## LECTURE 12 Unique factorization domains

## PAVEL RŮŽIČKA

ABSTRACT. We define Gaussian monoids as commutative cancellative monoids with unique decomposition into products of irreducibles and we characterize them via the existence of greatest common divisors and decreasing chain conditions of the relation of divisibility. We apply this characterization to prove that principal ideal domains are unique factorization domains. Finally we characterize primes in the ring  $\mathbf{Z}[i]$  of Gaussian integers and show some applications.

12.1. **Gaussian monoids.** Let M be a commutative monoid. We say that factorizations  $a = b_1 \cdot b_2 \cdots b_m$  and  $a = c_1 \cdot c_2 \cdots c_n$  of an element  $a \in M$  as a product of elements of M are associated if m = n and there is a permutation  $\sigma$  of the set  $\{1, 2, \ldots, n\}$  such that  $b_i \sim c_{\sigma(i)}$  for all  $i \in \{1, 2, \ldots, n\}$ .

A commutative cancellative monoid M is Gaussian if every non-invertible element  $a \in M$  has a factorization  $a = q_1 \cdot q_2 \cdots q_n$  as a product of irreducible elements and all such factorizations of the element a are associated.

**Lemma 12.1.** Let M be a Gaussian monoid. Let  $a \mid b$  in M and

$$b = q_1^{\beta_1} \cdot q_2^{\beta_2} \cdots q_n^{\beta_n},$$

where  $q_1, q_2, \ldots, q_n$  are pairwise non-associated irreducible elements and  $0 \le \alpha_i$ , for all  $i = 1, 2, \ldots, n$ . Then

$$a = u \cdot q_1^{\alpha_1} \cdot q_2^{\alpha_2} \cdots q_n^{\alpha_n},$$

where u is invertible and  $0 \le \alpha_i \le \beta_i$ , for all i = 1, 2, ..., n.

*Proof.* Since  $a \mid b$ , there is  $c \in M$  such that  $b = a \cdot c$ . We prove the statement by induction on the sum  $\beta_1 + \beta_2 + \cdots + \beta_n$ . If  $\beta_1 + \beta_2 + \cdots + \beta_n = 0$ , then b is invertible and since  $a \mid b$ , a is invertible as well, and the statement holds true. Suppose that a is not invertible and let q be an irreducible element dividing a; let  $a = q \cdot a'$ . Then  $b = q \cdot a' \cdot c$ . Since M is a Gaussian monoid,  $a' \cdot c$  is either invertible or it has a unique, up to being associated, factorization into the product of irreducible elements  $p_1 \cdot p_2 \cdots p_m$ . Compering the factorizations

$$b = q_1^{\beta_1} \cdot q_2^{\beta_2} \cdots q_n^{\beta_n} = q \cdot p_1 \cdot p_2 \cdots p_m,$$

The Lecture and the tutorial took place in Malá strana, room S11, on January 8, 2019.

we infer that there is  $i \in \{1, 2, ..., n\}$  such that  $q \sim q_i$ . Therefore there is an invertible  $v \in M$  such that  $q = q_i \cdot v$ . We can without loss of generality assume that i = 1. Then, applying the cancellativity of the monoid M, we get that

$$v \cdot a' \cdot c = q_1^{\beta_1 - 1} \cdot q_2^{\beta_2} \cdots q_n^{\beta_n}.$$

From the induction hypothesis, we get that

$$(12.1) v \cdot a' = u \cdot q_1^{\alpha_1 - 1} \cdot q_2^{\alpha_2} \cdots q_n^{\alpha_n},$$

for some  $0 \le \alpha_i \le \beta_i$ , i = 1, 2, ..., n, and some invertible  $u \in M$ . Multiplying both sides of equation (12.1) by  $q_1$ , we conclude that

$$a = u \cdot q_1^{\alpha_1} \cdot q_2^{\alpha_2} \cdots q_n^{\alpha_n}.$$

Corollary 12.2. Let M be a Gaussian monoid,  $a, b \in M$  and

$$b \sim q_1^{\beta_1} \cdot q_2^{\beta_2} \cdots q_n^{\beta_n},$$

where  $q_1, q_2, \ldots, q_n$  are pairwise non-associated irreducible elements and  $0 \le \alpha_i$ , for all  $i = 1, 2, \ldots, n$ . Then  $a \mid b$  if and only if

$$a \sim q_1^{\alpha_1} \cdot q_2^{\alpha_2} \cdots q_n^{\alpha_n},$$

for some  $0 \le \alpha_i \le \beta_i$ ,  $i = 1, 2, \ldots, n$ .

