21. If R is a ring with unit element 1 and ϕ is a homomorphism of R into an integral domain R' such that $I(\phi) \neq R$, prove that $\phi(1)$ is the unit element of R'.

3.5 More Ideals and Quotient Rings

We continue the discussion of ideals and quotient rings.

Let us take the point of view, for the moment at least, that a field is the most desirable kind of ring. Why? If for no other reason, we can divide in a field, so operations and results in a field more closely approximate our experience with real and complex numbers. In addition, as was illustrated by Problem 2 in the preceding problem set, a field has no homomorphic images other than itself or the trivial ring consisting of 0. Thus we cannot simplify a field by applying a homomorphism to it. Taking these remarks into consideration it is natural that we try to link a general ring, in some fashion, with fields. What should this linkage involve? We have a machinery whose component parts are homomorphisms, ideals, and quotient rings. With these we will forge the link.

But first we must make precise the rather vague remarks of the preceding paragraph. We now ask the explicit question: Under what conditions is the homomorphic image of a ring a field? For commutative rings we give a complete answer in this section.

Essential to treating this question is the converse to the result of Problem 2 of the problem list at the end of Section 3.4.

LEMMA 3.5.1 *Let R be a commutative ring with unit element whose only ideals are (0) and R itself. Then R is a field.*

Proof. In order to effect a proof of this lemma for any $a \neq 0 \in R$ we must produce an element $b \neq 0 \in R$ such that $ab = 1$.

So, suppose that $a \neq 0$ is in R. Consider the set $Ra = \{xa \mid x \in R\}$. We claim that Ra is an ideal of R. In order to establish this as fact we must show that it is a subgroup of R under addition and that if $u \in Ra$ and $r \in R$ then ru is also in Ra. (We only need to check that ru is in Ra for then ur also is since $ru = ur$.)

Now, if $u, v \in Ra$, then $u = r_1 a$, $v = r_2 a$ for some $r_1, r_2 \in R$. Thus $u + v = r_1 a + r_2 a = (r_1 + r_2)a \in Ra$; similarly $-u = -r_1 a = (-r_1)a \in Ra$. Hence Ra is an additive subgroup of R. Moreover, if $r \in R$, $ru = r(r_1 a) = (rr_1)a \in Ra$. Ra therefore satisfies all the defining conditions for an ideal of R, hence is an ideal of R. (Notice that both the distributive law and associative law of multiplication were used in the proof of this fact.)

By our assumptions on R, $Ra = (0)$ or $Ra = R$. Since $0 \neq a = 1a \in Ra$, $Ra \neq (0)$; thus we are left with the only other possibility, namely that $Ra = R$. This last equation states that every element in R is a multiple of

a by some element of R. In particular, $1 \in R$ and so it can be realized as a multiple of a; that is, there exists an element $b \in R$ such that $ba = 1$. This completes the proof of the lemma.

DEFINITION An ideal $M \neq R$ in a ring R is said to be a *maximal ideal* of R if whenever U is an ideal of R such that $M \subset U \subset R$, then either $R = U$ or $M = U$.

In other words, an ideal of R is a maximal ideal if it is impossible to squeeze an ideal between it and the full ring. Given a ring R there is no guarantee that it has any maximal ideals! If the ring has a unit element this can be proved, assuming a basic axiom of mathematics, the so-called axiom of choice. Also there may be many distinct maximal ideals in a ring R; this will be illustrated for us below in the ring of integers.

As yet we have made acquaintance with very few rings. Only by considering a given concept in many particular cases can one fully appreciate the concept and its motivation. Before proceeding we therefore examine some maximal ideals in two specific rings. When we come to the discussion of polynomial rings we shall exhibit there all the maximal ideals.

Example 3.5.1 Let R be the ring of integers, and let U be an ideal of R. Since U is a subgroup of R under addition, from our results in group theory, we know that U consists of all the multiples of a fixed integer n_0; we write this as $U = (n_0)$. What values of n_0 lead to maximal ideals?

We first assert that if p is a prime number then $P = (p)$ is a maximal ideal of R. For if U is an ideal of R and $U \supset P$, then $U = (n_0)$ for some integer n_0. Since $p \in P \subset U$, $p = mn_0$ for some integer m; because p is a prime this implies that $n_0 = 1$ or $n_0 = p$. If $n_0 = p$, then $P \subset U = (n_0) \subset P$, so that $U = P$ follows; if $n_0 = 1$, then $1 \in U$, hence $r = 1r \in U$ for all $r \in R$ whence $U = R$ follows. Thus no ideal, other than R or P itself, can be put between P and R, from which we deduce that P is maximal.

Suppose, on the other hand, that $M = (n_0)$ is a maximal ideal of R. We claim that n_0 must be a prime number, for if $n_0 = ab$, where a, b are positive integers, then $U = (a) \supset M$, hence $U = R$ or $U = M$. If $U = R$, then $a = 1$ is an easy consequence; if $U = M$, then $a \in M$ and so $a = rn_0$ for some integer r, since every element of M is a multiple of n_0. But then $n_0 = ab = rn_0b$, from which we get that $rb = 1$, so that $b = 1$, $n_0 = a$. Thus n_0 is a prime number.

In this particular example the notion of maximal ideal comes alive—it corresponds exactly to the notion of prime number. One should not, however, jump to any hasty generalizations; this kind of correspondence does not usually hold for more general rings.

