

$\downarrow$ Switch rows 1 + 2.

$$\begin{matrix} \cdot & 0 & + & + \\ 0 & \cdot & - & + \\ + & + & + & + \end{matrix}$$

$\downarrow$ Switch col.'s 1 + 2.

$$\begin{matrix} 0 & \cdot & + & + \\ \cdot & 0 & - & + \\ + & + & + & + \end{matrix}$$

Equivalent

$\downarrow$ Rows 1 + 3 are $\perp$:

$$\begin{matrix} 0 & - & + & + \\ \cdot & 0 & - & + \\ + & + & + & + \end{matrix}$$

$$\begin{matrix} 0 & - & + & + \\ + & 0 & - & + \\ + & + & + & + \end{matrix} \qquad \begin{matrix} 0 & - & + & + \\ - & 0 & - & + \\ + & + & + & + \end{matrix}$$

$\downarrow$ Multiply row 2 by (-1)

$$\begin{matrix} 0 & - & + & + \\ + & 0 & + & - \\ + & + & + & + \end{matrix}$$

$\downarrow$ switch col.'s 3 + 4.

$$\begin{matrix} 0 & - & + & + \\ + & 0 & - & + \\ + & + & + & + \end{matrix}$$

Equivalent.

So there's just $\begin{array}{cccc} 0 & - & + & + \\ + & 0 & - & + \\ + & + & + & + \end{array}$

Now compare col.'s 3 + 4:

$\begin{array}{cccc} 0 & - & + & + \\ + & 0 & - & + \\ + & + & + & + \end{array}$ $\qquad$ $\begin{array}{cccc} 0 & - & + & + \\ + & 0 & \boxed{-} & \boxed{+} \\ + & + & + & + \end{array}$

Now apply our rules: Every row must contain at least one square; no 2 squares in a column:

$\begin{array}{cccc} 0 & - & + & + \\ \boxed{+} & 0 & - & + \\ + & + & + & + \end{array}$ $\qquad$ $\begin{array}{cccc} 0 & \boxed{-} & + & + \\ + & 0 & \boxed{-} & \boxed{+} \\ \boxed{+} & + & + & + \end{array}$

Compare col.'s 1 + 3: V_{33} must be a circle.

Now compare columns 1 + 3. Since V_{21} is a circle, and V_{11} is 0, and $V_{21}V_{23}$ and $V_{31}V_{33}$ have opposite signs, then V_{23} must be a circle. (Contradiction!)

$\begin{array}{cccc} 0 & - & + & + \\ \boxed{+} & 0 & - & + \\ + & + & + & + \end{array}$

$\begin{array}{cccc} 0 & - & + & + \\ \boxed{+} & 0 & - & + \\ + & \boxed{+} & + & + \end{array}$ $\qquad$ $\begin{array}{cccc} 0 & \boxed{-} & + & + \\ \boxed{+} & 0 & - & + \\ + & + & + & + \end{array}$ (2 possibilities for col. 2)

Compare col.'s 2 + 4 – V_{34} must be a circle – no good! (row 3 is all circles)

Compare columns 2 + 3: V_{13} is a circle.
Compare columns 2 + 4: V_{14} is a circle.
Row 1 is all zeros and circles – no good.

So we've eliminated 2 zeros in a diagonal.

2 zeros in a row.

$$\begin{matrix} A & B & C & D \\ F & G & H & J \\ 0 & 0 & M & P \end{matrix} \quad \text{and snug. } (K=L=0.)$$

$$\begin{matrix} + & + & + & + \\ \cdot & \cdot & \cdot & + \\ 0 & 0 & \cdot & + \end{matrix} \longrightarrow \begin{matrix} + & + & + & + \\ \cdot & \cdot & + & + \\ 0 & 0 & - & + \end{matrix}$$

because rows 1 & 3 are $\perp$, and 2 & 3 are $\perp$.

$$\begin{matrix} + & + & + & + \\ + & - & + & + \\ 0 & 0 & - & + \end{matrix} \xrightarrow{\text{Col's } 3+4} \begin{matrix} + & + & + & + \\ + & - & + & + \\ 0 & 0 & \boxed{-} & \boxed{+} \end{matrix} \longrightarrow \begin{matrix} + & + & \oplus & \oplus \\ + & - & \oplus & \oplus \\ 0 & 0 & \boxed{-} & \boxed{+} \end{matrix}$$

Compare col's 1+2. Either F and G are both circles or A and B are. Either way, 1 row is all circles. (No good)

$$\begin{matrix} + & + & + & + \\ - & - & + & + \\ 0 & 0 & - & + \end{matrix}$$

$$\begin{matrix} A & B & x\cos\theta & x\sin\theta \\ F & G & y\cos\theta & y\sin\theta \\ 0 & 0 & -z\sin\theta & z\cos\theta \end{matrix}$$

(x, y, z all positive, $0° < \theta < 90°$, A and B pos, F and G neg.)

$A|F| + B|G| = xy$ (since rows 1+2 are $\perp$)

$F = -(1-A)$, $G = -(1-B)$

$A(1-A) + B(1-B) = xy$

$A^2 + B^2 + x^2 = (1-A)^2 + (1-B)^2 + y^2 = z^2 \ (= \mathcal{L}^2)$

$(x+y)\cos\theta + z\sin\theta = (x+y)\sin\theta + z\cos\theta \ (= 1)$

$(x+y-z)\cos\theta = (x+y-z)\sin\theta$

$\theta = 45°$ or $x+y=z$ or both.

$z(\sin\theta + \cos\theta) = 1$

$\mathcal{L} = z = \dfrac{1}{\sin\theta + \cos\theta}$, but $\sin\theta + \cos\theta > 1$, so $\mathcal{L} < 1$.

(Not optimal). So $\theta = 45°$. $\sin\theta = \cos\theta$ so

$C = D$, $H = J$, and $M = -P$.

so we have
$$\begin{matrix} A & B & C & C \\ -|F| & -|G| & H & H \\ 0 & 0 & -p & p \end{matrix} \qquad K=L=0.$$

The Tmod matrix is

$$\begin{bmatrix} aB+fG & aC+fH & aD+fJ & 0 & 0 & 0 \\ 0 & M & p & 0 & 0 & 0 \\ -bA-gF & 0 & 0 & bC+gH & bD+gJ & 0 \\ 0 & 0 & 0 & M & p & 0 \\ 0 & -cA-hF & 0 & -cB-hG & 0 & cD+hJ+mP \\ 0 & 0 & -dA-jF & 0 & -dB-jG & -dC-jH-pM \end{bmatrix}$$

$$= \begin{bmatrix} B-G & C-H & D-J & 0 & 0 & 0 \\ 0 & M & p & 0 & 0 & 0 \\ -A+F & 0 & 0 & C-H & D-J & 0 \\ 0 & 0 & 0 & M & p & 0 \\ 0 & -A-F & 0 & -B-G & 0 & D+J-P \\ 0 & 0 & -A-F & 0 & -B-G & -C-H-M \end{bmatrix}$$

$$= \begin{bmatrix} B+|G| & C-H & C-H & 0 & 0 & 0 \\ 0 & -p & p & 0 & 0 & 0 \\ -A-|F| & 0 & 0 & C-H & C-H & 0 \\ 0 & 0 & 0 & -p & p & 0 \\ 0 & -A+|F| & 0 & -B+|G| & 0 & 1-2P \\ 0 & 0 & -A+|F| & 0 & -B+|G| & -(1-2P) \end{bmatrix}$$

$$\sum_{i=1}^{6} a_i \times \text{row } i \text{ of matrix} = 0.$$

From last col. of matrix, $a_5 = a_6$.

Also, $A + |F| = B + |G| = 1$ so $a_1 = a_3$, from col. 1 of Tmod.

From comparing col.'s 2+3, $a_2 = 0$.

" " " 4+5, $a_4 = 0$.

$$a_1(C-H) + a_5(-A+|F|) = 0$$

$$a_3(C-H) + a_5(-B+|G|) = 0 \quad \text{and } a_1 = a_3.$$

So $a_1(C-H) + a_5(-B+|G|) = 0$.

$$a_5(-A+|F|) = a_5(-B+|G|)$$

$a_5 = 0$ or $-A+|F| = -B+|G|$ or both.

$-1+2|F| = -1+2|G|$

$F = G$

$A = B$

Columns 1 and 2 of V are identical (not optimal)

$a_1(C-H) = 0$

$C = H$

$$\begin{pmatrix} A & B & C & C \\ -(1-A) & -(1-B) & C & C \\ 0 & 0 & -P & P \end{pmatrix}$$

$a_1 = 0$

$a_3 = 0$

$-A+|F|$ from GPT

so

$a_5 = 0$

All a_i are 0

(no good)

$$\mathcal{L}^2 = A^2 + B^2 + 2C^2 = (1-A)^2 + (1-B)^2 + 2C^2$$

$$0 = 2 - 2A - 2B$$

$$A + B = 1$$

$$\begin{pmatrix} A & 1-A & C & C \\ -(1-A) & -A & C & C \\ 0 & 0 & -p & p \end{pmatrix}$$

Rows 1 + 2 are $\perp$: $2C^2 = 2A(1-A)$

$$C^2 = A(1-A)$$

$$\begin{aligned} \mathcal{L}^2 &= 2C^2 + A^2 + (1-A)^2 \\ &= 2A(1-A) + A^2 + (1-A)^2 \\ &= [A + (1-A)]^2 \\ &= 1. \end{aligned}$$

$\mathcal{L} = 1$ - not optimal.

