Frequently Used Notation

$\|A\|$	the order of the set A
$\|x\|$	the order of the element x
$\mathbb{Z}$, $\mathbb{Z}^+$	the integers, the positive integers
$\mathbb{Q}$, $\mathbb{Q}^+$	the rational numbers, the positive rational numbers
$\mathbb{R}$, $\mathbb{R}^+$	the real numbers, the positive real numbers
$\mathbb{C}$	the complex numbers
$\mathbb{Z}/n\mathbb{Z}$	the integers modulo n
$(\mathbb{Z}/n\mathbb{Z})^\times$	the (multiplicative group of) invertible integers modulo n
$A \times B$	the direct or Cartesian product of A and B
$A \oplus B$	the direct sum of A and B
X^r	the direct or Cartesian product of r copies of X
$f^{-1}(A)$	the inverse image or preimage of A under f
(a, b)	the greatest common divisor of a, b also the ideal generated by a, b
$a\|b$	a divides b
$H \leq G$	H is a subgroup of G
$H < G$	H is a proper subgroup of G
Z_n	the cyclic group of order n
D_{2n}	the dihedral group of order $2n$
S_Ω	the symmetric group on the set Ω
S_n	the symmetric group on n letters
A_n	the alternating group on n letters
Q_8	the quaternion group of order 8
V_4	the Klein 4-group
QD_{16}	the quasidihedral group of order 16
$\mathbb{F}_N$	the finite field of N elements
$GL_n(F)$	the general linear group
$SL_n(F)$	the special linear group
$\mathrm{Aut}(G)$	the automorphism group of G
$C_G(A)$	the centralizer in G of A
$N_G(A)$	the normalizer in G of A
$Z(G)$	the center of G
G_s	the stabilizer in G of s
x^g, H^g	the (right) conjugate of x (respectively, H) by g

$\langle A \rangle$	the group generated by the set A
$\langle x \rangle$	the group generated by the element x
$G = \langle . \ . \ . \mid . \ . \ . \rangle$	generators and relations (presentation) for G
$\ker \varphi$	the kernel of φ
$\text{im } \varphi$	the image of φ
$N \trianglelefteq G$	N is normal in G
$\lvert G:H \rvert$	the index of H in G
H char G	H is characteristic in G
$Syl_p(G)$	the Sylow p-subgroups of G
n_p	the number of Sylow p-subgroups of G
$[x, y]$	the commutator of x, y
$H \rtimes K$	the semidirect product of H and K
$F[x]$	polynomials in x with coefficients in F
$M_n(R)$	$n \times n$ matrices over R
$M_{n\times m}(R)$	$n \times m$ matrices over R
$M_{\mathcal{B}}^{\mathcal{E}}(\varphi)$	the matrix of the linear transformation φ with respect to bases $\mathcal{B}, \mathcal{E}$
$m_T(x)$	the minimal polynomial of the linear transformation T
$c_T(x)$	the characteristic polynomial of the linear transformation T
$\text{ch}(F)$	the characteristic of F
K/F	the field K is an extension of the field F
$[K:F]$	the degree of the field extension K/F
$F(\alpha)$, $F(\alpha, \beta)$, etc.	the field generated over F by α or α, β, etc.
$m_{\alpha,F}(x)$	the minimal polynomial of α over F
$\text{Aut}(K)$	the group of automorphisms of a field K
$\text{Aut}(K/F)$	the group of automorphisms of a field K fixing the field F
$\text{Gal}(K/F)$	the Galois group of the extension K/F
$\text{N}_{K/F}(\alpha)$	the norm of α from K to F
$\text{Tr}_{K/F}(\alpha)$	the trace of α from K to F
$tr(A)$	the trace of the matrix A
FG	the group ring of G over F
$GL(V)$	the group of invertible linear transformations of V
(θ, ψ)	the inner product of the characters θ, ψ
$\lVert \theta \rVert$	the norm of the character θ
$\text{Ind}_H^G(\psi)$	character of the induced representation

Dummit, David Steven.
Abstract algebra / David S. Dummit, Richard M. Foote.
p. cm.
ISBN 0-13-004771-6
1. Algebra, Abstract. I. Foote, Richard M., 1950-
II. Title.
QA162.D85 1991
512'.02—dc20 90-35456
CIP

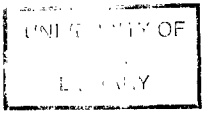
Editorial Production/Supervision: **Ed Thomas**
Cover design: **George Cornell**
Manufacturing buyer: **Paula Massenaro**

