7.3.3. Deduce from Exercise 7.3.2 that $g(z) = \overline{\overline{g(z)}}$ is of the form $c\bar{z} + d$ with $|c| = 1$.

7.3.4. Conclude from the preceding results that the Euclidean isometries are precisely the functions $f(z) = cz + d$ and $\overline{f}(z) = c\bar{z} + d$ with $|c| = 1$.

This characterization makes it easy to see the difference between

1. the orientation-preserving isometries, which are those of the form $f(z) = cz + d$ (because these are the translations and rotations), and
2. the orientation-reversing isometries, which are those of the form $\overline{f}(z) = c\bar{z} + d$ (because these are the rest).

It also gives an easier way to prove the result of Exercise 3.6.6*, that any orientation-reversing isometry is a glide reflection. The idea is to rotate and translate the coordinate system until the isometry looks like $\overline{h}(z) = \bar{z} + a$, with a real, which is a glide reflection along the x-axis.

7.3.5. Show that if $z' = (\cos\phi + i\sin\phi)z$ is taken as the new coordinate of the point z, then the x'- and y'-axes of this new coordinate system are the result of rotating the x- and y-axes through $-\phi$.

Now suppose that $\overline{f}(z) = (\cos\theta + i\sin\theta)\bar{z} + d$ is an orientation-reversing isometry. Thus, in the old coordinate system, the isometry sends z to $(\cos\theta + i\sin\theta)\bar{z} + d$.

7.3.6. If $z' = (\cos(-\theta/2) + i\sin(-\theta/2))z$, show that the point with new coordinate z'

- has old coordinate $(\cos(\theta/2) + i\sin(\theta/2))z'$,
- which is sent to the point with old coordinate

$$(\cos(\theta/2) + i\sin(\theta/2))\overline{z'} + d,$$

- which has new coordinate $\overline{z'} + d$.

Conclude that the isometry is given by $\overline{g}(z') = \overline{z'} + d$ in the new coordinate system.

Finally, suppose that $d = a + ib$, where a and b are real. Replace the coordinate z' (which is now called the *old coordinate*) by a new coordinate z'' defined by $z'' = z' - ib/2$.

7.3.7. Show the point with new coordinate z''

- has old coordinate $z'' + ib/2$,
- which is sent to the point with old coordinate $\overline{z''} + a + ib/2$,
- which has new coordinate $\overline{z''} + a$.

Conclude that the isometry is given by $\overline{h}(z'') = \overline{z''} + a$ in the new coordinate system and hence is a glide reflection.

In case you are wondering about functions of the form $f(z) = cz + d$, where c is *not* required to have absolute value 1, see the following.

7.3.8. Show that any function of the form $f(z) = cz + d$, where $c \neq 0$, is a composite of a translation or rotation with a *dilatation* —a function of the form $g(z) = rz$ where r is real.

As was mentioned in Section 3.10, these functions are precisely the *similarities*, and they are the only mappings of the plane that preserve angles.

7.4 The Gaussian Integers

In the complex numbers, the counterparts of the integers are called the *Gaussian integers*. They are the complex numbers of the form $a + ib$ where a and b are in $\mathbb{Z}$, and the set of them is denoted by $\mathbb{Z}[i]$. Like $\mathbb{Z}$, $\mathbb{Z}[i]$ is a ring and has notions of divisor and prime. For this reason alone, it is interesting to investigate the arithmetic of $\mathbb{Z}[i]$, but even more interesting is the insight it gives into $\mathbb{Z}$ itself. In a sense, $\mathbb{Z}[i]$ refines our understanding of $\mathbb{Z}$ by allowing ordinary integers to be analyzed in finer detail.

