The Decomposition of Integers as the Sums of Squares in General Number Fields

Aaron Tievsky

under the direction of
Payman Kassaei
Massachusetts Institute of Technology

Research Science Institute
August 5, 1998

Abstract

In this paper, we looked to characterize which elements of integer rings of quadratic and more general number fields could be expressed as the sums of squares. This is a well-known problem in $\mathbb{Z}$, but it has not been solved in more complex cases. We completely characterized these elements in the cases of $\mathbb{Z}[i]$ and $\mathbb{Z}[\frac{1+\sqrt{-3}}{2}]$. For a general number field K, we were able to determine a general form for elements expressible as the sums of two squares by considering the ideals of the maximal order of K and $K(i)$.

1 Introduction

Expressing an integer as the sum of squares is a classical problem of number theory. It has long been known that there exist integers x and y such that $x^2+y^2 = p$ for any prime p which is congruent to 1 modulo 4 and for $p = 2$. If $p \equiv 3 \pmod 4$, there is no such solution. Any non-negative integer which is a multiple of a prime congruent to 3 modulo 4 and contains an odd power of this prime cannot be expressed as the sum of two squares, while all other non-negative integers can. One can also show that every integer can be expressed as the sum of four squares [1]. It is not well-known, however, which elements of other integer rings can be expressed as the sum of squares.

A number field is a field extension of $\mathbb{Q}$ of finite dimension. The maximal order (integer ring) $\mathcal{O}_K$ of a number field K is the ring of elements α in the field such that $f(\alpha) = 0$ for some monic polynomial function f in $\mathbb{Z}[x]$. In this paper, we will analyze which integers in various number fields can be written as the sums of two squares.

First we will use the unique factorization of ideals into prime ideals to examine the case of a general number field which does not contain i. Then we will consider the example of $\mathbb{Z}[\frac{1+\sqrt{-3}}{2}]$. Lastly, we will look into the unique case when i is an element of the number field.

2 Statement of the Problem

The clasical problem is formulated as follows:

For which values of $a \in \mathbb{Z}$ does $a = x^2 + y^2$ have a solution for $x, y \in \mathbb{Z}$?

This is equivalent to finding all values of a such that $a = N(\alpha)$ for some $\alpha \in \mathbb{Z}[i]$. In our extension of the problem to a general number field K, it is convenient to express a as the norm of some element. We will use ideals in our investigation, and hence we must reformulate the problem:

For which ideals $A \subset \mathcal{O}_K$ does $A = N_{K(i)/K}(I)$ for some ideal I in $\mathcal{O}_{K(i)}$?

3 Decomposition when $i \notin K$

3.1 The General Case

In the general case, the maximal order $\mathcal{O}_K$ of field K as well as $\mathcal{O}_{K(i)}$ fail to have unique factorization into primes. Thus we will consider the ideals of $\mathcal{O}_K$ and $\mathcal{O}_{K(i)}$. We intend to find the image of $N_{K(i)/K} : D_{K(i)} \to D_K$ where D_L is the semigroup of ideals of $\mathcal{O}_L$ (see appendix A).

Theorem 3.1.1: If $\wp$ is a prime ideal in $\mathcal{O}_K$ then $\wp = N_{K(i)/K}(\mathbf{I})$ for some ideal $\mathbf{I}$ in $\mathcal{O}_{K(i)}$ iff $\wp\mathcal{O}_{K(i)}$ is not a prime ideal of $\mathcal{O}_{K(i)}$.

Proof: Assume $\wp\mathcal{O}_{K(i)}$ is not a prime ideal in $\mathcal{O}_{K(i)}$. Then $\wp\mathcal{O}_{K(i)} = \mathbf{IJ}$ where $\mathbf{I}$ and $\mathbf{J}$ are non-unit ideals of $\mathcal{O}_{K(i)}$. $N_{K(i)/K}(\wp\mathcal{O}_{K(i)}) = \wp^{dim(K(i)/K)} = \wp^2$. Since $\mathbf{I}$ and $\mathbf{J}$ are not units, $N_{K(i)/K}(\mathbf{I}) = N_{K(i)/K}(\mathbf{J}) = \wp$. $\mathbf{I}$ and $\mathbf{J}$ cannot be the product of non-unit ideals because then their norms could not be prime. Thus $\wp$ is the norm of a prime ideal of $\mathcal{O}_{K(i)}$. Now assume $\wp = N_{K(i)/K}(\mathbf{I})$. $N_{K(i)/K}(\mathbf{I})\mathcal{O}_{K(i)} = \mathbf{I}\overline{\mathbf{I}}$, and therefore $\wp\mathcal{O}_{K(i)}$ is not a prime ideal. This completes the proof.

Lemma 3.1.1: If $\wp$ is a prime ideal of $\mathcal{O}_K$, then $\wp| < x^2+1 >_{\mathcal{O}_K}$ for some $x \in \mathcal{O}_K$ iff $N_{K/\mathbb{Q}}(\wp) =< m >_{\mathbb{Z}}$ for some $m \in \mathbb{Z}_+$ such that $m \equiv 1 \pmod 4$ or $2|m$.

Proof: If $\wp$ is a prime ideal in $\mathcal{O}_K$, then $\mathcal{O}_K/\wp$ is a field and hence $\mathcal{O}_K/\wp - \{0\}$ is a cyclic multiplicative group. If $N_{K/\mathbb{Q}}(\wp) =< m >_{\mathbb{Z}}$ then the order of this group is $m-1$. If $m \equiv 1 \pmod 4$, then take $x = g^{\frac{N_{K/\mathbb{Z}}(\wp)-1}{4}}$ where g is a representative of the class which generates $\mathcal{O}_K/\wp - \{0\}$. Then $x^2+1 \in \wp$, and so $\wp| < x^2+1 >_{\mathcal{O}_K}$. If $2|m$ then take $x = 1$. To prove the converse, assume that $\wp| < x^2+1 >_{\mathcal{O}_K}$. If $-1 \not\equiv 1$, then 4 divides the order of x. Since the order of x must divide the size of $\mathcal{O}_K/\wp - \{0\}$, $4|(m-1)$ and so $m \equiv 1 \pmod 4$. If $-1 \equiv 1$ then $2 \in \wp$ and $N_{K/\mathbb{Q}}(\wp) = 2^f$ for some $f > 0$. Hence $2|m$. Thus the proof is complete.

