SUMS OF SQUARES

Lecture at Vanier College, Sept. 23, 2011.

Eyal Goren, McGill University.

email: eyal.goren@mcgill.ca

Two Squares.

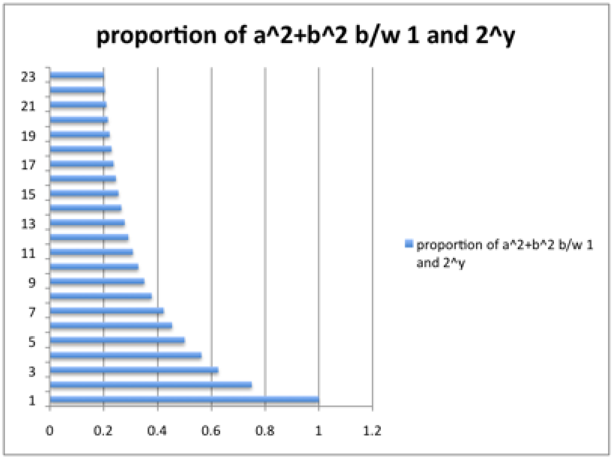
Which numbers are a sum of two squares? More precisely, which positive integers are the sum of squares of two integers?

Here is a table:

1	2	**3**	4	5	6	**7**	8	9	10	**11**	12	13	14	**15**	16	17	18	**19**	20	21	22	**23**	24	25	26	**27**
Y	Y	N	Y	Y	N	N	Y	Y	Y	N	N	Y	N	N	Y	Y	Y	N	Y	N	N	N	N	Y	Y	N

Here is some statistics

Figure 1. Statistics



The data raises many questions. For one, **the plot strongly suggest an asymptotic. But what exactly *is* this asymptotic?** what is the type of convergence?

A theorem of Edmund Landau (1908) tells us that *the proportion of numbers between* 0 *and* 2^y *that are sum of two squares is assymptotically* $\frac{K}{\sqrt{\log(2)\cdot\sqrt{y}}}$, *where* $K \sim 0.764$ *is a constant, called the Landau-Ramanujan constant.*[1] A lot is also known about the type of convergence.

Another question is whether there is **a pattern to which numbers are a sum of 2 squares?** Consider for example the case of n a prime number. Can you tell whether it is a sum of 2 squares or not?

Recall that an integer A is said to be congruent to a modulo n if A leaves a residue of a when divided by n. We write $A \equiv a \pmod{n}$. For example: $13 \equiv 1 \pmod{4}$, because $13 = 3 \times 4 + 1$. There are the following facts: if $A \equiv a \pmod{n}$, $B \equiv b \pmod{n}$, then

$$A + B \equiv a + b \pmod{n}, \quad A - B \equiv a - b \pmod{n}, \quad AB \equiv ab \pmod{n}.$$

We can in fact make the residues modulo n into a new system of numbers: to add or multiply two residues, add or multiply them as usual integers and then pass to the residue modulo n. For example, to multiply $3 \pmod{5}$ with $4 \pmod{5}$ we first multiply 3 and 4 to get 12, which gives the residue $2 \pmod{5}$. Therefore, $3 \times 4 = 2 \pmod{5}$. In that way we get a system of numbers that satisfies all the usual identities (such as $a(b + c) = ab + ac$, and so on). The implication:

$$xy = 0 \quad \Longrightarrow \quad x = 0 \text{ or } y = 0,$$

is not always true[2], but is true if n is a prime.

Theorem.(Pierre de Fermat, 1640)*An integer* n *is a sum of two squares if and only if for every prime* q *congruent to* 3 *modulo* 4, q *divides* n *to an even power.*

Example. $2^3 \cdot 5 \cdot 7^2 \cdot 13 = 25480$ and so, according to the theorem, 25480 is a sum of two squares, Indeed, $25480 = 42^2 + 154^2$.

One can get some insight as to the meaning and proof of the theorem by reasoning as follows: If the theorem is true, then it has the following consequences:

- *Any power of* 2 *is a sum of two squares.* This is easy to check: $2^{2a} = (2^a)^2$. $2^{2a+1} = (2^a)^2 + (2^a)^2$.
- *A prime congruent to* 1 *modulo* 4 *is a sum of two squares.*
- *An integer congruent to* 3 *modulo* 4 *is not a sum of two squares.*
- *The product of two numbers that are each a sum of* 2 *squares is a sum of two squares.*

Some of these are easy to prove: Note that $(2x)^2 = 4x^2$ is congruent to zero modulo 4 and $(2x + 1)^2 = 4x^2 + 4x + 1$ is congruent to 1 modulo 4. If n is a sum of two squares then n is therefore congruent to $0, 1$ or 2 modulo 4.