A simple example is the Diophantus identity

$$(a_1^2 + b_1^2)(a_2^2 + b_2^2) = (a_1a_2 - b_1b_2)^2 + (b_1a_2 + a_1b_2)^2$$

for a_1, b_1, a_2, and b_2 in $\mathbb{Z}$. As already suggested in Exercise 7.1.1, this identity is more understandable in $\mathbb{Z}[i]$, where we have the factorizations

$$a_1^2 + b_1^2 = (a_1 + ib_1)(a_1 - ib_1),$$
$$a_2^2 + b_2^2 = (a_2 + ib_2)(a_2 - ib_2).$$

If we rearrange these factors of $(a_1^2 + b_1^2)(a_2^2 + b_2^2)$ as

$$(a_1 + ib_1)(a_2 + ib_2)(a_1 - ib_1)(a_2 - ib_2)$$

and then combine the first two and the last two, we get

$$[(a_1a_2 - b_1b_2) + i(b_1a_2 + a_1b_2)][(a_1a_2 - b_1b_2) - i(b_1a_2 + a_1b_2)],$$

which is a Gaussian integer factorization of $(a_1a_2 - b_1b_2)^2 + (b_1a_2 + a_1b_2)^2$.

We noted in Section 7.1 that Diophantus' identity shows that the absolute value function $|a+ib|$ is multiplicative. Even more directly, it shows that the function $|a + ib|^2$ is multiplicative. The latter is a very useful function on $\mathbb{Z}[i]$, called the *norm* of $a + ib$ and written $N(a+ib)$. Diophantus' identity is precisely the *multiplicative property of the norm*:

$$N((a_1 + ib_1)(a_2 + ib_2)) = N(a_1 + ib_1)N(a_2 + ib_2).$$

The norm is useful because:

- It is an ordinary integer and hence reduces some questions about $\mathbb{Z}[i]$ to questions about $\mathbb{Z}$.
- It is multiplicative and hence the norm of a factor divides the norm of a product.

In particular, the norm draws our attention to the *units* $1, -1, i, -i$ of $\mathbb{Z}[i]$, the members of norm 1. These are the numbers that divide every Gaussian integer and hence can be regarded as redundant factors (like 1 and -1 in $\mathbb{Z}$). When unit factors are disregarded, each Gaussian integer can be split into finitely many factors

$$a + ib = (a_1 + ib_1)(a_2 + ib_2) \cdots (a_k + ib_k),$$

which are *Gaussian primes* in the sense that $a_j + ib_j$ has no divisors of smaller norm except units. It follows that $a_j + ib_j$ has no divisors at all except units and multiples of itself by units.

Gaussian prime factorizations exist in $\mathbb{Z}[i]$ for much the same reason that prime factorizations exist in $\mathbb{Z}$: *each Gaussian integer has a Gaussian prime divisor* (compare with Section 1.3). If $a + ib$ has no nonunit divisor of smaller norm, then $a+ib$ itself is a Gaussian prime. Otherwise, take a nonunit divisor $a' + ib'$ of smaller norm, and see whether $a' + ib'$ has a nonunit divisor $a'' + ib''$ of still smaller norm, and so on. Because the norms are natural numbers, this process ends in a finite number of steps, necessarily with a Gaussian prime divisor $a_1 + ib_1$. We then repeat the process on the Gaussian integer $(a + ib)/(a_1 + ib_1)$, which has smaller norm than $a + ib$, and so on.

Exercises

The norm sometimes enables us to recognize Gaussian primes.

7.4.1. Find some Gaussian integers whose norms are prime.

7.4.2. A Gaussian integer with prime norm is a Gaussian prime. Why?

However, ordinary primes are not necessarily Gaussian primes.

7.4.3. Show that 2 is not a Gaussian prime. Also find an odd prime that is not a Gaussian prime.

Your odd prime should be of the form $4n+1$, because ordinary primes of the form $4n + 3$ *are* Gaussian primes. This is proved with the help of conjugation.

7.4.4. Suppose that p is an ordinary prime and $p = (a+ib)c$ is a Gaussian factorization without units. Show in turn that

- $p = (a - ib)\bar{c}$
- $p^2 = (a^2 + b^2)|c|^2$
- $p = a^2 + b^2$
- p is not of the form $4n + 3$.

In Exercise 1.3.5 it was proved that there are infinitely many primes of the form $4n + 3$, so it follows from Exercise 7.4.4 that there are infinitely many Gaussian primes. The same result can also be proved directly, in the manner of Euclid, once we clarify the idea of division with remainder in $\mathbb{Z}[i]$. This will be done in the next section.