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

ISBN 0-13-004771-6

PRENTICE-HALL INTERNATIONAL (UK) LIMITED, *London*
PRENTICE-HALL OF AUSTRALIA PTY, LIMITED, *Sydney*
PRENTICE-HALL CANADA INC., *Toronto*
PRENTICE-HALL HISPANOAMERICANA, S.A., *Mexico*
PRENTICE-HALL OF INDIA PRIVATE LIMITED, *New Delhi*
PRENTICE-HALL OF JAPAN, INC., *Tokyo*
SIMON & SCHUSTER ASIA PTE. LTD., *Singapore*
EDITORA PRENTICE-HALL DO BRASIL, LTDA., *Rio de Janeiro*



Contents

Preface

This book evolved out of notes by the authors from courses given at various universities over a period of about thirteen years. The backgrounds of students in these courses were quite diverse, ranging from freshman and sophomore undergraduates to beginning graduate students, and included computer science and other science and engineering majors in addition to mathematics majors. The expectations of the students were quite varied. For some, the algebra course was to be an introduction with no follow up (some only took one semester of the full year sequence) and for some the course was intended to be the foundation for future work in mathematics. We felt that all of these students should be given some insight into the main themes in abstract algebra and some understanding of how these themes provide a unifying framework for the study of the basic algebraic structures: groups, rings, and fields. One of our guiding principles has therefore been to introduce as early as possible notions such as homomorphisms, isomorphisms, actions, classifications, and the notion of the structure of an algebraic object. We point out how these themes recur as we pass from the study of one algebraic system to another and provide a natural flow for the development of basic abstract algebra. To a certain extent this has influenced our choice of material as well. We have made no attempt at being encyclopedic, but rather we have delved more deeply into a few areas in order that students may achieve a better grasp of how the structure of algebraic objects is unravelled in greater depth. Wherever possible we have attempted to broaden students' horizons by pointing out connections of one subject to other areas of algebra and mathematics and by mentioning results at the forefront of research. We also introduce new material in a form which students do not subsequently have to dislodge in favor of more

sophisticated versions. So, for example, permutations are introduced in cycle notation format and quotients are developed as fibers of homomorphisms. In this way students are immediately exposed to concepts the way an algebraist (or mathematician) envisions and works with them.

The book contains a wealth of examples and exercises. The exercises range from routine computations to fairly sophisticated theoretical ones. All the exercises, however, are within the scope of the text and hints are given [in brackets] where we felt they were needed to make them accessible. The main results in the text do not rely on material in the exercises, although some details in proofs in the text are occasionally left to the reader and some examples use results from the exercises. In many instances, new material is introduced first in the exercises (often a few sections before it appears in the text) so that students may obtain an easier introduction to it by doing these exercises (e.g. Lagrange's Theorem appears in the exercises in Section 1.7 and in the text in Section 3.2). The exercises also contain a considerable amount of material not covered in the text. Some exercises lead through a series of steps to new results such as Wedderburn's Theorem on finite division rings or Hilbert's Basis Theorem. There sometimes appear sequences of exercises on one topic; these problems are prefaced with some remarks and are structured so that they may be read as text without actually doing the exercises. Some of these sequences are reviews (such as the exercises on Gauss–Jordan elimination in Chapter 11) and others contain new material relating to another area (such as the exercises on the exponential of a matrix and its application to systems of differential equations in Section 12.3). In addition, a number of exercises describe useful computational techniques (for example, Berlekamp's Algorithm for factoring polynomials mod p in Section 14.3).

Notationally, Proposition 2 in Chapter 4 is referred to simply as Proposition 2 within Chapter 4 and as Proposition 4.2 in other chapters.

It is not usually possible to cover the entire text in one academic year, although the authors have been able to cover the bulk of the material in one year, including all of Part III and much of Part V. A basic introductory (one year) course should include Chapters 1 to 3, Sections 4.1 to 4.3, Sections 5.1 to 5.3, Chapters 7 to 9, Sections 10.1, 11.1 and Chapters 13 and 14, returning to pick up Sections 4.4 to 4.6 if they are not covered initially. Parts III and V may also be used as a well integrated one semester course. We note that linear algebra is not assumed to be a prerequisite although knowledge of it does increase the repertoire of examples in group theory. The basic linear algebra covered in Chapter 11 provides the background mainly for the theory of canonical forms in Chapter 12.

Finally, we would like to thank D. Dorman, H. Kisilevsky, J. Sands, L. Simons and our students over the past few years for helpful suggestions and comments.