Now we must characterize which prime ideals of $\mathcal{O}_K$ remain prime in $\mathcal{O}_{K(i)}$.

Theorem 3.1.2: Let $\wp$ be a prime ideal of $\mathcal{O}_K$. $\wp\mathcal{O}_{K(i)}$ is a prime ideal of $\mathcal{O}_{K(i)}$ iff $N_{K/\mathbb{Q}}(\wp) =< m >_{\mathbb{Z}}$ where $m \equiv 3 \pmod 4$.

Proof: Assume $N_{K/\mathbb{Q}}(\wp) =< m >_{\mathbb{Z}}$ where $m \not\equiv 3 \pmod 4$. By lemma 4.2, one can find $x \in \mathcal{O}_K$ such that $\wp| < x^2+1 >_{\mathcal{O}_K}$ and so $\wp\mathcal{O}_{K(i)}| < x+i >_{\mathcal{O}_{K(i)}} < x-i >_{\mathcal{O}_{K(i)}}$. Suppose that $\wp\mathcal{O}_{K(i)}$ is a prime ideal of $\mathcal{O}_{K(i)}$. Then $\wp\mathcal{O}_{K(i)}| < x \pm i >_{\mathcal{O}_{K(i)}}$. This means that $(x \pm i) \in \wp\mathcal{O}_{K(i)}$. If $\wp$ has generators $\alpha_1, \alpha_2, ...\alpha_n$, then $x \pm i = (c_1\alpha_1 + c_2\alpha_2 + ... + c_n\alpha_n) + (d_1\alpha_1 + d_2\alpha_2 + ... + d_n\alpha_n)i$ where all c_i and d_j are in $\mathcal{O}_K$. But then $(d_1\alpha_1 + d_2\alpha_2 + ... + d_n\alpha_n) = \pm 1 \in \wp$ and $\wp = \mathcal{O}_K$, which cannot be true since $\mathcal{O}_K$ is not prime. This creates a contradiction, and so $\wp\mathcal{O}_{K(i)}$ cannot be prime.

Now assume $N_{K/\mathbb{Q}}(\wp) =< m >_{\mathbb{Z}}$ where $m \equiv 3 \pmod 4$ and $m > 0$, and suppose that $\wp\mathcal{O}_{K(i)}$ is not a prime ideal of $\mathcal{O}_{K(i)}$. By Theorem 3.1.1, $\wp = N_{K(i)/K}(\mathbf{I})$ where $\mathbf{I}$ is a

prime ideal of $\mathcal{O}_{K(i)}$. Then $N_{K(i)/\mathbb{Q}}(\mathbf{I}) = N_{K/\mathbb{Q}}(N_{K(i)/K}(\mathbf{I})) = N_{K/\mathbb{Q}}(\wp) =< m >_{\mathbb{Z}}$. Also, $\mathbb{Q} \subset \mathbb{Q}(i) \subset K(i)$ and hence by the same argument $N_{K(i)/\mathbb{Q}}(\mathbf{I}) = N_{\mathbb{Q}(i)/\mathbb{Q}}(N_{K(i)/\mathbb{Q}(i)}(\mathbf{I}))$.

By definition, $N_{K(i)/\mathbb{Q}(i)}(\mathbf{I}) = P^f$ for some prime ideal P of $\mathbb{Z}[i]$ and $f \in \mathbb{N}$. Therefore $N_{K(i)/\mathbb{Q}}(\mathbf{I}) = (N_{\mathbb{Q}(i)/\mathbb{Q}}(P))^f$. $N_{\mathbb{Q}(i)/\mathbb{Q}}(P) =< d >_{\mathbb{Z}}$ where $d > 0$. If P lies above a prime p in Z which is congruent to 3 (mod 4), $d = p^2 \equiv 1 \pmod 4$. Thus in any case $d \equiv 1 \pmod 4$ or $2|d$. Therefore $N_{K(i)/\mathbb{Q}}(\mathbf{I}) =< d^f >_{\mathbb{Z}}$ and $d^f \equiv 1 \equiv m \pmod 4$. This contradicts the initial assumption, and so $\wp\mathcal{O}_{K(i)}$ is a prime ideal of $\mathcal{O}_{K(i)}$ and the proof is complete.

Theorem 3.1.3: (Characterization of the Image of $N_{K(i)/K}$) Let $\mathbf{A} \subset \mathcal{O}_K$ be an ideal. Then $\mathbf{A} = N_{K(i)/K}(\mathbf{I})$ for some $\mathbf{I}$ of $\mathcal{O}_{K(i)}$ if and only if

$$\mathbf{A} = \prod_{i=1}^{k} \mathbf{P}_i^{\alpha_i} \prod_{j=1}^{\ell} \mathbf{Q}_j^{2\beta_j},$$

where $\mathbf{P}_i$ is a prime ideal of $\mathcal{O}_K$ such that $N_{K/\mathbb{Q}}(\mathbf{P}_i) =< m_i >_{\mathbb{Z}}$ and $m_i \equiv 1 \pmod 4$ or m_i is even ($m_i > 0$), and $\mathbf{Q}_i$ is a prime ideal of $\mathcal{O}_K$ such that $N_{K/\mathbb{Q}}(\mathbf{Q}_i) =< n_i >_{\mathbb{Z}}$ and $n_i \equiv 3 \pmod 4$ ($n_i > 0$).