However, before going more deeply into $\mathbb{Z}[i]$, it should be pointed out that $\mathbb{Z}[i]$ is not the only ring of "integers" in $\mathbb{C}$. Unlike $\mathbb{R}$, which has $\mathbb{Z}$ as its only integers, $\mathbb{C}$ has many subrings that can reasonably be regarded as integers. Another example is the set

$$\mathbb{Z}[\sqrt{-2}] = \{a + b\sqrt{-2} : a, b \in \mathbb{Z}\}.$$

This set is a ring because the sum and product of any two of its members are also members, whence it inherits the ring properties from $\mathbb{C}$. As on $\mathbb{Z}[i]$, the square of the absolute value gives a norm on $\mathbb{Z}[\sqrt{-2}]$, which is integer-valued and multiplicative. The "integers" in $\mathbb{Z}[i]$ and $\mathbb{Z}[\sqrt{-2}]$ are called *quadratic integers* because they satisfy quadratic equations with rational coefficients. We shall say a little more about quadratic integers in general in Section 7.8.

7.4.5. Show that $N(a+b\sqrt{-2}) = a^2 + 2b^2$. Use this norm to show that 5 is a "prime" in $\mathbb{Z}[\sqrt{-2}]$, and that 1 and -1 are the only units in $\mathbb{Z}[\sqrt{-2}]$.

7.5 Unique Gaussian Prime Factorization

We have now come to the point where further progress in the arithmetic of $\mathbb{Z}[i]$ depends on a uniqueness theorem for Gaussian prime factorization. At the same point in ordinary arithmetic (Section 1.6), we derived unique prime factorization from the fact that $\gcd(a,b) = ma+nb$ for some integers m and n, which follows in turn from the fact that $\gcd(a,b)$ is obtainable by the Euclidean algorithm.

The same argument applies in $\mathbb{Z}[i]$, except that there is no subtraction form of the Euclidean algorithm. We have to use division with remainder, which depends on the following.

Division property of $\mathbb{Z}[i]$. *If α and β are Gaussian integers with $\beta \neq 0$, then there are Gaussian integers μ and ρ with*

$$\alpha = \mu\beta + \rho \quad \textit{and} \quad N(\rho) < N(\beta).$$

Proof Because the norm N is the square of the absolute value, it suffices to find ρ with $\alpha = \mu\beta + \rho$ and $|\rho| < |\beta|$.

Consider the set of all Gaussian integer multiples of β. The points in this set lie at the corners of a grid of squares, namely the translates by multiples of β of the square with corners 0, β, $i\beta$, and $(1+i)\beta$. (Figure 7.4; the grid is square because multiplication by i rotates through a right angle.)

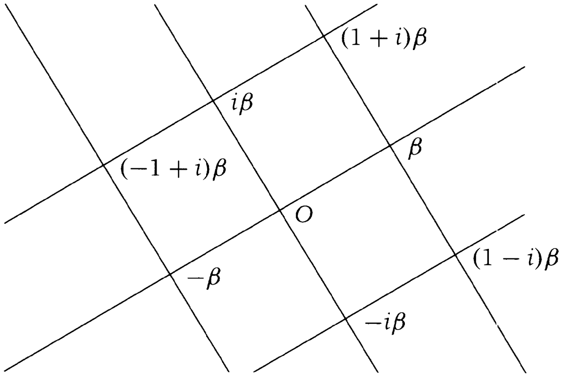

FIGURE 7.4 Multiples of a Gaussian integer.

Let $\mu\beta$ be the corner nearest to α, so $|\alpha - \mu\beta|$ is the distance between them. This distance is the hypotenuse of a right-angled triangle with sides $\leq |\beta|/2$ (Figure 7.5), hence $|\alpha - \mu\beta| < |\beta|$ by the triangle inequality. Thus if we let $\rho = \alpha - \mu\beta$ we have $\alpha = \mu\beta + \rho$ with $|\rho| < |\beta|$, as required. □

Thanks to the division property, the successive divisions in the Euclidean algorithm produce remainders with strictly decreasing norms. Because the norms are natural numbers, the algorithm terminates, and it produces the gcd for the same reason as in $\mathbb{Z}$: if the algorithm starts on α and β, *all* the divisors of α and β persist as divisors of all the numbers produced by the algorithm. The gcd of α and β is not only "greatest" in the sense that all common divisors divide it; it is also greatest in norm, by the norm multiplicative property.

