Sum of four squares via the Hurwitz quaternions

Robert Hines

September 21, 2016

None of the following is original, in fact it's mostly ripped from wikipedia (edited for ease of exposition).

1 Sum of two squares

Let's warm up with deciding which positive integers can be written as a sum of two squares. The key to answering this question is noting its mulitiplicative nature. Specifically,

$$(a^2+b^2)(x^2+y^2) = (ax-by)^2 + (ay+bx)^2 = (ax+by)^2 + (ay-bx)^2$$

which we may think of in terms of norms of complex numbers

$$|a+bi|^2|x+yi|^2 = |(a+bi)(x+yi)|^2$$

or mulitiplicativity of the norm in the Gaussian integers, or as a composition of binary quadratic forms. In any case, this (basically) reduces the problem to writing primes as a sum of two squares. There is an obvious congruence obstruction, namely that

$$x^2+y^2 \equiv 0, 1, 2(4),$$

i.e. a prime $p \equiv 3(4)$ is never a sum of two squares. We obviously have $1 = 1^2 + 0^2$ and $2 = 1^2 + 1^2$ so we focus our attention on odd primes $p \equiv 1(4)$. We will use the fact that the Gaussian integers, $\mathbb{Z}[\sqrt{-1}]$, have unique factorization. It is a euclidean domain with respect the usual absolute value on the complex numbers. To see this note that the condition

Given any $a, b \in \mathbb{Z}[i]$, $b \neq 0$, there are $q, r \in \mathbb{Z}[i]$ such that

$$a = bq + r, \ |r| < |b|,$$

is equivalent (dividing by b) to the condition

For any $a/b \in \mathbb{Q}(i)$ there is a $q \in \mathbb{Z}[i]$ such that

$$|a/b - q| < 1.$$

You can convince yourself of the latter by drawing circles of radius 1 around lattice points in $\mathbb{Z}^2$. [Exercise: show that the covering radius around integers is ≤ 1 in $\mathbb{Q}(\sqrt{d})$, $d < 0$ square free, iff $d = -1, -2, -3, -7, -11$.] One other fact we need is that -1 is a square modulo p for $p \equiv 1(4)$ since

$$(-1)^{\frac{p-1}{2}} \equiv 1(p)$$

(if -1 weren't a square, we'd have $(-1)^{\frac{p-1}{2}} \equiv -1(p)$, recalling $(\mathbb{Z}/(p))^\times$ is cyclic of order $p-1$). Hence there is an m such that $p|m^2+1$. Over the Gaussian integers, this gives

$$p|m^2+1 = (m+i)(m-i).$$

Now p divides neither of the factors on the right hand side so that p is not a Gaussian prime. Hence p has a non-trivial factorization over the Gaussian integers, which must be of the form $p = (x+yi)(x-yi) = x^2+y^2$ since $|p|^2 = p^2$ so that $p = \alpha\beta$ with $|\alpha|^2 = |\beta|^2 = p$. Hence p is a sum of two squares.

To summarize, n is a sum of two squares iff for all $q|n$ with $q \equiv 3(4)$, q occurs with an even exponent in the prime factorization of n. One direction is clear from the above, and to see that the exponents of $q|n$, $q \equiv 3(4)$, must be even, we have the following lemma.

Lemma 1. *If $q|a^2+b^2$ and $q \equiv 3(4)$, then $q|a$ and $q|b$ (i.e. q^2 divides $n = a^2+b^2$).*

Proof. If q doesn't divide both a and b, say $(a,q) = 1$, let $a'a \equiv 1(q)$. Then we have

$$-a^2 \equiv b^2(q),\ -1 = (a'b)^2(q).$$

However, -1 is not a quadratic residue modulo q, a contradiction. $\square$

An example:

$$\begin{aligned} 4680 &= 2^3 \cdot 3^2 \cdot 5 \cdot 13 = (1^2+1^2)^3(3^2+0^2)(1^2+2^2)(2^2+3^2) = \dots \\ &= 18^2+66^2 = 42^2+54^2. \end{aligned}$$

The representation of a prime as a sum of two squares is unique (up to order and sign). More generally we have (including sign and order)

$$r_2(n) = 4(d_1(n) - d_3(n))$$

where $d_i(n)$ is the number of divisors of n congruent to i modulo 4 ($16 = 4(8-4)$ in the example above).

2 Hurwitz quaternions

To prove that *every* positive integer can be express as a sum of four squares, we follow the proof above but use properties of "unique" factorization in a non-commutative ring, the Hurwitz quaternions (a maximal order the quaternion algebra $\left(\frac{-1,-1}{\mathbb{Q}}\right)$).