Example 3.5.2 Let R be the ring of all the real-valued, continuous functions on the closed unit interval. (See Example 3.3.5.) Let

$$M = \{f(x) \in R \mid f(\tfrac{1}{2}) = 0\}.$$

M is certainly an ideal of R. Moreover, it is a maximal ideal of R, for if the ideal U contains M and $U \neq M$, then there is a function $g(x) \in U$, $g(x) \notin M$. Since $g(x) \notin M$, $g(\frac{1}{2}) = \alpha \neq 0$. Now $h(x) = g(x) - \alpha$ is such that $h(\frac{1}{2}) = g(\frac{1}{2}) - \alpha = 0$, so that $h(x) \in M \subset U$. But $g(x)$ is also in U; therefore $\alpha = g(x) - h(x) \in U$ and so $1 = \alpha\alpha^{-1} \in U$. Thus for any function $t(x) \in R$, $t(x) = 1t(x) \in U$, in consequence of which $U = R$. M is therefore a maximal ideal of R. Similarly if γ is a real number $0 \leq \gamma \leq 1$, then $M_\gamma = \{f(x) \in R \mid f(\gamma) = 0\}$ is a maximal ideal of R. It can be shown (see Problem 4 at the end of this section) that every maximal ideal is of this form. Thus here the maximal ideals correspond to the points on the unit interval.

Having seen some maximal ideals in some concrete rings we are ready to continue the general development with

THEOREM 3.5.1 *If R is a commutative ring with unit element and M is an ideal of R, then M is a maximal ideal of R if and only if R/M is a field.*

Proof. Suppose, first, that M is an ideal of R such that R/M is a field. Since R/M is a field its only ideals are (0) and R/M itself. But by Theorem 3.4.1 there is a one-to-one correspondence between the set of ideals of R/M and the set of ideals of R which contain M. The ideal M of R corresponds to the ideal (0) of R/M whereas the ideal R of R corresponds to the ideal R/M of R/M in this one-to-one mapping. Thus there is no ideal between M and R other than these two, whence M is a maximal ideal.

On the other hand, if M is a maximal ideal of R, by the correspondence mentioned above R/M has only (0) and itself as ideals. Furthermore R/M is commutative and has a unit element since R enjoys both these properties. All the conditions of Lemma 3.5.1 are fulfilled for R/M so we can conclude, by the result of that lemma, that R/M is a field.

We shall have many occasions to refer back to this result in our study of polynomial rings and in the theory of field extensions.

Problems

1. Let R be a ring with unit element, R not necessarily commutative, such that the only right-ideals of R are (0) and R. Prove that R is a division ring.

*2. Let R be a ring such that the only right ideals of R are (0) and R. Prove that either R is a division ring or that R is a ring with a prime number of elements in which $ab = 0$ for every $a, b \in R$.

3. Let J be the ring of integers, p a prime number, and (p) the ideal of J consisting of all multiples of p. Prove
   (a) $J/(p)$ is isomorphic to J_p, the ring of integers mod p.
   (b) Using Theorem 3.5.1 and part (a) of this problem, that J_p is a field.

**4. Let R be the ring of all real-valued continuous functions on the closed unit interval. If M is a maximal ideal of R, prove that there exists a real number γ, $0 \leq \gamma \leq 1$, such that $M = M_\gamma = \{f(x) \in R \mid f(\gamma) = 0\}$.

3.6 The Field of Quotients of an Integral Domain

Let us recall that an integral domain is a commutative ring D with the additional property that it has no zero-divisors, that is, if $ab = 0$ for some $a, b \in D$ then at least one of a or b must be 0. The ring of integers is, of course, a standard example of an integral domain.

The ring of integers has the attractive feature that we can enlarge it to the set of rational numbers, which is a field. Can we perform a similar construction for any integral domain? We will now proceed to show that indeed we can!

DEFINITION A ring R *can be imbedded* in a ring R' if there is an isomorphism of R into R'. (If R and R' have unit elements 1 and $1'$ we insist, in addition, that this isomorphism takes 1 onto $1'$.)

R' will be called an *over-ring* or *extension* of R if R can be imbedded in R'.

With this understanding of imbedding we prove

THEOREM 3.6.1 *Every integral domain can be imbedded in a field.*

Proof. Before becoming explicit in the details of the proof let us take an informal approach to the problem. Let D be our integral domain; roughly speaking the field we seek should be all quotients a/b, where $a, b \in D$ and $b \neq 0$. Of course in D, a/b may very well be meaningless. What should we require of these symbols a/b? Clearly we must have an answer to the following three questions:

1. When is $a/b = c/d$?
2. What is $(a/b) + (c/d)$?
3. What is $(a/b)(c/d)$?

In answer to 1, what could be more natural than to insist that $a/b = c/d$

if and only if $ad = bc$? As for 2 and 3, why not try the obvious, that is, define

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd} \quad \text{and} \quad \frac{a}{b}\frac{c}{d} = \frac{ac}{bd}.$$

In fact in what is to follow we make these considerations our guide. So let us leave the heuristics and enter the domain of mathematics, with precise definitions and rigorous deductions.

Let $\mathcal{M}$ be the set of all ordered pairs (a, b) where $a, b \in D$ and $b \neq 0$. (Think of (a, b) as a/b.) In $\mathcal{M}$ we now define a relation as follows:

$$(a, b) \sim (c, d) \text{ if and only if } ad = bc.$$

We claim that this defines an equivalence relation on $\mathcal{M}$. To establish this we check the three defining conditions for an equivalence relation for this particular relation.