**Lemma 12.3.** Let M be a Gaussian monoid. Then the greatest common divisor exits for every pair of elements of M.

*Proof.* Let  $a, b \in M$ . Since M is a Gaussian monoid, there are pairwise non-associated irreducible elements  $q_1, q_2, \ldots, q_n$  in M and integers  $0 \le \alpha_i, \beta_i, i = 1, 2, \ldots, n$ , such that

$$a \sim q_1^{\alpha_1} \cdot q_2^{\alpha_2} \cdots q_n^{\alpha_n}$$
 and  $b \sim q_1^{\beta_1} \cdot q_2^{\beta_2} \cdots q_n^{\beta_n}$ .

It follows readily from Corollary 12.2 that

$$(a,b) = [q_1^{\gamma_1} \cdot q_2^{\gamma_2} \cdots q_n^{\gamma_n}]_{\sim},$$

where  $\gamma_i = \min\{\alpha_i, \beta_i\}$ , for all  $i \in \{1, 2, ..., n\}$ .

Let M be a Gaussian monoid and  $a \in M$ . Let  $a = q_1^{\alpha_1} \cdot q_2^{\alpha_2} \cdots q_n^{\alpha_n}$ , where  $q_i$  are pairwise non-associated irreducible elements. We set  $h(a) := \alpha_1 + \alpha_2 + \cdots + \alpha_n$ . It follows readily from Corollary 12.2 that

$$(12.2) a \mid b \text{ and } b \nmid a \implies h(a) < h(b).$$

**Theorem 12.4.** Let M be a commutative cancellative monoid. The monoid M is Gaussian if an only if it satisfies the following two properties:

- (i) There is no infinite sequence  $a_1, a_2, \ldots$  of elements of M such that  $a_i \mid a_i$  if and only if  $i \leq j$ ;
- (ii) Every irreducible element of M is prime.

*Proof.* ( $\Rightarrow$ ) Suppose that the monoid M is Gaussian. An infinite sequence  $a_1, a_2, a_3, \ldots$  such that  $a_j \mid a_i$  if and only if  $i \leq j$  would induce an infinite strictly decreasing sequence  $h(a_1) > h(a_2) > \cdots$  of non-negative integers due to (12.2). Such a sequence does not exist. Therefore (i) holds true. Item (ii) follows from Theorem 11.12 and Lemma 12.3.

 $(\Leftarrow)$  Suppose that properties (i) and (ii) hold true. First we show that every non-invertible element of the monoid M is a product of irreducible elements.

Claim 1. Let  $\mathcal{U}$  be a non-empty subset of M. Then there is  $a \in \mathcal{U}$  such that

$$(12.3) b \mid a \implies a \sim b,$$

for all  $b \in \mathcal{U}$ .

Proof of Claim 1. Suppose otherwise, that is, there is a non-empty  $\mathcal{U} \subseteq M$  such that for every  $a \in \mathcal{U}$ , there is  $b \in \mathcal{U}$  such that  $b \mid a$  and  $a \nmid b$ . We can pick any  $a_1 \in \mathcal{U}$  and then construct inductively an infinite sequence  $a_1, a_2, \ldots$  such that  $a_{i+1} \mid a_i$  and  $a_i \nmid a_{i+1}$  for all  $i \in \mathbb{N}$ . It follows readily that  $a_j \mid a_i$  if and only if  $i \leq j$ , which violates (i).  $\square$  Claim 1.

Let us denote by  $\mathcal{U}$  the set of all non-invertible  $a \in M$  that are not products of irreducible elements, and suppose that  $\mathcal{U} \neq \emptyset$ . It follows from Claim 1 that there is  $a \in \mathcal{U}$  satisfying (12.3). Clearly none of the elements of  $\mathcal{U}$  is irreducible. It follows that  $a = b \cdot c$  for some  $b, c \in M$  with  $a \nmid b$  and  $a \nmid c$ . Since a satisfies (12.3), both  $b, c \notin \mathcal{U}$ . It follows that there are irreducible elements  $q_1, \ldots, q_n$  and  $p_1, \ldots, p_m$  such that  $b = q_1 \cdots q_n$  and  $c = p_1 \cdots p_m$ . From  $a = b \cdot c$ , we get that  $a = p_1 \cdots p_m \cdot q_1 \cdots q_n$ , which contradicts  $a \in \mathcal{U}$ .

