Journal of Number Theory **96**, 417–431 (2002)
doi:10.1006/jnth.2002.2779

Geometry of Numbers Proof of Götzky's Four-Squares Theorem

Jesse Ira Deutsch

Mathematics Department, University of Botswana, Private Bag 0022, Gaborone, Botswana
E-mail: deutschj@mopipi.ub.bw

Communicated by D. Goss

Received October 24, 2001; revised December 17, 2001

The totally positive algebraic integers of certain number fields have been shown to be the sums of four squares of integers from their respective fields. The case of $\mathbb{Q}(\sqrt{5})$ was demonstrated by Götzky and the cases of $\mathbb{Q}(\sqrt{2})$ and $\mathbb{Q}(\sqrt{3})$ were demonstrated by Cohn. In the latter situation, only those integers with even coefficient on the radical term could possibly be represented by sums of squares. These results utilized modular functions in order to get the exact number of representations. Here a method of Grace is adapted to show the existence of a four-squares representation for $\mathbb{Q}(\sqrt{5})$ without, however, obtaining the number of these. Also, results about the representation of primes by sums of two squares are obtained for $\mathbb{Q}(\sqrt{5})$. © 2002 Elsevier Science (USA)

Key Words: four-squares theorem; two-squares theorem; convex body; lattice.

1. INTRODUCTION

Two-squares and four-squares representation theorems for the rational integers were asserted by Fermat in the 1600s and first publicly proved by the efforts of Euler and Lagrange in the 1700s. Fermat's claim that all positive rational integers were sums of three triangular numbers, four squares, five pentagonal numbers, etc., was established by Cauchy in the early 19th century. In 1928, Götzky proved that all totally positive integers in $\mathbb{Q}(\sqrt{5})$ were the sum of four squares of algebraic integers from that field. In particular, Götzky showed that the number of such representations for totally positive integral $\alpha \in \mathbb{Q}(\sqrt{5})$ is

$$8\sum_{(v)|\alpha} |N(v)| - 4 \sum_{2|(v)|\alpha} |N(v)| + 8 \sum_{4|(v)|\alpha} |N(v)|, \tag{1.1}$$

where N stands for the field norm and (v) runs through the ideals dividing α. About 1960, Cohn extended this work to show the existence of representations by sums of squares for $\mathbb{Q}(\sqrt{2})$ and $\mathbb{Q}(\sqrt{3})$. Further, it was

0022-314X/02 $35.00

shown that three squares are sufficient in a number of cases. The theorems of Götzky and Cohn were demonstrated by means of the theory of modular functions for two complex variables and related theta functions. While the results of Götzky and Cohn required knowledge of the fundamental regions of the associated modular group and careful estimates of the values of the theta functions, they were able to obtain exact values for the number of such representations. See [1, 2, 5, 8] for further details.

An alternative demonstration showing the existence of the representation of a rational integer as the sum of four squares used the convex body theorem of Minkowski's geometry of numbers (see [6]). Also the classical two-squares theorem can be shown with very similar techniques. In this paper, analogs of these theorems are proven using the convex body technique for representation by sums of squares in the case of the quadratic field $\mathbb{Q}(\sqrt{5})$. While the formula for the number of representations is not recovered, the existence of such representations is demonstrated via a technique alternate to the standard modular function and theta function methods.

2. RATIONAL INTEGER CASE

Let us now review Minkowski's convex body method as it pertains to the classical case of representation by two and four rational integer squares. The reference for this is Hardy and Wright [8], and Grace [6]. We start with Minkowski's theorem as stated in [10].

THEOREM 1. *Let Λ be an n-dimensional lattice in $\mathbb{R}^n$ and let E be a convex, measurable, centrally symmetric subset of $\mathbb{R}^n$ such that*

$$\operatorname{vol}(E) > 2^n \operatorname{vol}(\mathbb{R}^n/\Lambda). \tag{2.1}$$

Then E contains some nonzero point of Λ. If E is also compact, then the strict inequality can be weakened to $\geqslant$.

For two squares the key is the following theorem.

THEOREM 2. *Any rational prime p for which -1 is a quadratic residue has a representation $p = x^2 + y^2$ where $x, y \in \mathbb{Z}$.*

Proof. Following the notation of Hardy and Wright, let l satisfy $l^2 \equiv -1 \pmod{p}$ and form the lattice in $\mathbb{R}^2$ with basis $\{(1, l), (0, p)\}$. The volume of this lattice is p. Consider a circle of radius r centered at the origin. We need $\pi r^2 > 4p$ in order to have Minkowski's theorem guarantee the existence of a point inside the circle. Thus it is merely necessary to fix $r^2 = 1.28p$. Let the nonzero lattice point inside the circle be $m(1, l) + n(0, p)$ where m and n are integers. Then $x = m$ and $y = lm + np$.

$$x^2 + y^2 \equiv m^2 + l^2m^2 \equiv m^2(1 + l^2) \equiv 0 \pmod{p} \tag{2.2}$$

while

$$0 < x^2 + y^2 < 1.3p < 2p. \tag{2.3}$$

Hence, x and y are integers that satisfy $x^2 + y^2 = p$. ∎

With some additional elementary number theory, one may deduce Fermat's theorem that every prime of the form $4n + 1$ is the sum of two squares (see [8]).