1. If $(a, b) \in \mathcal{M}$, then $(a, b) \sim (a, b)$ since $ab = ba$.
2. If $(a, b), (c, d) \in \mathcal{M}$ and $(a, b) \sim (c, d)$, then $ad = bc$, hence $cb = da$, and so $(c, d) \sim (a, b)$.
3. If $(a, b), (c, d), (e, f)$ are all in $\mathcal{M}$ and $(a, b) \sim (c, d)$ and $(c, d) \sim (e, f)$, then $ad = bc$ and $cf = de$. Thus $bcf = bde$, and since $bc = ad$, it follows that $adf = bde$. Since D is commutative, this relation becomes $afd = bed$; since, moreover, D is an integral domain and $d \neq 0$, this relation further implies that $af = be$. But then $(a, b) \sim (e, f)$ and our relation is transitive.

Let $[a, b]$ be the equivalence class in $\mathcal{M}$ of (a, b), and let F be the set of all such equivalence classes $[a, b]$ where $a, b \in D$ and $b \neq 0$. F is the candidate for the field we are seeking. In order to create out of F a field we must introduce an addition and a multiplication for its elements and then show that under these operations F forms a field.

We first dispose of the addition. Motivated by our heuristic discussion at the beginning of the proof we define

$$[a, b] + [c, d] = [ad + bc, bd].$$

Since D is an integral domain and both $b \neq 0$ and $d \neq 0$ we have that $bd \neq 0$; this, at least, tells us that $[ad + bc, bd] \in F$. We now assert that this addition is well defined, that is, if $[a, b] = [a', b']$ and $[c, d] = [c', d']$, then $[a, b] + [c, d] = [a', b'] + [c', d']$. To see that this is so, from $[a, b] = [a', b']$ we have that $ab' = ba'$; from $[c, d] = [c', d']$ we have that $cd' = dc'$. What we need is that these relations force the equality of $[a, b] + [c, d]$ and $[a', b'] + [c', d']$. From the definition of addition this boils down to showing that $[ad + bc, bd] = [a'd' + b'c', b'd']$, or, in equivalent terms, that $(ad + bc)b'd' = bd(a'd' + b'c')$. Using $ab' = ba'$, $cd' = dc'$

this becomes: $(ad + bc)b'd' = adb'd' + bcb'd' = ab'dd' + bb'cd' = ba'dd' + bb'dc' = bd(a'd' + b'c')$, which is the desired equality.

Clearly $[0, b]$ acts as a zero-element for this addition and $[-a, b]$ as the negative of $[a, b]$. It is a simple matter to verify that F is an abelian group under this addition.

We now turn to the multiplication in F. Again motivated by our preliminary heuristic discussion we define $[a, b][c, d] = [ac, bd]$. As in the case of addition, since $b \neq 0$, $d \neq 0$, $bd \neq 0$ and so $[ac, bd] \in F$. A computation, very much in the spirit of the one just carried out, proves that if $[a, b] = [a', b']$ and $[c, d] = [c', d']$ then $[a, b][c, d] = [a', b'][c', d']$. One can now show that the nonzero elements of F (that is, all the elements $[a, b]$ where $a \neq 0$) form an abelian group under multiplication in which $[d, d]$ acts as the unit element and where

$$[c, d]^{-1} = [d, c] \text{ (since } c \neq 0, [d, c] \text{ is in } F).$$

It is a routine computation to see that the distributive law holds in F. F is thus a field.

All that remains is to show that D can be imbedded in F. We shall exhibit an explicit isomorphism of D into F. Before doing so we first notice that for $x \neq 0, y \neq 0$ in D, $[ax, x] = [ay, y]$ because $(ax)y = x(ay)$; let us denote $[ax, x]$ by $[a, 1]$. Define $\phi:D \to F$ by $\phi(a) = [a, 1]$ for every $a \in D$. We leave it to the reader to verify that ϕ is an isomorphism of D into F, and that if D has a unit element 1, then $\phi(1)$ is the unit element of F. The theorem is now proved in its entirety.

F is usually called the *field of quotients* of D. In the special case in which D is the ring of integers, the F so constructed is, of course, the field of rational numbers.

Problems

1. Prove that if $[a, b] = [a', b']$ and $[c, d] = [c', d']$ then $[a, b][c, d] = [a', b'][c', d']$.
2. Prove the distributive law in F.
3. Prove that the mapping $\phi:D \to F$ defined by $\phi(a) = [a, 1]$ is an isomorphism of D into F.
4. Prove that if K is any field which contains D then K contains a subfield isomorphic to F. (*In this sense F is the smallest field containing D.*)

*5. Let R be a commutative ring with unit element. A nonempty subset S of R is called a multiplicative system if
   1. $0 \notin S$.
   2. $s_1, s_2 \in S$ implies that $s_1 s_2 \in S$.

Let $\mathscr{M}$ be the set of all ordered pairs (r, s) where $r \in R$, $s \in S$. In $\mathscr{M}$ define $(r, s) \sim (r', s')$ if there exists an element $s'' \in S$ such that

$$s''(rs' - sr') = 0.$$

(a) Prove that this defines an equivalence relation on $\mathscr{M}$.

Let the equivalence class of (r, s) be denoted by $[r, s]$, and let R_S be the set of all the equivalence classes. In R_S define $[r_1, s_1] + [r_2, s_2] = [r_1s_2 + r_2s_1, s_1s_2]$ and $[r_1, s_1][r_2, s_2] = [r_1r_2, s_1s_2]$.

(b) Prove that the addition and multiplication described above are well defined and that R_S forms a ring under these operations.

(c) Can R be imbedded in R_S?

