UNIVERSAL FORMS: THE FOUR–SQUARE THEOREM AND ITS GENERALIZATIONS

Mark B. Beintema and Azar N. Khosravani

1. Introduction. The problem of writing a positive integer as a sum of squares has been a source of fascination for centuries, attracting the attention of some of the finest mathematical minds throughout history. One well–known result along these lines is the "Four–Square Theorem."

Theorem 1. If n is a positive integer, then there exist integers x, y, z, and w such that $n = x^2 + y^2 + z^2 + w^2$.

As many readers are no doubt aware, the first correct proof of the theorem is due to Lagrange. However, the history of this theorem, and its impact on the development of modern number theory, are less well–known. In Section 1 of this paper we give a historical account of the Four–Square Theorem and present a proof due to Euler. In Section 2, we present a generalization due to Ramanujan. For references we list the most readily accessible, most of which can be found in any college library. For biographical information on the mathematicians involved, we recommend [13 or 14]. Finally, although we attempt to follow the thread of reasoning of the original work, where convenient we use modern terminology.

2. The Four–Square Theorem. The first explicit statement of the Four–Square Theorem was given in 1621 by Claude Bachet, who in that year published a Latin translation of Diophantus' *Arithmetic*. In his commentary, Bachet stated that he had verified the theorem for all numbers up to 325, and remarked that Diophantus himself seemed to have known the theorem. The reason for doing so was that, although Diophantus gives necessary conditions for an integer to be the sum of two squares, or of three squares, no precondition is given for an integer to be the sum of four squares. Among the mathematicians to read Bachet's text was Fermat, and it was in the margins of this book that he wrote many of his most famous results and conjectures (including his fabled "last theorem"). In the 1630's, Fermat began a long and fruitful correspondence with Mersenne in which he formulated many results and conjectures concerning the represenation of integers by quadratic forms [13]. In a letter from 1636 [7], Fermat stated that he had a proof of the theorem, presumably by his method of descent. Like Bachet, Fermat also credited knowledge of the theorem to Diophantus.

In the same letter containing the statement of the Four–Square Theorem, Fermat states the following result concerning triangular numbers. An integer n is triangular if there exists a positive integer a such that $n = a(a+1)/2$.

<u>Theorem 2</u>. Every positive integer is the sum of three triangular numbers.

The theorem on triangular numbers was rediscovered independently by Gauss, resulting in one of the most famous entries of his celebrated "mathematical diary"

$$\text{EYPHKA!} \quad \text{Num} = \Delta + \Delta + \Delta.$$

Although Fermat never published proofs of the Four–Square Theorem or the Triangular Number Theorem, his remarks inspired Euler, who subsequently worked on the problem for nearly forty years. Learning of Fermat's work through correspondence with Goldbach, Euler set out to prove many of the former's results and conjectures. Over the years 1730 to 1770, he proved a number of results related to the Four–Square Theorem. In 1730, Euler stated that the Four–Square Theorem would follow from Fermat's Theorem on triangular numbers. The proof that follows is essentially due to Gauss [8].

<u>Proof</u>. We begin with the observation (also due to Euler) that every integer of the form $8k+3$ is the sum of three squares. Assuming Theorem 2, we write

$$k = \sum_{i=1}^{3} c_i(c_i+1)/2 = \frac{1}{2}\sum_{i=1}^{3} c_i^2 + c_i$$

hence,

$$8k+3 = \sum_{i=1}^{3} (2c_i+1)^2.$$

Adding 1 to each side, we see that $8k+4$ is a sum of four odd squares for any k, say $8k+4 = \sum x_i^2$. Now we have

$$4k+2 = \frac{1}{2}x_i^2 = \left(\frac{x_1+x_2}{2}\right)^2 + \left(\frac{x_1-x_2}{2}\right)^2 + \left(\frac{x_3+x_4}{2}\right)^2 + \left(\frac{x_3-x_4}{2}\right)^2.$$

Rewriting as $4k+2 = \sum z_i^2$, and considering congruences modulo 4, we see that two of the summands are even and two are odd, say z_1, z_2 are even and z_3, z_4 are odd. Then we have

$$2k+1 = \frac{1}{2}z_i^2 = \left(\frac{z_1+z_2}{2}\right)^2 + \left(\frac{z_1-z_2}{2}\right)^2 + \left(\frac{z_3+z_4}{2}\right)^2 + \left(\frac{z_3-z_4}{2}\right)^2.$$

It follows that every odd integer is the sum of four squares. Now, if $n = \sum x_i^2$, then easily $4n = \sum(2x_i)^2$; thus, if n is the sum of four squares, so is $4n$. The result now follows by noting that every even integer can be written either in the form $4^a(2k+1)$ or $4^a(4k+2)$.