2 zeros in a column.

$$\begin{array}{cccc} O & B & C & D \\ O & G & H & J \\ K & L & M & P \end{array} \qquad \begin{array}{cccc} 0 & \cdot & \cdot & + \\ 0 & \cdot & \cdot & + \\ + & + & + & + \end{array} \nearrow \begin{array}{cccc} 0 & + & - & + \\ 0 & - & + & + \\ + & + & + & + \end{array} \qquad \searrow \begin{array}{cccc} 0 & + & - & + \\ 0 & - & - & + \\ + & + & + & + \end{array}$$

$\downarrow$

Multiply 2nd row by (−1):

$$\begin{array}{cccc} 0 & + & - & + \\ 0 & + & + & - \\ + & + & + & + \end{array}$$

Switch columns 2 + 4; this is equivalent to the first one.

So we have
$$\begin{array}{cccc} 0 & + & - & + \\ 0 & - & + & + \\ + & + & + & + \end{array}$$

Columns 2+3:

$$\swarrow \begin{array}{cccc} 0 & + & - & + \\ 0 & - & + & + \\ + & + & + & + \end{array} \qquad \searrow \begin{array}{cccc} 0 & + & - & + \\ 0 & - & + & + \\ + & \boxed{+} & \boxed{+} & + \end{array} \rightarrow \begin{array}{cccc} 0 & + & - & \boxed{+} \\ 0 & - & + & \boxed{+} \\ + & \boxed{+} & \boxed{+} & + \end{array}$$

2 squares in same column - no good.

Columns 3+4

$$\swarrow \begin{array}{cccc} 0 & + & - & + \\ 0 & - & + & + \\ + & + & + & + \end{array} \qquad \searrow \begin{array}{cccc} 0 & + & \boxed{-} & \boxed{+} \\ 0 & - & + & + \\ + & + & + & + \end{array} \searrow \begin{array}{cccc} 0 & + & \boxed{-} & \boxed{+} \\ 0 & \boxed{-} & + & + \\ + & + & + & + \end{array}$$

Compare col's 1 + 3 - no good.

Columns 2+3

$$\swarrow \begin{array}{cccc} 0 & + & - & + \\ 0 & \boxed{-} & + & \boxed{+} \\ + & + & + & + \end{array} \qquad \searrow \begin{array}{cccc} 0 & + & - & + \\ 0 & - & + & + \\ + & + & + & + \end{array} \searrow \begin{array}{cccc} 0 & \boxed{+} & - & + \\ 0 & - & \boxed{+} & + \\ + & + & + & + \end{array}$$

B would be a circle - no good.

So we have
$$\begin{matrix} 0 & \boxed{+} & - & + \\ 0 & - & \boxed{+} & + \\ 1 & + & + & + \end{matrix} \quad = \quad \begin{matrix} 0 & B & C & D \\ 0 & G & H & J \\ 1 & L & M & P \end{matrix}$$

Add a 4th row: w, x, y, z to get a 4×4 matrix W which is $\mathcal{L} \times$ a 4×4 orthogonal matrix. Let $w < 0$.

Then $w = -\sqrt{\mathcal{L}^2 - 1}$. We must choose x, y, z so that col's 2, 3, and 4 are all $\perp$ to col. 1.

Result: $x = \dfrac{L}{\sqrt{\mathcal{L}^2-1}}$, $y = \dfrac{M}{\sqrt{\mathcal{L}^2-1}}$, $z = \dfrac{P}{\sqrt{\mathcal{L}^2-1}}$.

x, y, and z are all > 0.

$$W = \begin{pmatrix} 0 & B & C & D \\ 0 & G & H & J \\ 1 & L & M & P \\ w & x & y & z \end{pmatrix}$$

The 3rd and 4th columns of W are $\perp$

so $|C|\,D = HJ + MP + yz$

$$= HJ + MP\left(1 + \frac{1}{\mathcal{L}^2 - 1}\right) = HJ + MP\frac{\mathcal{L}^2}{\mathcal{L}^2-1}$$

The 2nd and 4th columns of W are $\perp$ so

$$|G|\,J = BD + LP\frac{\mathcal{L}^2}{\mathcal{L}^2-1}$$

Also, $D + J + P = 1$

$$D^2 + J^2 + P^2 < D + J + P = 1 \quad \text{so} \quad D^2 + J^2 + P^2 < 1$$

And, $D^2 + J^2 + P^2 + z^2 = \mathcal{L}^2$

$$\mathcal{L}^2 - z^2 = D^2 + J^2 + P^2 < 1.$$

$$z = \frac{P}{\sqrt{\mathcal{L}^2-1}} \longrightarrow z^2 = \frac{P^2}{\mathcal{L}^2-1}$$

$$\mathcal{L}^2 - \frac{p^2}{\mathcal{L}^2-1} < 1.$$

$$(\mathcal{L}^2-1)\mathcal{L}^2 - p^2 < \mathcal{L}^2-1$$

$$\mathcal{L}^4 - \mathcal{L}^2 - p^2 < \mathcal{L}^2 - 1$$

$$\mathcal{L}^4 - 2\mathcal{L}^2 + 1 < p^2$$

$$(\mathcal{L}^2-1)^2 < p^2$$

$$\mathcal{L}^2 - 1 < p \quad \text{so} \quad \frac{p}{\mathcal{L}^2-1} > 1.$$

$$HJ + Mp\frac{\mathcal{L}^2}{\mathcal{L}^2-1} > HJ + M\mathcal{L}^2$$

$$BD + Lp\frac{\mathcal{L}^2}{\mathcal{L}^2-1} > BD + L\mathcal{L}^2$$

$H > \frac{1}{2}$ so $HJ + M\mathcal{L}^2 > \frac{1}{2}J + M\mathcal{L}^2$

$B > \frac{1}{2}$ so $BD + L\mathcal{L}^2 > \frac{1}{2}D + L\mathcal{L}^2$

But $|C|D = HJ + Mp\frac{\mathcal{L}^2}{\mathcal{L}^2-1}$, so $|C|D > \frac{1}{2}J + M\mathcal{L}^2$

Similarly, $|G|J > \frac{1}{2}D + L\mathcal{L}^2$

Adding these, $|C|D + |G|J > \frac{1}{2}(D+J) + (L+M)\mathcal{L}^2$

$$\left(\tfrac{1}{2} - |C|\right)D + \left(\tfrac{1}{2} - |G|\right)J + (L+M)\mathcal{L}^2 < 0.$$

$|C|$ and $|G|$ are $< \frac{1}{2}$

so we find that a positive number is negative – Contradiction!

One zero and snug-fitting.

$$\begin{matrix} O & B & C & D \\ F & G & H & J \\ K & L & M & P \end{matrix} = \begin{matrix} 0 & + & + & + \\ + & & & \\ + & & & \end{matrix}$$

$$\begin{matrix} 0 & + & + & + \\ + & + & + & - \\ + & & & \end{matrix} \qquad \begin{matrix} 0 & + & + & + \\ + & + & - & - \\ + & & & \end{matrix}$$

$$\begin{matrix} 0 & + & + & + \\ + & + & + & - \\ + & + & + & \end{matrix}$$

No good- can't make rows $\perp$.

$$\begin{matrix} 0 & + & + & + \\ + & + & + & - \\ + & + & - & \end{matrix}$$

P may be + or −.

$$\begin{matrix} 0 & + & + & + \\ + & + & + & - \\ + & - & - & \end{matrix}$$

P is +. (so rows 1 & 3 can be $\perp$)

$$\begin{matrix} 0 & + & + & + \\ + & + & - & - \\ + & & + & + \end{matrix}$$

L must be −

$$\begin{matrix} 0 & + & + & + \\ + & + & - & - \\ + & & + & - \end{matrix}$$

L is + or −.

$$\begin{matrix} 0 & + & + & + \\ + & + & - & - \\ + & & - & - \end{matrix}$$

No good.

So we have

1. $\begin{matrix} 0 & + & + & + \\ + & + & + & - \\ + & + & - & + \end{matrix}$

2. $\begin{matrix} 0 & + & + & + \\ + & + & + & - \\ + & + & - & - \end{matrix}$

3. $\begin{matrix} 0 & + & + & + \\ + & + & + & - \\ + & - & - & + \end{matrix}$

4. $\begin{matrix} 0 & + & + & + \\ + & + & - & - \\ + & - & + & + \end{matrix}$

5. $\begin{matrix} 0 & + & + & + \\ + & + & - & - \\ + & + & + & - \end{matrix}$

6. $\begin{matrix} 0 & + & + & + \\ + & + & - & - \\ + & - & + & - \end{matrix}$

2. and 5. are equivalent, so there's just 1, 2, 3, 4, and 6.

(1.) $\begin{array}{cccc} 0 & + & + & + \\ + & + & + & - \\ + & + & - & + \end{array} \longrightarrow \begin{array}{cccc} 0 & + & + & + \\ \oplus & + & + & - \\ \boxplus & + & - & + \end{array} \longrightarrow \begin{array}{cccc} 0 & + & + & + \\ \oplus & \boxplus & \oplus & \ominus \\ \boxplus & + & - & + \end{array}$

Start with column 1.

