A Rational Function Whose Integral Values Are Sums of Two Squares

Sam Vandervelde

March 23, 2013

Abstract

A problem from the 1988 IMO asserts that for positive integers a and b the set of integral values assumed by $(a^2+b^2)/(ab+1)$ is exactly the set of positive squares. We present an extension of this result involving a rational function in three variables whose integral values consist of precisely those numbers expressible as a sum of two positive squares. This immediately implies that a certain Diophantine equation has no solutions in positive integers, a result that we also prove directly. We conclude with an extension to four variables.

Keywords: Diophantine equation, Vieta jumping, Pell equation, 1988 IMO
MSC2010: 11D09

1 History

In July 1988 the twenty-ninth International Mathematical Olympiad was held in Canberra, Australia. Due to a relatively dismal performance on a qualifying test, this author, who had entertained hopes of participating in the event, remained home and thus passed up the chance to work on one of the most legendary IMO problems ever posed. The final question set for the exam read as follows.

1988 IMO, Problem 6. Let a and b be positive integers such that $ab+1$ divides a^2+b^2. Show that

$$\frac{a^2+b^2}{ab+1}$$

is the square of an integer.

According to Arthur Engel [1], the jury ultimately decided to include this question on the exam despite the fact that it had stymied every member of the problem committee. It has since gained the reputation of being one of the most difficult problems ever to appear on an IMO. Of the 268 contestants present that day only fourteen earned more than three of the available seven points for their attempts to solve it.

The problem combines the simplicity of statement, elegance of result, and surprising level of difficulty that lures mathematicians to the field of number theory. Naturally it has spawned a variety of similarly intriguing results. In [6] Robert Weber points out that when $a, b \in \mathbb{N}$ the set of integral values for

$$\frac{a^2 + b^2 + 1}{ab + 1} \tag{1}$$

consists of exactly those integers of the form $t^2 + 1$ for $t \in \mathbb{N}$, while the set of integral values for

$$\frac{a^2 + b^2}{ab - 1} \tag{2}$$

contains only the single number 5. As shown in [2], when $a, b, c \in \mathbb{N}$ the set of integral values for the expression

$$\frac{(a + b + c)^2}{abc} \tag{3}$$

is precisely $\{1, 2, 3, 4, 5, 6, 8, 9\}$. Furthermore, at the end of [3] the reader is invited to investigate integral values assumed by the rational function

$$\frac{a^2 + b^2 + c^2}{ab + ac + bc + 1}, \tag{4}$$

while the paper [7] presents a method for finding all integral values taken on by the rational function

$$\frac{a^2 + b^2 + c^2}{abc}. \tag{5}$$

Blended together, these last two examples provide us with the rational function to which we now turn our attention.

Theorem 1.1 *Let a, b and c be positive integers such that $abc + 1$ divides $a^2 + b^2 + c^2$. Then*

$$\frac{a^2 + b^2 + c^2}{abc + 1}$$

is the sum of two positive squares.

Contrary to what one might expect, the author stumbled across this result while casting about for triples of integers satisfying interesting divisibility properties, and only gradually realized that it was a generalization of the IMO problem he had missed the opportunity to grapple with over two decades previously. Nonetheless, the discovery of a proof held a certain redemptive quality. The purpose of this note is to present the proof.

2 Main Result

A computer search turns up a total of 112 triples (a, b, c) in the range $1 \le a \le b \le c \le 99$ for which $a^2 + b^2 + c^2$ is a multiple of $abc + 1$. Nearly one hundred of them belong to the infinite family $a = 1$, $b = t$, $c = t + 1$ for $t \in \mathbb{N}$. Indeed, one quickly verifies that

$$\frac{1^2 + t^2 + (t+1)^2}{(1)(t)(t+1)+1} = 2, \tag{6}$$

which is a sum of two positive squares, as claimed. However, fourteen more exotic triples also surface, including

$$\frac{1^2 + 4^2 + 68^2}{(1)(4)(68)+1} = 17, \quad \frac{2^2 + 3^2 + 78^2}{(2)(3)(78)+1} = 13, \quad \frac{5^2 + 6^2 + 59^2}{(5)(6)(59)+1} = 2. \tag{7}$$

Note that perfect squares do not appear as quotients except via Pythagorean triples, since quotients must be the sum of two positive squares. So no quotient is equal to 16, but we can obtain 25, for instance when $a = 3$, $b = 4$, $c = 300$.

We now commence with the proof, which employs a technique that has come to be known as "Vieta jumping," combined with a descent argument.