[1] In fact, Landau's result is more general.

[2] For example $2 \times 3 = 6 = 0 \pmod{6}$, but neither 2 nor 3 are zero, even when viewed as residues $\pmod{6}$.

To explain the consequence that the sum of numbers each a sum of two squares is itself a sum of two squares, we recall the **complex numbers**. It is not necessary, but it is enlightening and it offers a perspective that can be used in many similar situations.

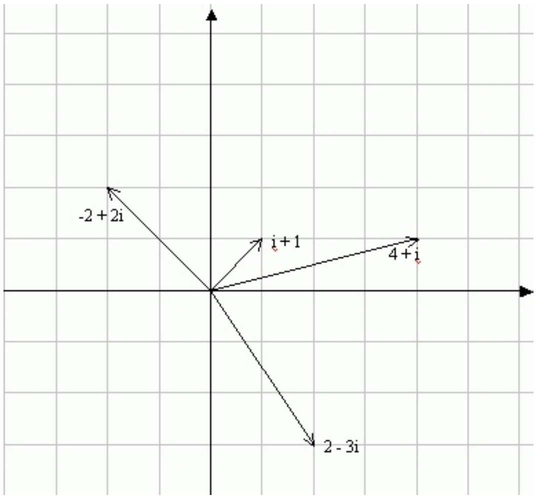
To construct the complex numbers, a symbol i is added to the real numbers and it is decreed that

$$\boxed{i^2 = -1}$$

The complex numbers are then "numbers" of the form $z = x + yi$, where x and y are real numbers. These numbers can be added and multiplied much like usual numbers. For example:

$$(1+i)(3+2i) = 1 \times 3 + i \times 3 + 1 \times 2i + i \times 2i = 3 + 3i + 2i - 2 = 1 + 5i.$$

We can depict the complex numbers in the (x, y) plane,



and then define the **absolute value** of z as

$$|z| = \sqrt{x^2 + y^2}.$$

It is precisely the length of the line connecting the origin to the point z, i.e., to the point (x, y). Multiplication, as one can check by direct calculation, satisfies

$$|z_1 z_2| = |z_1| \cdot |z_2|. \tag{1}$$

This tells us the following: since

$$(c + di)(e + fi) = ce - df + (cf + de)i,$$

after taking the square of the absolute values, we have by equation (1) the identity

$$(ce - df)^2 + (cf + de)^2 = (c^2 + d^2)(e^2 + f^2).$$

To make sure you understand the trick, express 5707026 as a sum of two squares (note that $9 = 3^2$ is associated with complex number 3. So, for example, to express 90 as a sum of squares, we write it as $2 \times 5 \times 9$. 2 is associated with $1 + i$, 5 with $1 + 2i$, 9 with 3 and so 90 is associated with $(1 + i) \times (1 + 2i) \times 3 = -3 + 9i$ giving us that $72 = (-3)^2 + 9^2 = 3^2 + 9^2$.

It took quite a long time for mathematicians to accept the existence of a number whose square is -1. We tend to think about numbers as depicting a reality – one plus one is two, etc. – and, as such, we accept the real numbers, numbers of the form 13.2768, and even 2.333333. . . as, well. . . , real! From this perspective, it was difficult to accept that there is a new number, called i, that satisfies $i^2 = -1$. And, at first, mathematicians found it more comfortable to think of i as an **operator**; namely, a symbol that denotes a concrete operation. The operation here is that i is the transformation of the plane taking (x, y) to $(-y, x)$; that is, rotation counterclockwise by 90°. This interpretation follows directly from the equation

$$i(x + yi) = -y + xi.$$

Then, i^2 sends (x, y) to $(-x, -y)$ and so is like multiplication by -1. We can therefore interpret i^2 as -1.