Euler continued working on the Four–Square Theorem over the ensuing decades, trying a number of approaches [2]. The following identity, established by Euler in 1748, proved to be essential to the eventual proof of the theorem.

Lemma 3. $(a^2+b^2+c^2+d^2)(\alpha^2+\beta^2+\gamma^2+\delta^2) = x^2+y^2+z^2+w^2$, where

$$x = a\alpha + b\beta + c\gamma + d\delta \qquad y = a\beta - b\alpha - c\gamma + d\delta$$

$$z = a\gamma + b\delta - c\alpha - d\beta \qquad w = a\delta - b\gamma + c\beta - d\alpha.$$

This lemma shows that if two integers can each be written as the sum of four squares, then their product can be written as a sum of four squares as well. The proof of the theorem is thereby reduced to proving that every prime number can be written as the sum of four squares. In 1749, Euler came one step closer to a proof.

Theorem 4. If N is prime, then there exist four squares, none of which is divisible by N, whose sum is divisible by N.

Proof. To prove this theorem, he first establishes some facts about what are now called *quadratic residues*. Recall that an integer x is a quadratic residue modulo N if, for some integer a, $x \equiv a^2 \pmod{N}$. Euler quickly proves the following: (1) the product of two quadratic residues is again a quadratic residue, (2) the product of a residue and non–residue is a non-residue, and (3) the product of two non-residues is a quadratic residue.

Assume that it is not possible to find three squares, not all divisible by N, whose sum is divisible by N. Then for any x, y relatively prime to N, we know that N does not divide the sum x^2+y^2. It follows that if a is a quadratic residue modulo N, then $-a$ is a non–residue; in particular, -1 is a non–residue. Let a be any nonzero quadratic residue modulo N. Then there exists x, y so that $x^2+y^2 \equiv a+1$

(mod N). Now if $-(a+1)$ was a quadratic residue, there would be three squares whose sum was divisible by N; by assumption, this is not the case. Thus, $-(a+1)$ is a non–residue. Recalling that -1 is a non–residue, and invoking (3) above, we see that $a+1$ is a quadratic residue whenever a is. Letting $a=1$, we deduce that all integers are quadratic residues modulo N, an obvious contradiction.

The Four–Square Theorem would now follow if it could be shown that every divisor of a sum of four squares is itself a sum of four squares. Despite extended efforts, Euler was unable to complete this part of the proof.

Joseph Louis Lagrange corresponded extensively with Euler on a number of mathematical subjects. He succeeded in proving many of Fermat's conjectures concerning binary quadratic forms–indeed, at this time Lagrange was developing a very general theory on such forms. In 1770, building upon Euler's work and using results from the theory of binary forms, he succeeded in proving that all primes can be written as a sum of four squares, thus, obtaining the first correct proof of the Four–Square Theorem. Ironically, just two years later, Euler succeeded in proving the final step that was missing from his earlier efforts [6]. We present Euler's simpler proof below.

Theorem 5. If N is a divisor of the sum of four squares, no one of which is divisible by N, then N is the sum of four squares.

Proof. The strategy is a classic "descent" argument. Show that if N divides a sum of four squares, then there is a strictly smaller sum of four squares for which N is also a divisor. By repeated applications of the argument, one finally deduces that N itself is a sum of four squares. Begin by writing $Nn = p^2+q^2+r^2+s^2$, where

$$p = a+n\alpha, \quad q = b+n\beta, \quad r = c+n\gamma, \quad s = d+n\delta. \tag{1}$$

We can assume that each of a, b, c, and d has absolute value less than or equal to $n/2$.