Proof: Any ideal $\mathbf{A}$ over $\mathcal{O}_K$ can be decomposed as the product of prime ideals as follows:

$$\mathbf{A} = \prod_{i=1}^{k} \mathbf{P}_i^{\alpha_i} \prod_{j=1}^{\ell} \mathbf{Q}_j^{2\beta_j+\epsilon_j},$$

with $\epsilon_i = 0$ or 1. Suppose $\mathbf{A} = N_{K(i)/K}(\mathbf{I})$ for some $\mathbf{I}$ and $\epsilon_k = 1$ for some k. $\mathbf{Q}_k^{2\beta_k+1}\mathcal{O}_{K(i)}|\mathbf{I}\overline{\mathbf{I}}$. Since $\mathbf{Q}_k\mathcal{O}_{K(i)}$ is prime, it follows that $\mathbf{Q}_k^{\beta_k+1}\mathcal{O}_{K(i)}|\mathbf{I}$ or $\overline{\mathbf{I}}$. Without loss of generality, choose $\mathbf{I}$. Then $\overline{\mathbf{Q}_k^{\beta_k+1}\mathcal{O}_{K(i)}} = \mathbf{Q}_k^{\beta_k+1}\mathcal{O}_{K(i)}|\overline{\mathbf{I}}$ and $\mathbf{Q}_k^{2\beta_k+2}|\mathbf{A}$. This result is contrary to the prime factorization of $\mathbf{A}$, and so $\epsilon_i = 0$ for all i.

If all $\epsilon_i = 0$, we notice that $\mathbf{P}_i = N(I_i)$ for some ideal I_i and $\mathbf{Q}_i^2 = N(\mathbf{Q}_i \mathcal{O}_{K(i)})$. Since the norm is multiplicative, $\mathbf{A}$ must be the norm of some ideal in $\mathcal{O}_{K(i)}$ and the theorem is proven.

3.2 Decomposition when $\mathcal{O}_{K(i)}$ is a Principal Ideal Domain

Now assume $\mathcal{O}_{K(i)}$ is a principal ideal domain (i.e. all ideals in $\mathcal{O}_{K(i)}$ are principal). Let $a \in \mathcal{O}_K$ and $< a >= N(\mathbf{I})$ for some $\mathbf{I}$ of $\mathcal{O}_{K(i)}$. We distinguish two different cases:

1) $< 2 >$ is unramified in $\mathcal{O}_K$. Then all elements of $\mathcal{O}_{K(i)}$ can be written as $x + yi$ where $x, y \in \mathcal{O}_K$ (see appendix B for proof), and

2) $< 2 >$ is ramified in $\mathcal{O}_K$. Then all elements of $\mathcal{O}_{K(i)}$ can be written as $\frac{x}{2} + \frac{y}{2}i$ where $x, y \in \mathcal{O}_K$ and $x \equiv y \pmod 2$ (see appendix B for proof).

Theorem 3.2.1:

1) If $< 2 >$ is unramified in $\mathcal{O}_K$, then $< a >= N(\mathbf{I})$ iff ϵa is the sum of two squares where ϵ is some unit in $\mathcal{O}_K$.

2) If $< 2 >$ is ramified in $\mathcal{O}_K$, then $< a >= N(\mathbf{I})$ iff $4\epsilon a$ is the sum of two squares where ϵ is some unit in $\mathcal{O}_K$.

Proof:

1) By assumption, $\mathbf{I} =< \alpha >_{\mathcal{O}_{K(i)}}$ where $\alpha = x + yi$ for $x, y \in \mathcal{O}_K$. Hence $< a >_{\mathcal{O}_K} =< N_{K(i)/K}(\alpha) >_{\mathcal{O}_K} =< x^2 + y^2 >_{\mathcal{O}_K}$. This implies that there exists a unit ϵ of $\mathcal{O}_K$ such that $\epsilon a = x^2 + y^2$. Now suppose that $\epsilon a = x^2 + y^2$. Then $< a >_{\mathcal{O}_K} =< N(x + yi) >_{\mathcal{O}_K} = N(< x + yi >_{\mathcal{O}_{K(i)}})$ and $x + yi \in \mathcal{O}_{K(i)}$. Thus $< a >$ is the norm of an ideal in $\mathcal{O}_{K(i)}$.

2) By assumption, $\mathbf{I} =< \alpha >_{\mathcal{O}_{K(i)}}$ where $\alpha = \frac{x}{2} + \frac{y}{2}i$ for $x, y \in \mathcal{O}_K$. Hence $< a >_{\mathcal{O}_K} =<$

$N_{K(i)/K}(\alpha) >_{\mathcal{O}_K} =< \frac{x^2+y^2}{4} >_{\mathcal{O}_K}$. This shows that there exists a unit ϵ of $\mathcal{O}_K$ such that $4\epsilon a = x^2 + y^2$. Now suppose that $4\epsilon a = x^2 + y^2$. Then $< a >_{\mathcal{O}_K} =< N(\frac{x}{2} + \frac{y}{2}i) >_{\mathcal{O}_K} = N(< \frac{x}{2} + \frac{y}{2}i >_{\mathcal{O}_{K(i)}})$. Since $4|(x^2 + y^2), \frac{x}{2} + \frac{y}{2}i \in \mathcal{O}_{K(i)}$. Thus $< a >$ is the norm of an ideal in $\mathcal{O}_{K(i)}$.