(d) Prove that the mapping $\phi: R \to R_s$ defined by $\phi(a) = [as, s]$ is a homomorphism of R into R_S and find the kernel of ϕ.

(e) Prove that this kernel has no element of S in it.

(f) Prove that every element of the form $[s_1, s_2]$(where $s_1, s_2 \in S$) in R_S has an inverse in R_S.

6. Let D be an integral domain, $a, b \in D$. Suppose that $a^n = b^n$ and $a^m = b^m$ for two relatively prime positive integers m and n. Prove that $a = b$.

7. Let R be a ring, possibly noncommutative, in which $xy = 0$ implies $x = 0$ or $y = 0$. If $a, b \in R$ and $a^n = b^n$ and $a^m = b^m$ for two relatively prime positive integers m and n, prove that $a = b$.

3.7 Euclidean Rings

The class of rings we propose to study now is motivated by several existing examples—the ring of integers, the Gaussian integers (Section 3.8), and polynomial rings (Section 3.9). The definition of this class is designed to incorporate in it certain outstanding characteristics of the three concrete examples listed above.

DEFINITION An integral domain R is said to be a *Euclidean ring* if for every $a \neq 0$ in R there is defined a nonnegative integer $d(a)$ such that

1. For all $a, b \in R$, both nonzero, $d(a) \leq d(ab)$.
2. For any $a, b \in R$, both nonzero, there exist $t, r \in R$ such that $a = tb + r$ where either $r = 0$ or $d(r) < d(b)$.

We do not assign a value to $d(0)$. The integers serve as an example of a Euclidean ring, where $d(a) =$ absolute value of a acts as the required function. In the next section we shall see that the Gaussian integers also form a Euclidean ring. Out of that observation, and the results developed in this part, we shall prove a classic theorem in number theory due to

Fermat, namely, that every prime number of the form $4n + 1$ can be written as the sum of two squares.

We begin with

THEOREM 3.7.1 *Let R be a Euclidean ring and let A be an ideal of R. Then there exists an element $a_0 \in A$ such that A consists exactly of all a_0x as x ranges over R.*

Proof. If A just consists of the element 0, put $a_0 = 0$ and the conclusion of the theorem holds.

Thus we may assume that $A \neq (0)$; hence there is an $a \neq 0$ in A. Pick an $a_0 \in A$ such that $d(a_0)$ is minimal. (Since d takes on nonnegative integer values this is always possible.)

Suppose that $a \in A$. By the properties of Euclidean rings there exist $t, r \in R$ such that $a = ta_0 + r$ where $r = 0$ or $d(r) < d(a_0)$. Since $a_0 \in A$ and A is an ideal of R, ta_0 is in A. Combined with $a \in A$ this results in $a - ta_0 \in A$; but $r = a - ta_0$, whence $r \in A$. If $r \neq 0$ then $d(r) < d(a_0)$, giving us an element r in A whose d-value is smaller than that of a_0, in contradiction to our choice of a_0 as the element in A of minimal d-value. Consequently $r = 0$ and $a = ta_0$, which proves the theorem.

We introduce the notation $(a) = \{xa \mid x \in R\}$ to represent the ideal of all multiples of a.

DEFINITION An integral domain R with unit element is a *principal ideal ring* if every ideal A in R is of the form $A = (a)$ for some $a \in R$.

Once we establish that a Euclidean ring has a unit element, in virtue of Theorem 3.7.1, we shall know that a Euclidean ring is a principal ideal ring. The converse, however, is false; there are principal ideal rings which are not Euclidean rings. [See the paper by T. Motzkin, *Bulletin of the American Mathematical Society*, Vol. 55 (1949), pages 1142–1146, entitled "The Euclidean algorithm."]

COROLLARY TO THEOREM 3.7.1 *A Euclidean ring possesses a unit element.*

Proof. Let R be a Euclidean ring; then R is certainly an ideal of R, so that by Theorem 3.7.1 we may conclude that $R = (u_0)$ for some $u_0 \in R$. Thus every element in R is a multiple of u_0. Therefore, in particular, $u_0 = u_0c$ for some $c \in R$. If $a \in R$ then $a = xu_0$ for some $x \in R$, hence $ac = (xu_0)c = x(u_0c) = xu_0 = a$. Thus c is seen to be the required unit element.

DEFINITION If $a \neq 0$ and b are in a commutative ring R then a is said to *divide* b if there exists a $c \in R$ such that $b = ac$. We shall use the symbol

$a \mid b$ to represent the fact that a divides b and $a \not\mid b$ to mean that a does not divide b.

The proof of the next remark is so simple and straightforward that we omit it.

REMARK 1. *If $a \mid b$ and $b \mid c$ then $a \mid c$.*
2. *If $a \mid b$ and $a \mid c$ then $a \mid (b \pm c)$.*
3. *If $a \mid b$ then $a \mid bx$ for all $x \in R$.*

DEFINITION If $a, b \in R$ then $d \in R$ is said to be a *greatest common divisor* of a and b if

1. $d \mid a$ and $d \mid b$.
2. Whenever $c \mid a$ and $c \mid b$ then $c \mid d$.

We shall use the notation $d = (a, b)$ to denote that d is a greatest common divisor of a and b.

LEMMA 3.7.1 *Let R be a Euclidean ring. Then any two elements a and b in R have a greatest common divisor d. Moreover $d = \lambda a + \mu b$ for some $\lambda, \mu \in R$.*

Proof. Let A be the set of all elements $ra + sb$ where r, s range over R. We claim that A is an ideal of R. For suppose that $x, y \in A$; therefore $x = r_1a + s_1b$, $y = r_2a + s_2b$, and so $x \pm y = (r_1 \pm r_2)a + (s_1 \pm s_2)b \in A$. Similarly, for any $u \in R$, $ux = u(r_1a + s_1b) = (ur_1)a + (us_1)b \in A$.