Proof: Given positive integers $a \le b \le c$ for which $abc+1$ divides $a^2+b^2+c^2$, let their quotient be k. Replacing c with the variable x, we see that the equations

$$\frac{a^2 + b^2 + x^2}{abx + 1} = k \qquad \Longleftrightarrow \qquad x^2 - (kab)x + (a^2 + b^2 - k) = 0 \tag{8}$$

have one solution $x = c$. Since the sum of the roots of the quadratic on the right is kab, the other root must be the integer $kab - c$. Therefore replacing c by $kab - c$ yields a new triple of integers that also satisfy the conditions of Theorem 1.1 and give the same quotient k as before. To utilize a descent we are interested in bounding the value of this second root, in order to show that the new triple is smaller than the original one in some respect.

We immediately deduce that $kab - c \ge 0$, since the expression on the left in (8) will be negative for $x \le -1$ (or possibly undefined when $a = b = 1$ and $x = -1$), and hence not equal to k.

We next show that the smaller root of the quadratic equation on the right in (8) is less than b. Since $b \le c$ this root cannot be c, so it must be $kab - c$, and it will follow that $kab - c < b$. According to the quadratic formula the smaller root is

$$\tfrac{1}{2}\left(kab - \sqrt{k^2a^2b^2 - 4a^2 - 4b^2 + 4k}\right). \tag{9}$$

This root is smaller than b precisely when

$$(ka - 2)b < \sqrt{k^2a^2b^2 - 4a^2 - 4b^2 + 4k}. \tag{10}$$

If $ka - 2 < 0$ then the inequality is clearly true. Otherwise both sides are nonnegative, so the following statements are equivalent:

$$\begin{aligned} kab - 2b &< \sqrt{k^2a^2b^2 - 4a^2 - 4b^2 + 4k}, \\ k^2a^2b^2 - 4ab^2k + 4b^2 &< k^2a^2b^2 - 4a^2 - 4b^2 + 4k, \\ a^2 - k &< b^2(ka - 2). \end{aligned} \tag{11}$$

When $ka - 2 > 0$ the right-hand side of the last line will be at least b^2, while the left-hand side will be less than a^2. But $a \leq b$, so the inequality holds in this case as well.

It only remains to consider $ka = 2$. When $a = 1$ and $k = 2$ the last line reduces to $-1 < 0$, which is true once again. The other possibility is $a = 2$ and $k = 1$, in which case equation (8) reduces to $x^2 - 2bx + b^2 + 3 = 0$, or $(x - b)^2 = -3$. But this equation has no real solutions, and we are assuming the existence of the solution $x = c$, so this case never actually arises.

In summary, we have shown that given positive integers $a \leq b \leq c$ for which $abc+1$ divides $a^2+b^2+c^2$ with quotient k, then the triple of integers a, b, $kab-c$ also satisfies this condition, having the same quotient k with $0 \leq kab - c < b$. As long as $kab-c > 0$ we obtain a new triple of positive integers whose maximal element is smaller than before. (Hypothetically if $b = c$ we would need to apply this step twice to reduce the maximal element, but in reality one can show that this situation does not occur.) Clearly we cannot repeat this Vieta jump indefinitely, so eventually we produce a triple $A \leq B \leq C$ with $kAB - C = 0$. Once we reach this stage we obtain the triple 0, A, B at the next step, having quotient

$$\frac{0^2 + A^2 + B^2}{(0)(A)(B) + 1} = k \qquad \Longrightarrow \qquad k = A^2 + B^2. \tag{12}$$

But the value of the quotient k is invariant throughout this process, so we have shown that k is equal to a sum of two positive squares, as desired. ■

It is not hard to see that every sum of two positive squares must occur as a value of k. For suppose that $k = A^2 + B^2$. Running the descent process in reverse on the triple 0, A, B yields the numbers A, B, $A^3B + AB^3$, a trio of positive integers with quotient $A^2 + B^2$, as one can directly verify.

The method of proof even provides a means of organizing all solutions. As we have seen, if $a \leq b \leq c$ are positive integers satisfying the statement of Theorem 1.1, having quotient k, then the integers a, b, $kab - c$ also work. But the same argument shows that the triples a, c, $kac - b$ and b, c, $kbc - a$ are solutions as well. In general the numbers $kac - b$ and $kbc - a$ are distinct and larger than c. So we are able to run the descent in reverse to find two triples that could reduce to a, b, c via the process described in the proof, implying that the set of all solutions having a particular quotient $k = A^2 + B^2$ may be pictured as a binary tree, whose root triple is A, B, $A^3B + AB^3$. For instance, the smallest triple with $k = 5$ is $a = 1$, $b = 2$, $c = 10$. Above this come the triples 1, 10, 48 and 2, 10, 99, and so forth.