Compare col.'s 1+3, 1+4, then V_{22} must be □. Compare col.'s 2+4 – no good since V_{22} is □ and V_{24} is O.

$\downarrow$

$\begin{array}{cccc} 0 & + & + & + \\ \boxplus & + & + & - \\ \oplus & + & - & + \end{array}$

$\downarrow$

$\begin{array}{cccc} 0 & + & + & + \\ \boxplus & + & + & - \\ \oplus & \boxplus & \ominus & \oplus \end{array}$

Compare col.'s 1+3, 1+4, then V_{32} is □; compare 2+3 – no good. So we've eliminated (1.)

We will look at (2.) later.

(3.) $\begin{array}{cccc} 0 & + & + & + \\ + & + & + & - \\ + & - & - & + \end{array} \longrightarrow \begin{array}{cccc} 0 & + & + & + \\ \oplus & \oplus & \oplus & \ominus \\ \boxplus & - & - & + \end{array}$

Col. 1, 1+2, 1+3, 1+4. 4 O's in a row – no good.

$\downarrow$

$\begin{array}{cccc} 0 & + & + & + \\ \boxplus & + & + & - \\ \oplus & \ominus & \ominus & \oplus \end{array}$

Col. 1, 1+2, 1+3, 1+4. Also no good.

(4.) $\begin{array}{cccc} 0 & + & + & + \\ + & + & - & - \\ + & - & + & + \end{array}$

No good for same reason as (3.)

(6.) $\begin{array}{cccc} 0 & + & + & + \\ + & + & - & - \\ + & - & + & - \end{array} \longrightarrow \begin{array}{cccc} 0 & + & + & + \\ \boxplus & + & - & - \\ \oplus & \ominus & \oplus & \boxminus \end{array}$

Col. 1, 1+2, 1+3, V_{34} is □, 3+4 – no good since V_{33} is O and V_{34} is □.

$\searrow$ $\begin{array}{cccc} 0 & + & + & + \\ \oplus & \oplus & \ominus & \boxminus \\ \boxplus & - & + & - \end{array}$

Col. 1, 1+2, 1+3, V_{24} is □, 2+4 – no good.

The only case left is ②: $\begin{matrix} 0 & + & + & + \\ + & + & + & - \\ + & + & - & - \end{matrix}$

Multiply 4th col. by (-1): $\begin{matrix} 0 & + & + & - \\ + & + & + & + \\ + & + & - & + \end{matrix}$

Switch col.'s 3+4 and then 2+4: $\begin{matrix} 0 & + & - & + \\ + & + & + & + \\ + & - & + & + \end{matrix}$

Switch rows 2+3: $\begin{matrix} 0 & + & - & + \\ + & - & + & + \\ + & + & + & + \end{matrix}$ (This is the form we originally used for this case.)

$$\begin{matrix} 0 & + & - & + \\ \boxed{+} & - & + & + \\ + & + & + & + \end{matrix} \qquad\qquad \begin{matrix} 0 & + & - & + \\ + & - & + & + \\ \boxed{+} & + & + & + \end{matrix}$$

We started with column 1; then:

$$\begin{matrix} 0 & \boxed{+} & - & + \\ \boxed{+} & - & + & + \\ + & + & + & \boxed{+} \end{matrix} \qquad\qquad \begin{matrix} 0 & \boxed{+} & - & + \\ + & - & \boxed{+} & + \\ \boxed{+} & + & + & + \end{matrix}$$

Col.'s 1+2, 2+3, V_{34}, $V_{14}+V_{24}$, col.'s 3+4, V_{12}, V_{22}.

Col.'s 1+2, 2+4, V_{23}, $V_{13}+V_{33}$, col.'s 3+4, V_{12}, V_{32}.

However, these 2 sub-cases are equivalent: Start with $\begin{matrix} 0 & \boxed{+} & - & + \\ + & - & \boxed{+} & + \\ \boxed{+} & + & + & + \end{matrix}$ and interchange rows 2+3, then col.'s 3+4,

to obtain $\begin{matrix} 0 & \boxed{+} & + & - \\ \boxed{+} & + & + & + \\ + & - & + & \boxed{+} \end{matrix}$

Then multiply the 2nd col. and the 1st row by (-1), to get the first sub-case.

So we have just one subcase, with $A=0$, C and G negative, and the others positive, and with

$$\begin{matrix} 0 & B & C & D \\ F & G & H & J \\ K & L & M & P \end{matrix}$$

$$\begin{matrix} 0 & \boxed{+} & - & + \\ \boxed{+} & - & + & + \\ + & + & + & \boxed{+} \end{matrix}$$

We don't know yet whether H is $> \frac{1}{2}$ or $< \frac{1}{2}$.

$$T_{mod} = \begin{bmatrix} G+L & H+M & J+P & 0 & 0 & 0 \\ B & C & D & 0 & 0 & 0 \\ F-K & 0 & 0 & C-H+M & D-J+P & 0 \\ 0 & -F-K & 0 & B-G-L & 0 & -D+J+P \\ 0 & 0 & -F-K & 0 & -B-G-L & -C-H-M \end{bmatrix}$$

$$C-H+M = -|C|-H+M = -1+2M$$
$$D-J+P = 1-2J$$
$$B-G-L = B+|G|-L = 1-2L$$
$$-D+J+P = 1-2D$$
$$-B-G-L = -B+|G|-L = -1+2|G|$$
$$-C-H-M = |C|-H-M = -1+2|C|$$

$$F-K = 1-2K$$
$$-F-K = -1$$
$$G+L = -|G|+L$$

We may change the first row of T_{mod} to $(B+G+L,\ C+H+M,\ D+J+P,\ 0, 0, 0)$ without affecting the rank.

$$B+G+L = 1-2|G|$$
$$C+H+M = 1-2|C|$$
$$D+J+P = 1$$

T_{mod} becomes

$$\begin{bmatrix} 1-2|G| & 1-2|C| & 1 & 0 & 0 & 0 \\ B & -|C| & D & 0 & 0 & 0 \\ 1-2K & 0 & 0 & -(1-2M) & 1-2J & 0 \\ 0 & -1 & 0 & 1-2L & 0 & 1-2D \\ 0 & 0 & -1 & 0 & -(1-2|G|) & -(1-2|C|) \end{bmatrix}$$

For optimality, this matrix must not have maximum rank: some linear combination of the rows must be zero, with not all of the coefficients zero. (Call them a_1, a_2, a_3, a_4, a_5.)

$\boxed{24}$

$a_3(1-2M) = a_4(1-2L)$
$a_3(1-2J) = a_5(1-2|G|)$
$a_4(1-2D) = a_5(1-2|C|)$

(one or more)
If any of a_3, a_4, or a_5 is zero, then they are all zero. This leads to
$a_1(1-2|G|) + a_2 B = 0$
$a_1(1-2|C|) - a_2|C| = 0$
$a_1 + a_2 D = 0$
If one of a_1 and a_2 is zero, then both are zero. (Then all the a_i are zero.)
So we may assume $a_1 \neq 0$, $a_2 \neq 0$.
But $|G| < \frac{1}{2}$, $|C| < \frac{1}{2}$, so $\frac{a_1}{a_2}$ would have to be both pos. & neg. (contradiction)
So a_3, a_4, and a_5 are all nonzero.
$\frac{a_3}{a_4}\frac{a_4}{a_5}\frac{a_5}{a_3} = 1$. So $\left(\frac{1-2L}{1-2M}\right)\left(\frac{1-2|C|}{1-2D}\right)\left(\frac{1-2J}{1-2|G|}\right) = 1$
$(1-2L)(1-2|C|)(1-2J) = (1-2M)(1-2D)(1-2|G|)$.

$a_1 + a_2 D = a_5$
$a_1(1-2|C|) - a_2|C| = a_4$
$a_1(1-2|G|) + a_2 B = -(1-2K)a_3$

$$a_3 = \frac{a_1(1-2|G|)+a_2 B}{-(1-2K)}$$

$$\frac{a_4}{a_5} = \frac{a_1(1-2|C|) - a_2|C|}{a_1 + a_2 D}$$ but from before,

$$\frac{a_4}{a_5} = \frac{1-2|C|}{1-2D}.$$

$(a_1 + a_2 D)(1-2|C|) = (1-2D)[a_1(1-2|C|) - a_2|C|]$
Can a_2 be zero? If so, then $a_1(1-2|C|) = a_1(1-2D)(1-2|C|)$

$\boxed{25}$

$$a_1(1-2|C|)\underbrace{(1-1+2D)}_{2D}=0$$

So either $a_1=0$, or $D=0$ (impossible) or $|C|=\frac{1}{2}$ (impossible.) If a_1 and $a_2=0$ then $a_5=0$, so all the a_i are zero. So $a_2\neq 0$.