Theorem 3.2.2: Suppose $\mathcal{O}_{K(i)}$ is a principal ideal domain. Let $a \in \mathcal{O}_K$ and $< a >= P_1^{\alpha_1} P_2^{\alpha_2} ... P_k^{\alpha_k} Q_1^{\beta_1} Q_2^{\beta_2} ... Q_\ell^{\beta_\ell}$ where $N(P_i) =< m_i >$ for positive m_i and m_i is even or $\equiv 1 \pmod 4$, and $N(Q_i) =< n_i >$ for positive n_i and $n_i \equiv 3 \pmod 4$. Then:

1) If $< 2 >$ is unramified in $\mathcal{O}_K$, for some unit ϵ in $\mathcal{O}_K$ the quantity ϵa is representable as the sum of two squares iff $\beta_1, \beta_2, ...\beta_\ell$ are all even.

2) If $< 2 >$ is ramified in $\mathcal{O}_K$, for some unit ϵ in $\mathcal{O}_K$ the quantity $4\epsilon a$ is representable as the sum of two squares iff $\beta_1, \beta_2, ...\beta_\ell$ are all even.

Proof: By Theorem 3.1.3, $< a >= N_{K(i)/K}(I)$ for some ideal I in $\mathcal{O}_{K(i)}$ iff all β_i are even. By Theorem 3.2.1, $< a >= N_{K(i)/K}(I)$ iff ϵa is the sum of two squares in the case that $< 2 >$ is unramified or $4\epsilon a$ is the sum of two squares in the case that it is ramified. Thus the theorem is proven.

3.3 Example: $\mathbb{Z}[\frac{1+\sqrt{-3}}{2}]$

Consider the quadratic number field $\mathbb{Q}(\sqrt{-3})$. For the sake of convenience, let $\rho = \frac{1+\sqrt{-3}}{2}$. One can show that the maximal order of this number field is $\mathbb{Z}[\rho][3]$ and that the maximal order of $\mathbb{Q}(\sqrt{-3}, i)$ is $\mathbb{Z}[\rho, i]$ (see appendix C). It can also be shown that both $\mathbb{Z}[\rho]$ and $\mathbb{Z}[\rho, i]$ are Euclidean and are therefore principal ideal domains and that $< 2 >$ is unramified in $\mathbb{Z}[\rho]$ (see appendix C).

We can easily show that the units in $\mathbb{Z}[\rho]$ are $\pm 1, \pm\rho$, and $\pm\overline{\rho}$, and all of these units are expressible in $\mathbb{Z}[\rho]$ (see appendix C). From this result we can prove the following theorem:

Theorem 3.3.1: If $a \in \mathbb{Z}[\rho]$ and ϵ is a unit in $\mathbb{Z}[\rho]$, ϵa is the sum of two squares if and only if a is the sum of two squares.

Proof: Let $\epsilon a = x^2 + y^2$ for $x, y \in \mathbb{Z}[\rho]$. One can express a as $\epsilon^{-1}(x^2 + y^2)$. Since ϵ^{-1} is a unit of $\mathbb{Z}[\rho]$, it is expressible as the sum of two squares and (by Theorem A.6) so is a.

Therefore Theorem 3.2.2 gives the following corollary:

Corollary 3.3.1: (Characterization of expressible elements in $\mathbb{Z}[\rho]$) Let $a \in \mathbb{Z}[\rho]$. If $a = p_1^{\alpha_1} p_2^{\alpha_2} ... p_k^{\alpha_k} q_1^{\beta_1} q_2^{\beta_2} ... q_\ell^{\beta_\ell}$ where p_i are prime elements of $\mathbb{Z}[\rho]$ such that $N(p_i) \equiv 1 \pmod 4$ or $p = 2$ and q_i are prime elements of $\mathbb{Z}[\rho]$ such that $N(q_i) \equiv 3 \pmod 4$, then a is representable as the sum of two squares iff all β_i are even.

4 Decomposition when $i \in K$

4.1 The General Case

When $i \in K$, K can be expressed as $K'(i)$ for some number field K'. Thus any element α of $\mathcal{O}_K$ can be written as $\alpha_0 + \alpha_1 i$ where $\alpha_0, \alpha_1 \in \mathcal{O}_{K'}$. For the sake of convenience, we call an element of $\mathcal{O}_K$ *even* if 2 divides the element and *odd* otherwise. We will call an element *expressible* if it can be written as the sum of two squares.

One can show that all irreducible elements of $\mathcal{O}_K$ which are of the form $(2k + \epsilon^2) + 2\ell i$, where $k, \ell \in \mathcal{O}_K$ and ϵ is a unit of $\mathcal{O}_{K'}$, can be expressed as a sum of two squares (see appendix D for proof).

In the special case of $K' = \mathbb{Q}$, the division algorithm in $\mathbb{Z}$ suggests that every odd element of $\mathcal{O}_{K'}$ can be written as $(2k + \epsilon^2)$ and so

Corollary 4.1.1: An irreducible element π in $\mathbb{Z}[i]$ is expressible iff π is an odd integer plus an even integer times i.

Now we will look to characterize reducible elements of $\mathcal{O}_K$ which are the sums of squares. Since the product of two expressible elements is also expressible (see appendix A), we have shown that

Theorem 4.1.1: For a number field K', an element a of $\mathcal{O}_{K'(i)}$ expressible if

$$a = \prod_{j=1}^{n} ((2k_j + \epsilon_j^2) + 2\ell_j i)$$

where k_j and ℓ_j are in $\mathcal{O}_{K'}$ and ϵ_j is a unit in $\mathcal{O}_{K'}$. Note: there may also be other elements of $\mathcal{O}_K$ which are the sum of two squares.

Lemma 4.1.1: For any $\alpha \in \mathcal{O}_{K'}$, $4|\alpha$ implies that α is the sum of two squares.

Proof: Let $\alpha = 4x$. Since $4x = (x + 1)^2 + (ix - i)^2$, α is the sum of two squares.