Since A is an ideal of R, by Theorem 3.7.1 there exists an element $d \in A$ such that every element in A is a mutiple of d. By dint of the fact that $d \in A$ and that every element of A is of the form $ra + sb$, $d = \lambda a + \mu b$ for some $\lambda, \mu \in R$. Now by the corollary to Theorem 3.7.1, R has a unit element 1; thus $a = 1a + 0b \in A$, $b = 0a + 1b \in A$. Being in A, they are both multiples of d, whence $d \mid a$ and $d \mid b$.

Suppose, finally, that $c \mid a$ and $c \mid b$; then $c \mid \lambda a$ and $c \mid \mu b$ so that c certainly divides $\lambda a + \mu b = d$. Therefore d has all the requisite conditions for a greatest common divisor and the lemma is proved.

DEFINITION Let R be a commutative ring with unit element. An element $a \in R$ is a *unit* in R if there exists an element $b \in R$ such that $ab = 1$.

Do not confuse a unit with a unit element! A unit in a ring is an element whose inverse is also in the ring.

LEMMA 3.7.2 *Let R be an integral domain with unit element and suppose that for $a, b \in R$ both $a \mid b$ and $b \mid a$ are true. Then $a = ub$, where u is a unit in R.*

Proof. Since $a \mid b$, $b = xa$ for some $x \in R$; since $b \mid a$, $a = yb$ for some $y \in R$. Thus $b = x(yb) = (xy)b$; but these are elements of an integral domain, so that we can cancel the b and obtain $xy = 1$; y is thus a unit in R and $a = yb$, proving the lemma.

DEFINITION Let R be a commutative ring with unit element. Two elements a and b in R are said to be *associates* if $b = ua$ for some unit u in R.

The relation of being associates is an equivalence relation. (Problem 1 at the end of this section.) Note that in a Euclidean ring any two greatest common divisors of two given elements are associates (Problem 2).

Up to this point we have, as yet, not made use of condition 1 in the definition of a Euclidean ring, namely that $d(a) \leq d(ab)$ for $b \neq 0$. We now make use of it in the proof of

LEMMA 3.7.3 *Let R be a Euclidean ring and $a, b \in R$. If $b \neq 0$ is not a unit in R, then $d(a) < d(ab)$.*

Proof. Consider the ideal $A = (a) = \{xa \mid x \in R\}$ of R. By condition 1 for a Euclidean ring, $d(a) \leq d(xa)$ for $x \neq 0$ in R. Thus the d-value of a is the minimum for the d-value of any element in A. Now $ab \in A$; if $d(ab) = d(a)$, by the proof used in establishing Theorem 3.7.1, since the d-value of ab is minimal in regard to A, every element in A is a multiple of ab. In particular, since $a \in A$, a must be a multiple of ab; whence $a = abx$ for some $x \in R$. Since all this is taking place in an integral domain we obtain $bx = 1$. In this way b is a unit in R, in contradiction to the fact that it was not a unit. The net result of this is that $d(a) < d(ab)$.

DEFINITION In the Euclidean ring R a nonunit π is said to be a *prime element* of R if whenever $\pi = ab$, where a, b are in R, then one of a or b is a unit in R.

A prime element is thus an element in R which cannot be factored in R in a nontrivial way.

LEMMA 3.7.4 *Let R be a Euclidean ring. Then every element in R is either a unit in R or can be written as the product of a finite number of prime elements of R.*

Proof. The proof is by induction on $d(a)$.

If $d(a) = d(1)$ then a is a unit in R (Problem 3), and so in this case, the assertion of the lemma is correct.

We assume that the lemma is true for all elements x in R such that $d(x) < d(a)$. On the basis of this assumption we aim to prove it for a. This would complete the induction and prove the lemma.

If a is a prime element of R there is nothing to prove. So suppose that $a = bc$ where neither b nor c is a unit in R. By Lemma 3.7.3, $d(b) < d(bc) = d(a)$ and $d(c) < d(bc) = d(a)$. Thus by our induction hypothesis b and c can be written as a product of a finite number of prime elements of R; $b = \pi_1\pi_2 \cdots \pi_n$, $c = \pi'_1\pi'_2 \cdots \pi'_m$ where the π's and π''s are prime elements of R. Consequently $a = bc = \pi_1\pi_2 \cdots \pi_n\pi'_1\pi'_2 \cdots \pi'_m$ and in this way a has been factored as a product of a finite number of prime elements. This completes the proof.

DEFINITION In the Euclidean ring R, a and b in R are said to be *relatively prime* if their greatest common divisor is a unit of R.

Since any associate of a greatest common divisor is a greatest common divisor, and since 1 is an associate of any unit, if a and b are relatively prime we may assume that $(a, b) = 1$.

LEMMA 3.7.5 *Let R be a Euclidean ring. Suppose that for $a, b, c \in R$, $a \mid bc$ but $(a, b) = 1$. Then $a \mid c$.*

Proof. As we have seen in Lemma 3.7.1, the greatest common divisor of a and b can be realized in the form $\lambda a + \mu b$. Thus by our assumptions, $\lambda a + \mu b = 1$. Multiplying this relation by c we obtain $\lambda ac + \mu bc = c$. Now $a \mid \lambda ac$, always, and $a \mid \mu bc$ since $a \mid bc$ by assumption; therefore $a \mid (\lambda ac + \mu bc) = c$. This is, of course, the assertion of the lemma.