Substituting the values (1), we have

$$Nn = a^2+b^2+c^2+d^2+2n(a\alpha+b\beta+c\gamma+d\delta)+n^2(\alpha^2+\beta^2+\gamma^2+\delta^2). \tag{2}$$

Immediately we see that n divides $a^2+b^2+c^2+d^2$, say

$$nn' = a^2+b^2+c^2+d^2. \tag{3}$$

Dividing both sides of (2) by n, we get $N = n'+2x+n(\alpha^2+\beta^2+\gamma^2+\delta^2)$, where $x = a\alpha+b\beta+c\gamma+d\delta$. Thus,

$$Nn' = (n')^2+2xn'+nn'(\alpha^2+\beta^2+\gamma^2+\delta^2). \tag{4}$$

By the lemma, and (3) above, we have

$$\begin{aligned} nn'(\alpha^2+\beta^2+\gamma^2+\delta^2) &= (a^2+b^2+c^2+d^2)(\alpha^2+\beta^2+\gamma+\delta^2 \\ &= x^2+y^2+z^2+w^2, \quad \text{with} \quad x \quad \text{as above.} \end{aligned}$$

Thus, we rewrite (4) to get

$$Nn' = (n')^2 + 2xn' + x^2 + y^2 + z^2 + w^2 = (n'+x)^2 + y^2 + z^2 + w^2.$$

Our assumption on the absolute values of a, b, c, and d implies $a^2+b^2+c^2+d^2 < n^2$, hence, $n' < n$. Repeating the argument, one obtains a decreasing sequence of integers $Nn', Nn'', \ldots$ each a sum of four squares. It follows that N is a sum of four squares.

3. Universal Forms. Efforts to prove theorems concerning sums of squares gave rise to the modern theory of arithmetic quadratic forms. It will be useful to introduce some terminology. A quadratic form is a homogeneous polynomial of degree two with integer coefficients; i.e. an expression of the form

$$q(\mathbf{x}) = \sum a_{ij} x_i x_j, \quad a_{ij} \in \mathbb{Z}.$$

A form is called binary (respectively ternary, quaternary) if it involves two (respectively three, four) variables. We say that a quadratic form q represents an integer n if the equation $q(\mathbf{x}) = n$ has a solution with each x an integer. A form which represents all positive integers is called *universal* in the current literature. Thus, an alternate statement to the Four–Square Theorem is that the form $x^2 + y^2+z^2+w^2$ is universal. A natural question arises: are there other universal forms? To answer this question requires representation theorems for ternary forms. Fermat and Euler each had found some elementary results concerning which integers could be represented as a sum of three squares [2, 13]. The definitive theorem is due to Legendre, who proved the following in 1798.

Theorem 6. A positive integer can be written as the sum of three squares if and only if it is not of the form $4^a(8k+7)$.

In the Disquisitiones Arithmetica (art. 293), Gauss used this theorem to prove Fermat's theorem on triangular numbers, obtaining the Four–Square Theorem as a corollary (in an argument very similar to that presented previously). In the 19th century, Liouville, Pepin, and Jacobi found a number of forms of the type $ax^2+by^2+cz^2+dw^2$ which represent all positive integers — that is, universal quaternary forms. However, their interest was more in the enumeration of representations than in the discovery and classification of univeral forms [2]. It was Ramanujan who undertook a complete classification of such forms. In 1916, he wrote what was to become a very influential paper in the theory of quadratic forms [13]. Using Legendre's Theorem (which he incorrectly attributes to Cauchy) and similar results, Ramanujan gave the following generalization of the Four–Square Theorem.

Theorem 7. There are 54 forms $q = ax^2 + by^2 + cz^2 + dw^2$ which represent all positive integers.

(Initially Ramanujan claimed to have found 55 such forms. It was later pointed out by Dickson [3] that one of the forms fails to be universal.)

Sketch of Proof. Without loss of generality, assume $a \leq b \leq c \leq d$. Now $a = 1$ or else 1 is not represented. Similarly, $1 \leq b \leq 2$ or else 2 is not represented. Now, by cases:

$$x^2 + y^2 + cz^2 + dw^2 \ : \ \text{Then } 1 \leq c \leq 3, \ \text{since if } c > 3, \ 3 \text{ is not represented.}$$

$$x^2 + 2y^2 + cz^2 + dw^2 \ : \ \text{Then } 2 \leq c \leq 5, \ \text{since if } c > 5, \ 5 \text{ is not represented.}$$

Continuing this way, Ramanujan arrives at the following allowable values of (a, b, c, d).

