MATRIX GENERATION OF THE DIOPHANTINE SOLUTIONS TO SUMS OF $3 \leq n \leq 9$ SQUARES THAT ARE SQUARE (PREPRINT)

JORDAN O. TIRRELL
Student, Lafayette College
Easton, PA 18042 USA
e-mail: tirrellj@lafayette.edu

CLIFFORD A. REITER
Department of Mathematics, Lafayette College
Easton, PA 18042 USA
e-mail: reiterc@lafayette.edu

Abstract

Pythagorean Triples are well known examples of integer solutions to sums of two squares giving another square. It is well known that Pythagorean Triples may be generated parametrically. It is somewhat less well known that they may also be generated via matrices. In this note we describe how matrix generators may be used to produce all the Diophantine solutions of a square being a sum of squares when the number of squares in the sum is between 3 and 9. For $3 \leq n \leq 8$ all the Diophantine solutions may be obtained via matrix multiplication from a single type of initial solution. For $n = 9$ two different types of initial solutions are required.

1. Introduction

Diophantine equations requiring a sum of squares to be a square correspond to integer solutions of equations of the form

$$x_1^2 + ... + x_n^2 = r^2. \tag{1}$$

When $n = 2$, the nontrivial, positive solutions correspond to Pythagorean Triples. Generating Pythagorean Triples parametrically is a standard topic in elementary number theory. Generation of Pythagorean Triples via matrix generators has been known for some time, but recently have been popularized [2]. When $n = 3$ or $n = 4$ the parametric generation of

2000 Mathematics Subject Classification: 11D09, 20H99
Keywords: Sums of Squares, Barning Tree.
This work was supported in part by a Lafayette College EXCEL grant.

the Diophantine solutions to Equation (1) is also known [3,4]. Recently, matrix generators for the case $n = 3$ were determined [3]. Our interest in Diophantine solutions to sums of squares being square resulted from our search for Perfect Parallelepipeds [3]. A Perfect Parallelepiped is a relaxed version of a Perfect Cuboid. More specifically, a Perfect Parallelepiped is a parallelepiped in $\Re^3$ with all three edges, all six face diagonals and all four body diagonals having an integer length [1]. Whether such a parallelepiped exists is an open question. While the edge coordinates need not be integer, our searches for Perfect Parallelepipeds in $\Re^3$ and $\Re^4$ used integer coordinates. Hence our interest in Diophantine solutions to sums of squares that are square.

It is convenient to express integer solutions to $x_1^2 + ... + x_n^2 = r^2$ as vectors since we will be generating solutions via matrix arithmetic. Diophantine solutions of Equation (1) correspond to integer vectors $\vec{u} =< x_1, ..., x_n >$ where $||\vec{u}|| = |r|$ is an integer. Thus, solutions to Equation (1) arise from integer length integer vectors. Two-dimensional integer length positive integer vectors correspond to Pythagorean Triples. These triples can all be generated from $< 1, 0, 1 >$ by coordinate interchanges, sign changes, and multiplication by the matrix

$$B = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 3 \end{pmatrix}.$$

That gives rise to the algebraic structure on Pythagorean Triples known as the Barning Tree [2]. In [3] it is seen that

$$J_3 = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 2 \end{pmatrix}$$

gives rise to the 3-dimensional integer length integer vectors in a similar manner, although negative entries are both allowed and utilized there. The main result in this note shows that matrices related to J_3 generalize these results for $3 \leq n \leq 9$. While the $n = 3$ case is not new, the argument used here is simpler; in particular, it does not require the case analysis used in [2,3].

2. Notation and Examples

It will be convenient for our purposes to express Diophantine solutions to Equation (1) in the form $\vec{u}^{+} =< x_n, ..., x_1, r >$ which we call extended coordinates. Extended coordinate vectors are demarked with a superscript "plus" on their name. Since $\vec{u}^{+}$ gives a solution to Equation (1), the length of the vector $\vec{u} =< x_n, ..., x_1 >$ is $|r|$. The coordinates $x_n, ...x_1$ are called the ordinary coordinates and we view $\vec{u}^{+}$ as being in Z^n even though we formally list $n+1$ coordinates. In particular, when describing the length of an extended coordinate vector, we mean the length of its ordinary coordinates, so we can write $||\vec{u}^{+}|| = |r|$. Notice that we have listed the ordinary coordinates with indices in descending order. This will be convenient for the proofs of our theorems. We illustrate this terminology and notations as follows. The Pythagorean triple $\vec{u}^{+} =< 3, 4, 5 >$ is viewed as an integer length integer vector in Z^2 and we write $||\vec{u}^{+}|| = 5$. For another example, the integer length integer

vector $\vec{u} =< 1, 2, 2 >$ can be written in extended coordinates as $\vec{u}^{+} =< 1, 2, 2, 3 >$ or $\vec{u}^{+} =< 1, 2, 2, -3 >$.