We wish to show that prime elements in a Euclidean ring play the same role that prime numbers play in the integers. If π in R is a prime element of R and $a \in R$, then either $\pi \mid a$ or $(\pi, a) = 1$, for, in particular, (π, a) is a divisor of π so it must be π or 1 (or any unit). If $(\pi, a) = 1$, one-half our assertion is true; if $(\pi, a) = \pi$, since $(\pi, a) \mid a$ we get $\pi \mid a$, and the other half of our assertion is true.

LEMMA 3.7.6 *If π is a prime element in the Euclidean ring R and $\pi \mid ab$ where $a, b \in R$ then π divides at least one of a or b.*

Proof. Suppose that π does not divide a; then $(\pi, a) = 1$. Applying Lemma 3.7.5 we are led to $\pi \mid b$.

COROLLARY *If π is a prime element in the Euclidean ring R and $\pi \mid a_1a_2 \cdots a_n$ then π divides at least one $a_1, a_2, \ldots, a_n$.*

We carry the analogy between prime elements and prime numbers further and prove

THEOREM 3.7.2 (UNIQUE FACTORIZATION THEOREM) *Let R be a Euclidean ring and $a \neq 0$ a nonunit in R. Suppose that $a = \pi_1\pi_2 \cdots \pi_n = \pi'_1\pi'_2 \cdots \pi'_m$ where the π_i and π'_j are prime elements of R. Then $n = m$ and each π_i, $1 \leq i \leq n$ is an associate of some π'_j, $1 \leq j \leq m$ and conversely each π'_k is an associate of some π_q.*

Proof. Look at the relation $a = \pi_1\pi_2 \cdots \pi_n = \pi'_1\pi'_2 \cdots \pi'_m$. But $\pi_1 \mid \pi_1\pi_2 \cdots \pi_n$, hence $\pi_1 \mid \pi'_1\pi'_2 \cdots \pi'_m$. By Lemma 3.7.6, π_1 must divide some π'_i; since π_1 and π'_i are both prime elements of R and $\pi_1 \mid \pi'_i$ they must be associates and $\pi'_i = u_1\pi_1$, where u_1 is a unit in R. Thus $\pi_1\pi_2 \cdots \pi_n = \pi'_1\pi'_2 \cdots \pi'_m = u_1\pi_1\pi'_2 \cdots \pi'_{i-1}\pi'_{i+1} \cdots \pi'_m$; cancel off π_1 and we are left with $\pi_2 \cdots \pi_n = u_1\pi'_2 \cdots \pi'_{i-1}\pi'_{i+1} \cdots \pi'_m$. Repeat the argument on this relation with π_2. After n steps, the left side becomes 1, the right side a product of a certain number of π' (the excess of m over n). This would force $n \leq m$ since the π' are not units. Similarly, $m \leq n$, so that $n = m$. In the process we have also showed that every π_i has some π'_i as an associate and conversely.

Combining Lemma 3.7.4 and Theorem 3.7.2 we have that *every nonzero element in a Euclidean ring R can be uniquely written (up to associates) as a product of prime elements or is a unit in R.*

We finish the section by determining all the maximal ideals in a Euclidean ring.

In Theorem 3.7.1 we proved that any ideal A in the Euclidean ring R is of the form $A = (a_0)$ where $(a_0) = \{xa_0 \mid x \in R\}$. We now ask: What conditions imposed on a_0 insure that A is a maximal ideal of R? For this question we have a simple, precise answer, namely

LEMMA 3.7.7 *The ideal $A = (a_0)$ is a maximal ideal of the Euclidean ring R if and only if a_0 is a prime element of R.*

Proof. We first prove that if a_0 is not a prime element, then $A = (a_0)$ is not a maximal ideal. For, suppose that $a_0 = bc$ where $b, c \in R$ and neither b nor c is a unit. Let $B = (b)$; then certainly $a_0 \in B$ so that $A \subset B$. We claim that $A \neq B$ and that $B \neq R$.

If $B = R$ then $1 \in B$ so that $1 = xb$ for some $x \in R$, forcing b to be a unit in R, which it is not. On the other hand, if $A = B$ then $b \in B = A$ whence $b = xa_0$ for some $x \in R$. Combined with $a_0 = bc$ this results in $a_0 = xca_0$, in consequence of which $xc = 1$. But this forces c to be a unit in R, again contradicting our assumption. Therefore B is neither A nor R and since $A \subset B$, A cannot be a maximal ideal of R.

Conversely, suppose that a_0 is a prime element of R and that U is an ideal of R such that $A = (a_0) \subset U \subset R$. By Theorem 3.7.1, $U = (u_0)$. Since $a_0 \in A \subset U = (u_0)$, $a_0 = xu_0$ for some $x \in R$. But a_0 is a prime element of R, from which it follows that either x or u_0 is a unit in R. If u_0

is a unit in R then $U = R$ (see Problem 5). If, on the other hand, x is a unit in R, then $x^{-1} \in R$ and the relation $a_0 = xu_0$ becomes $u_0 = x^{-1}a_0 \in A$ since A is an ideal of R. This implies that $U \subset A$; together with $A \subset U$ we conclude that $U = A$. Therefore there is no ideal of R which fits strictly between A and R. This means that A is a maximal ideal of R.