Lemma 4.1.2: For any $x \in \mathcal{O}_{K'}$, $2|(x + 1)$ implies that x is the sum of two squares.

Proof: If $2|(x + 1)$, then $\frac{x+1}{2} \in \mathcal{O}_{K'}$. Since $x = (\frac{x+1}{2})^2 + (\frac{xi-i}{2})^2$, x is the sum of two squares.

Now let us consider the special case where $K' = \mathbb{Q}$. In this case we can obtain a stronger result.

4.2 Decomposition in $\mathbb{Z}[i]$

If $\alpha \in \mathbb{Z}[i]$ and $\alpha = a+bi$ where a is odd and b is even, $2|((a+bi)+1)$. Therefore by Lemma 4.1.2, we obtain the following theorem:

Theorem 4.2.1: Gaussian integer $a+bi$ where a is odd and b is even can be written as the sum of two squares.

Now consider a Gaussian integer $\alpha = a+bi$ where a and b are both even. This can be expressed as $2^m x + 2^n yi$ where x and y are odd. One can show that α is the sum of two squares if $m \neq n$ (see appendix D).

In the case where $m = n$, first consider $m = n = 1$. If $\alpha = (r_1 + r_2 i)^2 + (r_3 + r_4 i)^2$, then $\alpha = (r_1^2 + r_3^2) - (r_2^2 + r_4^2) + 2(r_1 r_2 + r_3 r_4)i$. Since $r_1 r_2 + r_3 r_4$ must be odd, we may assume without loss of generality that $r_1, r_2 \equiv 1 \pmod 2$. This means that $r_1^2 \equiv r_2^2 \pmod 4$, and so $a \equiv r_3^2 - r_4^2 \pmod 4$. Since $m = 2$, $a \equiv 2 \pmod 4$, and so $r_3^2 - r_4^2 \equiv 2 \pmod 4$. This equation has no solutions, so $m = n = 1$ implies that α is not the sum of two squares.

If $m = n \geq 2$, $4|\alpha$. Therefore by Lemma 4.1.1 α is the sum of two squares.

One can extend these arguments to prove that any Gaussian integer α with an even coefficient of i is the sum of three squares (see appendix D for proof). Thus we have shown

Theorem 4.2.2: Gaussian integer $\alpha = a+bi$ is expressible if and only if b is even and $a \equiv b \pmod 4$ implies that $a \not\equiv 2 \pmod 4$. Moreover, α is the sum of three squares iff b is even.

5 Discussion / Conclusions

If $i \in K$, integers which are a product of irreducible elements of the form $(\epsilon^2 + 2k) + 2\ell i$ can all be written as the sum of two squares. In the special case of $\mathbb{Z}[i]$, all integers of the form $a + 2bi$ where $a, b \in \mathbb{Z}$ are expressible as long as a and $2b$ are not both congruent to 2 (mod 4). All integers $a + 2bi$ can be written as the sum of three squares. If the maximal order of $K(i)$ is a principal ideal domain, a unit multiple of an element or 4 times a unit multiple is the sum of two squares if and only if it contains even powers of any prime whose norm is congruent to 3 (mod 4). In the specific case of $\mathbb{Z}[\frac{1+\sqrt{-3}}{2}]$, this result can be simplified and the units can be ignored. For general integer ring $\mathcal{O}_K$, a similar but weaker result holds: any element which is the sum of two squares must meet the criteria described when $\mathcal{O}_{K(i)}$ is a principal ideal domain, but not all such elements are the sum of two squares (e.g. a product of irreducible elements of $\mathcal{O}_{K[i]}$). In my future research into this problem, I will attempt to characterize more fully which elements of $\mathcal{O}_K$ are irreducible but not prime in $\mathcal{O}_{K(i)}$ and I will try to find a set in which the maximal order of the quaternion algebra is Euclidean. This will enable me to continue my investigation of this problem and expand it to the problem of expressing integers as the sums of four squares.

6 Ackowledgments

I would like to thank Payman Kassaei for his help and guidance, Professor Hartley Rogers for overseeing my mentorship, Dr. John Rickert for his assistance, and the Center for Excellence in Education for providing me with this opportunity.

References

[1] Niven, I. and Zuckerman, H.: *An Introduction to the Theory of Numbers*. New York, Wiley 1960.

[2] Burton, David M.: *A First Course in Rings and Ideals*. Reading, MA: Addison-Wesley, 1970.

[3] Samuel, P. *Theorie Algebrique des Nombres*. Paris: Hermann, 1971.

A Principles of Number Fields

For proofs, see [3].

Definition A.1: A *number field* is a field extension of $\mathbb{Q}$ of finite dimension.

Definition A.2: $I \subset \mathcal{O}_K$ is an *ideal* if it is closed under addition and $x \in \mathcal{O}_K, a \in I$ implies that $ax \in I$.

Definition A.3: The ideal generated by the set $S \subset \mathcal{O}_K$ is $\{a_1x_1 + a_2x_2 + ... + a_nx_n | a_i \in S, x_i \in \mathcal{O}_K\}$. If $S = \{a_1, a_2, ...a_n\}$, then this ideal is denoted by $< a_1, a_2, ...a_n >_{\mathcal{O}_K}$. We say an ideal is *principal* if $I =< a >$ for some $a \in \mathcal{O}_K$.

Definition A.4: If $K \subset L$ is a number field and $I \subset \mathcal{O}_K$ is an ideal, define $I\mathcal{O}_L$ to be the ideal generated by I in $\mathcal{O}_L$.

Definition A.5: Ideal $P \subset \mathcal{O}_K$ is called a *prime ideal* if $xy \in P$ implies that x or y is in P. If P is a prime ideal generated by p, then p is called a *prime element* of $\mathcal{O}_K$.