Some illustrations producing familiar Pythagorean triples using B include the following.

$$B\begin{pmatrix}1\\0\\1\end{pmatrix} = \begin{pmatrix}3\\4\\5\end{pmatrix}$$

and

$$B\begin{pmatrix}3\\-4\\5\end{pmatrix} = \begin{pmatrix}5\\12\\13\end{pmatrix}.$$

When $n > 3$, we extend the J_3 matrix by leaving it in the lower right as a block, putting an $n-3$ by $n-3$ identity matrix, I_{n-3}, in the upper left and zeros elsewhere. That is, J_n is the partitioned matrix

$$J_n = \begin{pmatrix} I_{n-3} & 0 \\ 0 & J_3 \end{pmatrix}$$

In the following examples we see that when we multiply an integer length integer vector by J_n, the result is an integer length integer vector. That is, we see that multiplication by these matrices of an extended coordinate vector giving a solution to Equation (1) yields a solution to Equation (1).

$$J_3\begin{pmatrix}1\\0\\0\\1\end{pmatrix} = \begin{pmatrix}1\\2\\2\\3\end{pmatrix}, \quad J_4\begin{pmatrix}1\\1\\1\\1\\2\end{pmatrix} = \begin{pmatrix}1\\4\\4\\4\\7\end{pmatrix}, \quad J_5\begin{pmatrix}2\\2\\2\\2\\3\\5\end{pmatrix} = \begin{pmatrix}2\\2\\10\\10\\9\\17\end{pmatrix}$$

3. Matrix Generators

In this section we show that for each $3 \leq n \leq 9$, multiplication by J_n, together with ordinary coordinate interchanges and sign changes, generates all of the solutions to Equation (1) in these dimensions. We begin by establishing some useful properties.

Proposition 1. *Let $n \geq 3$ be an integer and let $\vec{u}^{+} =< x_n, ..., x_1, r >$ give a Diophantine solution to Equation (1). Then the following hold.*

(a) For all choices of plus or minus sign, $< \pm x_n, ..., \pm x_1, \pm r >$ also give solutions to Equation (1).

(b) Any permutation of ordinary coordinates of $\vec{u}^{+}$ gives a solution to Equation (1).

(c) $J_n\vec{u}^{+}$ also gives a solution to Equation (1).

Proof. Part (a) follows since the square of a number is the same as the square of its opposite. Part (b) follows since the order of terms does not affect the sum of the terms.

For part (c), direct computation yields that $J_n\vec{u}^+ =< x_n, ..., x_4, r+x_1+x_2, r+x_1+x_3, r+x_2+x_3, 2r+x_1+x_2+x_3 >$. We can directly check that $||J_n\vec{u}^+||^2 = x_n^2 + ... + x_4^2 + (r+x_1+x_2)^2 + (r+x_1+x_3)^2 + (r+x_2+x_3)^2 = (2r+x_1+x_2+x_3)^2$ where we need to use the fact that $x_1^2 + ... + x_n^2 = r^2$ since we have assumed Equation (1) holds.

Proposition 2. *(a)* $J_3^{-1} = \begin{pmatrix} 0 & 1 & 1 & -1 \\ 1 & 0 & 1 & -1 \\ 1 & 1 & 0 & -1 \\ -1 & -1 & -1 & 2 \end{pmatrix}$

(b) $J_3^{-1} = SJ_3S$ *where* $S = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

(c) $J_n^{-1} = \begin{pmatrix} I_{n-3} & 0 \\ 0 & J_3^{-1} \end{pmatrix} = \begin{pmatrix} I_{n-3} & 0 \\ 0 & S \end{pmatrix} \begin{pmatrix} I_{n-3} & 0 \\ 0 & J_3^{-1} \end{pmatrix} \begin{pmatrix} I_{n-3} & 0 \\ 0 & S \end{pmatrix}$

Proof. Parts (a) and (b) may be verified by direct computation. Part (c) follows from (a) and (b) using partitioned matrices.