$$a_1\left[(1-2|C|)-(1-2|C|)(1-2D)\right]=-a_2|C|(1-2D)-a_2D(1-2|C|)$$

$$a_1(1-2|C|)(2D)=-a_2(|C|+D-4|C|D)$$

$$\frac{a_1}{a_2}=\frac{|C|+D-4|C|D}{-2D(1-2|C|)}$$

Also, $a_1(1-2|C|)-a_2|C|=a_4$

$$\frac{a_1}{a_2}(1-2|C|)-|C|=\frac{a_4}{a_2}$$

From before, $a_3=\dfrac{a_1(1-2|G|)+a_2B}{-(1-2K)}$

$$a_5=a_1+a_2D$$

$$\frac{a_3}{a_5}=\frac{a_1(1-2|G|)+a_2B}{-(1-2K)(a_1+a_2D)}$$

But also, $\dfrac{a_3}{a_5}=\dfrac{1-2|G|}{1-2J}$

$$\frac{a_3}{a_5}=\frac{\frac{a_1}{a_2}(1-2|G|)+B}{-(1-2K)\left(\frac{a_1}{a_2}+D\right)}=\frac{1-2|G|}{1-2J}$$

so $-(1-2|G|)(1-2K)\left(\frac{a_1}{a_2}+D\right)=(1-2J)\left[\frac{a_1}{a_2}(1-2|G|)+B\right]$

26

$$B(1-2J) + D(1-2|G|)(1-2K)$$

$$= \left(\frac{a_1}{a_2}\right)\left[-(1-2|G|)(1-2K) - (1-2J)(1-2|G|)\right]$$

$$= -(1-2|G|)\left[(1-2K)+(1-2J)\right]\left(\frac{a_1}{a_2}\right)$$

so $$\frac{a_1}{a_2} = \frac{B(1-2J) + D(1-2|G|)(1-2K)}{-(1-2|G|)\left[(1-2K)+(1-2J)\right]}$$

Comparing this with the other expression for $\frac{a_1}{a_2}$, we obtain

$$2D(1-2|C|)\left[B(1-2J)+D(1-2|G|)(1-2K)\right]$$

$$= (1-2|G|)\left[(1-2K)+(1-2J)\right]\left[|C|(1-2D)+D(1-2|C|)\right]$$

$$2DB(1-2J)(1-2|C|) + 2D^2(1-2|C|)(1-2|G|)(1-2K)$$

$$= \left[D(1-2|C|)+|C|(1-2D)\right](1-2|G|)(1-2K)$$

$$+ \left[D(1-2|C|)+|C|(1-2D)\right](1-2|G|)(1-2J)$$

$$\left\{2D^2(1-2|C|) - \left[D(1-2|C|)+|C|(1-2D)\right]\right\}(1-2|G|)(1-2K)$$

$$= (1-2|G|)(1-2J)\left[|C|(1-2D)+D(1-2|C|)\right] - 2DB(1-2J)(1-2|C|)$$

$$\text{L.H.S.} = (1-2|G|)(1-2K)\left[\underbrace{(2D^2-D)}_{=-D(1-2D)}(1-2|C|) - |C|(1-2D)\right]$$

$$\text{L.H.S.} = (1-2|G|)(1-2K)(1-2D)(-|C|-D+2|C|D)$$

$\boxed{27}$

R.H.S of same equation is

$$(1-2J)\underbrace{\left\{(1-2|G|)\left[|C|(1-2D)+D(1-2|C|)\right]-2DB(1-2|C|)\right\}}_{=\,|C|(1-2D)(1-2|G|)+D(1-2|C|)(1-2|G|-2B)}$$

but $B+|G|=1-L$

$$2(B+|G|)=2-2L$$

$$1-2|G|-2B=-1+2L$$

so $\text{R.H.S.}=(1-2J)\left[|C|(1-2D)(1-2|G|)-D(1-2|C|)(1-2L)\right]$

$$=(1-2J)|C|(1-2D)(1-2|G|)-D\underbrace{(1-2|C|)(1-2L)(1-2J)}_{=(1-2D)(1-2G)(1-2M)}$$

$\text{R.H.S.}=\cancel{(1-2D)(1-2J)\left[|C|(1-2|G|)-D(1-2M)(1-2G)\right]}$

$(1-2D)(1-2|G|)\left[|C|(1-2J)-D(1-2M)\right]$

~~We can remove a factor of (1-2D) from both L.H.S + R.H.S:~~

$\cancel{(1-2|G|)(1-2K)(|C|+D-2|C|D)=(1-2J)(|C|-D+2|C||G|+2DM)}$

$\cancel{\text{R.H.S.}=(1-2D)(1-2J)(1-2|G|)\left[|C|-D(1-2M)\right]}$

$\cancel{\text{so }(1-2K)(|C|+D-2|C|D)=(1-2J)(|C|-D+2MD)}$

so the equation becomes $-(1-2K)(|C|+D-2|C|D)$

$$=|C|(1-2J)-D(1-2M)$$

$$(1-2K)(|C|+D-2|C|D)=D-|C|-2MD+2|C|J$$

$$|C|+D-2|C|D-2K(|C|+D-2|C|D)=D-|C|-2MD+2|C|J$$

$$2|C|-2|C|D-2K(|C|+D-2|C|D)=2|C|J-2DM$$

$\boxed{28}$

$$|C| - |C|D - K|C| - KD + 2|C|DK - |C|J + DM = 0.$$

$$|C|(1-K-J) + |C|D(-1+2K) + D(-K+M) = 0.$$

$$|C|(1-K-J-D) = D(K-M-2|C|K)$$

$$D+J+P=1$$

$$|C|(P-K) = D(K-M-2|C|K)$$

85 (1)

Tues., 4/30/96

$\boxed{29}$

$$|C|(P-K) = D(K-M-2|C|K)$$

$$\begin{array}{cccc} O & B & C & D \\ F & G & H & J \\ K & L & M & P \\ w & x & y & z \end{array} \qquad \begin{array}{cccc} O & \boxplus & \ominus & \oplus \\ \boxplus & \ominus & + & \oplus \\ \oplus & \oplus & \oplus & \boxplus \end{array}$$

Add 4th row

$P-K$ is positive, so $K-M-2|C|K > 0$

$$K(1-2|C|) > M$$
$$|C|+H+M=1$$
$$1-2|C| = H+M-|C|$$
$$K(H+M-|C|)>M$$
$$KM+K(H-|C|)>M$$
$$K(H-|C|) > M(1-K) \qquad K<1 \text{ so } H-|C|>0.$$
$$F+K=1$$
$$K(H-|C|) > MF > \tfrac{1}{2}M \qquad \cancel{\text{so } H>|C|}$$
$$K(H-|C|) < \tfrac{1}{2}(H-|C|)$$
$$\tfrac{1}{2}(H-|C|) > \tfrac{1}{2}M$$
$$H-|C| > M$$
$$H > M+|C|$$
$$H > 1-H$$
$$2H>1$$
$$H > \tfrac{1}{2}.$$

In 4th row, let w be negative, so $F^2+K^2 = L^2 - w^2$,

$$w = -\sqrt{L^2-F^2-K^2},$$

30

$$x = \frac{FG+KL}{\sqrt{\mathcal{L}^2 - F^2 - K^2}}, \qquad y = \frac{FH+KM}{\sqrt{\mathcal{L}^2 - F^2 - K^2}};$$

$$z = \frac{FJ+KP}{\sqrt{\mathcal{L}^2 - F^2 - K^2}}$$

y and z are positive.

$$CD + HJ + MP + yz = 0.$$

$$HJ + MP + yz = |C| D$$

$$HJ + MP + yz > \tfrac{1}{2} J + \tfrac{1}{2} M + yz$$

so $\tfrac{1}{2}(J+M) + yz < |C| D$

$$BD + GJ + LP + xz = 0.$$

$$BD + LP + xz = |G| J$$

$$BD + LP + xz > \tfrac{1}{2}(D+L) + xz$$

$$\tfrac{1}{2}(J+M) + \tfrac{1}{2}(D+L) + xz + yz < |C| D + |G| J$$

$$\tfrac{1}{2}(D+L+J+M) + xz + yz < |C| D + |G| J$$

$$\tfrac{1}{2} D - |C| D + \tfrac{1}{2} J - |G| J + \tfrac{1}{2}(L+M) + (x+y) z < 0.$$

$$\underbrace{\left(\tfrac{1}{2} - |C|\right) D + \left(\tfrac{1}{2} - |G|\right) J + \tfrac{1}{2}(L+M)}_{\text{positive}} + (x+y) z < 0$$

$\boxed{31}$

$(x+y)z < 0$

$z > 0$ so $x+y<0.$

$x+y<0$ so $FG+KL+FH+KM<0$

$-F|G|+KL+FH+KM<0$

$F(\underbrace{H-|G|}_{\text{positive}})+K(L+M)<0$

We have a contradiction.

33

Fri., 4/5/96

Important theorems
for a cube in a tesseract, concerning
plane rotations (involving 2 columns of V)

$$\begin{array}{cc} A & D \\ B & E \\ C & F \end{array}$$

↑ ↑ 2 columns of V. (Col 1 + Col 2)

If Col 1 → (Col 1)(cos θ) + (Col 2)(sin θ)
and Col 2 → (Col 2) cos θ − (Col 1) sin θ

Let $a =$ sign A, $b =$ sign B, etc., so that $a = 1$ for A positive, $a = -1$ for A negative, and $a = 0$ for $A = 0$.