D. Dummit
R. Foote

Preliminaries

Some results and notation that are used throughout the text are collected in this chapter for convenience. Students may wish to review this chapter quickly at first and then read each section more carefully again as the concepts appear in the course of the text.

0.1 BASICS

The basics of set theory: sets, $\cap, \cup, \in$, etc. should be familiar to the reader. Our notation for subsets of a given set A will be

$$B = \{a \in A \mid \ldots \text{ (conditions on a) } \ldots\}.$$

The *order* or *cardinality* of a set A will be denoted by $|A|$. If A is a finite set the order of A is simply the number of elements of A.

It is important to understand how to test whether a particular $x \in A$ lies in a subset B of A (cf. Exercises 1 to 4). The *Cartesian product* of two sets A and B is the collection $A \times B = \{(a, b) \mid a \in A, b \in B\}$, of ordered pairs of elements from A and B.

We shall use the following notation for some common sets of numbers:

(1) $\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \ldots\}$ denotes the *integers* (the $\mathbb{Z}$ is for the German word for numbers: "Zahlen").

(2) $\mathbb{Q} = \{a/b \mid a, b \in \mathbb{Z}, b \neq 0\}$ denotes the *rational numbers* (or *rationals*).

(3) $\mathbb{R} = \{$ all decimal expansions $\pm d_1 d_2 \ldots d_n . a_1 a_2 a_3 \ldots \}$ denotes the *real numbers* (or *reals*).

(4) $\mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}, i^2 = -1\}$ denotes the *complex numbers*.

(5) $\mathbb{Z}^+, \mathbb{Q}^+$ and $\mathbb{R}^+$ will denote the positive (nonzero) elements in $\mathbb{Z}, \mathbb{Q}$ and $\mathbb{R}$, respectively.

We shall use the notation $f : A \to B$ or $A \xrightarrow{f} B$ to denote a function f from A to B and the value of f at a is denoted $f(a)$ (i.e. we shall apply all our functions on the left). We use the words *function* and *map* interchangeably. The set A is called the *domain* of f and B is called the *codomain* of f. The notation $f : a \mapsto b$ or $a \mapsto b$ if f is understood indicates that $f(a) = b$, i.e. the function is being specified on *elements*.

If the function f is not specified on elements it is important in general to check that f is *well defined*, i.e. is unambiguously determined. For example, if the set A is the union of two subsets A_1 and A_2 then one can try to specify a function from A to the set $\{0, 1\}$ by declaring that f is to map everything in A_1 to 0 and is to map everything in A_2 to 1. This unambiguously defines f unless A_1 and A_2 have elements in common (in which case it is not clear whether these elements should map to 0 or to 1). Checking that this f is well defined therefore amounts to checking that A_1 and A_2 have no interseetion.

The set

$$f(A) = \{b \in B \mid b = f(a), \text{ for some } a \in A\}$$

is a subset of B, called the *range* or *image* of f (or the *image of A under f*). For each subset C of B the set

$$f^{-1}(C) = \{a \in A \mid f(a) \in C\}$$

consisting of the elements of A mapping into C under f is called the *preimage* or *inverse image* of C under f. For each $b \in B$, the preimage of $\{b\}$ under f is called the *fiber* of f over b. Note that f^{-1} is not in general a function and that the fibers of f generally contain many elements since there may be many elements of A mapping to the element b.

If $f : A \to B$ and $g : B \to C$, then the composite map $g \circ f : A \to C$ is defined by

$$(g \circ f)(a) = g(f(a)).$$

Let $f : A \to B$.

(1) f is *injective* or is an *injection* if whenever $a_1 \neq a_2$, then $f(a_1) \neq f(a_2)$.

(2) f is *surjective* or is a *surjection* if for all $b \in B$ there is some $a \in A$ such that $f(a) = b$, i.e. the image of f is *all* of B. Note that since a function always maps onto its range (by definition) it is necessary to specify the codomain B in order for the question of surjectivity to be meaningful.

(3) f is *bijective* or is a *bijection* if it is both injective and surjective. If such a bijection f exists from A to B, we say A and B are in *bijective correspondence*.

(4) f has a *left inverse* if there is a function $g : B \to A$ such that $g \circ f : A \to A$ is the identity map on A, i.e. $(g \circ f)(a) = a$, for all $a \in A$.

(5) f has a *right inverse* if there is a function $h : B \to A$ such that $f \circ h : B \to B$ is the identity map on B.