Note that when $a = b$ the values of $kac - b$ and $kbc - a$ are not distinct, so the binary tree does not branch there. This only occurs for a root triple of the form A, A, $2A^4$. Also, when k can be written as a sum of two positive squares in more than one way, as in $65 = 8^2 + 1^2 = 7^2 + 4^2$, then there is more than one tree corresponding to that value of k.

We also observe that when $k = 3$ there is a striking similarity between our equation $a^2 + b^2 + c^2 = 3abc + 3$ and the Markov equation $a^2 + b^2 + c^2 = 3abc$, discussed by Markov in [5]. Our method of solution is based on the one used to analyze the Markov equation, and the collection of all solutions may be organized via a binary tree in an analogous manner.

3 A Diophantine Equation

To appreciate the utility of Vieta jumping, we next consider the following corollary to Theorem 1.1.

Corollary 3.1 *There are no solutions to the equation*

$$a^2 + b^2 + c^2 = abc + 1. \tag{13}$$

in positive integers a, b and c.

Proof: A solution would correspond to a triple with quotient $k = 1$, but 1 is not equal to the sum of two positive squares. ■

However, once we expand our domain to all integers we do find solutions to (13); namely, $a = 0$, $b = 0$, $c = \pm 1$ or any permutation of these values. (It is not hard to see that these are the only integral solutions.) In fact, there are even positive rational solutions, such as $a = \frac{5}{2}$, $b = \frac{19}{6}$, $c = \frac{10}{3}$. We next provide an alternate proof to Corollary 3.1, both for sake of comparison and because it is appealing in its own right.

Proof: Suppose to the contrary that there do exist positive integers satisfying (13). We will need the fact that two of them have to be multiples of 4. To this end, a parity check reveals that at exactly one of a, b and c must be odd, say a. Considering (13) mod 8 leads to $b^2 + c^2 \equiv abc \bmod 8$, and checking the possibilities $b, c \equiv 0, 2 \bmod 4$ shows that the only option is $b, c \equiv 0 \bmod 4$. Therefore we can write $c = 4n$ for some $n \geq 1$.

Now rearrange (13) to obtain

$$(a - 2bn)^2 - (4n^2 - 1)b^2 = -(16n^2 - 1). \tag{14}$$

Defining $x = a - 2bn$, $y = b$ yields the Pell-like equation

$$x^2 - (2n - 1)(2n + 1)y^2 = -(4n - 1)(4n + 1), \tag{15}$$

which is like the one considered in [4], inspired by the same IMO problem.

We next observe that at least one of the numbers $2n - 1$, $2n + 1$, $4n - 1$ or $4n + 1$ is congruent to 2 mod 3; just consider $n \equiv 0, 1, 2 \bmod 3$ in turn. To

begin, suppose that $2n - 1 \equiv 2 \bmod 3$. Then $2n - 1$ must have an odd prime divisor p with $p \equiv 2 \bmod 3$. (Here we use the fact that $n \geq 1$; this step fails when $n = 0$.) Reducing (15) mod p yields

$$x^2 \equiv -(4n-1)(4n+1) \equiv -(1)(3) \bmod p. \tag{16}$$

But quadratic reciprocity implies that $(\frac{-3}{p}) = -1$ for p odd, $p \equiv 2 \bmod 3$, yielding a contradiction. The same reasoning applies when $2n + 1 \equiv 2 \bmod 3$; in this case (15) reduces to $x^2 \equiv -(-1)(-3) \equiv -3 \bmod p$ and we again reach a contradiction. It now becomes clear why we needed $c = 4n$ rather than just taking c to be even: otherwise we would not have been able guarantee that the prime divisor with $p \equiv 2 \bmod 3$ was odd, in which case the claim that $(\frac{-3}{p}) = -1$ would no longer be valid.

On the other hand, suppose that $4n - 1 \equiv 2 \bmod 3$. In this case we require the slightly finer observation that the prime factorization of $4n-1$ must include an odd prime $p \equiv 2 \bmod 3$ raised to an odd power. We once again reduce (15) mod p to obtain

$$x^2 - \left(-\tfrac{1}{2}\right)\left(\tfrac{3}{2}\right)y^2 \equiv 0 \bmod p \qquad \Longrightarrow \qquad (2x)^2 \equiv -3y^2 \bmod p. \tag{17}$$

If $y \not\equiv 0 \bmod p$, then the latter congruence implies that -3 is a square mod p, contradicting the fact that $(\frac{-3}{p}) = -1$, noted above. Therefore y is divisible by p, hence x is also, and we have a factor of p^2 in $x^2 - (2n-1)(2n+1)y^2$. Writing $x = px'$, $y = py'$ and then dividing through by p^2, we may repeat this argument as long as the quantity is divisible by p. We conclude that the prime factorization of $x^2-(2n-1)(2n+1)y^2$ involves an even power of p, contradicting the fact that $(4n-1)(4n+1)$ contains an odd power of p. Virtually identical reasoning applies when $4n + 1 \equiv 2 \bmod 3$, so we reach a contradiction in every situation, which completes the proof. ■

4 More Variables

It is natural to wonder whether Theorem 1.1 can be extended to four or more variables. The answer turns out to be yes, for the most part.