$|A| = a|A| \to a(A\cos\theta + D\sin\theta)$ for θ sufficiently small and $a \neq 0$.

$$\frac{d}{d\theta}|A| = a(-A\sin\theta + D\cos\theta)$$

at $\theta = 0$, this is $\left.\frac{d}{d\theta}|A|\right|_{\theta=0} = aD = ad|D|$.

Similarly, if $D \neq 0$, then

$$|D| = dD \to d(D\cos\theta - A\sin\theta)$$

$$\left.\frac{d}{d\theta}|D|\right|_{\theta=0} = -dA = -ad|A|.$$

If A, B, C, D, E, F are all nonzero, then

$$\left.\frac{d}{d\theta}(R_1)\right|_{\theta=0} = \left.\frac{d}{d\theta}(|A|+|B|+|C|)\right|_{\theta=0} = ad|D| + be|E| + cf|F|.$$

This cannot be zero by GPT. (Generalization of Phil's theorem), or the cube is not optimal.

4/5/96

Similarly, $\frac{d}{d\theta}(R_2)\Big|_{\theta=0} = -ad|A| - be|B| - cf|C|$.

Since $\frac{d}{d\theta}(R_1)\Big|_{\theta=0} \neq 0$ for an optimal cube, we can increase or decrease R_1 by ~~change~~ a small rotation, if we don't care what happens to R_2.

Now, suppose $R_1 = R_2$. When is it possible to decrease R_1 and R_2 simultaneously? It is impossible if $\frac{dR_1}{d\theta}$ and $\frac{dR_2}{d\theta}$ have opposite signs. When do they have opposite signs?

Suppose $ad = 1$ and be and cf are -1. Then $\frac{dR_1}{d\theta} = |D| - |E| - |F|$ and

$$\frac{dR_2}{d\theta} = -|A| + |B| + |C|.$$

Assume we have normalized V so that $R_1 = R_2 = 1$. Then $|A| + |B| + |C| = |D| + |E| + |F| = 1$.

$$|D| - |E| - |F| = -1 + 2|D| = -(1 - 2|D|)$$
$$-|A| + |B| + |C| = 1 - 2|A|$$

If $|A|$ and $|D|$ are both $> \frac{1}{2}$ or both $< \frac{1}{2}$, then $\frac{dR_1}{d\theta}$ and $\frac{dR_2}{d\theta}$ have opposite signs, and conversely.

If a cube is optimal and it is impossible to decrease R_1 and R_2 simultaneously, then $|A|$ and $|D|$ are both $> \frac{1}{2}$ or both $< \frac{1}{2}$.

This is ~~is~~ if ad has one sign and be and cf have the other.

35

4/5/96

Now suppose that $A=0$, and all others are nonzero.

$$A \to A\cos\theta + D\sin\theta$$
$$= D\sin\theta$$
$$|A| \to |D\sin\theta| = |D|\sin\theta$$

for positive θ, ~~or~~ and $|D|(-\sin\theta)$ for negative θ.

For pos. θ, $\left.\dfrac{d|A|}{d\theta}\right|_{\theta=0} = |D|$.

For neg. θ, $\left.\dfrac{d|A|}{d\theta}\right|_{\theta=0} = -|D|$.

So $\dfrac{dR_1}{d\theta} = \pm|D| + be|E| + cf|F|$ ($+$ for θ pos., $-$ for θ neg.).

$$\frac{dR_2}{d\theta} = -ad|A| - be|B| - cf|C|$$
$$= -be|B| - cf|C|.$$

If the cube is optimal and we cannot decrease R_1, then

$$|D| + be|E| + cf|F| > 0$$
and $-|D| + be|E| + cf|F| < 0$.

If $be = 1$ and $cf = -1$, then

$|D| + |E| - |F| > 0$ and $-|D| + |E| - |F| < 0$.

$|F| < |D| + |E|$ and $|E| < |D| + |F|$.

If $R_1 = R_2$ and if the cube is optimal and we cannot decrease R_1 and R_2 simultaneously, then $|D| + be|E| + cf|F|$ and $-be|B| - cf|C|$ ~~have opposite signs, and~~ are not both negative, and $-|D| + be|E| + cf|F|$ and $-be|B| - cf|C|$ are not both positive.

4/5/96.

If $be=1$ and $cf=-1$, then
~~$|D|+|E|-|F|$ and $|B|-|C|$ are not both <0~~
~~and $-|D|+|E|-|F|$ and $|B|-|C|$ are not both >0.~~
~~If $|B|<|C|$, then $|D|+|E| > |F|$.~~
~~If $|B|>|C|$, then~~

If we can decrease R_1 with θ pos., then
$$|D|+|E|-|F|<0.$$
Then we cannot decrease R_2 with θ pos., so
$$-|B|+|C|>0.$$
So if $|F|>\frac{1}{2}$, then $|C|>\frac{1}{2}$.
If $|C|<\frac{1}{2}$, then $|F|<\frac{1}{2}$.

Similarly, if we can decrease R_1 with θ neg., then
$$-|D|+|E|-|F|>0. \quad (|E|>\tfrac{1}{2}.).$$
Then we cannot decrease R_2 with θ neg., so
$$-|B|+|C|<0. \quad (|B|>\tfrac{1}{2}.)$$
So if $|E|>\frac{1}{2}$ then $|B|>\frac{1}{2}$
and if $|B|<\frac{1}{2}$ then $|E|<\frac{1}{2}$.

Mon., Oct. 30, 1995

Mistake in cube paper – p. 8 of old single-spaced version – $\gamma' = 45^\circ$ should be $\alpha' = 45^\circ$.

$$\begin{vmatrix} s\alpha & c\alpha & 0 & 0 \\ c\alpha s\beta & -s\alpha s\beta & +c\beta & 0 \\ c\alpha c\beta s\gamma & -s\alpha c\beta s\gamma & -s\beta s\gamma & c\gamma \end{vmatrix} \qquad \alpha' = 45^\circ$$

$$\vec{V}_1 = \left(\frac{s\alpha}{c\gamma}, \frac{c\alpha}{c\gamma}, 0, 0 \right)$$

$$\vec{V}_2 = \left(\frac{c\alpha\, s\beta}{c\gamma}, -\frac{s\alpha\, s\beta}{c\gamma}, \frac{c\beta}{c\gamma}, 0 \right)$$

$$\vec{V}_3 = \left(c\alpha\, c\beta \tan\gamma, -s\alpha\, c\beta \tan\gamma, -s\beta \tan\gamma, 1 \right)$$

Choose $\alpha, \beta, \gamma = \alpha', \beta', \gamma'$. $\qquad c\gamma = \frac{1}{\sqrt{X_0}}$

$$\vec{V}_1 = \left(\frac{\sqrt{2}}{2} \underbrace{\sec\gamma}_{=\sqrt{X_0}}, \frac{\sqrt{2}}{2} \sec\gamma, 0, 0 \right)$$

$$\vec{V}_2 = \left(\frac{\sqrt{2}}{2} s\beta \sec\gamma, -\frac{\sqrt{2}}{2} s\beta \sec\gamma, c\beta \sec\gamma, 0 \right)$$

$$\vec{V}_3 = \left(\frac{\sqrt{2}}{2} c\beta \tan\gamma, -\frac{\sqrt{2}}{2} c\beta \tan\gamma, -s\beta \tan\gamma, 1 \right)$$

$$\vec{V}_1 + \vec{V}_2 = \left(\frac{\sqrt{2}}{2} \sqrt{X_0} (1 + s\beta), \frac{\sqrt{2}}{2} \sqrt{X_0} (1 - s\beta), c\beta \sqrt{X_0}, 0 \right)$$

$$\vec{V}_1 + \vec{V}_3 = \left(\frac{\sqrt{2}}{2} (\sqrt{X_0} + c\beta \tan\gamma), \frac{\sqrt{2}}{2} (\sqrt{X_0} - c\beta \tan\gamma), -s\beta \tan\gamma, 1 \right)$$

$$\vec{V}_2 + \vec{V}_3 = \left(\frac{\sqrt{2}}{2} (s\beta \sec\gamma + c\beta \tan\gamma), -\frac{\sqrt{2}}{2} (s\beta \sec\gamma + c\beta \tan\gamma), c\beta \sec\gamma - s\beta \tan\gamma \right.$$

$$\vec{V}_1 + \vec{V}_2 + \vec{V}_3 = \left(\frac{\sqrt{2}}{2} (\sqrt{X_0} + s\beta\sqrt{X_0} + c\beta \tan\gamma), \frac{\sqrt{2}}{2} (\sqrt{X_0} - s\beta\sqrt{X_0} - c\beta \tan\gamma), c\beta \sqrt{X_0} - s\beta \tan\gamma, 1 \right)$$

$$s\gamma = \sqrt{1-\frac{1}{x_0}}$$

$$\tan\gamma = s\gamma\sec\gamma = \sqrt{x_0-1}$$

$$\sec\gamma = \sqrt{x_0}$$

$$\vec{O} = (0,0,0,0)$$

$$c\gamma = c\beta + s\beta\, s\gamma$$

Let $c\beta = \sqrt{y}$.

$$c\gamma = \sqrt{y} + \sqrt{1-y}\; s\gamma$$

$$c^2\gamma = y + (1-y)s^2\gamma + 2\sqrt{y}\sqrt{1-y}\; s\gamma$$

$$c^2\gamma - y - (1-y^2)s^2\gamma = 2\sqrt{y}\sqrt{1-y}\; s\gamma$$

From old notes, p. 25, $S = \dfrac{\sqrt{2}\,R - 1}{\sqrt{2} - R}$

$$R = \frac{1}{\sqrt{x_0}} = c\gamma$$

$$S = c\beta$$

$$c\beta = \frac{\frac{\sqrt{2}}{\sqrt{x_0}} - 1}{\sqrt{2} - \frac{1}{\sqrt{x_0}}} = \frac{\sqrt{2} - \sqrt{x_0}}{\sqrt{2x_0} - 1} = \left(\frac{\sqrt{2}-\sqrt{x_0}}{\sqrt{2x_0}-1}\right)\left(\frac{\sqrt{2x_0}+1}{\sqrt{2x_0}+1}\right)$$

$$= \frac{2\sqrt{x_0} + \sqrt{2} - \sqrt{2}x_0 - \sqrt{x_0}}{2x_0 - 1} = \frac{\sqrt{2} + \sqrt{x_0} - \sqrt{2}\,x_0}{2x_0-1}$$

$$= \frac{\sqrt{x_0} - \sqrt{2}(x_0 - 1)}{2x_0 - 1}$$

Mon., May 6, 1996

Is $S<1$?