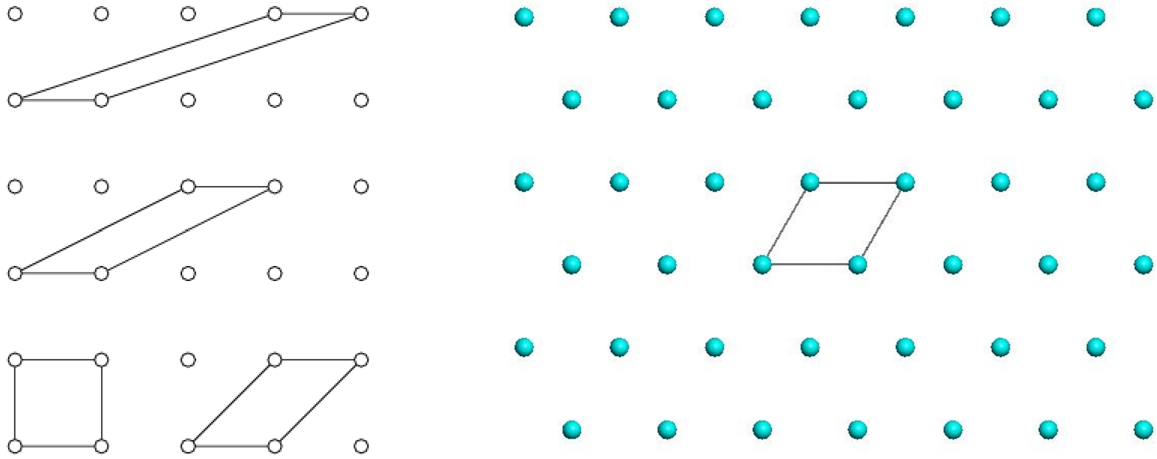
Based on the identities above, to prove "the if part" of the theorem, it is therefore enough to prove that every prime congruent to 1 modulo 4 is a sum of two squares. To prove that such a prime is a sum of squares, we will need the following fact:

Lemma. *There is an integer $0 < u < p$ such that u^2 leaves residue -1 modulo p.*

For example, $2^2 = 4 = -1 \pmod 5, 5^2 = 25 = -1 \pmod{13}, 12^2 = 144 = -1 \pmod{29}$. We assume the lemma. We mention that it does not hold for primes congruent to 3 modulo 4.

We now introduce a new idea: a **lattice** L in the plane is a collection of points such that the sum, or difference, of two points in L is again in L. For example the set of points (x, y) such that x and y are integers is a lattice. A lattice has a **fundamental parallelogram**, a parallelogram based at zero, whose vertices are points of L, that contains no point of L in its interior.

Figure 2. The lattice of integers and some fundamental parallelograms, and the hexagonal lattice



Minkowski's theorem. *If the fundamental parallelogram of L has area V then a disc of radius r such that $\pi r^2 \geq 4V$ must contain a point of L besides $(0,0)$.*

Application. We consider the lattice L of points (x, y) such that $y - ux$ is divisible by p. The parallelogram with vertices $(0,0), (0,p), (1,u), (1, p+u)$ is a fundamental parallelogram. Given a point (x, y) of the lattice, we can write it as $x(1,u) + \frac{y-xu}{p}(0,p)$. This shows that such a point cannot be inside the fundamental parallelogram. The area of this parallelogram is p.

Let us take a disc of radius $r = \sqrt{4\pi^{-1}p}$, a radius that satisfies Minkowski's condition "on the nose". There is a lattice point (x, y) in this disc, that is, a point such that $x^2 + y^2 < 4\pi^{-1}p < 2p$ and $y - ux$ is divisible by p. Now, y^2 is congruent to $u^2x^2 = -x^2$ modulo p and so,

$$p|(x^2 + y^2).$$

Since $x^2 + y^2 < 2p$ it follows that

$$p = x^2 + y^2.$$

Proof of Minkowski's theorem. Call the disc D and consider $d = \frac{1}{2}D$ - the disc shrunk by a factor of 2. Suppose that d is disjoint from $l + d$ for all $l \in L$. Then, we can translate the various parts of d lying in different parallelograms so that they all lie in the fundamental one, and they are all disjoint. Therefore, the area of d, which is $\frac{1}{4}\pi r^2$ is less than that of L, which is V. That is,

$$\frac{1}{4}\pi r^2 < V.$$

This is a contradiction, and so d is not disjoint from $l + d$ for some $l \neq 0$. That means, that for some $s, t \in D$ we have $\frac{1}{2}s = l + \frac{1}{2}t$. Thus,

$$l = \frac{1}{2}(s - t).$$

It follows that the distance of l from the origin is at most $\frac{1}{4}(r + r) < r$, and so l lies in D.

Four Squares.

Here we would like to discuss the following theorem.

Theorem. *Every positive integer is a sum of* 4 *squares.*

The strategy of showing that if x is a sum of 2 squares and y is a sum of 2 squares than so is xy was very useful. Can we prove such a statement for 4 squares?