To demonstrate the classical four-squares theorem, the following lemma is also needed.

LEMMA 3. *For any rational prime p, there exist rational integers x and y such that $x^2 + y^2 + 1 \equiv 0 \pmod{p}$.*

Proof. If -1 is a quadratic residue of p, then there exists an integer x such that $x^2 + 1 \equiv 0 \pmod{p}$. In the other case, -1 is not a quadratic residue modulo p. Consider the set $\{y^2 + 1\}$ as y run through the integers. If this set consists only of squares modulo p, then a square modulo p plus one is always a square modulo p. By induction every positive integer is a square modulo p. That is impossible, since only half the integers between 1 and $p - 1$ are. Thus there exists a y such that $y^2 + 1$ is not a square modulo p. But -1 is not a square either for the p under consideration. Hence, $-y^2 - 1$ is a quadratic residue, so there exists an integer x such that

$$x^2 = -y^2 - 1 \pmod{p}. \tag{2.4}$$

This proves the lemma. ∎

Information about the volume of the ball of radius r in $\mathbb{R}^n$ is crucial for the application of Minkowski's convex body technique.

LEMMA 4. *Let $B_m(r)$ be the closed ball of radius r in $\mathbb{R}^m$ centered at the origin. Its Euclidean volume is*

$$\mathrm{vol}(B_{2n}(r)) = \frac{\pi^n}{n!} r^{2n}, \qquad \mathrm{vol}(B_{2n-1}(r)) = \frac{2^n \pi^{n-1}}{1 \times 3 \times 5 \cdots (2n-1)} r^{2n-1}$$

for $n \in \mathbb{Z}$. ∎

See Ref. [9]. The proof of the main result of this section by convex body techniques now follows (see [6]).

THEOREM 5. *Every positive integer is the sum of four rational integer squares.*

Proof. From standard arguments, see [8], it is only necessary to prove the theorem for rational integer primes. Given a prime p, choose integers a and b such that $a^2 + b^2 + 1 \equiv 0 \pmod{p}$. Consider the space $\mathbb{R}^4$ with a typical point (x, y, z, w). Let the lattice under consideration have as basis

$$\{(1, 0, a, b), (0, 1, b, -a), (0, 0, p, 0), (0, 0, 0, p)\}. \tag{2.5}$$

Hence, a typical element of the lattice has integer coordinates x, y, z, and w that satisfy

$$\begin{aligned} z &\equiv ax + by \pmod{p}, \\ w &\equiv bx - ay \pmod{p}. \end{aligned} \tag{2.6}$$

This lattice has volume p^2 from consideration of the absolute value of the determinant of a matrix whose rows are a basis (see [10]). A ball of radius r in $\mathbb{R}^4$ that contains a nonzero point of this lattice by Minkowski's theorem must satisfy

$$\frac{\pi^2}{2} r^4 \geqslant 16p^2. \tag{2.7}$$

Any r such that $r^2 > 1.81p$ will do. Hence choosing $r^2 = 1.9p$ we find that a nonzero point of the lattice exists inside the ball with equation

$$x^2 + y^2 + z^2 + w^2 \leqslant 1.9p. \tag{2.8}$$

But for this choice of integer x, y, z and w we have

$$\begin{aligned} &x^2 + y^2 + z^2 + w^2 \\ &\quad \equiv x^2 + y^2 + a^2x^2 + 2abxy + b^2y^2 + b^2x^2 - 2abxy + a^2y^2 \pmod{p} \\ &\quad \equiv x^2(1 + b^2 + a^2) + y^2(1 + b^2 + a^2) \pmod{p} \\ &\quad \equiv 0 \pmod{p}. \end{aligned} \tag{2.9}$$

Hence, $x^2 + y^2 + z^2 + w^2$ must equal p. ∎

3. SUMS OF TWO SQUARES FOR $\mathbb{Q}(\sqrt{5})$

The analogous theorems for representation by sums of squares of algebraic integers of a quadratic field require the use of higher dimensional cases of Minkowski's theorem. Let $O(\sqrt{5})$ denote the ring of algebraic integers of $\mathbb{Q}(\sqrt{5})$, and let ε be the fundamental unit $\frac{1+\sqrt{5}}{2}$. Then $O(\sqrt{5})$ is the ring $\mathbb{Z}[1,\varepsilon]$. An overline is used to represent the conjugate of a number with respect to the ring $\mathbb{Q}(\sqrt{5})$. Since the sum of squares must be totally positive, given an algebraic integer prime ρ, it is sufficient to find some unit, κ, for which $\kappa\rho$ is represented by an appropriate sum of squares. For the analog of the two-squares theorem, we must consider a lattice in $\mathbb{R}^4$.

Theorem 6. *A prime ρ of $O(\sqrt{5})$ for which -1 is a quadratic residue has a representation $\kappa\rho = x^2 + y^2$ where $x, y \in O(\sqrt{5})$ and κ is a unit.*

Proof. Let the prime ρ lie under the prime p of $\mathbb{Z}$. There are three possibilities, namely, p can ramify, p can split, or p can remain inert.