$S = \dfrac{\sqrt{2}R-1}{\sqrt{2}-R} \approx$ .403776484 ; $\beta \approx 16.716750^\circ$

~~.406778562~~ ~~$\beta \approx 6.965256^\circ$~~

.421593691

$S =$ ~~.992619871~~ .957738439

Is $\sqrt{2}R-1 < \sqrt{2}-R$?

Is $(1+\sqrt{2})R < 1+\sqrt{2}$? Yes.

$\sqrt{X_0} \approx 1.007435$ $\quad \dfrac{1}{\sqrt{X_0}} = R \approx .992619871$

$R = \dfrac{1}{\sqrt{X_0}}$

$s\beta = .287640520$

$\gamma \approx 6.965255582^\circ$ $\quad c\beta = .957738446$

$\alpha = 45^\circ$ $\quad \tan\gamma = .122169060$

$s\gamma = .121267437$

$c\gamma = .992619871$

$\vec{O} = (0,0,0,0)$

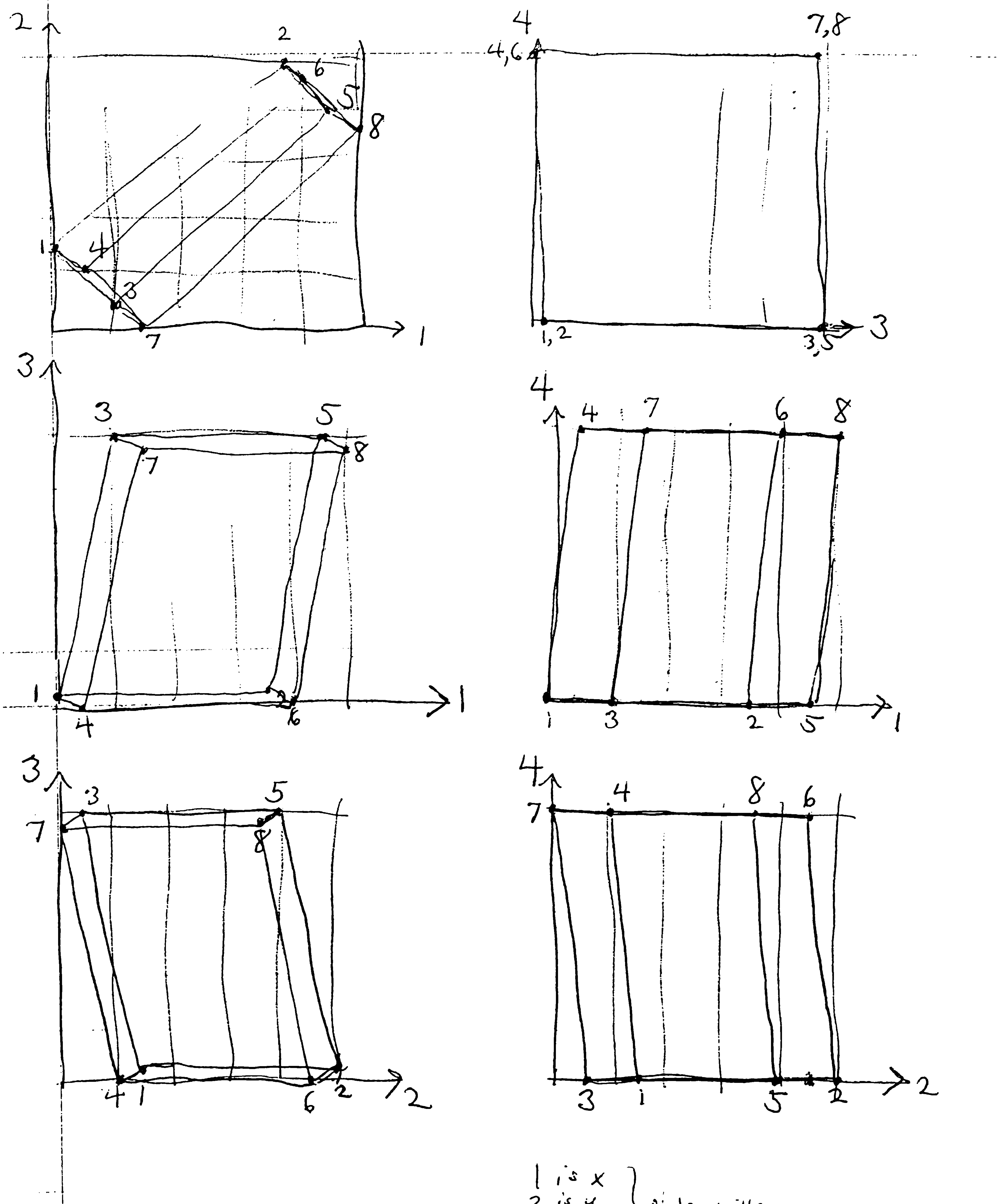
To 1st ~~row~~ col, add zero (leave unchanged). It already goes from 0 to 1.

Vertex no. 1	$\vec{O}_1 = 0$	.712364120	.287640526	.035140772	0
2	$\vec{V}_{11} = \frac{\sqrt{2}}{2}(1.007435) =$ ~~1.403776484~~		1	.035140772	0
3	$V_{21} =$ .204904786		.082735740	1	0
4	$V_{31} =$ .082735740		.204904786	0	1
5	$V_{11}+V_{21} =$ .917268906		.795099860	1	0
6	$V_{11}+V_{31} =$ .795099860		.917268906	0	1
7	$V_{21}+V_{31} =$ .287640526		0	.964859231	1
8	$V_{11}+V_{21}+V_{31} = 1.$		.712364120	.964859231	1

