Notes on Chapters 3 and 4 of Dedekind's *Theory of Algebraic Integers*

Jeremy Avigad

November 15, 2002

These notes are intended as a high-level overview of some of the central ideas of Dedekind's theory of ideals, as presented in Chapters 3 and 4.

We saw at the end of Chapter 2 (and in the last set of notes) that Dedekind's goal is to extend the unique factorization of ideals in $\mathbb{Z}[\sqrt{-5}]$ to the unique factorization of ideals in the ring of integers of an arbitrary number field, with "proofs based immediately on fundamental characteristics, rather than on calculation."

I find it useful to recast Dedekind's example in Chapter 2 in three components:

- First, introducing the ring of "integers," $\mathbb{Z}[\sqrt{-5}]$, as the set $\{a + b\sqrt{-5} \mid a, b \in \mathbb{Z}\}$.
- Second, using calculations and a particular representation of the ideals in $\mathbb{Z}[\sqrt{-5}]$ to derive properties of these ideals.
- Third, using these results (and additional calculations) to prove unique factorizations for the ring of ideals.

Dedekind's more general, "conceptual" treatment of the theory of algebraic integers can similarly separated into three components:

- First, introducing the "right" definition of the integers of an arbitrary number field.
- Second, using ideas from Galois theory (rather than explicit representations) to derive properties of the ideals in these rings of integers.
- Third, using these results and general properties of ideals to prove unique factorization.

In Section 1, I will review some of the necessary background on finite dimensional field extensions of $\mathbb{Q}$. In Sections 2–4, I will discuss the three components indicated above. Finally, in Section 5, we will revisit an old friend: the quadratic form, $x^2 + y^2$.

I will omit most of the proofs, and even many important definitions. For those of you who are interested in learning the mathematics, I recommend Stewart and Tall's *Algebraic Number Theory and Fermat's Last Theorem* (AK Peters has recently issued a third edition). Their treatment, however, uses fractional ideals, making it closer to Dedekind's final (1894) presentation of ideal theory. Two other textbooks that have presentations that are closer to the 1877 version are Pollard and Diamond, *The Theory of Algebraic Integers*, and Ireland and Rosen, *A Classical Introduction to Modern Number Theory*.

1 Background

Recall that an element θ of $\mathbb{C}$ is *algebraic* over $\mathbb{Q}$ if it is the solution to a polynomial equation $f(\theta) = 0$, where f is any element of $\mathbb{Q}[x]$, i.e. a polynomial with coefficients in $\mathbb{Q}$. By factoring f into irreducible components and taking the component of which θ is a root, we can assume that f is irreducible. By dividing through by the leading coefficient, we can assume that f is monic, i.e. has leading coefficient 1.

So, from now on, assume f is monic and irreducible, of degree n. The field $\mathbb{Q}(\theta)$ is defined to be the smallest field containing every element of $\mathbb{Q}$ as well as θ. We have seen that $\mathbb{Q}(\theta)$ is isomorphic to $\mathbb{Q}[x]/(f)$, where (f) is the principal ideal of $\mathbb{Q}[x]$ generated by f. One can also view $\mathbb{Q}(\theta)$ as a vector space over $\mathbb{Q}$, by forgetting the operation of multiplication in $\mathbb{Q}(\theta)$. This vector space has dimension n, and the set $\{1, \theta, \theta^2, \ldots, \theta^{n-1}\}$ forms a basis.

Over the complex numbers, $\mathbb{C}$, f will factor as

$$f(x) = (x - \theta_1)(x - \theta_2) \ldots (x - \theta_n)$$

where $\theta = \theta_1, \theta_2, \ldots, \theta_n$ are the n roots of f. From the point of view of $\mathbb{Q}$, all these roots looks the same; i.e. we know

$$\mathbb{Q}(\theta_1) \simeq \mathbb{Q}(\theta_2) \simeq \ldots \simeq \mathbb{Q}(\theta_n) \simeq \mathbb{Q}[x]/(f).$$

One can show that since f is irreducible over $\mathbb{Q}$, all these roots are distinct. We obtain n embeddings $\sigma_1, \ldots, \sigma_n$ of $\mathbb{Q}(\theta_1)$ into the complex numbers, where σ_i maps θ_1 to θ_i. The θ_i's are called the "conjugates" of θ_1 over $\mathbb{Q}$.