If p ramifies, then ρ must be $\sqrt{5}$ times a unit, but $\varepsilon\sqrt{5} = 1^2 + \varepsilon^2$.

If p splits, then let λ be an element of $O(\sqrt{5})$ such that $\lambda^2 \equiv -1 \pmod{\rho}$. Then $|\rho\bar{\rho}| = p$. We wish to have the following elements in the lattice under consideration:

$$(\alpha, \bar{\alpha}, \lambda\alpha + \mu\rho, \bar{\lambda}\bar{\alpha} + \overline{\mu\rho}), \tag{3.1}$$

where α and μ run through all of $O(\sqrt{5})$. The lattice with the following basis elements has precisely the above four-tuples.

$$\{(1, 1, \lambda, \bar{\lambda}), (\varepsilon, \bar{\varepsilon}, \lambda\varepsilon, \bar{\lambda}\bar{\varepsilon}), (0, 0, \rho, \bar{\rho}), (0, 0, \varepsilon\rho, \overline{\varepsilon\rho})\}. \tag{3.2}$$

The size of this lattice is the absolute value of the determinant of a basis, which is $|5\rho\bar{\rho}|$ or $5p$. A ball of radius r centered at the origin in $\mathbb{R}^4$ large enough to contain a nonzero point of the lattice by recourse to Minkowski's theorem must satisfy

$$\frac{\pi^2}{2} r^4 \geqslant 2^4 \times 5p \Rightarrow r^4 \geqslant \frac{160}{\pi^2} p = p\ 16.211\cdots \tag{3.3}$$

so $r^2 = 4.1\sqrt{p}$ suffices. A nonzero point in the lattice which exists inside this ball is of the form $(\alpha, \bar{\alpha}, \beta, \bar{\beta})$ with α and β elements of $O(\sqrt{5})$. Further

$$\alpha^2 + \beta^2 \equiv 0 \pmod{\rho} \tag{3.4}$$

so we may write for some $\kappa \in O(\sqrt{5})$

$$\alpha^2 + \beta^2 = \kappa\rho. \tag{3.5}$$

Thus

$$\alpha^2 + \bar{\alpha}^2 + \beta^2 + \bar{\beta}^2 < 4.1\sqrt{p},$$
$$\kappa\rho + \overline{\kappa\rho} < 4.1\sqrt{p}. \tag{3.6}$$

Since $\kappa\rho$ must be totally positive we may use the Arithmetic Mean–Geometric Mean inequality to conclude

$$2\sqrt{\kappa\rho\overline{\kappa\rho}} \leqslant \kappa\rho + \overline{\kappa\rho} < 4.1\sqrt{p},$$
$$\kappa\rho\overline{\kappa\rho} \leqslant (2.05)^2 p,$$
$$|\kappa\bar{\kappa}| \leqslant 4. \tag{3.7}$$

Let $\kappa = a + b\varepsilon$ where $a, b \in \mathbb{Z}$. Then

$$\begin{aligned}\kappa\bar{\kappa} &\equiv a^2 + ab - b^2 \ (\mathrm{mod}\ 5)\\ &\equiv -4a^2 - 4ab - b^2 \ (\mathrm{mod}\ 5)\\ &\equiv -(2a+b)^2 \ (\mathrm{mod}\ 5)\end{aligned} \tag{3.8}$$

but the only squares modulo 5 are $0, \pm 1$ so the only negatives of squares are again $0, \pm 1$. Hence the norm of κ, $\kappa\bar{\kappa}$, cannot equal $\pm 2, \pm 3$. Certainly the norm of κ cannot equal zero. If the norm is ± 1 then κ is a unit and we are done. Suppose it is equal to ± 4. Since 2 remains prime in $O(\sqrt{5})$, 2 must divide one of κ or $\bar{\kappa}$. This implies 2 divides each of these factors. We may write $\kappa = 2\mu$. Thus the norm of μ is ± 1 and μ is a unit. Therefore we have $\alpha^2 + \beta^2 = \kappa\rho = 2\mu\rho$. Note that

$$\begin{aligned}&\alpha^2 + \beta^2 \equiv 0 \ (\mathrm{mod}\ 2)\\ \Rightarrow\quad &(\alpha + \beta)^2 \equiv 0 \ (\mathrm{mod}\ 2)\\ \Rightarrow\quad &\alpha + \beta \equiv 0 \ (\mathrm{mod}\ 2)\end{aligned} \tag{3.9}$$

since, again, 2 is prime in $O(\sqrt{5})$. Similarly $\alpha - \beta \equiv 0 \ (\mathrm{mod}\ 2)$. Thus $(\alpha \pm \beta)/2$ are elements of $O(\sqrt{5})$. To finish, we utilize a technique of Euler.

$$\left(\frac{\alpha+\beta}{2}\right)^2 + \left(\frac{\alpha-\beta}{2}\right)^2 = \frac{\alpha^2+\beta^2}{2} = \mu\rho, \tag{3.10}$$

This gives the necessary representation.