To 2nd col, add $|V_{22}+V_{32}|$

To 3rd col, add $|V_{33}|$

Edges: 1-2, 1-3, 1-4, 5-8, 6-8, 7-8, 2-5, 2-6, 3-5, 3-7, 4-6, 4-7



1 is x
2 is y
3 is z
4 is w

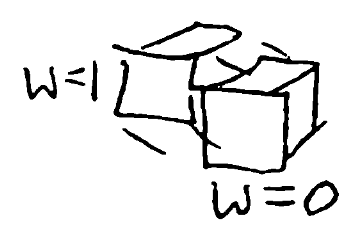
Note written on Mon, May 20

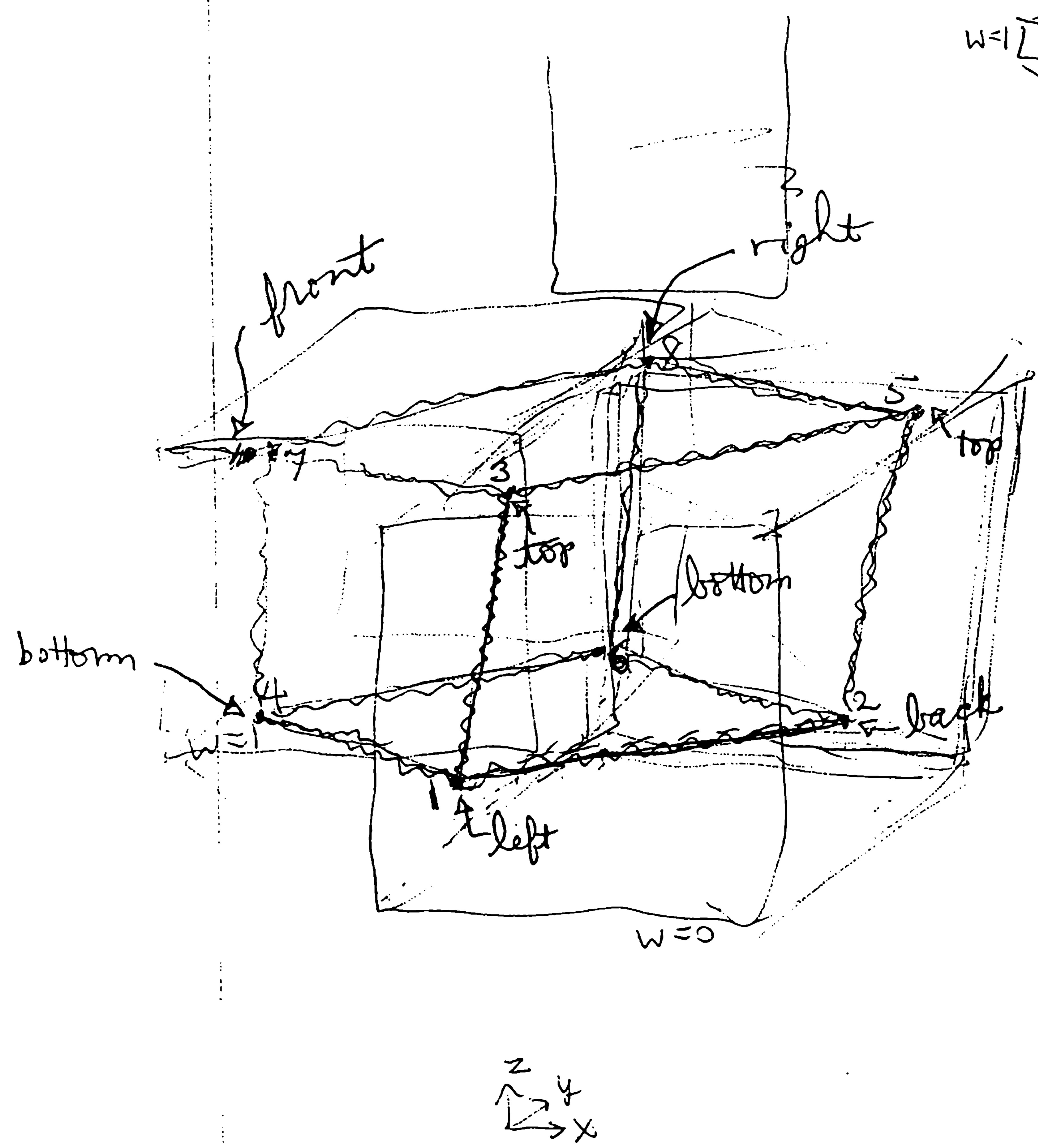
Tues, May 7, 1996

Angles corresponding to direction cosine matrix.

45°	45°	90°	90°
78.265°	101.735°	16.717°	90°
85.289°	94.711°	91.999°	6.965°

Mon., May 20, 1996





1 - left face of w=0 cube (x=0)

2 - back face of w=0 cube (y=1)

3 - top face of w=0 cube (z=1)

4 - bottom face of w=1 cube (z=0)

5 - top face of w=0 cube (z=1)

6 - bottom face of w=1 cube (z=0)

7 - front face of w=1 cube (y=0)

8 - right face of w=1 cube (x=1)

Acknowledgements

Kay Shultz would like to thank Philip DeVicci for showing her how to prove "Phil's Theorem," which makes it easier to find f(2,3), and which can be generalized to GPT, used in finding f(3,4). She would also like to thank the anonymous referees at Mathematics Magazine, who (in the early 1980s) showed her an easier way to find f(2,3) which can be modified to help find f(3,4). Both Kay Shultz and Greg Huber would like to thank Dr. Jing Yang of the Department of Physics, Haverford College, for his re-typesetting and production assistance. Greg Huber's work was partly supported by the National Science Foundation under grant no. NSF PHY11-25915.

Appendix A
Explanations

A Note to the Reader

Some parts of this paper are written as if I had not yet completed the analysis of the largest cube in a tesseract. This is because the paper was written in two stages, the first at a time when I didn't think I would be able to rule out (or in) the most difficult cases. However, both stages were completed in 1996. Only the Table of Contents and Appendices were added in 2013, and the abstract was updated slightly to reflect work which was actually completed in 1996.

There is a break in the paper between pages 62 and 63. This probably occurred because I realized that I would not be able to type the little squares and circles surrounding the plus and minus signs, such as on pages 63, 68, etc. This is why the rest of the paper is handwritten.

Appendix B
Theorems

Phil's Theorem states that for $m=2$, $n=3$, if $\pm\vec{v}_1 \pm \vec{v}_2$ (with any combination of the + and − signs) contains one or more zeros, then $\mathcal{L} \le 1$ (not optimal). Use the normalization where the n-cube has unit side. $\vec{v}_1 - \vec{v}_2$, say, fits into the cube, since it's a diagonal of a face. So its components have absolute values ≤ 1. If one or more components are zero, then the maximum squared length of $\vec{v}_1 - \vec{v}_2$ is $2 \; (= 1^2 + 1^2)$. So $|\vec{v}_1 - \vec{v}_2| \le \sqrt{2}$, and so $|\vec{v}_1| \le 1$.

GPT (Generalization of Phil's Theorem) states that for $m=3$, $n=4$, if $\pm\vec{v}_1 \pm \vec{v}_2 \pm \vec{v}_3$ has 1 or more zeros, or if $\pm\vec{v}_1 \pm \vec{v}_2$, or $\pm\vec{v}_1 \pm \vec{v}_3$, etc., has 2 or more zeros, or if $\vec{v}_1$, etc., has 3 or more 0s, then $\mathcal{L} \le 1$. The proof is similar: $|\pm\vec{v}_1 \pm \vec{v}_2 \pm \vec{v}_3| \le \sqrt{1^2+1^2+1^2} = \sqrt{3}$ so $|\vec{v}_1| \le 1$.

$|\pm\vec{v}_1 \pm \vec{v}_2| \le \sqrt{1^2+1^2} = \sqrt{2}$ so $|\vec{v}_1| \le 1$.

$|\pm\vec{v}_1| \le \sqrt{1^2} = \sqrt{1}$ so $|\vec{v}_1| \le 1$. (That's obvious anyway.)

The Zero Theorem

For a cube in a tesseract, suppose we add a 4th row to V to convert it into a 4x4 orthogonal matrix. Call the 4x4 matrix W. We shall prove that if one or more elements in the 4th row are zero, then the cube is not optimal.

Since all elements of W have absolute values ≤ 1, it follows that $w_{ij}^2 \leq |w_{ij}|$ for all (i, j). Assume that $w_{4j} = 0$. Then $R_j = \sum_{i=1}^{3} |w_{ij}| \geq \sum_{i=1}^{3} w_{ij}^2 = 1$. Since $R_j \geq 1$ for at least one value of j, $R \geq 1$ and the cube is not optimal.

Appendix C
Additions

On pp. 9–10, there's an easier way to rule out α', β', and $\gamma' = 90°$. Since $R_1 = s\alpha' + c\alpha' s\beta' + c\alpha' c\beta' s\gamma'$ (p. 8), then since $\sin 90° = 1$, $\alpha' = 90°$ implies $R_1 \geq 1$. If $\sin\beta' = 1$, then $R_2 = c\alpha' + s\alpha' + s\alpha' c\beta' s\gamma' \geq c\alpha' + s\alpha' \geq 1$. If $s\gamma' = 1$, then $R_3 = c\beta' + s\beta' \geq 1$. This method seems to work even in higher dimensions, for an m-cube in an $(m+1)$-cube.

And yet another way to prove that the minimum of R does not occur on the boundary:

On p. 8, we will add a 4th row to the matrix near the middle of the page.* If this 4th row is $(w\ x\ y\ z)$, then we can take $z = s\gamma$. Then $y = s\beta\, c\gamma$, and $x = s\alpha\, c\beta\, c\gamma$, and $w = -c\alpha\, c\beta\, c\gamma$. If $s\alpha = 0$, then $x = 0$. If $s\beta = 0$, then $y = 0$. If $s\gamma = 0$, then $z = 0$. If $c\alpha = 0$, then $w = 0$. If $c\beta = 0$, then w and x are 0. And if $c\gamma = 0$, then w, x, and y are 0. In all cases here, we can use the Zero Theorem,** proving that the cube is not optimal because $R \geq 1$.

* (so as to produce a 4×4 orthogonal matrix)

** See Appendix B, p. B2.

On p.46, we consider $\begin{matrix} 0 & X & X & X \\ 0 & X & X & X \\ X & 0 & X & X \end{matrix}$ where the X's may be zero or nonzero. There is a parametrization near the bottom of the page. I don't know if the parametrization is correct. However, it doesn't matter, because this case is easily ruled out by the Zero Theorem,* which states that if we add a 4th row to a 3x4 matrix so as to produce a 4x4 orthogonal matrix, and the 4th row contains one or more zeros, then the case is not optimal. Here, let us call the 4th row (w x y z). Since the first two columns must be orthogonal, w or x or both must be zero.

* See Appendix B, p. B2.

On p. 37, we decided to omit the analysis of two cases. Here we describe them, for the sake of completeness.