Definition A.6: If I and J are ideals of $\mathcal{O}_K$, we define IJ to be the ideal generated by the set $\{ij | i \in I, j \in J\}$.

Lemma A.1: If P is a prime ideal, $IJ \subset P$ implies that $I \subset P$ or $J \subset P$.

Definition A.7: The set of all ideals of $\mathcal{O}_K$ forms a semigroup under multiplication with the unit ideal $\mathcal{O}_K$. This is called the *semigroup of ideals* of $\mathcal{O}_K$.

Theorem A.1: (Unique factorization in the semigroup of ideals of $\mathcal{O}_K$) Any ideal $I \subset \mathcal{O}_K$ can be uniquely factored into a product of prime ideals. This is the analogue of the unique factorization theorem in $\mathbb{Z} = \mathcal{O}_{\mathbb{Q}}$.

Definition A.8: Let $K \subset L$ be an extension of number fields of degree n. Any $\alpha \in \mathcal{O}_L$ defines a linear transformation of the finite dimensional vector space L/K. The determinant

of this transformation is called $N_{L/K}(\alpha)$.

Theorem A.2: For $\alpha, \beta \in \mathcal{O}_L$

1) $N_{L/K}(\alpha) \in \mathcal{O}_K$.

2) $N_{L/K}(\alpha\beta) = N_{L/K}(\alpha)N_{L/K}(\beta)$.

3) If $\alpha \in \mathcal{O}_K$, $N_{L/K}(\alpha) = \alpha^n$.

Theorem A.3: Let P be a prime ideal of $\mathcal{O}_K$. $\mathcal{O}_K/P$ is a field.

Theorem / Definition A.9: Let $K \subset L$ be an extension of number fields of degree n. Let $\wp$ be a prime ideal of $\mathcal{O}_L$. Then $P = \wp \cap \mathcal{O}_K$ is a prime ideal of $\mathcal{O}_K$ and $\mathcal{O}_L/\wp$ is a finite extension of $\mathcal{O}_K/\wp$ of degree f. Define $N_{L/K}(\wp) = P^f$. If $I \subset \mathcal{O}_L$ is an ideal and $I = \wp_1^{\alpha_1}\wp_2^{\alpha_2}...\wp_r^{\alpha_r}$ define $N_{L/K}(I) = (N_{L/K}(\wp_1))^{\alpha_1}(N_{L/K}(\wp_2))^{\alpha_2}...(N_{L/K}(\wp_r))^{\alpha_r}$.

Theorem A.4:

1) $N_{L/K}(< \alpha >_{\mathcal{O}_L}) =< N_{L/K}(\alpha) >_{\mathcal{O}_K}$.

2) $N_{L/K}$ is multiplicative.

3) $N_{L/K}(I\mathcal{O}_L) = I^n$.

4) If $L = K(i)$ and $I \subset \mathcal{O}_{K(i)}$, then $N_{L/K}(I)\mathcal{O}_{K(i)} = I\overline{I}$.

5) $N_{L/\mathbb{Q}}(I) = N_{K/\mathbb{Q}}(N_{L/K}(I))$.

6) If $P \subset \mathcal{O}_K$ is a prime ideal, then $N_{K/\mathbb{Q}}(P) =< p^f >_{\mathbb{Z}}$ and $\mathcal{O}_K/P$ is a finite field of p^f elements.

Theorem A.5: Any Euclidean ring is a principal ideal domain (see appendix C for the definition of a Euclidean ring).

Theorem A.6: If a and b are expressible elements of $\mathcal{O}_K$, ab is also expressible.

Proof: By the initial assumption, one can write a as $a_0^2 + a_1^2$ and b as $b_0^2 + b_1^2$. Thus

$a = (a_0 + a_1 i)(a_0 - a_1 i)$ and $b = (b_0 + b_1 i)(b_0 - b_1 i)$. Then $ab = ((a_0 b_0 - a_1 b_1) + (a_0 b_1 + a_1 b_0)i)((a_0 b_0 - a_1 b_1) - (a_0 b_1 + a_1 b_0)i) = (a_0 b_0 - a_1 b_1)^2 + (a_0 b_1 + a_1 b_0)^2$ and is thus expressible.

B Characterization of the Elements of $\mathcal{O}_{K(i)}$

Definition B.1: An ideal I in $\mathcal{O}_K$ is *unramified* if $J|I$ implies that $J^2 \not| I$ for all ideals J in $\mathcal{O}_K$. Otherwise it is said to be *ramified*.

Lemma B.1: If a is an unramified element of $\mathcal{O}_K$, $a|x^2$ implies that $a|x$ for $x \in \mathcal{O}_K$.

If $\alpha = x + yi \in \mathcal{O}_{K(i)}$, then $N(\alpha), tr(\alpha) \in \mathcal{O}_K$. This means that $2x, 2y \in \mathcal{O}_K$ and $x^2 + y^2 \in \mathcal{O}_K$. Let $x = m/2$ and $y = n/2$ for some $m, n \in \mathcal{O}_K$. Then $x + yi \in \mathcal{O}_{K(i)}$ iff $4|(m^2 + n^2)$ in $\mathcal{O}_{K(i)}$.

Theorem B.1: $\mathcal{O}_{K(i)} = \mathcal{O}_K[i]$ iff $< 2 >$ is unramified in $\mathcal{O}_K$.

1) If $< 2 >$ is unramified, $4|(m^2 + n^2)$ implies that $2|(m + n)^2$ since $2|2mn$. This in turn implies that $2|(m + n)$ (by Lemma B.1) and $4|(m + n)^2$. This indicates that $4|2mn$ and so $2|mn$. Since $< 2 >$ is unramified, $2|m$ or $2|n$. Since 4 must divide $m^2 + n^2$, both m and n must be divisible by 2. Thus $x, y \in \mathcal{O}_K$ and $\mathcal{O}_{K(i)} = \mathcal{O}_K$.