If you multiply through the equation for f above, you get

$$f(x) = x^n - (\theta_1 + \ldots + \theta_n)x^{n-1} + \sum_{i \neq j}(\theta_i\theta_j) - \ldots \\ + (-1)^n \theta_1\theta_2 \cdots \theta_n.$$

Each of these coefficients is said to be a *symmetric polynomial* in the θ_i's, which is to say, permuting the θ_i's does not change the value of the expression. Because f is a polynomial with rational coefficients, all these values are rational. Ignoring the minus signs these coefficients, viewed for the moment as polynomials in the variables $\theta_1, \ldots, \theta_n$, are called the *elementary symmetric functions* in $\theta_1, \ldots, \theta_n$. One can show that *every* symmetric polynomial can be written as a polynomial combination of elementary symmetric functions. So, in the case at hand, we have the following:

Proposition 1.1 *Let θ_1 be any algebraic number, and let $\theta_2, \ldots, \theta_n$ be its conjugates. Then any symmetric polynomial of $\theta_1, \ldots, \theta_n$ is an element of $\mathbb{Q}$.*

This fact lies at the heart of Galois' theory of equations, which aims to account for when the roots $\theta_1, \ldots, \theta_n$ can be written as expressions in the coefficients of f using quotients and radicals. I will provide a short excerpt from Stillwell's *Mathematics and its History* which provides a nice summary of the main idea behind the theory.

These ideas are useful more generally, though, in that they tell us how we can "get a handle" on the elements of a number field in terms of ordinary rationals, i.e. the elements of $\mathbb{Q}$. For example, fix a number field $\mathbb{Q}(\theta)$ as above, and let α be any element of $\mathbb{Q}(\theta)$. (So, α can be expressed as a polynomial in θ.) Let $\sigma_1, \ldots, \sigma_n$ be the n embeddings of $\mathbb{Q}(\theta)$ into $\mathbb{C}$, as above. Define the *norm* of α (relative to $\mathbb{Q}(\theta)$) to be

$$N(\alpha) = \sigma_1(\alpha)\sigma_2(\alpha) \cdots \sigma_n(\alpha)$$

and define the *trace* of α to be

$$Tr(\alpha) = \sigma_1(\alpha) + \ldots + \sigma_n(\alpha).$$

The proposition above tells us that $N(\alpha)$ and $Tr(\alpha)$ are elements of $\mathbb{Q}$. Furthermore, it is easy to see that for any α and β in $\mathbb{Q}(\theta)$, $N(\alpha\beta) = N(\alpha)N(\beta)$, and $Tr(\alpha + \beta) = Tr(\alpha) + Tr(\beta)$.

2 The algebraic integers

In analogy to the Gaussian integers, one might try to define the set integers of an algebraic number field $\mathbb{Q}(\alpha)$ of degree n to be the subring $\mathbb{Z}[\alpha]$ of the complex numbers generated by α. This amounts to the set of integral linear combinations of $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$. This is what Kummer did in defining the integers of a cyclotomic extension of $\mathbb{Q}$, which correspond to the case where α is a primitive pth root of 1, where p is prime; and in, the case where $\alpha = \sqrt{-2}$, this provides the appropriate unique factorization domain for the analysis of the Diophantine equation $x^3 = y^2 + 2$.

Recall, however, the Dedekind found two problems with this definition:

1. The definition is dependent on the presentation of $\mathbb{Q}(\alpha)$. For example, let ω be a primitive cube root of 1,

$$\omega = \frac{-1 \pm \sqrt{-3}}{2}.$$

Then $\mathbb{Q}(\sqrt{-3}) = \mathbb{Q}(\omega)$, but $\mathbb{Z}[\sqrt{-3}]$ is different from $\mathbb{Z}[\omega]$.

2. Some instances of the definition have undesirable properties. For example, there is no way of extending Kummer's theory of ideal divisors to $\mathbb{Z}[\sqrt{-3}]$.

A good theory of the algebraic integers will therefore have to begin by finding the right definition. Here it is:

Definition 2.1 *A complex number α is an algebraic integer if it is the solution to a monic polynomial with coefficients in $\mathbb{Z}$.*

It turns out that this is equivalent to saying that the *minimal* monic polynomial satisfied by α has (rational) integer coefficients.