2 zeros in the first column, with $R = R_1 > R_2, R_3, R_4$. Assume that $F = K = 0$.

$$V = \begin{bmatrix} A & B & C & D \\ 0 & G & H & J \\ 0 & L & M & P \end{bmatrix}$$

$$T_{mod} = \begin{bmatrix} aB & aC & aD & 0 & 0 & 0 \\ G & H & J & 0 & 0 & 0 \\ L & M & P & 0 & 0 & 0 \end{bmatrix}, \text{ with } a = \pm 1.$$

Suppose $a_1 a(B,C,D) + a_2(G,H,J) + a_3(L,M,P) = 0$. Then, since (B,C,D) and (G,H,J) and (L,M,P) are all mutually orthogonal, we can multiply by (B,C,D) to get $a_1 a(B^2+C^2+D^2) = 0$. Since $B=C=D=0$ would lead to a non-optimal cube, it must be that $a_1 = 0$. Similarly, multiplying by (G,H,J) yields either $a_2 = 0$ or $G=H=J=0$. But the latter means $\vec{V}_2 = 0$, an impossibility. Thus, $a_2 = 0$. Then similarly again, $a_3 = 0$. Thus T_{mod} has maximum rank and the cube is not optimal.

Two zeros in a row: $K = L = 0$ and $R = R_1 = R_2 > R_3, R_4$.

$$V = \begin{bmatrix} A & B & C & D \\ F & G & H & J \\ 0 & 0 & M & P \end{bmatrix}$$

$$T_{mod} = \begin{bmatrix} aB+fG & aC+fH & aD+fJ & 0 & 0 & 0 \\ 0 & M & P & 0 & 0 & 0 \\ -bA-gF & 0 & 0 & bC+gH & bD+gJ & 0 \\ 0 & 0 & 0 & M & P & 0 \end{bmatrix}$$

If $a_1 \neq 0$ and $a_2, a_3,$ and $a_4 = 0$, then use col.'s 2 and 3 of T_{mod}: $a_1 \underbrace{[a(C,D) + f(H,J)]}_{\text{must be 0, but that would be non-optimal by GPT. So } a_1 = 0 \text{ (contradiction).}} = 0$.

If only $a_2 \neq 0$: use col.'s 2 and 3: $M = P = 0$ or $a_2 = 0$. But $M = P = 0$ means $\vec{V}_3 = 0$. (Not allowed.) So $a_2 = 0$ (contradiction).

If only $a_3 \neq 0$, use col.'s 4 and 5 and proceed similarly to $a_1 \neq 0$.

$a_4 \neq 0$: use col.'s 4 and 5 and it's like $a_2 \neq 0$.

If only a_1 and $a_2 \neq 0$, use col.'s 2+3:

$$a_1[a(C,D) + f(H,J)] + a_2(M,P) = 0.$$

Multiply through by (M,P) to get $a_2(M^2+P^2) = 0$.

Since $a_2 \neq 0$, M and P are 0. (Not allowed.)

Only a_1 and $a_3 \neq 0$: Use col.'s 2 and 3. It's like the case where only $a_1 \neq 0$.

a_1 and a_4: similar.

a_2 and a_3: similar to just a_2.

a_2 and a_4: similar to just a_2 or just a_4.

a_3 and a_4: use col.'s 4 and 5 and it's like $a_1 + a_2$.

$a_1, a_2, + a_3$: use col.'s 4 + 5 and it's like just a_3.

$a_1, a_2, + a_4$: like a_4.

a_1, a_3, a_4: like a_1.

a_2, a_3, a_4: like a_2.

$a_1, a_2, a_3, + a_4$: like $a_1 + a_2$.

Thus this case is also ruled out.

On p. 38, we are discussing the case of 2 zeros in a diagonal: $A=G=0$, with $R=R_1=R_2>R_3, R_4$. I just want to show how to do a_3 + a_4 nonzero, $a_1=a_2=0$. Use col.'s 1, 4, + 5.

$$a_3(-lK,\ bC+lM,\ bD+lP)+a_4(-F,H,J)=0.$$

$$a_3\left[b(0,C,D)+l(-K,M,P)\right]+a_4(-F,H,J)=0.$$

Since $V=\begin{bmatrix}0 & B & C & D\\ F & 0 & H & J\\ K & L & M & P\end{bmatrix}$, it follows that

$(-F,H,J)$ is orthogonal to both $(0,C,D)$ and $(-K,M,P)$. Thus we can multiply through by $(-F,H,J)$ to get $a_4(F^2+H^2+J^2)=0$. So again we get a contradiction.

Note: For $a_1, a_2, a_3 \neq 0$, $a_4=0$, use columns 4 and 5 of T_{mod} to rule it out.

For $a_1, a_2, a_4 \neq 0$, $a_3=0$, use 4 and 5 as well.

For $a_1, a_3, a_4 \neq 0$, $a_2=0$, use 2 and 3.

For $a_2, a_3, a_4 \neq 0$, $a_1=0$, use 2 and 3.

Appendix D
Clarifications

Snug case with no zeros

The proof that the snug case with no zeros is not optimal is somewhat fragmented, so we will clarify it here.

$$V = \begin{bmatrix} A & B & C & D \\ F & G & H & J \\ K & L & M & P \end{bmatrix}$$

On p. 43, we assume that only J, M, and P are negative. On p. 44, we assume that some linear combination of the rows of T is zero, where T is given on p. 32. On p. 45, we find that 2 columns of V are identical. On pp. 58-59, we prove the Identical Column Theorem, which shows that V is not optimal because it has 2 identical columns. On pp. 59-60, we show that for the sub-case in which 2 columns have the same pattern of signs, there was no loss of generality in assuming that J, M, and P were the negative entries.

On pp. 60 and 63, we consider the other subcase, where no two columns have the same pattern of signs, and we rule it out, using results from pp. 60 and 89.

Appendix E
Corrections

On p. 20 (largest square in an odd-dimensional cube), I don't know why I wrote "all equal" above and below the matrix. Also, when I was working on the proof, I don't think I realized that, in rotating a maximal column with both entries nonzero with a non-maximal column with a zero, the zero would become nonzero. However, this does not affect the validity of the proof.

On p. 29, where it says "If $w_{1j} \geq 0$ for some $j, \ldots$ for all $i \leq m/2$" and "If $w_{1j} \leq 0 \ldots$ for all $i \leq m/2$.", it should say,

"If $w_{1j} > 0 \ldots$ then $v_{ij} \geq 0 \ldots$"
"If $w_{1j} = 0 \ldots$ then $v_{ij} = 0 \ldots$"
"If $w_{1j} < 0 \ldots$ then $v_{ij} \leq 0 \ldots$"

The paragraph beginning with "However" is correct. The rest of p. 29, beginning with "Therefore," and the top part of p. 30, before "General Infinitesimal Rotations", should be replaced by the following:

Therefore, in the upper half of V, all of the matrix elements are zero except for one in each column equal to the corresponding w_{1j}, and if that w_{1j} is zero, then they are all zero. Similarly for the lower half of V and w_{2j}.

Thus, all v_{ij} can only be $0, \pm 1, \pm\frac{1}{4}$, and $\pm\frac{3}{4}$. Now the squared length must be $\frac{4n-3}{4m}$, which can never be an integer, since it is $\frac{\text{odd}}{\text{even}}$. But only four of the v_{ij} have non-integral squared length: $\pm\frac{3}{4}$ appears twice and $\pm\frac{1}{4}$ also appears twice. So if $m > 4$, some rows will contain only zeros and $\pm$ones, with an integral squared length. The only hope is $m=4$, where each row must contain one fraction. But then two rows will have a squared length of an integer plus $\frac{9}{16}$, and the other two will have an integer plus $\frac{1}{16}$, so these cannot be equal.

This is a pity, because we have not found any more values of $f(m,n)$. However, we have proved that for m even and ≥ 4, and n odd, $f(m,n)$ is <u>strictly</u> less than $\sqrt{\frac{4n-3}{4m}}$.

A Mystery

At the bottom of p. 41, it says that we omit the proof for the snug case with exactly 2 zeros in the same row of V. However, we do prove it, starting on p. 69. And, it's not shown on p. 48, toward the bottom of the page, where the other cases are shown.

Why I did this is a mystery to me.

On p. 60, on the 6th line from the bottom, V_{33} should be V_{32}.

On p. 71, near the left edge of the page, near the bottom, where it says $-A+|F|$ from GPT, it should be $-A+|F| \neq 0$ by GPT.

On p. 73, near bottom right, "col's 1+3" should be 2+4. Same with "columns 1+3" near bottom left.

On p. 36, just before $(B,C,D) \perp (L,M,P)$, it should say "Since $\vec{v}_1 \perp \vec{v}_3$,". Also, $(G,H,J) \pm (L,M,P)$ should be $\pm(G,H,J) \pm (L,M,P)$ and $\vec{v}_2 \pm \vec{v}_3$ should be $\pm\vec{v}_2 \pm \vec{v}_3$, in all 3 places.

On p. 61, second line, "square" should be "cube."

Appendix F
Other

Lower bound for f(m, m+1)

On p. 15 there is a formula for a lower bound for $f(m, m+1)$. Unfortunately, I no longer have my derivation of this formula, and I have not succeeded in re-deriving it.

I know that I used the idea in the middle of p. 9, showing that $f(3,4) > 1$, and assuming that it generalizes to higher dimensions, but I just haven't been able to get a closed-form formula like the one on p. 15.

I also know that I used the formula $\dfrac{v+w}{1+\frac{vw}{c^2}}$, with $c=1$. This is the velocity addition formula from special relativity. If $v < c$ and $w < c$, then $\dfrac{v+w}{1+\frac{vw}{c^2}}$ is larger than v, larger than w, and less than c. (Assuming $v, w > 0$.) It's a nice substitute for the "max" function.