If p remains inert in $O(\sqrt{5})$ then there are two cases besides the trivial case of 2. When $p \equiv 1 \ (\mathrm{mod}\ 4)$ then the theorem follows easily from the corresponding theorem for the rational integers. When $p \equiv -1 \pmod 4$ then there are some times when p is the sum of two squares in $\mathbb{Q}(\sqrt{5})$, for example $\varepsilon^2 + \bar{\varepsilon}^2 = 3$. Since p is assumed prime, the order of the field $O(\sqrt{5})/(p)$ is p^2 so it is reasonable that a square root of -1 will show up in certain cases. If -1 is a square modulo p then so is -5 in the ring of integers under consideration. Thus it makes sense to consider the problem of representation of these p by the quadratic form $x^2 + 5y^2$ over the rational integers.

Let v satisfy $v^2 \equiv -5 \pmod p$. We wish to show that v can be chosen as a rational integer in $O(\sqrt{5})/(p)$. Since $v^2 \equiv -5 \pmod p$ then also $\bar{v}^2 \equiv -5 \pmod p$ and hence $v^2 \equiv \bar{v}^2 \pmod p$. This implies $(v + \bar{v})(v - \bar{v}) \equiv 0 \pmod p$. Since p is assumed prime at least one of the above factors is congruent to 0 mod p. If $v + \bar{v} \equiv 0 \pmod p$ then writing $v = a + b\varepsilon$ for some rational integers a and b we have

$$a + b\varepsilon + a + b\bar{\varepsilon} \equiv 0 \pmod p \tag{3.11}$$

which gives $2a + b \equiv 0 \pmod p$. Consequently, $v \equiv a + b\varepsilon \equiv b\sqrt{5}/2 \pmod p$. However $v^2 \equiv -5 \pmod p$ so

$$\begin{aligned} -5 &\equiv \left(\frac{b\sqrt{5}}{2}\right)^2 \pmod p \\ \Rightarrow -5 &\equiv b^2 \frac{5}{4} \pmod p \\ \Rightarrow -4 &\equiv b^2 \pmod p \end{aligned} \tag{3.12}$$

which implies that -1 is a quadratic residue of $\mathbb{Z}/p\mathbb{Z}$. This is a contradiction as we have assumed that $p \equiv -1 \pmod 4$. We may conclude that $v \equiv \bar{v} \pmod p$. Thus $b \equiv 0 \pmod p$ and $v \equiv a \pmod p$, the desired result.

Let l be the multiplicative inverse of v with respect to the field $O(\sqrt{5})/(p)$. Then l can also be chosen as a rational integer modulo p. Form the lattice in $\mathbb{R}^2$ with basis $\{(1, l), (0, p)\}$. The volume of this lattice is p. We must now consider an ellipse of the form $x^2 + 5y^2 = r^2$ centered at the origin. The area of the ellipse is π times the product of the semiaxes or $\pi r^2/\sqrt{5}$. We need $\pi r^2/\sqrt{5} > 4p$ for Minkowski's theorem to apply. This necessitates $r^2 > 4\sqrt{5}p/\pi$ or $r^2 > 2.85p$.

Fix $r^2 = 2.9p$. Let the nonzero lattice point inside the circle be $m(1, l) + n(0, p)$ where m and n are integers. Then $x = m$ and $y = lm + np$.

$$x^2 + 5y^2 \equiv m^2 + 5l^2m^2 \equiv m^2(1 + 5(-5)^{-1}) \equiv 0 \pmod p \tag{3.13}$$

while

$$0<x^2+5y^2<2.9p<3p. \tag{3.14}$$

Thus for this choice of rational integral x and y we have x^2+5y^2 equal to p or $2p$. If the first, then we may write $p=x^2+(y\sqrt{5})^2$.

For the last case we have $2p=x^2+5y^2$. Taking this modulo two we see that $0\equiv x+y \pmod 2$. Let $b=y$, then $x=2a+b$ for a, b rational integers. Then

$$\begin{aligned}(a+b\varepsilon)^2+(a+b\bar{\varepsilon})^2 &= 2a^2+2ab(\varepsilon+\bar{\varepsilon})+b^2(\varepsilon^2+\bar{\varepsilon}^2)\\ &= 2a^2+2ab+3b^2\\ &= \tfrac{1}{2}(4a^2+4ab+6b^2)\\ &= \tfrac{1}{2}((2a+b)^2+5b^2)\\ &= \tfrac{1}{2}(x^2+5y^2)\\ &= \tfrac{1}{2}2p=p. \end{aligned} \tag{3.15}$$

In all cases, the prime (times an appropriate unit if necessary) is the sum of two squares from the ring of algebraic integers $O(\sqrt{5})$. ∎

It should be remarked that given a fixed $\kappa\rho$ in the theorem above, the representation as a sum of squares is essentially unique up to changes in sign and order of the x and y. The proof follows the rational case closely (see [3, pp. 119–120]) except for the step where the argument has 1 equaling the sum of two squares which immediately implies one of those squares is zero. Suppose $1=v^2+\lambda^2$ in $O(\sqrt{5})$. Write $v=a+b\varepsilon$ and $\lambda=c+d\varepsilon$ where a,b,c,d are rational integers. Then the rational part of v^2 is $a^2+ab+3b^2/2$ so we obtain

$$\begin{aligned}1 &= a^2+ab+\frac{3b^2}{2}+c^2+cd+\frac{3d^2}{2}\\ &= \left(a+\frac{b}{2}\right)^2+\frac{5b^2}{4}+\left(c+\frac{d}{2}\right)^2+\frac{5d^2}{4}\end{aligned} \tag{3.16}$$

which immediately gives $b=d=0$. Thus one of a or c is zero, so one of v or λ must be zero.