Given a field F ($\mathrm{char}(F) \neq 2$) and $a, b \in F^\times$, there is a four-dimensional (unital, associative) F-algebra $A = \left(\frac{a,b}{F}\right)$ with basis $1, i, j, ij = k$ determined by

$$i^2 = a,\ j^2 = b,\ ij = -ji.$$

A is either isomorphic to $M_2(F)$ or is a division algebra (skew fiield, non-commutative field), according as $ax^2+by^2 = 1$ has a solution $(x,y) \in F^2$. [This is the Clifford algebra for the quadratic form ax^2+by^2.]

The familiar real quaternions, or Hamiltonians $\mathbb{H} = \left(\frac{-1,-1}{\mathbb{R}}\right)$ is a division algebra as follows. We introduce the order two anti-isomorphism ("conjugation")

$$x = x_0 + x_1 i + x_2 j + x_3 k,\ \bar{x} = x_0 - x_1 i - x_2 j - x_3 k$$

and note that $N(x) := x\bar{x} = x_0^2 + x_1^2 + x_2^2 + x_3^2 \in \mathbb{R}$ with $x\bar{x} \geq 0$ with equality if and only if $x = 0$. Hence $x^{-1} = \frac{\bar{x}}{x\bar{x}}$ for $x \neq 0$. The multiplicativity of the norm once again allows us to reduce the problem of representing an integer as a sum of four squares to that of representing a prime as a sum of four squares

$$\begin{aligned} N(x)N(y) =& (x_0^2 + x_1^2 + x_2^2 + x_3^2)(y_0^2 + y_1^2 + y_2^2 + y_3^2) \\ N(xy) =& (x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3)^2 + (x_0y_1 + x_1y_0 + x_2y_3 - x_3y_2)^2 \\ &+ (x_0y_2 - x_1y_3 + x_2y_0 + x_3y_1)^2 + (x_0y_3 + x_1y_2 - x_2y_1 + x_3y_0)^2. \end{aligned}$$

You may also recognize the “real” and “imaginary” parts of

$$\begin{aligned} &(x_1i + x_2j + x_3k)(y_1i + y_2j + y_3k) \\ &= -(x_1y_1 + x_2y_2 + x_3y_3) + (x_2y_3 - x_3y_2)i + (x_3y_1 - x_1y_3)j + (x_1y_2 - x_2y_1)k \end{aligned}$$

as the dot and cross products of (x_1, x_2, x_3) and (y_1, y_2, y_3). The Hamiltonians have a representation as a real subalgebra of $M_2(\mathbb{C})$

$$\mathbb{H} \cong \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix},\ a, b \in \mathbb{C} \right\}$$

(quaternion conjugation is the adjoint/conjugate transpose) and the group of norm one elements is topologically S^3 ($\det(x) = a\bar{a} + b\bar{b} = 1$ describes the unit sphere), isomorphic to SU_2, a double cover of $SO_3(\mathbb{R}) \cong P^3(\mathbb{R})$.

As we used $\mathbb{Z}[i] \subseteq \mathbb{Q}(i)$ above, we want to take advantage of a subring of the division algebra $A := \left(\frac{-1,-1}{\mathbb{Q}}\right)$ which has similar properties. An obvious candidate for “integers” in the quaternions is the ring $\{x_0 + x_1i + x_2j + x_3k : x_i \in \mathbb{Z}\}$, but this isn't large enough (just barely: the only points of $\mathbb{R}^4$ not covered by open balls of radius 1 around $\mathbb{Z}^4$ are $(\mathbb{Z} + 1/2)^4$ since the length of the diagonal of the unit cube in four dimensions is $\sqrt{1^2 + 1^2 + 1^2 + 1^2} = 2$). So instead we'll throw in those middle points and consider the Hurwitz quaternions

$$\mathcal{O} := \{x_0 + x_1i + x_2j + x_3k : \{x_i\} \subseteq \mathbb{Z} \text{ or } \{x_i\} \subseteq \mathbb{Z} + 1/2\}$$

(all x_i integers or half-integers). One can verify that this is a ring, and that all norms are integers

$$N((x_0 + 1/2) + (x_1 + 1/2)i + (x_2 + 1/2)j + (x_3 + 1/2)k) = \sum_{i=1}^{4}(x_i + 1/2)^2 = 1 + \sum_{i=1}^{4} x_i(x_i + 1).$$

For later use, note that $\mathcal{O}^\times = \{x \in \mathcal{O} : N(x) = 1\}$. Some magical junk about $\mathcal{O}$:

- The 24 vectors of norm 1, $\mathcal{O}^\times$, are $\pm 1, \pm i, \pm j, \pm k$, and $(\pm\frac{1}{2} \pm \frac{i}{2} \pm \frac{j}{2} \pm \frac{k}{2})$ with all combinations of $\pm$. $\mathcal{O}^\times$ is the vertex set of the 24-cell, a self-dual convex regular polytope in four dimensions which tessellates $\mathbb{R}^4$.