As is usual, an integer vector is primitive if the greatest common divisor of its coordinates is 1.

Proposition 3. *Suppose* $\vec{u}^+$ *and* $\vec{v}^+$ *are solutions to Equation (1) and that* $\vec{u}^+$ *may be obtained from* $\vec{v}^+$ *by a sequence of coordinate sign changes, rearrangements of ordinary coordinates, and multiplications by* J_n*. Then the process is reversible in the sense that* $\vec{v}^+$ *may be obtained from* $\vec{u}^+$ *by a sequence of coordinate sign changes, rearrangements of ordinary coordinates, and multiplications by* J_n*. Moreover, the greatest common divisor of the coordinates of* $\vec{u}^+$ *and* $\vec{v}^+$ *will be the same. In particular,* $\vec{u}^+$ *is primitive if and only if* $\vec{v}^+$ *is primitive.*

Proof. First notice that coordinate sign changes and ordinary coordinate interchanges are reversible processes. In light of Proposition 2c, the reverse of multiplication by J_n is sign changes followed by multiplication by J_n followed by sign changes. The reversibility result follows. Next we remark that since $\vec{u}^+$ gives a solution to Equation (1), the greatest common divisor of the ordinary coordinates is the same as the greatest common divisor of the extended coordinates. Also, sign changes and ordinary coordinate rearrangements preserve the greatest common divisor. Since the coordinates of $J_n\vec{u}^+$ are an integer linear combination of the coordinates of $\vec{u}^+$, it is clear that any common divisor of $\vec{u}^+$ is a common divisor of $J_n\vec{u}^+$. In light of Proposition 2, we can multiply J_n^{-1} to see the converse is true.

We are now prepared to prove our main theorem for dimensions $3 \leq n \leq 8$.

Theorem 4. *Any primitive n-dimensional integer length integer vector in extended coordinates, where* $3 \leq n \leq 8$*, can be generated by starting with the trivial vector in extended coordinates*

$$< 0, ..., 0, 1, 1 >$$

and performing a sequence of ordinary coordinate interchanges, sign changes, and multiplications by J_n.

Proof. We will begin by assuming that we have a primitive Diophantine solution to Equation (1), $\vec{u}^{+} =< x_n, ..., x_1, r >$, with $3 \leq n \leq 8$ and that the solution is nontrivial in the sense that at least two ordinary coordinates are nonzero. We will show that we can construct a shorter vector that also gives a solution to Equation (1). Since a sequence of such length reductions must be finite, we see we will eventually construct a vector with only one nonzero ordinary coordinate, which can be changed to $< 0, ..., 0, 1, 1 >$ via sign changes and ordinary coordinate interchanges. Given the reversibility in Proposition 3, that will complete our proof.

First, we note that by making sign changes and rearranging ordinary coordinates, we may obtain a vector of the same length with ordered coordinates. Namely, we may assume that $\vec{u}^{+} =< x_n, ..., x_1, r >$ satisfies $0 \leq x_n \leq ... \leq x_4 \leq x_3 \leq x_2 \leq x_1 < r$. The nontriviality condition means that $x_1 > 0$ and $x_2 > 0$. Now let $\vec{v}^{+} =< x_n, ..., x_4, -x_3, -x_2, -x_1, r >$ be the solution with the last three ordinary coordinates having changed signs. We claim that $J_n\vec{v}^{+}$ is a solution with a strictly smaller length. By direct computation we see that

$$J_n\vec{v}^{+} =< x_n, ..., x_4, r - x_1 - x_2, r - x_1 - x_3, r - x_2 - x_3, 2r - x_1 - x_2 - x_3 >$$

which yields a vector with length $|2r - x_1 - x_2 - x_3|$. It suffices to show that this is less than r. If $2r - x_1 - x_2 - x_3 < 0$, then showing the length is less than r amounts to showing $x_1 + x_2 + x_3 < 3r$ which is true since for all k, $x_k < r$. If $2r - x_1 - x_2 - x_3 > 0$, then showing the length is less than r amounts to showing $r < x_1 + x_2 + x_3$. We know that $x_4^2 + ... + x_n^2 < 2x_1x_2 + 2x_2x_3 + 2x_1x_3$ because the left side has at most five terms (since $n \leq 8$), each of which is less than or equal to each of the six terms of the form x_ix_j on the right, due to the ordering of the coordinates and where we are certain the inequality is strict since there is an extra sixth term $x_1x_2 > 0$ on the right hand side. Therefore,

$$r^2 = x_1^2 + x_2^2 + x_3^2 + x_4^2 + ... + x_n^2 < x_1^2 + x_2^2 + x_3^2 + 2x_1x_2 + 2x_2x_3 + 2x_1x_3 = (x_1 + x_2 + x_3)^2$$

and hence

$$r < x_1 + x_2 + x_3$$

as required to complete the proof.