4. SUMS OF FOUR SQUARES FOR $\mathbb{Q}(\sqrt{5})$

Whereas we needed to consider a lattice in $\mathbb{R}^4$ to prove the two-squares theorem for the algebraic integers in $\mathbb{Q}(\sqrt{5})$ we must consider a lattice in $\mathbb{R}^8$ for the analogous four-squares theorem. It turns out that the eight-dimensional sphere does not give good enough bounds under Minkowski's criteria, so another region must be considered. To compute a needed volume the following lemma is useful.

LEMMA 7. *Let $B_4(r_1)$ be the closed ball of radius r_1 centered at the origin in $\mathbb{R}^4$. Let $r = \sqrt{x_1^2 + x_2^2 + x_3^2 + x_4^2}$ be the distance from a point of $\mathbb{R}^4$ to the origin. Let g be a continuous function of one variable and $G(x_1, x_2, x_3, x_4) = g(r)$. Then*

$$\int_{B_4(r_1)} G = 2\pi^2 \int_{r=0}^{r_1} g(r) r^3\, dr. \tag{4.1}$$

Proof. The four-dimensional analog of spherical coordinates is appropriate to use (see [9, pp. 427–428]),

$$\begin{cases} x_1 = r\cos(\theta_1), \\ x_2 = r\sin(\theta_1)\cos(\theta_2), \\ x_3 = r\sin(\theta_1)\sin(\theta_2)\cos(\theta_3), \\ x_4 = r\sin(\theta_1)\sin(\theta_2)\sin(\theta_3), \end{cases} \tag{4.2}$$

with $0<\theta_1, \theta_2<\pi$, $0<\theta_3<2\pi$ and $r > 0$. The Jacobian is $r^3\sin^2(\theta_1)\sin(\theta_2)$ (see [9, p. 428]). Thus

$$\begin{aligned} \int_{B_4(r_1)} G &= \int_{r=0}^{r_1} \int_{\theta_1=0}^{\pi} \int_{\theta_2=0}^{\pi} \int_{\theta_3=0}^{2\pi} g(r) r^3 \sin^2(\theta_1)\sin(\theta_2)\, dr\, d\theta_1\, d\theta_2\, d\theta_3 \\ &= 2\pi \times 2 \times \frac{\pi}{2} \int_{r=0}^{r_1} g(r) r^3\, dr \end{aligned} \tag{4.3}$$

as the integrals with respect to the thetas are simple exercises. ∎

Next, the particular region in $\mathbb{R}^8$ is described.

LEMMA 8. *The region $\mathscr{C} = \mathscr{C}(r)$ in $\mathbb{R}^8 = \{(x_1, x_2, \ldots, x_8) \mid \forall i\, x_i \in \mathbb{R}\}$ defined by*

$$\sqrt{x_1^2 + x_2^2 + x_3^2 + x_4^2} + \sqrt{x_5^2 + x_6^2 + x_7^2 + x_8^2} \leqslant r \tag{4.4}$$

is a centrally symmetric convex region with volume $\frac{\pi^4}{280} r^8$.

Proof. It is clear that this region is centrally symmetric. To show it is convex, start with the triangle inequality in $\mathbb{R}^4$ between the points $(0, 0, 0, 0)$, (w_1, w_2, w_3, w_4) and $(w_1 + z_1, w_2 + z_2, w_3 + z_3, w_4 + z_4)$. We find

$$\sqrt{(w_1 + z_1)^2 + (w_2 + z_2)^2 + (w_3 + z_3)^2 + (w_4 + z_4)^2} \\ \leqslant \sqrt{w_1^2 + w_2^2 + w_3^2 + w_4^2} + \sqrt{z_1^2 + z_2^2 + z_3^2 + z_4^2}. \tag{4.5}$$