Note that the set of *all* algebraic integers is not a unique factorization domain. Indeed, one can't even factor algebraic integers into irreducibles:

$$2 = \sqrt{2}\sqrt{2} = \sqrt{2}\sqrt[4]{2}\sqrt[4]{2} = \sqrt{2}\sqrt[4]{2}\sqrt[8]{2}\sqrt[8]{2} = \ldots$$

and all the above are integers. But given any finite extenion of the rational $\mathbb{Q}(\alpha)$, define set of integers $\mathfrak{o}$ of $\mathbb{Q}(\alpha)$ to be the set of algebraic integers that happen to be $\mathbb{Q}(\alpha)$. Then it *is* the case that every integer in $\mathbb{Q}(\alpha)$ can be factored into irreducible elements of that number field.

The definition is somehow mysterious, however. One can show that the set of algebraic integers of any number field form a ring, which is to say, sums and products of algebraic integers are again algebraic integers. In his introduction, Dedekind remarks that this is "immediate," but it strikes me as being far from obvious. Dedekind's elegant proof of this fact appears in Section 13, in Chapter 3; see also the excerpt from Stewart and Tall.

(Here is one way to test your understanding of the proof. Both $\sqrt{2}$ and $\sqrt{3}$ are algebraic integers, being roots of the polynomials $x^2 - 2$ and $x^2 - 3$ respectively. By the theorem, $\sqrt{2} + \sqrt{3}$ is an algebraic integer. Can you find the monic polynomial of which the latter is a root?)

One way of motivating the definition of the algebraic integers is to enumerate some properties that one would expect a reasonable set of "integers" to have:

1. They form a ring.
2. The integers of $\mathbb{Q}$ are exactly $\mathbb{Z}$.
3. If α is an integer and α' is a conjugate of α then α' is an integer as well.

The last clause corresponds to the intuition that conjugates look just the same from the point of view of $\mathbb{Q}$. These three clauses do not determine the set of algebraic integers uniquely; for example, they are satisfied by $\mathbb{Z}$ alone. But the algebraic integers are the *largest* set of algebraic numbers satisfying 1–3, which is to say, any set of algebraic numbers satisfying them has to be a subset of the algebraic integers. It turns out that the set of all algebraic numbers is the exactly field of fractions of the algebraic integers, just as $\mathbb{Q}$ is the field of fractions of $\mathbb{Z}$. In fact, we have the following even stronger property:

4. If α is any algebraic number, then for some rational integer n, $n\alpha$ is an algebraic integer.

One can show that if α is a primitive pth root of unity, then the set of integers of the cyclotomic field $\mathbb{Q}(\alpha)$ is exactly $Z[\alpha]$. So Kummer was lucky: he got it right. One can also characterize the integers of quadratic number fields $\mathbb{Q}(\sqrt{d})$, where d is a rational integer. Note that if d has any square factors, they can be factored out without changing $\mathbb{Q}(\sqrt{d})$, so it suffices to focus on d that are squarefree. In that case, the integers of $\mathbb{Q}(\sqrt{d})$ are $\mathbb{Z}[\sqrt{d}]$ is d is not congruent to 1 mod 4; but if d *is* congruent to 1 mod 4, the integers are

$$\mathbb{Z}[\frac{1}{2} + \frac{1}{2}\sqrt{d}].$$

So, in particular, the integers of $\mathbb{Q}(\omega)$, where ω is a primitive cube root of 1, are $\mathbb{Z}[\omega]$, and not $\mathbb{Z}[\sqrt{-3}]$.

Summing up, these are some of the factors that support Dedekind's definition as being the right one:

- It is a natural definition, satisfying requirements 1–3 above.
- It is the *largest* ring of algebraic numbers satisfying these requirements.
- Defining the integers of a number field as, simply, the integers that happen to be in the number field, makes the definition independent of the way the number field is represented; this solves the first problem associated with previous definitions.
- It agrees with Kummers' definition of the intgers of a cyclotomic field, and, at the same time, explains why $\mathbb{Z}[\sqrt{-3}]$ is *not* the right notion of integrality for $\mathbb{Q}(\omega)$. So it supports the previous successes of the theory, and explains the previous failures.
- It leads to a satisfying general theory of ideal divisors arbitrary number fields. In particular, it allows one to prove the unique factorization of ideals.

3 Discriminants and integral bases

One might argue that the definition of the algebraic integers in the previous section fits well with Dedekind's program, because rather than describing them as represented by a particular basis, one is rather describing them in terms of a characteristic *property.* I won't push this point too hard, however, since one can argue that, really, each algebraic integer is represented instead by its minimal polynomial; in other words, one is just using a different form of representation. But the definition means that, at least initially, proofs involving the algebraic integers will have to depend on their characterization as roots of polynomials, rather than in terms of a particular basis.