Notice that the theorem is constructive. That is, process in the proof can be repeatedly applied to the obtain a sequence of vectors of decreasing lengths ending in the trivial vector. If we let $\vec{u}^{+} \sim \vec{v}^{+}$ denote that the vectors are the same up to changing signs and rearranging ordinary coordinates. Then consider the following example. The vectors $< 22, 38, 40, 89, 107 > \sim < 22, -38, -40, -89, 107 >$ give solutions to Equation (1). Then we see

$J_4 < 22, -38, -40, -89, 107 > = < 22, -22, -20, -29, 47 > \sim < 20, 22, 22, 29, 47 >$
$J_4 < 20, -22, -22, -29, 47 > = < 20, -4, -4, 3, 21 > \sim < 3, 4, 4, 20, 21 >$
$J_4 < 3, -4, -4, -20, 21 > = < 3, -3, -3, 13, 14 > \sim < 3, 3, 3, 13, 14 >$

$J_4 < 3,-3,-3,-13,14 >=< 3,-2,-2,8,9 >\sim< 2,2,3,8,9 >$
$J_4 < 2,-2,-3,-8,9 >=< 2,-2,-1,4,5 >\sim< 1,2,2,4,5 >$
$J_4 < 1,-2,-2,-4,5 >=< 1,-1,-1,1,2 >\sim< 1,1,1,1,2 >$
$J_5 < 1,-1,-1,-1,2 >=< 1,0,0,0,1 >\sim< 0,0,0,1,1 >$ as expected.

Notice that rearranging ordinary coordinates can be accomplished by multiplying by a permutation matrix that leaves the last coordinate fixed and changing signs can be accomplished by multiplying by a diagonal matrix with -1 at the coordinates to be changed and 1 at the other diagonal positions. Those observations along with the properties of J_n, that we have previously shown, prove the following.

Corollary 5. *For dimensions $3 \leq n \leq 8$ there is a finite collection, S, of matrices consisting of J_n, the permutation of ordinary coordinate matrices and the sign change matrices such that every primitive solution to Equation (1) may be produced from $< 0,...0,1,1 >$ by a sequence of multiplications from matrices in S.*

Corollary 6. *Let $3 \leq n \leq 8$ and S be as in Corollary 5. Then every Diophantine solution to Equation (1) may be produced from $< 0,...0,s,s >$ by a sequence of multiplications from matrices in S. For nonzero solutions, the number s is the greatest common divisor of the coordinates.*

Proof. We have observed that the greatest common divisor is preserved by multiplication by J_n. The same is true for ordinary coordinate interchanges and sign changes. Thus, the Corollary follows from recognizing the zero solution and dividing nonzero solutions by their greatest common divisor, s, applying Corollary 5 to the resulting primitive solutions and then multiplying by s.

Lastly we turn to the case when $n = 9$ where we begin to see the complications that occur for higher dimensions.

Theorem 7. *Any primitive 9-dimensional integer length integer vector in extended coordinates can be generated by starting with either the vector $< 0,0,0,0,0,0,0,0,1,1 >$, or $< 1,1,1,1,1,1,1,1,1,3 >$, and performing a sequence of ordinary coordinate interchanges, sign changes, and multiplications by J_9.*

Proof. The outline of the proof is the same as for Theorem 4. We will begin by assuming that we have a primitive Diophantine solution to Equation (1) with at least two nonzero ordinary coordinates; namely $\vec{u}^{+} =< x_9,...,x_1,r >$ with $x_9 \leq x_8 \leq ...x_2 \leq x_1 < r$ where the last inequality is strict since we have assumed there at least two nonzero coordinates. We claim that unless $\vec{u}^{+}$ is the particular vector $\vec{w}^{+} =< 1,1,1,1,1,1,1,1,1,3 >$ that we will be able to produce a shorter solution. Thus, we will be able to obtain a shorter solution unless we encounter $\vec{w}^{+}$ or there is only one nonzero ordinary coordinate, and in that case we encounter $< 0,0,0,0,0,0,0,0,1,1 >$. As before, proving this claim will suffice for the proof of the theorem due to the reversibility of the allowed operations.