Given two points in $\mathscr{C}$ we need only show the midpoint of the line between these points is also in $\mathscr{C}$. But if $(x_1, \ldots, x_8)$ and $(y_1, \ldots, y_8)$ are in $\mathscr{C}$, then

$$\begin{aligned} &\sqrt{\left(\frac{x_1 + y_1}{2}\right)^2 + \cdots + \left(\frac{x_4 + y_4}{2}\right)^2} + \sqrt{\left(\frac{x_5 + y_5}{2}\right)^2 + \cdots + \left(\frac{x_8 + y_8}{2}\right)^2} \\ &\leqslant \sqrt{\left(\frac{x_1}{2}\right)^2 + \cdots + \left(\frac{x_4}{2}\right)^2} + \sqrt{\left(\frac{y_1}{2}\right)^2 + \cdots + \left(\frac{y_4}{2}\right)^2} \\ &\quad + \sqrt{\left(\frac{x_5}{2}\right)^2 + \cdots + \left(\frac{x_8}{2}\right)^2} + \sqrt{\left(\frac{y_5}{2}\right)^2 + \cdots + \left(\frac{y_8}{2}\right)^2} \\ &\leqslant \tfrac{1}{2}\sqrt{x_1^2 + x_2^2 + x_3^2 + x_4^2} + \tfrac{1}{2}\sqrt{x_5^2 + x_6^2 + x_7^2 + x_8^2} \\ &\quad + \tfrac{1}{2}\sqrt{y_1^2 + y_2^2 + y_3^2 + y_4^2} + \tfrac{1}{2}\sqrt{y_5^2 + y_6^2 + y_7^2 + y_8^2} \\ &\leqslant \tfrac{1}{2}r + \tfrac{1}{2}r = r. \end{aligned} \tag{4.6}$$

This demonstration of convexity is similar to an exercise in [10, p. 147].

To find the volume, consider the problem with reference to the coordinates

$$r_1 = \sqrt{x_1^2 + x_2^2 + x_3^2 + x_4^2}, \qquad r_2 = \sqrt{x_5^2 + x_6^2 + x_7^2 + x_8^2}. \tag{4.7}$$

Then the region $\mathscr{C}(r)$ is defined by $r_1 + r_2 \leqslant r$. For each point at radial distance r_1 from the origin in the four dimensional space generated by the first four x-coordinates, there is a region of volume $\mathrm{vol}(B_4(r - r_1))$ in $\mathscr{C}(r)$. Thus the volume of $\mathscr{C}(r)$ is

$$\int_{B_4(r)} \mathrm{vol}(B_4(r - r_1)) \tag{4.8}$$

where we are integrating over a four dimensional ball of radius r in the space generated by the first four x-coordinates corresponding to r_1. By the previous lemma this volume equals

$$2\pi^2 \int_{r_1=0}^{r} \frac{\pi^2}{2}(r - r_1)^4 r_1^3 \, dr_1 = \frac{\pi^4}{280} r^8. \quad \blacksquare \tag{4.9}$$

THEOREM 9. *For any prime ρ of $O(\sqrt{5})$ there exists a unit κ and algebraic integers x, y, z, w such that $\kappa\rho = x^2 + y^2 + z^2 + w^2$ where $\kappa, x, y, z, w \in O(\sqrt{5})$.*

Proof. The demonstration breaks into cases as in the two-squares theorem. Let the prime ρ lie under the prime p of $\mathbb{Z}$. If p remains inert in $O(\sqrt{5})$, then the theorem follows trivially from the corresponding theorem for the rational integers.

If p ramifies, then ρ must be $\sqrt{5}$ times a unit, and we have already shown a representation as a sum of two squares.

Suppose p splits, then $|\rho\bar{\rho}| = p$. By Lemma 3 there exist rational integers a and b such that $a^2 + b^2 + 1 \equiv 0 \pmod{p}$. Hence $a^2 + b^2 + 1 \equiv 0 \pmod{\rho}$.

We wish to have the following elements in the lattice under consideration:

$$(\alpha, \bar{\alpha}, \beta, \bar{\beta}, a\alpha + b\beta + \mu\rho, a\bar{\alpha} + b\bar{\beta} + \overline{\mu\rho}, b\alpha - a\beta + \nu\rho, b\bar{\alpha} - a\bar{\beta} + \overline{\nu\rho}), \tag{4.10}$$

where α, β, μ and ν run through all of $O(\sqrt{5})$. The lattice with the following basis elements has precisely the above eight-tuples:

$$\left\{\begin{array}{l} (1,1,0,0,a,a,b,b), (\varepsilon,\bar{\varepsilon},0,0,a\varepsilon,a\bar{\varepsilon},b\varepsilon,b\bar{\varepsilon}), \\ (0,0,1,1,b,b,-a,-a), (0,0,\varepsilon,\bar{\varepsilon},b\varepsilon,b\bar{\varepsilon},-a\varepsilon,-a\bar{\varepsilon}), \\ (0,0,0,0,\rho,\bar{\rho},0,0), (0,0,0,0,\varepsilon\rho,\overline{\varepsilon\rho},0,0), \\ (0,0,0,0,0,0,\rho,\bar{\rho}), (0,0,0,0,0,0,\varepsilon\rho,\overline{\varepsilon\rho}) \end{array}\right\}. \tag{4.11}$$