In particular, in contrast to the previous "definition" of the integers of $\mathbb{Q}(\sqrt{-5})$, it isn't even clear that the integers of a number field form a free module over $\mathbb{Z}$ of finite rank. Indeed, they do:

Proposition 3.1 *Let $\mathbb{Q}(\alpha)$ be a number field of degree n, and let $\mathfrak{o}$ be the integers of $\mathbb{Q}(\alpha)$. Then $\mathfrak{o}$ is a free $\mathbb{Z}$-module of rank n.*

In other words, $\mathfrak{o}$, as a $\mathbb{Z}$-module, has a basis consisting of integers $\omega_1, \ldots, \omega_n$. Such a basis is called, appropriately, an *integral basis*.

Dedekind proves this proposition in Section 18 of Chapter 3. The proof uses the notion of a *discriminant*, introduced in Section 17. Let $\alpha_1, \ldots, \alpha_n$ be any n elements of $\mathbb{Q}(\alpha)$. Let $\sigma_1, \ldots, \sigma_n$ be the n embeddings of $\mathbb{Q}(\alpha)$ into $\mathbb{C}$, and for each i and r, write $\alpha_i^{(r)}$ for $\sigma_r(\alpha_i)$. Form the matrix

$$\begin{matrix} \alpha_1^{(1)} & \alpha_2^{(1)} & \ldots & \alpha_n^{(1)} \\ \alpha_1^{(2)} & \alpha_2^{(2)} & \ldots & \alpha_n^{(2)} \\ \vdots & \vdots & \vdots & \vdots \\ \alpha_1^{(n)} & \alpha_2^{(n)} & \ldots & \alpha_n^{(n)} \end{matrix}$$

The square of the determinant of this matrix is called the *discriminant*.

The discriminant is a strange and wonderful thing. You can read about some of its properties either in Dedekind or in the excerpt from Stewart and Tall. In particular:

- If $\alpha_1, \ldots, \alpha_n$ is a (vector space) basis for $\mathbb{Q}(\alpha)$, then the discriminant is rational, and nonzero.

- If $\alpha_1, \ldots, \alpha_n$ is a (vector space) basis for $\mathbb{Q}(\alpha)$ and all of $\alpha_1, \ldots, \alpha_n$ are integers, then the discriminant is a rational integer.

By property 4 in the last section, we know that $\mathbb{Q}(\alpha)$ has a vector space basis consisting of integers: pick any vector space basis, and multiply by suitable rational integers. To finish off the proof, Dedekind shows that if $\alpha_1, \ldots, \alpha_n$ is any basis of $\mathbb{Q}(\alpha)$ consisting of integers, and there is an integer β in $\mathbb{Q}(\alpha)$ that is not in the $\mathbb{Z}$-module generated by $\alpha_1, \ldots, \alpha_n$, then there is another basis of $\mathbb{Q}(\alpha)$ consisting of integers, with a discriminant whose absolute value is strictly smaller. This implies that any basis of $\mathbb{Q}(\alpha)$ consisting of integers whose discrimant is *minimum* will be an integral basis of $\mathfrak{o}$. (See page 116 of the Dedekind text.)

I believe it is significant that, as Dedekind presents the proof, it is non-effective: it proves the existence of an integral basis without showing how to find one. The outline of an algorithm is implicit in the proof: start with any basis of $\mathbb{Q}(\alpha)$, and turn it into a basis consisting of integers. Then, as long as there are integers of $\mathbb{Q}(\alpha)$ that are not in the $\mathbb{Z}$-module generated by this basis, Dedekind's proof shows, explicitly, how to obtain a new basis of integers of smaller discriminant. The only catch is that the proof does not tell you how to determine whether, at any stage, you have gotten all the integers, and, if not, how to find a new one.

Hilbert's much more dramatic use of nonconstructive methods in proving his basis theorem appeared more than 10 years later, in 1888. The argument I have just describd is benign in comparison, since the requisite test and search procedure can be obtained in a straightforward way; see Section 2.6 of Stewart and Tall. Dedekind was no doubt aware of this, since all the relevant information is near at hand in his presentation. But the additional considerations needed are not obvious, and it is interesting that Dedekind does not deem it worthwhile to make the algorithmic information explicit. Of course, from a modern point of view, his proof stands complete even without an explicit algorithm. So, by de-emphasizing the algorithmic details, Dedekind is demonstrating a very modern sensibility. At the same time, he is providing the desired "conceptual" characterization of the integral bases of a number field: they are exactly the bases of integers that have the minimum possible discriminant.