Let  $a \in M$  and let  $a = p_1 \cdots p_n \sim q_1 \cdots q_m$  be two factorizations of a into a product of irreducible elements. We prove by induction on m that the two factorizations are associated. If m = 0, a is invertible, necessarily n = 0, and we are done. Since  $q_m$  is irreducible and therefore prime due to (ii),  $q_m \mid p_i$  for some  $i \in 1, 2, \ldots, n$ . Since  $p_i$  is irreducible, we infer that  $q_m \sim p_i$ . Since the monoid M is cancellative, we get that  $p_1 \cdots p_{i-1} \cdot p_{i+1} \cdots p_n \sim q_1 \cdots q_{m-1}$ . These two factorizations are associated due to the induction hypothesis.

12.2. Unique factorization domains. Let R be an integral domain. We say that R is a unique factorization domain<sup>1</sup> if the multiplicative monoid  $(R \setminus \{0\}, \cdot)$  of non-zero elements of R is a Gaussian monoid. This means, by the definition, that every non-invertible element of a unique factorization domain is a product of irreducible elements in a unique way up to the associated factorizations. It follows from Theorem 12.4 that unique factorization domains are characterized by the satisfaction of conditions (i) and (ii). Observe, applying equivalence (12.1), that property (i) is equivalent to

(i') There is no infinite strictly increasing chain of principal ideals.

<sup>&</sup>lt;sup>1</sup>Alterantively Gaussian domain

P. RŮŽIČKA

4

**Theorem 12.5.** Every principal ideal domain is a unique factorization domain.

*Proof.* Let R be a principal ideal domain. Suppose that there is an infinite sequence  $a_1, a_2, \ldots$  of elements of R such that  $(a_1) \subsetneq (a_2) \subsetneq \cdots$  and put  $I = \bigcup_{i \in \mathbb{N}} (a_i)$ . It is straightforward to see that a union of an increasing chain of ideal is an ideal, in particular I is an ideal of R. Since R is a principal ideal domain, the ideal I is principal, generated by, say  $b \in R$ . Since  $b \in I$ , there is  $n \in \mathbb{N}$  such that  $b \in (a_n)$ . It follows that  $I \subseteq (a_n) \subsetneq (a_{n+1}) \subseteq I$ , which is a contradiction. Therefore I satisfies (i'), hence (i). Property (ii) is due to Corollary 12.2.

**Example 12.6.** It is straightforward to verify that

$$Z[i\sqrt{3}] := \{a + i\sqrt{3}b \mid a, b \in Z\}$$

is an integral domain, indeed,

$$(a+i\sqrt{3}b)\cdot(c+i\sqrt{3}d) = (a\cdot c - 3\cdot b\cdot d) + i\sqrt{3}(b\cdot c + a\cdot d).$$

Let  $N(x+iy)=(x+iy)\cdot(x-iy)=x^2+y^2$  be a square of the complex norm. By (12.3) we have that  $N(\xi \cdot \eta)=N(\xi)\cdot N(\eta)$ . It follows that

(12.1) 
$$\alpha \mid \beta \implies N(\alpha) \mid N(\beta), \text{ for all } \alpha, \beta \in \mathbf{Z}[i\sqrt{3}].$$

It follows, that  $N(\alpha) = 1$ , for all invertible  $\alpha \in \mathbb{Z}[i\sqrt{3}]$ . On the other hand, if  $N(\alpha) = 1$ , we have that  $1 = \alpha \cdot \overline{\alpha}$ , and so  $\alpha$  is invertible. Therefore there are exactly two invertible elements in the domain, namely 1 and -1 and an element of  $\mathbb{Z}[i\sqrt{3}]$  is invertible if and only if its norm is 1. Observe that there is no  $\alpha = a + i\sqrt{3}b \in \mathbb{Z}[i\sqrt{3}]$  with  $N(\alpha) = 2$ . This is because we would have  $2 = a^2 + 3b^2$ , which is impossible. It follows that if  $\alpha \in \mathbb{Z}[i\sqrt{3}]$  satisfies  $N(\alpha) = 4$ , then  $\alpha$  is irreducible. We have that