2) If $< 2 >$ is ramified, then $< 2 >= P_1^{\alpha_1} P_2^{\alpha_2} ... P_r^{\alpha_r}$ for prime ideals P_i in $\mathcal{O}_K$ and for at least one value of j, $\alpha_j \geq 2$. Without loss of generality, let $j = 1$. Consider the ideal $I = P_1^{\alpha_1 - 1} P_2^{\alpha_2} ... P_r^{\alpha_r}$. In this case, $< 2 > | I^2$ but $< 2 > \not| I$. Thus there exists some element a in I such that $2|a^2$ but $2 \not| a$. $4|(a^2 + a^2)$, and so $a + ai \in \mathcal{O}_{K(i)}$ but $a \notin \mathcal{O}_K$. As a result, $\mathcal{O}_{K(i)} \neq \mathcal{O}_K[i]$.

C Miscellaneous Considerations in $\mathbb{Z}[\rho]$

Definition C.1: Ring $\Re$ is Euclidean iff $\exists \delta : \Re - \{0\} \mapsto \mathbb{R}_+$ such that

1) $\forall a, b \in \Re, \delta(ab) > \delta(a)$

2) $\forall a, b \in \Re (b \neq 0), \exists x \in \Re$ such that $\delta(a - bx) < \delta(b)$ or $\delta(a - bx) = 0$ (the division algorithm).

In the case of $\mathbb{Q}[\sqrt{-3}]$, we set $\delta(a + b\sqrt{-3})$ to be $N_{-3}(a + b\sqrt{-3})$. This is equal to $a^2 + 3b^2$. Note: since the norm N_{-3} is multiplicative and has integeral values, condition (1) automatically holds.

Theorem C.1: $\mathcal{O}_{\mathbb{Q}(\sqrt{-3})}$ is Euclidean.

Proof: In this case, a general element a of $\mathcal{O}_{\mathbb{Q}(\sqrt{-3})} = \mathbb{Z}[\frac{1+\sqrt{-3}}{2}]$ is $a_0 + a_1\rho$ for $a_0, a_1 \in \mathbb{Z}$. Let $x = x_0 + x + 1\rho$ be an element of $\mathbb{Z}[\rho]$ and $b \in \mathbb{Z}$. Then

$$N(a - bx) = N((a_0 - bx_0) + (a_1 - bx_1)\rho) = (\frac{2(a_0 + a_1) - b(2x_0 + x_1)}{2})^2 - d(\frac{a_1 - bx_1}{2})^2 \qquad (1)$$

The quantity $(\frac{a_1 - bx_1}{2})^2$ can be made less than or equal to $(b/4)^2$, but this places a restriction on the even-odd parity of x_1. Accordingly, $\frac{2(a_0+a_1)-b(2x_0+x_1)}{2}$ could have an absolute value as large as $b/2$ if the optimal value of x_1 is different in parity from the predetermined value. Since -3 is negative, the maximum value for $N(a - bx)$ is $\frac{7}{16}b^2$, which must be less than b^2.

Now consider all possible values of b in $\mathbb{Z}[\rho]$ with conjugate b^*. The quantity bb^* is an integer, so $\exists x \in \mathbb{Z}[\rho]$ such that $N(ab^* - (bb^*)x) \leq \frac{7}{16}N(bb^*)$. This means that $N(a - bx)N(b^*) \leq \frac{7}{16}N(b)N(b^*)$, and so $N(a - bx) \leq \frac{7}{16}N(b)$ and the division algorithm is proven for all $a, b \in \mathbb{Z}[\rho]$.

Theorem C.2: The units in $\mathbb{Z}[\rho]$ are $\pm 1, \pm\rho$, and $\pm\overline{\rho}$, and all of these units are expressible in $\mathbb{Z}[\rho]$.

Proof: If $a, b \in \mathbb{Z}$, $N(a+b\rho) = \frac{2a+b}{2} + \frac{b}{2}\sqrt{-3} = (\frac{2a+b}{2})^2 + 3(\frac{b}{2}\sqrt{-3})^2 = a^2 + ab + b^2$.

The units of $\mathbb{Z}[\rho]$ are all of the elements with norm 1[1]. Thus if $u = u_0 + u_1\rho$ is an element of $\mathbb{Z}[\rho], N(u) = u_0^2 + u_0u_1 + u_1^2 = 1$. This can only occur when $u_0 = \pm 1$ and $u_1 = 0$, when $u_0 = 0$ and $u_1 = \pm 1$, or when $u_0 = \pm 1$ and $u_1 = \mp 1$. This is only true for $\pm 1, \pm\rho$, and $\pm\overline{\rho}$, so $\pm$1these are the only units in $\mathbb{Z}[\rho]$. $1 = 1^2 + 0^2, -1 = \rho^2 + \overline{\rho}^2, \rho = 1^2 + \rho^2, -\rho = 0^2 + \overline{\rho}^2, \overline{\rho} = 1^2 + \overline{\rho}^2$, and $-\overline{\rho} = 0^2 + \rho^2$. Thus all units can be expressed as the sum of two squares in $\mathbb{Z}[\rho]$.

Theorem C.3: $\mathbb{Z}[\rho, i]$ is the maximal order of $\mathbb{Q}[\sqrt{-3}, i]$.

Proof: Suppose 2 is not prime in $\mathbb{Z}[\rho]$. Then $2 = ab$ where neither a nor b have norm 1. This implies that $4 = N(a)N(b)$ and so $N(a) = 2$. Let $a = a_0 + a_1\rho$ for $a, b \in \mathbb{Z}$. Since $N(a)$ can be written as $a_0^2 + a_0a_1 + a_1^2$, $2 = a_0^2 + a_0a_1 + a_1^2$. This has no solutions in $\mathbb{Z}$, and so 2 is prime. This means that 2 is unramified, and so $\mathcal{O}_{\mathbb{Q}(\sqrt{-3},i)} = \mathbb{Z}[\rho, i]$,

Theorem C.4: $\mathbb{Z}[\rho, i]$ is Euclidean.