4 The proof of unique factorization

Remember that in the case of $\mathbb{Z}[\sqrt{-5}]$, the proof of unique factorization was based on the following lemma:

Lemma 4.1 *If $\mathfrak{a}$ and $\mathfrak{b}$ are ideals and $\mathfrak{a} \supseteq \mathfrak{b}$, then there is a third ideal $\mathfrak{c}$ such that $\mathfrak{ac} = \mathfrak{b}$.*

Since the converse is obvious, the lemma asserts that "containing" and "dividing" amount to the same thing. It is easy to see, from Dedekind's definition of norm, that if $\mathfrak{a} \supsetneq \mathfrak{b}$, then $N(\mathfrak{a}) < N(\mathfrak{b})$. So, if we can establish the same lemma in this more general setting, unique factorization will follow just as before.

Dedekind's proof of the lemma for $\mathbb{Z}[\sqrt{-5}]$ relied, in turn, on the following:

Lemma 4.2 *If $\mathfrak{a}$ is any ideal, then there is another ideal $\mathfrak{d}$ such that $\mathfrak{da}$ is principal.*

This lemma is very helpful, because principal ideals have such a simple structure. To see how it implies Lemma 4.1, suppose $\mathfrak{a} \supseteq \mathfrak{b}$. Find an ideal $\mathfrak{d}$ such that $\mathfrak{da}$ is principal, say the ideal (η). Then we have

$$(\eta) = \mathfrak{da} \supseteq \mathfrak{db}.$$

This means, in particular, that every element of $\mathfrak{d}\mathfrak{b}$ is a multiple of η. Let $\mathfrak{c}$ be the result of dividing every element of $\mathfrak{d}\mathfrak{b}$ by η, i.e. let

$$\mathfrak{c} = \frac{1}{\eta}\mathfrak{d}\mathfrak{d}.$$

Then $\mathfrak{c}$ is an ideal, and

$$\mathfrak{a}\mathfrak{c} = \frac{1}{\eta}\mathfrak{a}\mathfrak{d}\mathfrak{b} = \frac{1}{\eta}(\eta)\mathfrak{b} = \mathfrak{b},$$

as required. One should take a moment to verify that these manipulations involving the fraction $1/\eta$ and ideals are really kosher, but this is not hard to do.

All the real work, then, is to be found in the proof of Lemma 4.2. Recall that in the case of $\mathbb{Z}[\sqrt{-5}]$, Dedekind's proof involved using an explicit representation of $\mathfrak{a}$, taking $\mathfrak{d}$ to be the conjugate of $\mathfrak{a}$, and showing that $\mathfrak{a}\mathfrak{d} = (N(\mathfrak{a}))$, via an explicit calculation. Here Dedekind needs to establish the same result, without invoking the specific representation.

As Dedekind notes, the proof of this lemma *has* to invoke the fact that we are considering the ring of *all* the integers in a number field, because the lemma is not generally true for arbitrary rings of integers. I will not go into the details of Dedekind's proof, other than to note that it relies on the following:

- The fact that the ring of integers has an integral basis.
- The fact that this basis is finite.
- Closure properties on the ring of algebraic integers of a number field.
- General facts about ideals and their multiplication.

In other words, Dedekind is able to avoid any use of the particular representation of the integers and their ideals. In modern texts, this endgame is typically played out in different ways, but all the presentations I have seen share these features.

5 The quadratic form, $x^2 + y^2$

The last four sections of Dedekind's monograph harvest some of the yields of the ideal-theoretic approach. For example, in Section 26, Dedekind generalizes Euler's function φ on the natural numbers to a function ψ on ideals, and derives some of its properties.For example, in analogy to standard facts about φ, one has:

- $\psi(\mathfrak{ab}) = \psi(\mathfrak{a})\psi(\mathfrak{b})$ when $\mathfrak{a}$ and $\mathfrak{b}$ are relatively prime.
- $\psi(\mathfrak{p}^m) = N(\mathfrak{p}^m)(1 - 1/N(\mathfrak{p}))$ when $\mathfrak{p}$ is a prime ideal.
- $N(\mathfrak{a}) = \sum_{\mathfrak{d}|\mathfrak{a}} \psi(\mathfrak{d})$.