$$(1+i\sqrt{3})\cdot(1-i\sqrt{3})=4=2.2,$$

and  $N(1+i\sqrt{3})=N(2)=4$ . The only elements of  $\mathbf{Z}[i\sqrt{3}]$  associated with 2 are its multiples by invertible elements, that is, 2 and -2. It follows that  $1+i\sqrt{3}\mid 2\cdot 2$  but  $1+i\sqrt{3}\nmid 2$ , and so  $1+i\sqrt{3}$  is irreducible (since  $N(1+i\sqrt{3})=4$ ) but not prime. Note also, that the elements  $2+i\sqrt{3}\cdot 2$  and 4 have no greatest common divisor in  $\mathbf{Z}[i\sqrt{3}]$ .

There is geometric reason, while we cannot prove that  $\mathbf{Z}[i\sqrt{3}]$  is an Euclidean domain in a similar way as we did for the domain  $\mathbf{Z}[i]$  of Gaussian integers. Elements of  $\mathbf{Z}[i\sqrt{3}]$  form a rectangular lattice in the complex plane consisting of rectangles with sides of length 1 and  $\sqrt{3}$  (see Figure 1). The distance of the center C of such a rectangle from each of the vertices is exactly 1. This is where the geometric argument successfully used for the Gaussian integers fails in case of the domain  $\mathbf{Z}[i\sqrt{3}]$ .

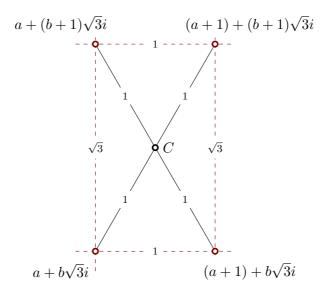


Figure 1. The critical point C

12.3. Primes in Gaussian integers. In the ring Z[i] is a unique factorization domain, in particular all irreducible elements of the ring are primes. We will describe all the primes in Z[i].

Similarly as in the case of  $Z[i\sqrt{3}]$ , it follows from (12.3) that

(12.1) 
$$\alpha \mid \beta \implies N(\alpha) \mid N(\beta), \text{ for all } \alpha, \beta \in \mathbf{Z}[i].$$

**Lemma 12.7.** An element  $\alpha \in \mathbb{Z}[i]$  is invertible if and only if  $N(\alpha) = 1$ . Here are four invertible elements in  $\mathbb{Z}[i]$ , namely 1, -1, i, -i.

*Proof.* If  $\alpha = a + ib \in \mathbf{Z}[i]$  is invertible, then  $\alpha \mid 1$ , hence  $N(\alpha) \mid 1$ , due to (12.1). Since  $N(\alpha) = a^2 + b^2$  is a positive integer, we conclude that  $N(\alpha) = 1$ . On the other hand, if  $N(\alpha) = \alpha \cdot \overline{\alpha} = 1$ , we readily see that  $\alpha$  is invertible (with the inverse  $\overline{\alpha}$ ).

From  $N(\alpha) = N(a+ib) = a^2 + b^2$ , we infer that  $N(\alpha) = 1$  if and only if either  $a \in \{-1, 1\}$  and b = 0 or a = 0 and  $b \in \{-1, 1\}$ . The four possible case give the four invertible elements listed in the lemma.

Corollary 12.8. Let a, b be positive integers. If  $a \sim b$  in  $\mathbb{Z}[i]$ , then a = b.

*Proof.* It follows from Lemma 11.4 that if  $a \sim b$ , then  $b = a \cdot \alpha$  for some invertible  $\alpha \in Z[i]$ . Since both a, b are positive integers, we get that  $\alpha = 1$ .

**Lemma 12.9.** If  $\alpha = a + ib \in \mathbb{Z}[i]$  is a prime, then  $\overline{\alpha} = a - ib$  is a prime as well.

*Proof.* It suffices to verify that  $\overline{\alpha}$  is irreducible. If  $\overline{\alpha}$  had a factorization  $\overline{\alpha} = \beta \cdot \gamma$  into the product of its proper divisors, then  $\alpha = \overline{\beta} \cdot \overline{\gamma} = \overline{\beta} \cdot \overline{\gamma}$  violate the irreducibility of  $\alpha$ , since  $\overline{\beta}$ ,  $\overline{\gamma}$  would be proper divisors of  $\alpha$ .  $\square$ 

P. RŮŽIČKA

Remark 12.10. The map given by  $\alpha \mapsto \overline{\alpha}$  is easily seen to be an automorphism of Z[i]. Any automorphism of an integral domain maps primes to primes.