To prove that, we introduce a generalization of complex numbers called Quaternion numbers. These are "numbers" of the shape $a + bi + cj + dk$ where a, b, c, d are real numbers and i, j, k are new symbols that satisfy

$$i^2 = j^2 = k^2 = -1, \quad ij = k = -ji.$$

Using these identities one can add and multiply quaternions in a formal way. Many familiar identities hold, for example $x(yz) = (xy)z$, but some identities do not hold. For example, $xy \neq yx$ in general. Indeed, $ij = -ji$. Still, we can define

$$|z| = \sqrt{a^2 + b^2 + c^2 + d^2},$$

and we have, by laborious calculation,

$$|z_1 z_2| = |z_1| \cdot |z_2|.$$

It now follows that if x is a sum of 4 squares and y is a sum of 4 squares then so is xy. And, in fact, the presentation of xy as a sum of 4 squares can be calculated from the presentation of x and of y, similar to the way we had done it for sum of two squares. Therefore, to show that every integer is a sum of 4 squares, it is enough to show that for primes congruent to 3 modulo 4.

The proof of this result is again an application of Minkowski's theorem, but now using lattices in 4-dimensional space. The concept of a lattice generalized easily and so is the concept of a ball and its volume. Minkowski considered lattices in a space of arbitrary dimension n and proved the following general result, whose proof is a generalization of the proof given for $n = 2$.

Minkowski's theorem 1889. *Let L be a lattice in n-dimensional space. If the fundamental parallelogram of L has volume V then a ball whose volume is at least $2^n V$ must contain a point of L besides $(0, 0)$.*

The choice of lattice in this case is much more tricky and we don't discuss this further.

Three Squares.

Also this problem admits a precise solution.

Legendre's theorem. *A positive integer is a sum of three squares, unless it is of the form $4^n(8m + 7)$.*

Note that $3 = 1^2 + 1^2 + 1^2$ is a sum of 3 squares, as is $5 = 0^2 + 1^2 + 2^2$, but $3 \cdot 5 = 15 = 8 \times 1 + 7$ is not a sum of 3 squares (as one can verify directly). Thus, the paradigm of our previous proofs fails. There is no 3-dimensional generalization of the integers, similar to the two dimensional generalization $a + bi$ (a, b integers) - the Gaussian integers, or the four dimensional generalization $a + bi + cj + dk$ (a, b, c, d integers) - the Hurwitz integers.

In fact, for many years, Sir William Rowan Hamilton tried to find generalizations of the complex numbers to three dimensional situation, namely, how to add and multiply "triples" $(a + bi + cj)$. Addition is not a problem, but multiplication yielded a situation where many identities we expect just fail. One novelty of Hamilton's approach was the treatment of the expressions $a + bi + cj$ as generalized numbers and not as operators. Indeed, from the point of view of operators it is hard to see why there shouldn't be such a multiplication rule. The problem was not trivial by any means, certainly not at the time. Hamilton's investigations came on the heels of breakthroughs achieved by the use of complex numbers (for example, Gauss's proof of the fundamental theorem of algebra in 1799) at a time when many mathematicians still had problems accepting the complex numbers as "numbers", namely, as purely algebraic quantities. In fact, Hamilton was the first to give a satisfactory definition of complex numbers. What Hamilton was doing was cutting-edge research.

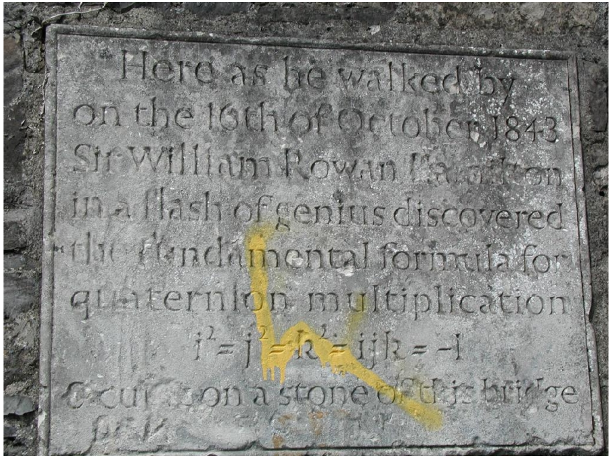
Hamilton was a prodigy. Born in Dublin at 1805 he could translate Latin, Greek and Hebrew at the age of 5. At twevlth he knew at least 9 languages and engaged in contests with the American 'calculating boy' Zerah Colburn (when Colburn was seven years old he took six seconds to give the number of hours in thirty-eight years, two months, and seven days). The competition was one of the events turning Hamilton's interest to mathematics. He entered Trinity college in Dublin at 1823, but was appointed as a professor while still an undergraduate, at 1827, at the age of 22. Thus, we can appreciate the difficulty of the problem of generalizing complex numbers in the light of the efforts of such an intellect; he had devoted 15 years to the problem.