Like before, we consider $J_9\vec{v}^{+}$ where $\vec{v}^{+} =< x_9,...,x_4,-x_3,-x_2,-x_1,r >$. The resulting length is $|2r-x_1-x_2-x_3|$. It suffices to show that this is less than r. If $2r-x_1-x_2-x_3 < 0$,

then showing the length is less than r amounts to showing $x_1 + x_2 + x_3 < 3r$ which again is true since for all k, $x_k < r$. If $2r - x_1 - x_2 - x_3 > 0$, then showing the length is less than r amounts to showing $r < x_1 + x_2 + x_3$. We know that

$$x_4^2 + ... + x_9^2 \leq 2x_1x_2 + 2x_2x_3 + 2x_1x_3 \tag{2}$$

because the left side has six terms, each of which is less than or equal to each of the six terms of the form x_ix_j on the right. However, by the ordering of the terms, we know $x_k \leq x_2$ for $k = 9, 8, ..., 4$. If any of those was a strict inequality, then $x_k^2 < x_1x_2$ and Equation (2) would also be a strict inequality. Otherwise, $x_9 = x_8 = ...x_4 = x_3 = x_2$. If $x_2 < x_1$ then $x_k^2 < x_1x_2$ and Equation (2) would also be a strict inequality. Thus, either Equation (2) is strict or $x_1 = x_2 = ... = x_9 = 1$ (since the solution is primitive). Thus, either the vector is $\vec{w}^+$ or $x_4^2 + ... + x_9^2 < 2x_1x_2 + 2x_2x_3 + 2x_1x_3$ and in that case we can produce a shorter solution in the same way as in Theorem 4. This completes the proof.

There are analogs for Corollaries 5 and 6 for the case when $n = 9$ so that nonprimitive solutions can be handled and, of course, there are two possible initial vectors that are used with a finite collection of matrices in this case.

We remark that $\vec{w}^+$ is stable under the reduction process of "change the last three signs" and then "multiply by J_9". However, an entire new family of solutions is produced by this vector when other combinations of ordering and sign changes are allowed. In particular, we observe the following.

$$J_9 < 1, 1, 1, 1, 1, 1, -1, -1, -1, 3 >=< 1, 1, 1, 1, 1, 1, 1, 1, 1, 3 >$$

$$J_9 < 1, 1, 1, 1, 1, 1, 1, 1, 1, 3 >=< 1, 1, 1, 1, 1, 1, 5, 5, 5, 9 >$$

It is straightforward to check that multiplication of an extended coordinate vector with all odd coordinates by J_9 gives rise to an extended coordinate vector with all odd coordinates, and likewise for coordinate interchanges and sign changes. In light of the reversiblity of these operations, we see that the family of primitive solutions arising from the two initial vectors $< 0, 0, 0, 0, 0, 0, 0, 0, 1, 1 >$ and $< 1, 1, 1, 1, 1, 1, 1, 1, 1, 3 >$ can be identified at a glance: the former have coordinates with mixed parity while the later have all coordinates with odd parity.

In higher dimensions the situation becomes more complicated. For example, when $n = 10$ we have extensions of the two previously seen initial vectors $< 0, 0, 0, 0, 0, 0, 0, 0, 1, 1 >$ and $< 0, 1, 1, 1, 1, 1, 1, 1, 1, 3 >$ but also see $< 7, 8, 8, 8, 8, 8, 8, 8, 8, 25 >$ which does not appear to be in the same class as the other initial conditions. In fact, with respect to the reduction process of ordering, changing the signs of the last three ordinary coordinates, and then multiplying by J_{10}, that vector leads to a two cycle where the other vector is $< 7, 8, 8, 8, 8, 8, 8, 9, 9, 9, 26 >$.

References

[1] R. Guy, Unsolved Problems in Number Theory, 2nd ed. Springer-Verlag, 1994.

[2] D. McCullough, Height and excess of Pythagorean triples, Mathematics Magazine, 78 (2005) 26-44.

[3] C. Reiter and J. Tirrell, Pursuing the perfect parallelepiped, JP Journal of Algebra, Number Theory, and Applications, 6 2 (2006) 279-294. Auxiliary Materials, http://www.lafayette.edu/∼reiterc/nt/ppllpd/index.html

[4] W. Sierpinski, Elementary Theory of Numbers, trans. by A. Hulanicki, Panstwowe Wydawnictwo Naukowe, 1964.