**Lemma 12.11.** Let  $\alpha \in \mathbb{Z}[i]$  be a prime. Then  $N(\alpha)$  is either prime integer or a square of a prime integer.

*Proof.* Since  $\mathbf{Z}[i]$  is an Euklidean domain due to Lemma 12.6, it is a unique factorization domain due to Lemma 12.4 and Theorem 12.5. Let p be a prime positive integer such that  $p \mid \alpha \cdot \overline{\alpha}$ . Since  $\mathbf{Z}[i]$  is a unique factorization domain, either  $p = N(\alpha)$  or p is associated to one of  $\alpha$  and  $\overline{\alpha}$ . In the latter case, since p is an integer, we infer that  $p \sim \alpha$ , hence  $N(\alpha) = p^2$ .

**Lemma 12.12.** Let p be a positive integer. If  $p = a^2 + b^2$ , for some integers a, b, then  $p \not\equiv 3 \pmod{4}$ .

*Proof.* It follows from  $a^2 \equiv 0 \pmod{4}$  or  $a^2 \equiv 1 \pmod{4}$ , for every  $a \in \mathbb{Z}$ .

**Theorem 12.13.** Let  $\alpha = a + ib$  be a prime in  $\mathbb{Z}[i]$ . Then one of the following cases holds true.

- (i)  $N(\alpha) = 2$  and  $\alpha \sim 1 + i$ .
- (ii)  $N(\alpha) = p$ , where p is a positive prime integer such that  $p \equiv 1 \pmod{4}$ . In this case  $\alpha$  is associated to no integer in  $\mathbf{Z}[i]$ .
- (iii)  $N(\alpha) = p^2$ , where p is a positive prime integer such that  $p \equiv 3 \pmod{4}$  and  $\alpha \sim p$ .

*Proof.* Since  $2 = (1+i) \cdot (1-i)$  and  $(1-i) = (-i) \cdot (1+i) \sim (1+i)$ , we get (i) in case that  $2 \mid N(\alpha)$ .

Suppose that  $N(\alpha) = a^2 + b^2 = p$  is an odd prime. It follows from Lemma 12.12 that  $p \equiv 1 \pmod{4}$ , that is, p = 4k + 1 for some positive integer k.

On the other hand, let p = 4k + 1 be such a positive prime. Applying Wilson's theorem (Lemma 8.18), we get that

$$(p-1)! \equiv -1 \pmod{p},$$

hence

6

$$(-1)^{2k}((2k)!)^2 \equiv 1 \cdot 2 \cdots 2k \cdot (p-2k) \cdots (p-1) \equiv -1 \pmod{p},$$

whence  $p \mid ((2k)!)^2 + 1 = ((2k)! + i)((2k)! - i)$ . If p was prime in  $\mathbb{Z}[i]$ , we would get that either  $p \mid (2k)! + i$  or (2k)! - i. But none of these is the case. Therefore p decomposes in  $\mathbb{Z}[i]$ , say  $p = \alpha \cdot \beta$ . It follows that

$$p^{2} = \alpha \cdot \beta \cdot \overline{\alpha \cdot \beta} = (\alpha \cdot \overline{\alpha}) \cdot (\beta \cdot \overline{\beta}).$$

Since p is a positive prime integer, we conclude that  $p = \alpha \cdot \overline{\alpha} = N(\alpha)$ .

Finally, let p be a positive prime integer such that  $p \equiv 3 \pmod{4}$ . There is no  $\alpha \in \mathbb{Z}[i]$  with  $N(\alpha) = p$ . Since  $N(p) = p^2$ , p has no a proper non-invertible divisor in  $\mathbb{Z}[i]$ . It follows that p is a prime both in  $\mathbb{Z}$  and  $\mathbb{Z}[i]$ .

Since  $\mathbf{Z}[i]$  is a unique factorization domain,  $p^2 = p \cdot p$  is the unique factorization of  $p^2$  as a product of irreducible elements. It follows that here are no other primes in  $\mathbf{Z}[i]$ .