Proposition 1. Let $f : A \to B$.

(1) The map f is injective if and only if f has a left inverse.

(2) The map f is surjective if and only if f has a right inverse.

(3) The map f is a bijection if and only if there exists $g : B \to A$ such that $f \circ g$ is the identity map on B and $g \circ f$ is the identity map on A.

(4) If A and B are finite sets with the same number of elements (i.e. $|A| = |B|$), then $f : A \to B$ is bijective if and only if f is injective if and only if f is surjective.

Proof: Exercise.

In the situation of part (3) of the proposition above the map g is necessarily unique and we shall say g is the *2-sided inverse* (or simply the *inverse*) of f.

A *permutation* of a set A is simply a bijection from A to itself.

If $A \subseteq B$ and $f : B \to C$, we denote the *restriction* of f to A by $f|_A$. When the domain we are considering is understood we shall occasionally denote $f|_A$ again simply as f even though these are formally different functions (their domains are different).

If $A \subseteq B$ and $g : A \to C$ and there is a function $f : B \to C$ such that $f|_A = g$, we shall say f is an *extension* of g to B (such a map f need not exist nor be unique).

Let A be a nonempty set.

(1) A *binary relation* on a set A is a subset R of $A \times A$ and we write $a \sim b$ if $(a, b) \in R$.

(2) The relation $\sim$ on A is said to be:

(a) *reflexive* if $a \sim a$, for all $a \in A$,

(b) *symmetric* if $a \sim b$ implies $b \sim a$ for all $a, b \in A$,

(c) *transitive* if $a \sim b$ and $b \sim c$ implies $a \sim c$ for all $a, b, c \in A$.

A relation is an *equivalence relation* if it is reflexive, symmetric and transitive.

(3) If $\sim$ defines an equivalence relation on A, then the *equivalence class* of $a \in A$ is defined to be $\{x \in A \mid x \sim a\}$. Elements of the equivalence class of a are said to be *equivalent* to a. If C is an equivalence class, any element of C is called a *representative* of the class C.

(4) A *partition* of A is any collection $\{A_i \mid i \in I\}$ of nonempty subsets of A (I some indexing set) such that

(a) $A = \cup_{i \in I} A_i$, and

(b) $A_i \cap A_j = \emptyset$, for all $i, j \in I$ with $i \neq j$,

i.e. A is the disjoint union of the sets in the partition.

The notions of an equivalence relation on A and a partition of A are the same:

Proposition 2. Let A be a nonempty set.

(1) If $\sim$ defines an equivalence relation on A then the set of equivalence classes of $\sim$ form a partition of A.

(2) If $\{A_i \mid i \in I\}$ is a partition of A then there is an equivalence relation on A whose equivalence classes are precisely the sets A_i, $i \in I$.

Proof: Omitted.

Finally, we shall assume the reader is familiar with proofs by induction.

EXERCISES

In Exercises 1 to 4 let $\mathcal{A}$ be the set of 2×2 matrices with real number entries. Recall that matrix multiplication is defined by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} p & q \\ r & s \end{pmatrix} = \begin{pmatrix} ap + br & aq + bs \\ cp + dr & cq + ds \end{pmatrix}.$$

Let

$$M = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

and let

$$\mathcal{B} = \{X \in \mathcal{A} \mid MX = XM\}.$$

1. Determine which of the following elements of $\mathcal{A}$ lie in $\mathcal{B}$:

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

2. Prove that if $P, Q \in \mathcal{B}$, then $P + Q \in \mathcal{B}$ (where $+$ denotes the usual sum of two matrices).
3. Prove that if $P, Q \in \mathcal{B}$, then $P \cdot Q \in \mathcal{B}$ (where $\cdot$ denotes the usual product of two matrices).
4. Find conditions on p, q, r, s which determine precisely when $\begin{pmatrix} p & q \\ r & s \end{pmatrix} \in \mathcal{B}$.
5. Determine whether the following functions f are well defined:
   (a) $f : \mathbb{Q} \to \mathbb{Z}$ defined by $f(a/b) = a$.
   (b) $f : \mathbb{Q} \to \mathbb{Q}$ defined by $f(a/b) = a^2/b^2$.
6. Determine whether the function $f : \mathbb{R}^+ \to \mathbb{Z}$ defined by mapping a real number r to the first digit to the right of the decimal point in a decimal expansion of r is well defined.
7. Let $f : A \to B$ be a surjective map of sets. Prove that the relation

   $$a \sim b \text{ if and only if } f(a) = f(b)$$

   is an equivalence relation whose equivalence classes are the fibers of f.