Theorem 4.1 *Let a, b, c and d be positive integers such that $abcd + 1$ divides $a^2 + b^2 + c^2 + d^2$. Then the quotient*

$$\frac{a^2 + b^2 + c^2 + d^2}{abcd + 1} \tag{18}$$

is either equal to 1, 2, or is the sum of three positive squares.

Proof: The argument proceeds in a nearly identical fashion to the proof of Theorem 1.1. Suppose that we have positive integers $a \leq b \leq c \leq d$ for which the quotient (18) is an integer k. Then replacing d by $kabc-d$ also yields a quadruple

of integers that give an integral quotient, and one shows that $kabc - d \geq 0$ as before.

The only significant deviation occurs in the process of bounding the value of $kabc - d$. This quantity will be less than c if and only if

$$(kab - 2)c < \sqrt{k^2a^2b^2c^2 - 4a^2 - 4b^2 - 4c^2 + 4k}. \tag{19}$$

When $kab - 2$ is negative the inequality is true, otherwise we can square and rearrange to obtain

$$a^2 + b^2 - k < (kab - 2)c^2. \tag{20}$$

If $kab - 2 \geq 2$ then the right-hand side is at least $c^2 + c^2$, which will exceed the left-hand side since $a, b \leq c$.

This leaves $kab = 2$ and $kab = 3$ to consider, which can only occur if $k \leq 3$. And in fact there are valid quadruples of integers that arise for each such value of k. For example, $a = 1$, $b = c = d = 3$ leads to the quotient $k = 1$, while $a = b = 1$, $c = d = 2$ gives a quotient $k = 2$ and $a = b = c = 1$, $d = 3$ has $k = 3$. Each example leads to an infinite family of solutions with the same k-value by reversing the descent operation, just as before.

Otherwise, when $k \geq 4$, we deduce that $kabc - d < c$, so this Vieta jump will decrease the size of the largest element, and we eventually reach a quadruple A, B, C, D for which $kABC - D = 0$, so $k = A^2 + B^2 + C^2$ is a sum of three positive squares, completing the proof in the same manner as before. ■

Of course, the density of positive integers that can be written as a sum of three positive squares is rather large:

$$1 - \frac{1}{8}\left(1 + \frac{1}{4} + \frac{1}{4^2} + \frac{1}{4^3} + \cdots\right) = \frac{5}{6}, \tag{21}$$

since Gauss has shown that only numbers of the form $4^t(8k + 7)$ for $k, t \in \mathbb{Z}_{\geq 0}$ fail to have such a representation. So in a sense Theorem 4.1 is not making nearly as grandiose a claim as Theorem 1.1. Thanks to Lagrange, there is even less incentive to pursue a statement involving five variables, so we will conclude our discussion at this point.

Acknowledgements

The author is grateful to John Robertson for describing much of the "folklore" presented in the first section regarding similar results, as well as for pointing out the binary tree organization of triples that give a common quotient k. Sincere thanks also to the referee, who suggested several improvements to the manuscript, including the connection with the Markov equation mentioned at the end of Section 2.

References

[1] A. Engel, Exploring Mathematics With Your Computer, Mathematical Association of America, 1993.

[2] R. Guy, R. Richberg, 'From Integers to Integers, But Not Very Many', *Amer. Math. Monthly*, **104** (1997), 278.

[3] I. Lauko, G. Pinter, L. Pinter, 'Another Step Further... on a Problem of the 1988 IMO', *Math. Mag.*, **79** (2006), 45–53.

[4] F. Luca, C. F. Osgood, P. Walsh, 'Diophantine Approximations and a Problem From the 1988 IMO', *Rocky Mountain J. Math.*, **36** (2006), 637–648.

[5] A. A. Markoff, 'Sur Les Formes Binaires Indéfinies', *Math. Ann.* **17** (1880), 379–399.

[6] I. Vidav, N. Komanda, 'Still Uniquely Fibonacci', *Amer. Math. Monthly*, **101** (1994), 279–280.

[7] I. M. Yaglom, The USSR Olympiad Problem Book, W. H. Freeman, 1962.