Proof: Let $\alpha \in \mathbb{Z}[\rho, i]$ and $b \in \mathbb{Z}$. We define the function $\delta(\alpha)$ (see Definition C.1) to be $N_{-3}(N_{-1}(\alpha))$. If $\alpha = a_0 + a_1 i$ and $a_0, a_1, x_0, x_1 \in \mathbb{Z}[\rho]$, consider the quantity $\alpha - b(x_0 + x_1 i) = r$. We want to show that $\exists x_0, x_1$ such that $\delta(r) < \delta(b)$, or $N_{-3}((a_0 - bx_0)^2 + (a_1 - bx_1)^2) < b^4$. We will make use of the following lemma:

Lemma C.1: $\forall y, z \in \mathbb{Z}[\rho], N_{-3}(y + z) \leq 2(N_{-3}(y) + N_{-3}(z))$.

Proof: The quantities y and z can be written as $y_0 + y_1\sqrt{-3}$ and $z_0 + z_1\sqrt{-3}$ where y_0, y_1, z_0, and z_1 are integers divided by 2. $N_{-3}(y+z) = (y_0^2+3y_1^2)+(z_0^2+3z_1^2)+2y_0z_0+6y_1z_1 =$

$N_{-3}(y) + N_{-3}(z) + 2y_0z_0 + 6y_1z_1$. Since $(y_0 - z_0)^2 \geq 0, 2y_0z_0 \leq y_0^2 + z_0^2$. Similarly, $6y_1z_1 \leq 3y_1^2+3z_1^2$. Thus $N_{-3}(y+z) \leq N\_3(y)+N_{-3}(z)+(y_0^2+3y_1^2)+(z_0^2+3z_1^2) = 2(N\_3(y)+N\_3(z))$.

Thus $\delta(r) \leq 2((N\_3(a_0-bx_0))^2+(N\_3(a_1-bx_1))^2$. This expression takes on a maximum when $N\_3(a_0 - bx_0)$ and $N\_3(a_1 - bx_1)$ are maximized. As shown in the proof of theorem 3.1.3, these values are both $\frac{7}{16}b^2$, and so $\delta(r) \leq \frac{196}{256}b^4$ and so the division algorithm holds.

The method for generalizing the divisor in the proof of Theorem C.1 can be used here to complete the proof.

D Miscellaneous Considerations when $i \in K$

Let us consider an irreducible element π of $\mathcal{O}_K$. If $\pi = A^2 + B^2$, it follows that $\pi = (A + Bi)(A - Bi)$ for A and B in $\mathcal{O}_K$. Since π is irreducible, $A + Bi$ or $A - Bi$ is a unit (i.e. ϵ or ϵi where ϵ is a unit in $\mathcal{O}_{K'}$). Without loss of generality, let $A + Bi$ be the unit. Suppose $A + Bi = \epsilon$. Then $A = \epsilon - Bi$. $A - Bi$ must be $\epsilon^{-1}\pi$, so $\epsilon^{-1}\pi = \epsilon - 2Bi$ and $\pi = \epsilon^2 - 2\epsilon Bi$. B can be expressed as $B_1 + B_2i$ where B_1 and B_2 are in $\mathcal{O}_{K'}$, and so $\pi = (\epsilon^2 - 2\epsilon B_1) - 2\epsilon B_2 i$. Note that every value of $B_1, B_2 \in \mathcal{O}_{K'}$ will yield a pi which is the sum of two squares, and thus we have proven

Theorem D.1: All irreducible elements of $\mathcal{O}_K$ which are of the form $(2k + \epsilon^2) + 2\ell i$, where $k, \ell \in \mathcal{O}_K$ and ϵ is a unit of $\mathcal{O}_{K'}$, can be expressed as a sum of two squares.

Note: Had we supposed that $A + Bi$ were ϵi, we would obtain $\epsilon i(A - Bi) = \pi$ and $-\epsilon^{-1}i(A + Bi) = 1$. These equations yield the same result.

Theorem D.2: If $\alpha \in \mathbb{Z}[i]$ and $\alpha = a + bi$ where a and b are even, α can be written as the sum of two squares if $m \neq n$.

Proof: α can be expressed as $2^m x + 2^n yi$ where x and y are odd. Suppose that $m < n$. Then $\alpha = 2^m(x + 2^{n-m}yi)$ Since x is odd and $2^{n-m}y$ is even, α is the product of integers expressable as the sum of squares, and is also expressible. If $m > n$, α can be written as $2^n(2^{m-n}x + yi)$, or $2^{n-1}(1-i)^2(-y + 2^{m-n}xi)$, which is the product of expressible Gaussian integers. Therefore all α where a and b are even can be written as the sum of two squares if $m \neq n$.

Theorem D.3: Any Gaussian integer α with an even coefficient of i is the sum of three squares.

Proof: Suppose that α is expressible. Let $\alpha = A^2 + B^2$. Then α can be written as the sum of three squares: $A^2 + B^2 + 0^2$. If α is not expressible but has an even coefficient of i, it must be of the form $2(x + yi)$ where x and y are odd. Then $\alpha = 2(x + (y-1)i) + 2i$. Since $y - 1$ is even, $x + (y-1)i$ is expressible and so is $2(x + (y-1)i)$. If one lets this quantity equal $X^2 + Y^2$, one obtains that $\alpha = X^2 + Y^2 + (1+i)^2$ and so α is the sum of three squares.