0.2 PROPERTIES OF THE INTEGERS

The following properties of the integers $\mathbb{Z}$ (many familiar from elementary arithmetic) will be proved in a more general context in the ring theory of Chapter 8, but it will be necessary to use them in Part I (of course, none of the ring theory proofs of these properties will rely on the group theory).

(1) (Well Ordering of $\mathbb{Z}$) If A is any nonempty subset of $\mathbb{Z}^+$, there is some element $m \in A$ such that $m \leq a$, for all $a \in A$ (m is called a *minimal element* of A).

(2) If $a, b \in \mathbb{Z}$ with $a \neq 0$, we say a *divides* b if there is an element $c \in \mathbb{Z}$ such that $b = ac$. In this case we write $a \mid b$; if a does not divide b we write $a \nmid b$.

(3) If $a, b \in \mathbb{Z} - \{0\}$, there is a unique positive integer d, called the *greatest common divisor of a and b* (or g.c.d. of a and b), satisfying:
   (a) $d \mid a$ and $d \mid b$ (so d is a common divisor of a and b), and
   (b) if $e \mid a$ and $e \mid b$, then $e \mid d$ (so d is the greatest such divisor).
   The g.c.d. of a and b will be denoted by (a, b). If $(a, b) = 1$, we say that a and b are *relatively prime*.

(4) If $a, b \in \mathbb{Z} - \{0\}$, there is a unique positive integer l, called the *least common multiple of a and b* (or l.c.m. of a and b), satisfying:

(a) $a \mid l$ and $b \mid l$ (so l is a common multiple of a and b), and
(b) if $a \mid m$ and $b \mid m$, then $l \mid m$ (so l is the least such multiple).
The connection between the greatest common divisor d and the least common multiple l of two integers a and b is given by $dl = ab$.

(5) The *Division Algorithm* : if $a, b \in \mathbb{Z} - \{0\}$, then there exist unique $q, r \in \mathbb{Z}$ such that

$$a = qb + r \quad \text{and} \quad 0 \leq r < |b|,$$

where q is the *quotient* and r the *remainder*. This is the usual "long division" familiar from elementary arithmetic.

(6) The *Euclidean Algorithm* is an important procedure which produces a greatest common divisor of two integers a and b by iterating the Division Algorithm: if $a, b \in \mathbb{Z} - \{0\}$, then we obtain a sequence of quotients and remainders

$$\begin{aligned} a &= q_0 b + r_0 && (0) \\ b &= q_1 r_0 + r_1 && (1) \\ r_0 &= q_2 r_1 + r_2 && (2) \\ r_1 &= q_3 r_2 + r_3 && (3) \\ &\vdots \\ r_{n-2} &= q_n r_{n-1} + r_n && (n) \\ r_{n-1} &= q_{n+1} r_n && (n+1) \end{aligned}$$

where r_n is the last nonzero remainder. Such an r_n exists since $|b| > |r_0| > |r_1| > \cdots > |r_n|$ is a decreasing sequence of strictly positive integers if the remainders are nonzero and such a sequence cannot continue indefinitely. Then r_n is the g.c.d. (a, b) of a and b.

Example

Suppose $a = 57970$ and $b = 10353$. Then applying the Euclidean Algorithm we obtain:

$$\begin{aligned} 57970 &= (5)10353 + 6205 \\ 10353 &= (1)6205 + 4148 \\ 6205 &= (1)4148 + 2057 \\ 4148 &= (2)2057 + 34 \\ 2057 &= (60)34 + 17 \\ 34 &= (2)17 \end{aligned}$$

which shows that $(57970, 10353) = 17$.

(7) One consequence of the Euclidean Algorithm which we shall use regularly is the following: if $a, b \in \mathbb{Z} - \{0\}$, then there exist $x, y \in \mathbb{Z}$ such that

$$(a, b) = ax + by$$

that is, *the g.c.d. of a and b is a $\mathbb{Z}$-linear combination of a and b*. This follows by recursively writing the element r_n in the Euclidean Algorithm in terms of the

previous remainders (namely, use equation (n) above to solve for $r_n = r_{n-2} - q_n r_{n-1}$ in terms of the remainders r_{n-1} and r_{n-2}, then use equation $(n-1)$ to write r_n in terms of the remainders r_{n-2} and r_{n-3}, etc., eventually writing r_n in terms of a and b).