One can then check that the remaining steps to unique prime factorization in $\mathbb{Z}$ can be imitated (with small changes) in $\mathbb{Z}[i]$:

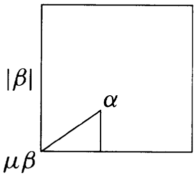

FIGURE 7.5 Distance to the nearest corner.

- $\gcd(\alpha, \beta) = \mu\alpha + \nu\beta$ for some Gaussian integers μ and ν.
- If a Gaussian prime ς divides $\alpha\beta$, then ς divides α or ς divides β (Gaussian prime divisor property).
- The factors in two Gaussian prime factorizations of a Gaussian integer agree up to order and unit factors.

The latter is the "unique prime factorization theorem" for Gaussian integers.

Before drawing conclusions from this theorem, a word of caution is in order: *watch out for units!* Remember that unique prime factorization was originally proved for natural numbers, where the factorization is unique up to order. In $\mathbb{Z}[i]$, the factors can also vary up to units, and this affects some of the conclusions we can draw. In fact, this already happens in $\mathbb{Z}$, where factors can vary in sign, due to the presence of the unit -1.

Take, for example, the theorem that relatively prime numbers a and b whose product is a square are themselves squares. This is true in the natural numbers (proved in Section 4.2), but in $\mathbb{Z}$ we can conclude only that a and b are either squares or the negatives of squares. The example $(-3^2)(-5^2) = 15^2$ shows we cannot do better. Similarly, if α and β are relatively prime Gaussian integers whose product is a square, we can conclude only that each of α and β is a unit times a square. The units i and $-i$ are not squares in $\mathbb{Z}[i]$, so α and β need not be squares.

However, things get better with cubes. In $\mathbb{Z}[i]$ all the units are cubes: $1 = 1^3$, $-1 = (-1)^3$, $i = (-i)^3$, and $-i = i^3$. Thus if α and β are relatively prime Gaussian integers whose product is a cube, then α and β are not merely units times cubes, but actual cubes, because a unit times a cube is a cube.

Exercises

As mentioned in the previous set of exercises, we can use division with remainder to give a direct proof that there are infinitely many Gaussian primes.

7.5.1. If $\alpha = \mu\beta + \rho$, with $0 < |\rho| < |\beta|$, show that α is not a multiple of β.

7.5.2. Use Exercise 7.5.1 to prove that there are infinitely many Gaussian primes.

The geometric argument used to prove the division property of $\mathbb{Z}[i]$ also applies to the ring $\mathbb{Z}[\sqrt{-2}]$ discussed in the previous set of exercises.

7.5.3. Show that the multiples of a number β in $\mathbb{Z}[\sqrt{-2}]$ lie at the corners of a grid of rectangles whose sides have lengths $|\beta|$ and $\sqrt{2}|\beta|$.

7.5.4. Deduce from Exercise 7.5.3 that $\mathbb{Z}[\sqrt{-2}]$ has division property like that of $\mathbb{Z}[i]$ and hence unique prime factorization.

In the exercises to Section 1.3 we mentioned that it is not known whether there are infinitely many primes of the form $p = n^2 + 1$.

7.5.5. Show that, if $p = n^2 + 1$ is prime, then $p = (n + i)(n - i)$ is a factorization into Gaussian primes.

7.5.6. Conversely, show that if n is a natural number and $n \pm i$ are Gaussian primes, then $n^2 + 1$ is prime.

(*Hint*: Suppose $n^2 + 1$ is not prime and use unique Gaussian prime factorization.)

It follows from the last two exercises that primes of the form $n^2 + 1$ correspond to Gaussian prime pairs of the form $n \pm i$. This calls to mind the famous *twin primes problem*, which is also unsolved: are there infinitely many pairs of ordinary primes of the form $(p, p + 2)$?

7.6 Fermat's Two Squares Theorem

Now is an appropriate time to recall the words of Diophantus quoted at the beginning of this chapter:

> 65 is naturally divided into two squares in two ways, namely into $7^2 + 4^2$ and $8^2 + 1^2$, which is due to the fact that 65 is the product of 13 and 5, each of which is the sum of two squares.

When Fermat read these words, he realized that the key to representing numbers as sums of two squares was to represent *primes*. In the example, 65 is the product of the primes 5 and 13, and the two representations of 65 as sums of two squares come from the unique