(1,1,1,1)	(1,2,3,5)	(1,2,4,8)*	(1,1,1,2)*	(1,2,4,5)
(1,2,5,8)	(1,1,2,2)	(1,2,5,5)**	(1,1,2,9)	(1,2,2,2)*
(1,1,1,6)	(1,2,3,9)	(1,1,1,3)	(1,1,2,6,)	(1,2,4,9)
(1,1,2,3)	(1,2,2,6)	(1,2,5,9)	(1,2,2,3)	(1,1,3,6)
(1,1,2,10)	(1,1,3,3)*	(1,2,3,6)*	(1,2,3,10)	(1,2,3,3)
(1,2,4,6)	(1,2,4,10)	(1,1,1,4)*	(1,2,5,6)	(1,2,5,10)*
(1,1,2,4)*	(1,1,1,7)	(1,1,2,11)	(1,2,2,4)*	(1,1,2,7)
(1,2,4,11)	(1,1,3,4)	(1,2,2,7)	(1,1,2,12)	(1,2,3,4)
(1,2,3,7)	(1,2,4,12)	(1,2,4,4)*	(1,2,4,7)	(1,1,2,13)
(1,1,1,5)	(1,2,5,7)	(1,2,4,13)	(1,1,2,5)	(1,1,2,8)*
(1,1,2,14)	(1,2,2,5)	(1,2,3,8)	(1,2,4,14)	(1,1,3,5)

(The forms marked * had previously been discovered by Pepin, Liouville, and Jacobi. Dickson showed that the form marked ** does not represent $n = 15$.

Having obtained the above list of candidates, Ramanujan proceeds to prove that each of the forms is universal. First, he shows that for $1 \leq d \leq 7$, the form $x^2 + y^2 + z^2 + dw^2$ is universal. This is accomplished by showing that for each d, there is a suitable w such that $N - dw^2$ is not of the type excluded by Legendre's Theorem, and hence, is a sum of three squares. Assume N can be written as $4^a(8k+7)$; otherwise the result follows from Legendre's Theorem by taking $w = 0$. Setting $w = 2^a$, we get $N - dw^2 = 4^a(8k + 7 - d)$. This is a sum of three squares for $d = 1, 2, 4, 5$ or 6, as well as the cases $d = 3$, $k = 0$, and $d = 7$, $0 \leq k \leq 2$. For the cases $d = 3$, $k > 0$ and $d = 7$, $k > 2$ we take $w = 2^{a+1}$ to obtain $N - dw^2 = 4^a(8k - 7 - 4d)$, which again is a sum of three squares.

Ramanujan handles the remaining forms by similar arguments, using results analogous to Legendre's Theorem, summarized in the following table.

Form	Integers Not Represented	Smallest Exception
$x^2 + y^2 + z^2$	$4^j(8k+7)$	7
$x^2 + y^2 + 2z^2$	$4^j(16k+14)$	14
$x^2 + y^2 + 3z^2$	$9^j(9k+6)$	6
$x^2 + 2y^2 + 2z^2$	$4^j(8k+7)$	7
$x^2 + 2y^2 + 3z^2$	$4^j(16k+10)$	10
$x^2 + 2y^2 + 4z^2$	$4^j(16k+14)$	14
$x^2 + 2y^2 + 5z^2$	$25^j(25k+10)$ or $25^j(25k+15)$	10

The first line is simply Legendre's theorem restated. The other results were later established by Dickson [3].

Ramanujan's paper sparked renewed interest in the representation theory of ternary forms. A particularly intriguing comment appears in a footnote in which Ramanujan discusses the form $x^2 + y^2 + 10z^2$.

> "The even numbers which are not of the form $x^2 + y^2 + 10z^2$ are the numbers $4^\lambda(16\mu + 6)$, while the odd numbers that are not of the form, $3, 7, 21, 31, 33, 43, 67, 79, 87, 133, 217, 219, 223, 253, 307, 391, \ldots,$ do not seem to obey any simple rule."