The story goes that Hamilton's little sons, aged eight and nine, had, during the climatic month of October 1843, greeted him at breakfast with ”Well, Papa, can you *multiply* triplets". Hamilton later wrote to his son: *Every morning in the early part of the above-cited month, on my coming down to breakfast, your (then) little brother William Edwin, and yourself, used to ask me: "Well, Papa, can you multiply triplets?" Whereto I was always obliged to reply, with a sad shake of the head: 'No, I can only add and subtract them".*

The inspiration that one can multiply quadruples $(a + bi + cj + dk)$ in a reasonable way came to Hamilton in a flash on October 16, 1843, while walking with his wife along the Royal Canal in Dublin, on their way to a concert. He describes it as "an electric circuit seemed to close and a spark flashed forth". He documented the insight by engraving the identities $i^2 = j^2 = k^2 = ijk = -1$ on a stone of the Brougham Bridge - performing also the first and most famous act of mathematical graffiti. The bridge has a plaque commemorating the event (Figure 3).

In fact, there are deep reasons as to why one cannot multiply triples. Suppose there was such a multiplication of triples (x, y, z) real numbers that preserved length and extended the multiplication of usual real numbers $(x, 0, 0)$. There is thus a multiplication rule for the sphere $x^2 + y^2 + z^2 = 1$ in three dimensional space. Take a tangent vector at the point $(1, 0, 0)$ and use the multiplication to transport it to a tangent vector at a point (x, y, z) ($(1, 0, 0)$ stands for the number one and so we want $(1, 0, 0) \times (x, y, z) = (x, y, z)$ for any (x, y, z), whatever the multiplication rule may be). This way, we get a field of tangent vectors consisting of unit vectors, and in particular, a field for which no vector is zero. We were able to "comb" the ball. But there is a theorem in topology

FIGURE 3. A plate commemorating Hamilton's discovery of the quaternions.



Here as he walked by
on the 16th of October 1843
Sir William Rowan Hamilton
in a flash of genius discovered
the fundamental formula for
quaternion multiplication
$i^2 = j^2 = k^2 = ijk = -1$
& cut it on a stone of this bridge

The hairy ball theorem. *One cannot comb the ball.*

The proof of this theorem uses totally different ideas from those we were pursuing so far and we shall not discuss the proof. It has the following consequence:

Application. *At any moment, somewhere on earth, the air stands perfectly still.*

Indeed, we may think about the direction of the wind as a field of tangent vectors; they cannot be all non-zero due to the hairy ball theorem.

Very Recent Developments.

Other number systems. Much as the complex numbers are a system of numbers in which we can do arithmetic, there are many others system of numbers (they are called number fields) in which we can do arithmetic. For example, we may consider the system of numbers of the form

$$a + b\sqrt{2}, \quad a, b \in \mathbb{Z}.$$

The sum and product of two such (real) numbers is a (real) number of the same sort and so we may ask, which numbers are sums of two, three, four, etc., squares in this system?

Legendre's theorem doesn't generalize easily. about 10 years ago, Cogdell, Sarnak and Piatetski - Shapiro gave an analogue that holds for every large enough number.

Conway-Schneeberger Theorem. The generalization of expressions of the form $x^2 + y^2, x^2 + y^2 + z^2$ are quadratic forms. They are expressions of the form $\sum_{ij} a_{ij} x_i x_j$ where the a_{ij} are constants and the x_i are variables. For instance, $x^2 + xy + 3y^3$, $2xz + y^2 - 5yz$, etc. Let n be an integer. If there is an assignment of integers values for the variables x_i such that $\sum_{ij} a_{ij} x_i x_j = n$, we say that the quadratic form represents n. For example, $x^2 + y^2$ represents 5, but does not represent 3. A natural question is when does a quadratic form represent every integer? We have the following striking theorem:

Theorem. (Bhargava, Conway-Schneeberger) *If a quadratic form $\sum_{ij} a_{ij} x_i x_j$, where a_{ij} are integers and are even integers if $i \neq j$, represents the integers $1, 2, \ldots, 15$ then it represents every integer.*