Example

Suppose $a = 57970$ and $b = 10353$, whose greatest common divisor we computed above to be 17. From the fifth equation (the next to last equation) in the Euclidean Algorithm applied to these two integers we solve for their greatest common divisor: $17 = 2057 - (60)34$. The fourth equation then shows that $34 = 4148 - (2)2057$, so substituting this expression for the previous remainder 34 gives the equation $17 = 2057 - (60)[4148 - (2)2057]$, i.e. $17 = (121)2057 - (60)4148$. Solving the third equation for 2057 and substituting gives $17 = (121)[6205 - (1)4148] - (60)4148 = (121)6205 - (181)4148$. Using the second equation to solve for 4148 and then the first equation to solve for 6205 we finally obtain

$$17 = (302)57970 - (1691)10353$$

as can easily be checked directly. Hence the equation $ax + by = (a, b)$ for the greatest common divisor of a and b in this example has the solution $x = 302$ and $y = -1691$. Note that it is relatively unlikely that this relation would have been found simply by guessing.

The integers x and y in (7) above are not unique. In the example with $a = 57970$ and $b = 10353$ we determined one solution to be $x = 302$ and $y = -1691$, for instance, and it is relatively simple to check that $x = -307$ and $y = 1719$ also satisfy $57970x + 10353y = 17$. The general solution for x and y is known (cf. the exercises below and in Chapter 8).

(8) An element p of $\mathbb{Z}^+$ is called a *prime* if $p > 1$ and the only positive divisors of p are 1 and p (initially, the word prime will refer only to positive integers). An integer $n > 1$ which is not prime is called *composite*. For example, 2,3,5,7,11,13,17,19, ... are primes and 4,6,8,9,10,12,14,15,16,18, ... are composite.

An important property of primes (which in fact can be used to *define* the primes (cf. Exercise 3)) is the following: if p is a prime and $p \mid ab$, for some $a, b \in \mathbb{Z}$, then either $p \mid a$ or $p \mid b$.

(9) The *Fundamental Theorem of Arithmetic* says: if $n \in \mathbb{Z}$, $n > 1$, then n can be factored uniquely into the product of primes, i.e. there are distinct primes $p_1, p_2, \ldots, p_s$ and positive integers $\alpha_1, \alpha_2, \ldots, \alpha_s$ such that

$$n = p_1^{\alpha_1} p_2^{\alpha_2} \ldots p_s^{\alpha_s}.$$

This factorization is unique in the sense that if $q_1, q_2, \ldots, q_t$ are any distinct primes and $\beta_1, \beta_2, \ldots, \beta_t$ positive integers such that

$$n = q_1^{\beta_1} q_2^{\beta_2} \ldots q_t^{\beta_t},$$

then $s = t$ and if we arrange the two sets of primes in increasing order, then $q_i = p_i$ and $\alpha_i = \beta_i$, $1 \leq i \leq s$. For example, $n = 1852423848 = 2^3 3^2 11^2 19^3 31$ and this decomposition into the product of primes is unique.

Suppose the positive integers a and b are expressed as products of prime powers:

$$a = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_s^{\alpha_s}, \quad b = p_1^{\beta_1} p_2^{\beta_2} \cdots p_s^{\beta_s}$$

where $p_1, p_2, \ldots, p_s$ are distinct and the exponents are ≥ 0 (we allow the exponents to be 0 here so that the products are taken over the same set of primes – the exponent will be 0 if that prime is not actually a divisor). Then the greatest common divisor of a and b is

$$(a, b) = p_1^{\min(\alpha_1, \beta_1)} p_2^{\min(\alpha_2, \beta_2)} \cdots p_s^{\min(\alpha_s, \beta_s)}$$

(and the least common multiple is obtained by instead taking the maximum of the α_i and β_i instead of the minimum).

Example

In the example above, $a = 57970$ and $b = 10353$ can be factored as $a = 2 \cdot 5 \cdot 11 \cdot 17 \cdot 31$ and $b = 3 \cdot 7 \cdot 17 \cdot 29$, from which we can immediately conclude that their greatest common divisor is 17. Note, however, that for large integers it is extremely difficult to determine their prime factorizations (several common codes in current use are based on this difficulty, in fact), so that this is not an effective method to determine greatest common divisors in general. The Euclidean Algorithm will produce greatest common divisors quite rapidly without the need for the prime factorization of a and b.

