SUMS OF TWO SQUARES AND LATTICES

KEITH CONRAD

One of the basic results of elementary number theory is Fermat's two-square theorem.

Theorem 1 (Fermat, 1640). *An odd prime p is a sum of two squares if and only if $p \equiv 1 \bmod 4$. Furthermore, a representation of a prime as a sum of two squares is unique up to the order of addition of the squares.*

That an odd prime which is a sum of two squares must be $1 \bmod 4$ follows from a calculation of squares modulo 4. To prove, conversely, that any prime $p \equiv 1 \bmod 4$ is a sum of two squares, there are several methods available: descent [6, Chap. 26] (this was Fermat's own approach, according to [7, p. 67]), factorization of p in the Gaussian integers [2, p. 120], Jacobi sums [2, p. 95], the pigeonhole principle [1, pp. 264–265], continued fractions [5, pp. 132–133], quadratic forms [3, pp. 163–164], and Minkowski's convex body theorem [3, pp. 454–455]. One of the virtues of the proof using Gaussian integers is that, thanks to unique factorization in $\mathbf{Z}[i]$, one simultaneously obtains the uniqueness of the representation of a prime $p \equiv 1 \bmod 4$ as a sum of two squares. This uniqueness can also be proved using simple congruence and divisibility arguments [1, pp. 265–266].

The question which motivated the present note is whether or not there is a proof of the uniqueness part of Theorem 1 using lattice methods, in the spirit of Minkowski's proof of the existence part of Theorem 1. We will give such a proof, as suggested by D. Clausen. Let p be an odd prime and assume $p = a^2 + b^2$ for some integers a and b. We want to show this is the only representation of p as a sum of two squares.

Since $a^2 + b^2 \equiv 0 \bmod p$, both a and b are nonzero modulo p, so dividing by b shows there is a solution to $k^2 + 1 \equiv 0 \bmod p$. For any integers x and y, $x^2 + y^2 \equiv 0 \bmod p$ if and only if $y \equiv \pm kx \bmod p$. Set

$$L = \{(x, y) \in \mathbf{Z}^2 : y \equiv kx \bmod p\} = \mathbf{Z}(1, k) + \mathbf{Z}(0, p),$$

which is a lattice in the plane whose fundamental parallelogram has area $\left|\begin{smallmatrix} 1 & k \\ 0 & p \end{smallmatrix}\right| = p$. (This is the lattice which appears in Minkowski's proof of the existence part of Theorem 1.) Let $C = \{(x, y) \in \mathbf{R}^2 : x^2 + y^2 = p\}$. The uniqueness in Theorem 1 amounts to showing C contains only 8 integral points (those coming from modifying a and b by order and sign). For each integral point (x, y) of C, exactly one of (x, y) or $(x, -y)$ is in L since $y \equiv \pm kx \bmod p$ and $k \not\equiv -k \bmod p$ (because $p \neq 2$). Therefore the total number of integral solutions to $x^2 + y^2 = p$ is $2\#(C \cap L)$.

Changing the signs on a and b if necessary, we may assume $b \equiv ka \bmod p$, so there are at least 4 points in $C \cap L$: $(a, b), (-a, -b), (-b, a)$, and $(b, -a)$. (There are four more integral points on C: $(a, -b), (-a, b), (b, a)$, and $(-b, -a)$, and they lie not on L but on the lattice $L' = \{(x, y) \in \mathbf{Z}^2 : y \equiv -kx \bmod p\} = \mathbf{Z}(1, -k) + \mathbf{Z}(0, p)$.) This same argument for other integral points on C shows $\#(C \cap L)$ is a multiple of 4.

We will now count $\#(C \cap L)$ in a different way, using areas. Construct the convex polygon whose vertices are the points in $C \cap L$. This polygon lies in C, so the area of the polygon is no larger than the area of C, which is πp. The area of the polygon can be given by an exact formula in terms of $\#(C \cap L)$ using Pick's theorem:

Theorem 2 (G. Pick, 1899). *Let $\Lambda \subset \mathbf{R}^2$ be a lattice and Π be a polygon with vertices on Λ. If Π is convex, or more generally has no self-intersections, then the area of Π is*

$(I+B/2-1)\Delta$, *where* I *is the number of interior points of the polygon in* L, B *is the number of boundary points of the polygon in* Λ, *and* Δ *is the area of a fundamental parallelogram for* Λ.

Often Pick's theorem is stated for polygons with vertices on the standard integral lattice $\mathbf{Z}^2$, but here the formulation with a more general lattice is relevant. This more general case can be reduced by linear algebra to the case of the standard integral lattice. A proof of Pick's theorem is in [4].

For the convex polygon whose vertices are $C \cap L$, the only point of L in the interior of C is the origin since (by the definition of L) each element of L has squared distance from $(0,0)$ equal to a multiple of p. Therefore $I = 1$. Since $B = \#(C \cap L)$ and $\Delta = p$, the area of the polygon is $(1 + B/2 - 1)p = \#(C \cap L)p/2$. Comparing this with the upper bound πp from before, we get $\#(C \cap L)p/2 < \pi p$, so $\#(C \cap L) < 2\pi \approx 6.2$. Since $\#(C \cap L)$ is a multiple of 4, we are left with $\#(C \cap L) = 4$, so the only integral solutions to $p = x^2 + y^2$ are the 8 choices coming from the pair (a, b) and changes in sign and order of the coordinates.

References

[1] D. M. Burton, "Elementary Number Theory," 6th ed., McGraw–Hill, New York, 2007.
[2] K. Ireland and M. Rosen, "A Classical Introduction to Modern Number Theory," 2nd ed., Springer–Verlag, New York, 1990.
[3] J. R. Goldman, "The Queen of Mathematics: A Historically Motivated Guide to Number Theory," A.K. Peters, Natick, MA, 2004.
[4] I. Niven and H. Zuckerman, *Lattice points and polygonal area*, Amer. Math. Monthly **74** (1967), 1195–1200.
[5] C. D. Olds, "Continued Fractions," Math. Assoc. America, Washington, D.C., 1963.
[6] J. H. Silverman, "A Friendly Introduction to Number Theory," 3rd ed., Prentice Hall, Upper Saddle River, NJ, 2006.
[7] A. Weil, "Number Theory: An Approach Through History from Hammurapi to Legendre," Birkhäuser, Boston, 1984.