The theorem has a surprising corollary:

**Corollary 12.14.** Every positive prime integer p such that  $p \equiv 1 \pmod{4}$  is uniquely written as a sum of two squares.

We show some applications of Gaussian integers.

**Example 12.15.** Prove that there not positive integers a, b, k such that

$$(12.2) a^2 = \frac{b}{103^k - b}.$$

Solution. By a simple computation we derive from (12.2) that

$$(12.3) (a^2+1) \cdot b = 103^k \cdot a^2$$

Since  $a^2$  and  $a^2 + 1$  are relatively prime, we infer from (12.3) that  $(a^2 + 1) \mid 103^k$ . Since 103 is a prime,  $a^2 + 1 = 103^l$  for some  $l \leq k$ . Decomposing in  $\mathbf{Z}[i]$ , we get that  $103^l = (a+i) \cdot (a-i)$  which is not the case, for  $\mathbf{Z}[i]$  is a unique factorization domain and  $103 \equiv 3 \pmod{4}$ .

Example 12.16. Find all integer solutions of the Diophantine equation

$$(12.4) a^2 + 4 = b^3.$$

Solution. In the ring Z[i], we can decompose

$$(12.5) b^3 = a^2 + 4 = (a+2i) \cdot (a-2i).$$

Claim 2. It follows from (12.4) that a + 2i is a cube in  $\mathbb{Z}[i]$ .

Proof of Claim 2. Since (a+2i,a-2i)=(a+2i,-4i)=(a+2i,4) and  $4\sim (1+i)^4$ , the greatest common divisor of a+2i and a-2i is a power of 1+i. Observe that  $1+i\sim 1-i$ , indeed,  $1-i=(-i)\cdot (1+i)$ . Since  $(1+i)^k\mid a+2i$  if and only if  $(1-i)^k\mid a-2i$ , for all  $k\in\mathbb{N}$ , we conclude that  $(1+i)^k\mid a+2i$  if and only if  $(1+i)^k\mid a-2i$  if and only if  $2^k\mid a^2+4=b^3$ . It follows that the maximal k such that  $(1+i)^k\mid a=2i$  is divisible by 3, say k=3m. Therefore  $a+2i=(1+i)^{3m}\cdot\alpha$ ,  $a-2i=(1+i)^{3m}\cdot\beta$  and  $(\alpha,\beta)=1$ . It follows that  $\alpha\sim\gamma^3$  in  $\mathbf{Z}[i]$ , and so  $\alpha=\gamma^3\cdot\nu$ , where  $\nu$  is invertible. Observe that all invertible elements of  $\mathbf{Z}[i]$  are cubes, indeed,  $1=1^3, -1=(-1)^3, i=(-i)^3, \text{ and } (-i)=i^3$ . It follows that  $\alpha$  is a cube in  $\mathbf{Z}[i]$  and so is a+2i.

Applying binomial expansion we infer from Claim 2 that

$$a + 2i = (x + iy)^3 = x \cdot (x^2 - 3y^2) + iy \cdot (3x^2 - y^2),$$

hence

(12.6) 
$$a = x \cdot (x^2 - 3y^2)$$
 and  $2 = y \cdot (3x^2 - y^2)$ .

8 P. RŮŽIČKA

Since x, y are integers, we conclude that  $y \in \{-2, -1, 1, 2\}$ . From this we get, by case checking, that the only integer solutions of (12.6) are  $y \in \{1, -2\}$  and  $x = \pm 1$ . These correspond to the only integer solutions  $a = \pm 2, b = 2$  and  $a = \pm 11, b = 5$  of (12.4).

## EXERCISES

**Exercise 12.1.** Prove that  $Z[\sqrt{5}] := \{a + \sqrt{5}b \mid a, b \in Z\}$  is not a unique factorization domain.

Exercise 12.2. Find all integer solutions of the Diophantine equation

$$a^2 + 49 = b^3.$$

Exercise 12.3. Prove that the Diophantine equation

$$a^2 + 1 = b^3.$$

has no integer solution.

Exercise 12.4. Find all integer solutions of the Diophantine equation

$$(12.7) a^2 + 2 = b^3.$$

[Hint: Show that  $\mathbf{Z}[i\sqrt{2}] := \{a + i\sqrt{2}b \mid a, b \in \mathbf{Z}\}$  is an Euklidean domain and decompose (12.7) in  $\mathbf{Z}[i\sqrt{2}]$ .]