The size of this lattice is the absolute value of the determinant of a basis, which is $|25\rho^2\bar{\rho}^2|$ or $25p^2$. Minkowski's convex body theorem using $\mathscr{C}(r)$ in $\mathbb{R}^8$ would require a parameter r such that

$$\frac{\pi^4}{280} r^8 \geqslant 2^8 \times 25p^2 \quad \Rightarrow \quad r^8 \geqslant \frac{2^8 \times 25 \times 280}{\pi^4} p^2 \tag{4.12}$$

so $r = 3.42p^{1/4}$ suffices. In this case we use the first, third, fifth and seventh coordinates as the four x-coordinates corresponding to r_1 in the previous lemma, and the remaining coordinates correspond to r_2. This again gives us algebraic integers such that

$$\alpha^2 + \beta^2 + \gamma^2 + \delta^2 = \kappa\rho, \tag{4.13}$$

where γ is the fifth coordinate and δ the seventh in the current set up. Also

$$\sqrt{\kappa\rho} + \sqrt{\overline{\kappa\rho}} \leqslant 3.42p^{1/4}, \tag{4.14}$$

so

$$
\begin{aligned}
2\sqrt{\sqrt{\kappa\rho\overline{\kappa\rho}}} &\leqslant \sqrt{\kappa\rho} + \sqrt{\overline{\kappa\rho}} < 3.42p^{1/4} \\
|\kappa\rho\overline{\kappa\rho}|^{1/4} &\leqslant (3.42/2)p^{1/4} \\
|\kappa\bar{\kappa}| &\leqslant 8.55. \qquad (4.15)
\end{aligned}
$$

Since $N(\kappa) = \kappa\bar{\kappa}$ cannot be congruent to 2 or 3 modulo 5, we are left with the possibilities ± 1, ± 4, ± 5 and ± 6 for $N(\kappa)$. However 2 and 3 remain prime in $O(\sqrt{5})$. If $N(\kappa) = \pm 6$ then 3 would divide κ and thus 9 would have to divide $\kappa\bar{\kappa}$, a contradiction.

If $N(\kappa)$ were congruent to ± 4, then the prime 2 would divide κ so κ would have to be 2 times a unit. Taking Eq. (4.13) modulo 2 we have the sum of four squares congruent to $0 \bmod 2$. If two of the variables are congruent modulo 2 then we may use Euler's technique again as in Eqs. (3.9) and (3.10). For example, if $\alpha \equiv \beta \pmod 2$ then $\alpha^2 + \beta^2 \equiv 0 \pmod 2$ so it follows that $\gamma^2 + \delta^2 \equiv 0 \pmod 2$. Thus by the argument used in Eq. (3.9), $(\alpha \pm \beta)/2$ and $(\gamma \pm \delta)/2$ are algebraic integers. Just like in (3.10)

$$
\begin{aligned}
&\left(\frac{\alpha+\beta}{2}\right)^2 + \left(\frac{\alpha-\beta}{2}\right)^2 + \left(\frac{\gamma+\delta}{2}\right)^2 + \left(\frac{\gamma-\delta}{2}\right)^2 \\
&\quad = \frac{\alpha^2+\beta^2}{2} + \frac{\gamma^2+\delta^2}{2} = \frac{\kappa}{2}\rho. \qquad (4.16)
\end{aligned}
$$

Thus ρ times a unit is a sum of four squares.

The other case is if the values of the four variables α, β, γ and δ are all distinct modulo 2. In this case since 2 is prime and the order of the field $O(\sqrt{5})/(2)$ is 4, we may take as a set of residue classes $\{0, 1, \varepsilon, \bar{\varepsilon}\}$. The four values can therefore be written as $2v_1 + 0$, $2v_2 + 1$, $2v_3 + \varepsilon$ and $2v_4 + \bar{\varepsilon}$ where all the v's are algebraic integers. Taking the sum of four squares gives

$$
\begin{aligned}
&4v_1^2 + (4v_2^2 + 4v_2 + 1) + (4v_3^2 + 4v_3\varepsilon + \varepsilon^2) + (4v_4^2 + 4v_4\bar{\varepsilon} + \bar{\varepsilon}^2) \\
&\quad \equiv 1^2 + \varepsilon^2 + \bar{\varepsilon}^2 \pmod 4 \equiv 0 \pmod 4. \qquad (4.17)
\end{aligned}
$$

This means that two times a unit times a prime is a multiple of 4 in $O(\sqrt{5})$. Therefore the prime itself must equal 2 times a unit, and we know 2 is the sum of four squares.

Finally the case of $N(\kappa)$ equaling ± 5 must be considered. In this case, κ is divisible by the prime $\sqrt{5}$, so κ must equal $\sqrt{5}$ times a unit. Thus we have the sum of four algebraic integer squares congruent to zero modulo the prime $\sqrt{5}$. The ring $O(\sqrt{5})/(\sqrt{5})$ is a field of order five, and there is an isomorphism between the residue class ring and $\mathbb{Z}/5\mathbb{Z}$ as there is only one field of any finite order up to isomorphism. The only squares in $\mathbb{Z}/5\mathbb{Z}$ are $0, 1, -1$. To have the sum of four squares equal to zero means that all four are zero, or the nonzero squares pair up with a -1 canceling each square that is one. In any event, we may break the four squares into two pairs where the sum of both squares in the pair is zero. Pulling back to $O(\sqrt{5})/(\sqrt{5})$ we have the four squares broken into two pairs each of whose sums is congruent to $0 \bmod \sqrt{5}$. Another result of Euler is now invoked (see [4, p. 46]):

Euler's Lemma. *Consider the problem of representations by sums of two squares in a ring of algebraic integers which is a principal ideal domain (and thus a unique factorization domain). If an algebraic integer which is a sum of two squares is divisible by a prime which is a sum of two squares, then the quotient is a sum of two squares.*

We have shown that $\varepsilon\sqrt{5}$ is a sum of two squares. Applying this lemma to each of the pairs of squares we just constructed, we find that $(\kappa/(\varepsilon\sqrt{5}))\rho$ is a sum of four squares. Thus ρ times an appropriate unit is a sum of four squares.