(10) The *Euler φ–function* is defined as follows: for $n \in \mathbb{Z}^+$ let $\varphi(n)$ be the number of positive integers $a \leq n$ with a relatively prime to n, i.e. $(a, n) = 1$. For example, $\varphi(12) = 4$ since 1, 5, 7 and 11 are the only positive integers less than or equal to 12 which have no factors in common with 12. Similarly, $\varphi(1) = 1$, $\varphi(2) = 1$, $\varphi(3) = 2$, $\varphi(4) = 2$, $\varphi(5) = 4$, $\varphi(6) = 2$, etc. For primes p, $\varphi(p) = p - 1$, and, more generally, for all $a \geq 1$ we have the formula

$$\varphi(p^a) = p^a - p^{a-1} = p^{a-1}(p - 1).$$

The function φ is *multiplicative* in the sense that

$$\varphi(ab) = \varphi(a)\varphi(b) \qquad \text{if } (a, b) = 1$$

(note that it is important here that a and b be relatively prime). Together with the formula above this gives a general formula for the values of φ : if $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_s^{\alpha_s}$, then

$$\begin{aligned} \varphi(n) &= \varphi(p_1^{\alpha_1})\varphi(p_2^{\alpha_2}) \cdots \varphi(p_s^{\alpha_s}) \\ &= p_1^{\alpha_1 - 1}(p_1 - 1) p_2^{\alpha_2 - 1}(p_2 - 1) \cdots p_s^{\alpha_s - 1}(p_s - 1). \end{aligned}$$

For example, $\varphi(12) = \varphi(2^2)\varphi(3) = 2^1(2 - 1)3^0(3 - 1) = 4$. The reader should note that we shall use the letter φ for many different functions throughout the text so when we want this letter to denote Euler's function we shall be careful to indicate this explicitly.

EXERCISES

1. For each of the following pairs of integers a and b, determine their greatest common divisor, their least common multiple, and write their greatest common divisor in the form $ax + by$ for some integers x and y.

(a) $a = 20, b = 13$.
(b) $a = 69, b = 372$.
(c) $a = 792, b = 275$.
(d) $a = 11391, b = 5673$.
(e) $a = 1761, b = 1567$.
(f) $a = 507885, b = 60808$.

2. Prove that if the integer k divides the integers a and b then k divides $as + bt$ for every pair of integers s and t.

3. Prove that if n is composite then there are integers a and b such that n divides ab but n does not divide either a or b.

4. Let a, b and N be fixed integers with a and b nonzero and let $d = (a, b)$ be the greatest common divisor of a and b. Suppose x_0 and y_0 are particular solutions to $ax + by = N$ (i.e. $ax_0 + by_0 = N$). Prove for any integer t that the integers

$$x = x_0 + \frac{b}{d}t \quad \text{and} \quad y = y_0 - \frac{a}{d}t$$

are also solutions to $ax + by = N$ (this is in fact the general solution).

5. Determine the value $\varphi(n)$ for each integer $n \leq 30$ where φ denotes the Euler φ-function.

6. Prove the Well Ordering Property of $\mathbb{Z}$ by induction and prove the minimal element is unique.

7. If p is a prime prove that there do not exist nonzero integers a and b such that $a^2 = pb^2$ (i.e. $\sqrt{p}$ is not a rational number).

8. Let p be a prime, $n \in \mathbb{Z}^+$. Find a formula for the largest power of p which divides $n! = n(n-1)(n-2)\ldots 2 \cdot 1$ (it involves the greatest integer function).

9. Write a computer program to determine the greatest common divisor (a, b) of two integers a and b and to express (a, b) in the form $ax + by$ for some integers x and y.

10. Prove for any given positive integer N there exist only finitely many integers n with $\varphi(n) = N$ where φ denotes Euler's φ-function. Conclude in particular that $\varphi(n)$ tends to infinity as n tends to infinity.

11. Prove that if d divides n then $\varphi(d)$ divides $\varphi(n)$ where φ denotes Euler's φ-function.

0.3 $\mathbb{Z}/n\mathbb{Z}$: THE INTEGERS MODULO *n*

Let n be a fixed positive integer. Define a relation on $\mathbb{Z}$ by

$$a \sim b \text{ if and only if } n \mid (b - a).$$

Clearly $a \sim a$, and $a \sim b$ implies $b \sim a$ for any integers a and b, so this relation is trivially reflexive and symmetric. If $a \sim b$ and $b \sim c$ then n divides $a - b$ and n divides $b - c$ so n also divides the sum of these two integers, i.e. n divides $(a-b)+(b-c) = a-c$, so $a \sim c$ and the relation is transitive. Hence this is an equivalence relation. Write $a \equiv b \pmod{n}$ (read: a is *congruent* to b mod n) if $a \sim b$. For any $k \in \mathbb{Z}$ we shall denote the equivalence class of a by $\bar{a}$ - this is called the *congruence class* or *residue class* of a mod n and consists of the integers which differ from a by an integral multiple of n, i.e.