Problems

1. In a commutative ring with unit element prove that the relation a is an associate of b is an equivalence relation.
2. In a Euclidean ring prove that any two greatest common divisors of a and b are associates.
3. Prove that a necessary and sufficient condition that the element a in the Euclidean ring be a unit is that $d(a) = d(1)$.
4. Prove that in a Euclidean ring (a, b) can be found as follows:

$$\begin{aligned} b &= q_0 a + r_1, \quad \text{where} \quad d(r_1) < d(a) \\ a &= q_1 r_1 + r_2, \quad \text{where} \quad d(r_2) < d(r_1) \\ r_1 &= q_2 r_2 + r_3, \quad \text{where} \quad d(r_3) < d(r_2) \\ &\vdots \qquad\qquad\qquad\qquad\qquad\qquad \vdots \\ r_{n-1} &= q_n r_n \end{aligned}$$

and $$r_n = (a, b).$$

5. Prove that if an ideal U of a ring R contains a unit of R, then $U = R$.
6. Prove that the units in a commutative ring with a unit element form an abelian group.
7. Given two elements a, b in the Euclidean ring R their *least common multiple* $c \in R$ is an element in R such that $a \mid c$ and $b \mid c$ and such that whenever $a \mid x$ and $b \mid x$ for $x \in R$ then $c \mid x$. Prove that any two elements in the Euclidean ring R have a least common multiple in R.
8. In Problem 7, if the least common multiple of a and b is denoted by $[a, b]$, prove that $[a, b] = ab/(a, b)$.

3.8 A Particular Euclidean Ring

An abstraction in mathematics gains in substance and importance when, particularized to a specific example, it sheds new light on this example. We are about to particularize the notion of a Euclidean ring to a concrete ring, the ring of Gaussian integers. Applying the general results obtained about Euclidean rings to the Gaussian integers we shall obtain a highly nontrivial theorem about prime numbers due to Fermat.

Let $J[i]$ denote the set of all complex numbers of the form $a + bi$ where a and b are integers. Under the usual addition and multiplication of complex numbers $J[i]$ forms an integral domain called the domain of *Gaussian integers.*

Our first objective is to exhibit $J[i]$ as a Euclidean ring. In order to do this we must first introduce a function $d(x)$ defined for every nonzero element in $J[i]$ which satisfies

1. $d(x)$ is a nonnegative integer for every $x \neq 0 \in J[i]$.
2. $d(x) \leq d(xy)$ for every $y \neq 0$ in $J[i]$.
3. Given $u, v \in J[i]$ there exist $t, r \in J[i]$ such that $v = tu + r$ where $r = 0$ or $d(r) < d(u)$.

Our candidate for this function d is the following: if $x = a + bi \in J[i]$, then $d(x) = a^2 + b^2$. The $d(x)$ so defined certainly satisfies property 1; in fact, if $x \neq 0 \in J[i]$ then $d(x) \geq 1$. As is well known, for any two complex numbers (not necessarily in $J[i]$) x, y, $d(xy) = d(x)d(y)$; thus if x and y are in addition in $J[i]$ and $y \neq 0$, then since $d(y) \geq 1$, $d(x) = d(x)1 \leq d(x)d(y) = d(xy)$, showing that condition 2 is satisfied. All our effort now will be to show that condition 3 also holds for this function d in $J[i]$. This is done in the proof of

THEOREM 3.8.1 *$J[i]$ is a Euclidean ring.*

Proof. As was remarked in the discussion above, to prove Theorem 3.8.1 we merely must show that, given $x, y \in J[i]$ there exists $t, r \in J[i]$ such that $y = tx + r$ where $r = 0$ or $d(r) < d(x)$.

We first establish this for a very special case, namely, where y is arbitrary in $J[i]$ but where x is an (ordinary) positive integer n. Suppose that $y = a + bi$; by the division algorithm for the ring of integers we can find integers u, v such that $a = un + u_1$ and $b = vn + v_1$ where u_1 and v_1 are integers satisfying $|u_1| \leq \frac{1}{2}n$ and $|v_1| \leq \frac{1}{2}n$. Let $t = u + vi$ and $r = u_1 + v_1 i$; then $y = a + bi = un + u_1 + (vn + v_1)i = (u + vi)n + u_1 + v_1 i = tn + r$. Since $d(r) = d(u_1 + v_1 i) = u_1{}^2 + v_1{}^2 \leq n^2/4 + n^2/4 < n^2 = d(n)$, we see that in this special case we have shown that $y = tn + r$ with $r = 0$ or $d(r) < d(n)$.

We now go to the general case; let $x \neq 0$ and y be arbitrary elements in $J[i]$. Thus $x\bar{x}$ is a positive integer n where $\bar{x}$ is the complex conjugate of x. Applying the result of the paragraph above to the elements $y\bar{x}$ and n we see that there are elements $t, r \in J[i]$ such that $y\bar{x} = tn + r$ with $r = 0$ or $d(r) < d(n)$. Putting into this relation $n = x\bar{x}$ we obtain $d(y\bar{x} - tx\bar{x}) < d(n) = d(x\bar{x})$; applying to this the fact that $d(y\bar{x} - tx\bar{x}) = d(y - tx)d(\bar{x})$ and $d(x\bar{x}) = d(x)d(\bar{x})$ we obtain that $d(y - tx)d(\bar{x}) < d(x)d(\bar{x})$. Since $x \neq 0$, $d(\bar{x})$ is a positive integer, so this inequality simplifies to $d(y - tx) < d(x)$. We represent $y = tx + r_0$, where $r_0 = y - tx$; thus t and r_0 are in

$J[i]$ and as we saw above, $r_0 = 0$ or $d(r_0) = d(y - tx) < d(x)$. This proves the theorem.