This form is often referred to in the current literature as simply the "Ramanujan Form". It has recently been proved that the set of integers not represented by this form is finite. Moreover, there is evidence to suggest that Ramanujan's list of exceptions is complete, save for two numbers: 679 and 2,719 [5, 11].

Our final remarks require a few more definitions. A form $q = \sum a_{ij}x_ix_j$ is diagonal if $a_{ij} = 0$ for $i \neq j$. It is called *classic* if a_{ij} is even for $i \neq j$. A form is called positive definite (or simply positive) if $q(\mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$. Finally, two quadratic forms are equivalent if one can be obtained from the other by a linear integral change of variable. It is easy to see that equivalent forms represent the same integers.

Following Ramanujan's landmark paper, attention turned to non–diagonal forms. Moreover, the emphasis was increasingly on classifying equivalence classes of forms. In 1948, "employing an extension of the method of Ramanujan," Margaret Willerding [16] showed that there are exactly 178 classes of universal classic positive quaternary quadratic forms. It is now well–known that there exist only finitely many universal quaternary quadratic forms; however, a complete classification of these forms (including forms with odd cross–terms) has yet to be found. The following recent result of Conway and Schneeberger [5], known as the "Fifteen Theorem," allows a computation approach to demonstrating the universality of classic forms.

Theorem 8. A positive classic form is univeral provided it represents each of the following integers: 1, 2, 3, 5, 6, 7, 10, 14, 15.

Amazingly, this list of integers is precisely the list of minimal exceptions in Ramanujan's case by case analysis.

References

1. L. E. Dickson, *Theory of Numbers Vol. II: Diophantine Analysis*, Chelsea, New York, 1952.

2. L. E. Dickson, *Theory of Numbers Vol. III: Quadratic and Higher Forms*, Chelsea, New York, 1952.

3. L. E. Dickson, "Integers Represented by Positive Ternary Quadratic Forms," *Bull. Amer. Math. Soc.*, 33 (1927), 63–70.

4. L. E. Dickson, "Quaternary Quadratic Forms Representing all Positive Integers," *Amer. J. of Math.*, 49 (1927), 39–56.

5. W. Duke, "Some Old Problems and New Results About Quadratic Forms," *Notes Amer. Math. Soc.*, 44 (1997), 190–196.

6. L. Euler, *Proof That Every Integer is a Sum of Four Squares, A Source Book in Mathematics*, (D. E. Smith, ed.), Dover, New York, 1959.

7. P. de Fermat, *Oeuvres Vol. II,* Gauthier–Villars and Sons, Paris, 1894.

8. C. F. Gauss, *Disquisitiones Arithmeticae*, Springer–Verlag, New York, 1986.

9. E. Grosswald, *Representations of Integers as Sums of Squares*, Springer-Verlag, New York, 1985.

10. J. L. Lagrange, "Demonstration d'un Theoreme d'Arithmetique," *Oeuvres Vol. III*, Gauthier–Villars, Paris, 1867.

11. K. Ono, "Ramanujan, Taxicabs, Birthdates, ZIP Codes and Twists," *Amer. Math. Monthly*, 104 (1997), 912–917.

12. S. Ramanujan, "On the Expression of a Number in the Form $ax^2 + by^2 + cz^2 + dw^2$," *Proc. Cambridge Phil. Soc.*, 19 (1917), 11–21.

13. W. Scharlau and H. Opolka, *From Fermat to Minkowski, Lectures on the Theory of Numbers and Its Historical Development*, Springer–Verlag, New York, 1985.

14. A. Weil, *Number Theory: An Approach Through History*, Birkhauser, Boston, 1983.

15. M. Willerding, "Determination of All Classes of Positive Quaternary Quadratic Forms Which Represent All (Positive) Integers," *Bull. Amer. Math. Soc.*, 54 (1948), 334–337.

Mark B. Beintema
Department of Mathematics
University of Wisconsin–Fox Valley
Menasha, WI 54952

Azar N. Khosravani
Department of Mathematics
University of Wisconsin–Oshkosh
Oshkosh, WI 54901
email: khosrava @uwosh.edu