$$\begin{aligned} \bar{a} &= \{a + kn \mid k \in \mathbb{Z}\} \\ &= \{a, a \pm n, a \pm 2n, a \pm 3n, \ldots\}. \end{aligned}$$

There are precisely n distinct equivalence classes mod n, namely

$$\bar{0}, \bar{1}, \bar{2}, \ldots, \overline{n-1}$$

determined by the possible remainders after division by n and these residue classes partition the integers $\mathbb{Z}$. The set of equivalence classes under this equivalence relation will be denoted by $\mathbb{Z}/n\mathbb{Z}$ and called the *integers modulo* n (or the *integers mod* n). The motivation for this notation will become clearer when we discuss quotient groups and quotient rings. Note that for different n's the equivalence relation and equivalence classes are different so we shall always be careful to fix n first before using the bar notation. The process of finding the equivalence class mod n of some integer a is often referred to as *reducing* a *mod* n. This terminology also frequently refers to finding the smallest nonnegative integer congruent to a mod n (the *least residue* of a mod n).

We can define an addition and a multiplication for the elements of $\mathbb{Z}/n\mathbb{Z}$, defining *modular arithmetic* as follows: for $\bar{a}, \bar{b} \in \mathbb{Z}/n\mathbb{Z}$, define their sum and product by

$$\bar{a} + \bar{b} = \overline{a+b} \qquad \text{and} \qquad \bar{a} \cdot \bar{b} = \overline{ab}.$$

What this means is the following: given any two elements $\bar{a}$ and $\bar{b}$ in $\mathbb{Z}/n\mathbb{Z}$, to compute their sum (respectively, their product) take *any representative* integer a in the *class* $\bar{a}$ and *any representative* integer b in the *class* $\bar{b}$ and add (respectively, multiply) the integers a and b as usual in $\mathbb{Z}$ and then take the equivalence class containing the result. The following Theorem 3 asserts that this is well defined, i.e. does not depend on the choice of representatives taken for the elements $\bar{a}$ and $\bar{b}$ of $\mathbb{Z}/n\mathbb{Z}$.

Example

Suppose $n = 12$ and consider $\mathbb{Z}/12\mathbb{Z}$, which consists of the twelve residue classes

$$\bar{0}, \bar{1}, \bar{2}, \ldots, \overline{11}$$

determined by the twelve possible remainders of an integer after division by 12. The elements in the residue class $\bar{5}$, for example, are the integers which leave a remainder of 5 when divided by 12 (the integers *congruent to* 5 mod 12). Any integer congruent to 5 mod 12 (such as 5, 17, 29, ... or -7, -19, ...) will serve as a representative for the residue class $\bar{5}$. Note that $\mathbb{Z}/12\mathbb{Z}$ consists of the twelve *elements* above (and each of these elements of $\mathbb{Z}/12\mathbb{Z}$ consists of an infinite number of usual integers).

Suppose now that $\bar{a} = \bar{5}$ and $\bar{b} = \bar{8}$. The most obvious representative for $\bar{a}$ is the integer 5 and similarly 8 is the most obvious representative for $\bar{b}$. Using *these* representatives for the residue classes we obtain $\bar{5} + \bar{8} = \overline{13} = \bar{1}$ since 13 and 1 lie in the same class modulo $n = 12$. Had we instead taken the representative 17, say, for $\bar{a}$ (note that 5 and 17 do lie in the same residue class modulo 12) and the representative -28, say, for $\bar{b}$, we would obtain $5 + \bar{8} = \overline{(17 - 28)} = \overline{-11} = \bar{1}$ and as we mentioned the result does not depend on the choice of representatives chosen. The product of these two classes is $\bar{a} \cdot \bar{b} = \overline{5 \cdot 8} = \overline{40} = \bar{4}$, also independent of the representatives chosen.

Theorem 3. The operations of addition and multiplication on $\mathbb{Z}/n\mathbb{Z}$ defined above are both well defined, that is, they do not depend on the choices of representatives for the classes involved. More precisely, if $a_1, a_2 \in \mathbb{Z}$ and $b_1, b_2 \in \mathbb{Z}$ with