Since $J[i]$ has been proved to be a Euclidean ring, we are free to use the results established about this class of rings in the previous section to the Euclidean ring we have at hand, $J[i]$.

LEMMA 3.8.1 *Let p be a prime integer and suppose that for some integer c relatively prime to p we can find integers x and y such that $x^2 + y^2 = cp$. Then p can be written as the sum of squares of two integers, that is, there exist integers a and b such that $p = a^2 + b^2$.*

Proof. The ring of integers is a subring of $J[\mathrm{i}]$. Suppose that the integer p is also a prime element of $J[i]$. Since $cp = x^2 + y^2 = (x + yi)(x - yi)$, by Lemma 3.7.6, $p \mid (x + yi)$ or $p \mid (x - yi)$ in $J[i]$. But if $p \mid (x + yi)$ then $x + yi = p(u + vi)$ which would say that $x = pu$ and $y = pv$ so that p also would divide $x - yi$. But then $p^2 \mid (x + yi)(x - yi) = cp$ from which we would conclude that $p \mid c$ contrary to assumption. Similarly if $p \mid (x - yi)$. Thus p is *not* a prime element in $J[i]$! In consequence of this,

$$p = (a + bi)(g + di)$$

where $a + bi$ and $g + di$ are in $J[i]$ and where neither $a + bi$ nor $g + di$ is a unit in $J[i]$. But this means that neither $a^2 + b^2 = 1$ nor $g^2 + d^2 = 1$. (See Problem 2.) From $p = (a + bi)(g + di)$ it follows easily that $p = (a - bi)(g - di)$. Thus

$$p^2 = (a + bi)(g + di)(a - bi)(g - di) = (a^2 + b^2)(g^2 + d^2).$$

Therefore $(a^2 + b^2) \mid p^2$ so $a^2 + b^2 = 1, p$ or p^2; $a^2 + b^2 \neq 1$ since $a + bi$ is not a unit, in $J[i]$; $a^2 + b^2 \neq p^2$, otherwise $g^2 + d^2 = 1$, contrary to the fact that $g + di$ is not a unit in $J[i]$. Thus the only feasibility left is that $a^2 + b^2 = p$ and the lemma is thereby established.

The odd prime numbers divide into two classes, those which have a remainder of 1 on division by 4 and those which have a remainder of 3 on division by 4. We aim to show that every prime number of the first kind can be written as the sum of two squares, whereas no prime in the second class can be so represented.

LEMMA 3.8.2 *If p is a prime number of the form $4n + 1$, then we can solve the congruence $x^2 \equiv -1 \bmod p$.*

Proof. Let $x = 1 \cdot 2 \cdot 3 \cdots (p - 1)/2$. Since $p - 1 = 4n$, in this product for x there are an even number of terms, in consequence of which

$$x = (-1)(-2)(-3) \cdots \left(-\left(\frac{p-1}{2}\right)\right).$$

But $p - k \equiv -k \bmod p$, so that

$$\begin{aligned} x^2 &\equiv \left(1 \cdot 2 \cdots \frac{p-1}{2}\right)(-1)(-2)\cdots\left(-\left(\frac{p-1}{2}\right)\right) \\ &\equiv 1 \cdot 2 \cdots \frac{p-1}{2}\,\frac{p+1}{2} \cdots (p-1) \\ &\equiv (p-1)! = -1 \bmod p. \end{aligned}$$

We are using here Wilson's theorem, proved earlier, namely that if p is a prime number $(p - 1)! \equiv -1(p)$.

To illustrate this result, if $p = 13$,

$$x = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 = 720 = 5 \bmod 13 \text{ and } 5^2 = -1 \bmod 13.$$

THEOREM 3.8.2 (FERMAT) *If p is a prime number of the form $4n + 1$, then $p = a^2 + b^2$ for some integers a, b.*

Proof. By Lemma 3.8.2 there exists an x such that $x^2 \equiv -1 \bmod p$. The x can be chosen so that $0 \leq x \leq p - 1$ since we only need to use the remainder of x on division by p. We can restrict the size of x even further, namely to satisfy $|x| \leq p/2$. For if $x > p/2$, then $y = p - x$ satisfies $y^2 \equiv -1 \bmod p$ but $|y| \leq p/2$. Thus we may assume that we have an integer x such that $|x| \leq p/2$ and $x^2 + 1$ is a multiple of p, say cp. Now $cp = x^2 + 1 \leq p^2/4 + 1 < p^2$, hence $c < p$ and so $p \nmid c$. Invoking Lemma 3.8.1 we obtain that $p = a^2 + b^2$ for some integers a and b, proving the theorem.

Problems

1. Find all the units in $J[i]$.
2. If $a + bi$ is not a unit of $J[i]$ prove that $a^2 + b^2 > 1$.
3. Find the greatest common divisor in $J[i]$ of
   (a) $3 + 4i$ and $4 - 3i$. (b) $11 + 7i$ and $18 - i$.
4. Prove that if p is a prime number of the form $4n + 3$, then there is no x such that $x^2 \equiv -1 \bmod p$.
5. Prove that no prime of the form $4n + 3$ can be written as $a^2 + b^2$ where a and b are integers.
6. Prove that there is an infinite number of primes of the form $4n + 3$.

*7. Prove there exists an infinite number of primes of the form $4n + 1$.

*8. Determine all the prime elements in $J[i]$.

*9. Determine all positive integers which can be written as a sum of two squares (of integers).