- $\mathcal{O}$ is the F_4 root lattice, the root system being the the union of the vertices of the 24-cell $\mathcal{O}^\times$ and its dual, all permutations of coordinates and choice of signs for $(\pm 1, \pm 1, 0, 0)$.

We need some euclidean property and lemma similar to $\left(\frac{-1}{p}\right) = 1$ for $p \equiv 1(4)$ to imitate our earlier proof.

Lemma 2. *For any $\alpha \in A$ there is an $x \in \mathcal{O}$ such that $N(\alpha - x) < 1$ (all rational quaternions are within unit distance of a Hurwitz quaternion).*

Proof. Choose x_0 so that $|\alpha_0 - x_0| < 1/4$ (which decides whether or not the x_i will be all integers or all half-integers), then follow through choosing $|x_i - \alpha_i| < 1/2$ for $i = 1, 2, 3$. Then $N(\alpha - x) < 1/16 + 1/4 + 1/4 + 1/4 = 13/16 < 1$ as desired. $\square$

This lemma is enough to show that all one-sided ideals of $\mathcal{O}$ are principal. Given a left ideal I, let $x \in I \setminus \{0\}$ have minimal norm. If $y \in I$ then there is a $q \in \mathcal{O}$ such that $N(yx^{-1} - q) < 1$, i.e. $N(y - qx) < N(x)$ impossible unless $y = qx$.

Lemma 3. *For any odd prime p, there are $a, b \in \mathbb{Z}$ such that $p | 1 + l^2 + m^2$ (i.e. -1 is a sum of two squares modulo p).*

Proof. There are $(p+1)/2$ distinct residues in $X = \left\{0^2, 1^2, \ldots, \left(\frac{p-1}{2}\right)^2\right\}$ (over a field $x^2 = y^2$ iff $x = \pm y$), and therefore $(p+1)/2$ distinct residues in $Y = \{-(1+x) : x \in X\}$. Hence $X \cap Y \neq \emptyset$, $a^2 \equiv -(1+b^2)(p)$ for some a, b, and $p | 1 + a^2 + b^2$ as desired. $\square$

3 Sum of four squares

Here we go!

Theorem 1. *Every positive integer is a sum of four squares.*

Proof. First note that 1 and 2 are both sums of four squares, so that we are left to show that every odd prime is a sum of four squares (via multiplicativity of the norm). For an odd prime p we have integers a, b such that

$$p | 1 + a^2 + b^2 = (1 + ai + bj)(1 - ai - bj)$$

and $p > 2$ divides neither of the factors on the right. Consider the (principal!) right ideal $p\mathcal{O} + (1 + ai + bj)\mathcal{O} = x\mathcal{O}$. This ideal contains p so we have a factorization $p = xy$ in $\mathcal{O}$. This factorization is non-trivial. If y were a unit then $1 + ai + bj \in x\mathcal{O} = p\mathcal{O}$, but $p \nmid 1 + ai + bj$. If x were a unit, then $\mathcal{O} = x\mathcal{O}$ and

$$1 - ai - bj \in (1 - ai - bj)(p\mathcal{O} + (1 + ai + bj)\mathcal{O}) \subseteq p\mathcal{O}.$$

This is impossible as $p \nmid 1 - ai - bj$.

So p factors non-trivially in $\mathcal{O}$, $p = xy$, with $p = N(x) = N(y)$, and $p = x_0^2 + x_1^2 + x_2^2 + x_3^2$. If all the x_i are integers, we've finished. If the x_i are all half-integers, let $\epsilon = \frac{1}{2}(\pm 1 \pm i \pm j \pm k)$ (a unit) such that $z = x + \epsilon$ has even integer coefficients. Then $p = x\epsilon\bar{\epsilon}\bar{x} = (\bar{z}\epsilon - 1)(\bar{\epsilon}z - 1)$, and $p = N(\bar{\epsilon}z - 1)$ where $\bar{\epsilon}z - 1$ has integer coefficients. $\square$