In the integers, when $\mathfrak{a}$ is a principal ideal (a), $N(\mathfrak{a}) = a$. So, the second and third properties really are generalizations of the corresponding facts for $\mathbb{Z}$. In analogy to Fermat's little theorem, we also have

- $\omega^{\psi(\mathfrak{a})} \equiv 1 \pmod{\mathfrak{a}}$

for any ideal $\mathfrak{a}$, for every integer ω. As with the integers, if ω is any integer relatively prime to $\mathfrak{a}$, the order of ω (i.e. the least k such that $\omega^k \equiv 1 \pmod{\mathfrak{a}}$) has to divide $\psi(\mathfrak{a})$. In the particular case where $\mathfrak{p}$ is a prime ideal, the generalization of Fermat's little theorem tells us

- $\omega^{N(\mathfrak{p})-1} \equiv 1 \pmod{\mathfrak{p}}$.

Again, in analogy with the ordinary integers, Dedekind shows that in the this case there will always be an element ω whose order mod $\mathfrak{p}$ is $N(\mathfrak{p}) - 1$. It is clear that, to Dedekind, a large part of the appeal of the new theory is that many of the proofs go "just as in the theory of rational numbers."

In Section 21, Dedekind had shown that if $\mathfrak{p}$ is a prime ideal dividing a principal ideal (p), where p is an ordinary prime of $\mathbb{Z}$, then the norm of $\mathfrak{p}$, $N(\mathfrak{p})$, is a power of p, say p^f. The *degree* of $\mathfrak{p}$ is defined to be the exponent, f. In Section 27, Dedekind considers the cyclotomic fields which were of central interest to Kummer. He shows that if f is the degree of $\mathfrak{p}$, then f is the least value such that $p^f \equiv 1 \pmod{m}$, and also the least value such that $\omega^f \equiv \omega \pmod{\mathfrak{p}}$ for every ω in the number field.

He soon arrives at the following theorem, which, he reports, is "the main theorem of Kummer's theory." Let p and m be distinct prime numbers, and let f be the order of p modulo m. In other words, f is the least positive exponent such that $p^f \equiv 1 \bmod m$. Recall that by Fermat's little theorem, f must divide $\varphi(m)$, say $\varphi(m) = ef$.

Theorem 5.1 *Let p, m, e, and f be as above, and let θ be a primitive mth root of unity. Then in the cyclotomic field $\mathbb{Q}(\theta)$, the principal ideal (p) generated by p factors as*

$$(p) = \mathfrak{p}_1 \dots \mathfrak{p}_e$$

where $\mathfrak{p}_1, \dots, \mathfrak{p}_e$ are distinct prime ideals of degree f.

Dedekind remarks that one obtains the generalization to arbitary m (i.e. composite ones also) in a similar way.

To demonstrate the power of these results, Dedekind shows how they lead to the law of quadratic reciprocity, thus obtained "from entirely general principles and without calculation." As another example, he considers Fermat's theorem on primes representable as sums of squares.

To get the latter, take $m = 4$ in Kummer's theorem. Then θ is a primitive 4th root of unity, i.e. $\theta = \pm i$. For concreteness, take $\theta = i$, and note that the integers of $\mathbb{Q}(\theta)$ are just the Gaussian integers. Because the Gaussian integers are a Euclidean domain, we know that every ideal is principal.

Remember that the hardest part of characterizing the values of $x^2 + y^2$ amounted to verifying Fermat's assertion that every prime number p congruent to 1 mod 4 can be written as a sum of squares. By Kummer's theorem, the principle ideal (p) can be factored as $\mathfrak{p}_1 \ldots \mathfrak{p}_e$, where the $\mathfrak{p}_i$ are prime ideals of degree f, and $ef = \varphi(4) = 2$. By the discussion above, f is the least value such that $\omega^{p^f} \equiv \omega \pmod{p}$ for every integer ω.

Now, if p is congruent to 1 mod 4, then i^{p-1} is congruent to 1 mod p, and so i^p is congruent to i mod p. Using the fact that $(a + bi)^p$ is congruent to $a^p + (bi)^p$ mod p (think about coefficients in the binomial expansion), we see that for every Gaussian integer ω, ω^p is congruent to ω mod p. So $f = 1$, and $e = 2$. This means that (p) is the product of two prime ideals, $(p) = \mathfrak{p}_1\mathfrak{p}_2$. In turn, $\mathfrak{p}_1$ and $\mathfrak{p}_2$ have to be principal, say $\mathfrak{p}_1 = (\alpha_0)$ and $\mathfrak{p}_2 = (\alpha_1)$. Then we have

$$p = p^f = N(\alpha_0) = N(a + bi) = a^2 + b^2,$$

as required.