This completes the proof of the theorem. ∎

Theorem 10. *Any totally positive integer of $O(\sqrt{5})$ is the sum of four squares of integers from that ring.*

Proof. Since $O(\sqrt{5})$ is a principal ideal domain, and a unique factorization domain, we may write an arbitrary totally positive integer α as a product of primes

$$\alpha = \rho_1 \rho_2 \cdots \rho_m. \tag{4.18}$$

If ρ_1 is totally positive, divide both sides by ρ_1 to get a totally positive integer on the left equal to a product of primes on the right. If ρ_1 and its conjugate are both less than zero, multiply ρ_1 and ρ_2 each by negative one to force ρ_1 to be totally positive and divide both sides by the new ρ_1. If ρ_1 is positive and its conjugate is negative, then multiply ρ_1 by ε while multiplying ρ_2 by ε^{-1}, and dividing as before. If the signs of ρ_1 and its conjugate are the reverse use $-\varepsilon$ as a factor.

In all events, we get a totally positive integer on the left equal to a product of $m-1$ primes on the right. Proceed by induction until there is only one prime remaining on the right. This prime must therefore be totally positive. Thus we may write α as a product of totally positive primes.

Given any adjusted prime ρ in our factorization of α, by the previous theorem we know there exists a unit κ such that $\kappa\rho$ is the sum of four squares. Thus $\kappa\rho$ and ρ are both totally positive, so κ must be a totally positive unit. Since all units of $O(\sqrt{5})$ are of the form $\pm\varepsilon^n$ and $\varepsilon > 0$ while $\bar{\varepsilon} < 0$, κ must be of the form ε^{2n}. Dividing the representation of $\kappa\rho$ as the sum of four squares by ε^{2n} we obtain a representation of ρ as the sum of four squares of integers from $O(\sqrt{5})$. Hence by the norm multiplication identity for quaternions, α itself can be written as the sum of four squares. ∎

5. OTHER QUADRATIC FIELDS

The convex body theorem is very dependent on the size of the lattices involved. These lattices themselves have sizes dependent upon the discriminant of the ring of integers of the underlying quadratic field. With fields of larger discriminants, use of the convex body theorem becomes problematic and involved. Furthermore, it is not possible for all the totally positive integers in quadratic fields like $\mathbb{Q}(\sqrt{2})$ or $\mathbb{Q}(\sqrt{3})$ to be representable as sums of squares since the coefficient of $\sqrt{2}$ or $\sqrt{3}$ must be even. It is uncertain if the convex body theorem can be applied in these cases in as straightforward a manner as in $\mathbb{Q}(\sqrt{5})$.

6. COMPUTATIONS

Certain auxiliary calculations, especially the computation of the determinants of lattices, were performed with PUNIMAX, a variant of the MAXIMA computer algebra system. The computer involved had 32 Mb of RAM, a Pentium 133 cpu and was running LINUX 2.0.35.

ACKNOWLEDGMENTS

The author expresses his appreciation to the Mathematics Department of the University of Botswana for the use of their computing facilities. The author thanks B. Haible, the maintainer of PUNIMAX, for permitting its free use for academic purposes [7], and also Harvey Cohn for many encouraging e-mail communications.

REFERENCES

1. H. Cohn, Decomposition into four integral squares in the fields of $2^{\frac{1}{2}}$ and $3^{\frac{1}{2}}$, *Amer. J. Math.* **82** (1960), 301–322.
2. H. Cohn, Calculation of class numbers by decomposition into three integral squares in the fields of $2^{\frac{1}{2}}$ and $3^{\frac{1}{2}}$, *Amer. J. Math.* **83** (1961), 33–56.
3. H. Davenport, "The Higher Arithmetic," Dover, New York, 1983.

4. H. Edwards, "Fermat's Last Theorem," Springer-Verlag, New York, 1977.
5. F. Götzky, Über eine zahlentheoretische Anwendung von Modulfunktionen zweier Veränderlicher, *Math. Ann.* **100** (1928), 411–437.
6. J. Grace, The four square theorem, *J. London Math. Soc.* **2** (1927), 3–8.
7. B. Haible, Private communication (1997).
8. G. Hardy, H. Wright, "Introduction to the Theory of Numbers," 4th ed., Oxford Univ. Press, London, 1971.
9. S. Lang, "Analysis I," Addison–Wesley, Reading, MA, 1968.
10. D. Marcus, "Number Fields," Springer-Verlag, New York, 1977.