Year	Subject Title	Sem.	Sub Code
2018 -19 Onwards	ABSTRACT ALGEBRA	IV	

OBJECTIVES:

To introduce concepts and algebraic structures of Groups and Rings with additional operations and axioms.

UNIT: I

SET THEORY: Definition - Mappings – The Integers.

(Chapter 1 – Sections: 1.1 to 1.3)

1.1 Set Theory

We shall not attempt a formal definition of a set nor shall we try to lay the groundwork for an axiomatic theory of sets. Instead we shall take the operational and intuitive approach that a set is some given collection of objects. In most of our applications we shall be dealing with rather specific things, and the nebulous notion of a set, in these, will emerge as something quite recognizable. For those whose tastes run more to the formal and abstract side, we can consider a set as a primitive notion which one does not define.

A few remarks about notation and terminology. Given a set S we shall use the notation throughout $a \in S$ *to read "a is an element of S."* In the same vein, $a \notin S$ will read "a is *not* an element of S." The set A will be said to be a *subset* of the set S if every element in A is an element of S, that is, if $a \in A$ implies $a \in S$. We shall write this as $A \subset S$ (or, sometimes, as $S \supset A$), which may be read "A is contained in S" (or, S contains A). This notation is not meant to preclude the possibility that $A = S$. By the way, what is meant by the equality of two sets? For us this will always mean that they contain the same elements, that is, every element which is in one is in the other, and vice versa. In terms of the symbol for the containing relation, the two sets A and B are equal, written $A = B$, if both $A \subset B$ and $B \subset A$. The standard device for proving the equality of two sets, something we shall be required to do often, is to demonstrate that the two opposite containing relations hold for them. A subset A of S will be called a *proper* subset of S if $A \subset S$ but $A \neq S$ (A is not equal to S).

The *null set* is the set having no elements; it is a subset of every set. We shall often describe that a set S is the null set by saying it is *empty*.

One final, purely notational remark: Given a set S we shall constantly use the notation $A = \{a \in S \mid P(a)\}$ to read "A is the set of all elements in S for which the property P holds." For instance, if S is the set of integers

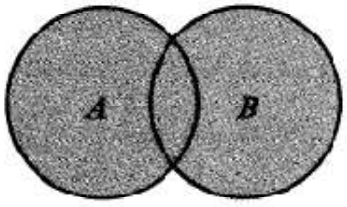
and if A is the subset of positive integers, then we can describe A as $A = \{a \in S \mid a > 0\}$. Another example of this: If S is the set consisting of the objects (1), (2), . . . , (10), then the subset A consisting of (1), (4), (7), (10) could be described by $A = \{(i) \in S \mid i = 3n + 1,\ n = 0, 1, 2, 3\}$.

Given two sets we can combine them to form new sets. There is nothing sacred or particular about this number two; we can carry out the same procedure for any number of sets, finite or infinite, and in fact we shall. We do so for two first because it illustrates the general construction but is not obscured by the additional notational difficulties.

DEFINITION The *union* of the two sets A and B, written as $A \cup B$, is the set $\{x \mid x \in A \text{ or } x \in B\}$.

A word about the use of "or." In ordinary English when we say that something is one or the other we imply that it is not both. The mathematical "or" is quite different, at least when we are speaking about set theory. *For when we say that x is in A or x is in B we mean x is in at least one of A or B, and may be in both.*

Let us consider a few examples of the union of two sets. For any set A, $A \cup A = A$; in fact, whenever B is a subset of A, $A \cup B = A$. If A is the set $\{x_1, x_2, x_3\}$ (i.e., the set whose elements are x_1, x_2, x_3) and if B is the set $\{y_1, y_2, x_1\}$, then $A \cup B = \{x_1, x_2, x_3, y_1, y_2\}$. If A is the set of all blonde-haired people and if B is the set of all people who smoke, then $A \cup B$ consists of all the people who either have blonde hair or smoke or both. Pictorially we can illustrate the union of the two sets A and B by

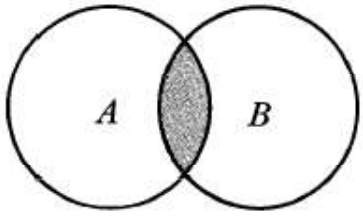


Here, A is the circle on the left, B that on the right, and $A \cup B$ is the shaded part.

DEFINITION The *intersection* of the two sets A and B, written as $A \cap B$, is the set $\{x \mid x \in A \text{ and } x \in B\}$.

The intersection of A and B is thus the set of all elements which are both in A and in B. In analogy with the examples used to illustrate the union of two sets, let us see what the intersections are in those very examples. For

any set A, $A \cap A = A$; in fact, if B is any subset of A, then $A \cap B = B$. If A is the set $\{x_1, x_2, x_3\}$ and B the set $\{y_1, y_2, x_1\}$, then $A \cap B = \{x_1\}$ (we are supposing no y is an x). If A is the set of all blonde-haired people and if B is the set of all people that smoke, then $A \cap B$ is the set of all blonde-haired people who smoke. Pictorially we can illustrate the intersection of the two sets A and B by



Here A is the circle on the left, B that on the right, while their intersection is the shaded part.

Two sets are said to be *disjoint* if their intersection is empty, that is, is the null set. For instance, if A is the set of positive integers and B the set of negative integers, then A and B are disjoint. Note however that if C is the set of nonnegative integers and if D is the set of nonpositive integers, then they are not disjoint, for their intersection consists of the integer 0, and so is not empty.

Before we generalize union and intersection from two sets to an arbitrary number of them, we should like to prove a little proposition interrelating union and intersection. This is the first of a whole host of such results that can be proved; some of these can be found in the problems at the end of this section.

PROPOSITION *For any three sets, A, B, C we have*

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

Proof. The proof will consist of showing, to begin with, the relation $(A \cap B) \cup (A \cap C) \subset A \cap (B \cup C)$ and then the converse relation $A \cap (B \cup C) \subset (A \cap B) \cup (A \cap C)$.

We first dispose of $(A \cap B) \cup (A \cap C) \subset A \cap (B \cup C)$. Because $B \subset B \cup C$, it is immediate that $A \cap B \subset A \cap (B \cup C)$. In a similar manner, $A \cap C \subset A \cap (B \cup C)$. Therefore

$$(A \cap B) \cup (A \cap C) \subset (A \cap (B \cup C)) \cup (A \cap (B \cup C)) = A \cap (B \cup C).$$

Now for the other direction. Given an element $x \in A \cap (B \cup C)$, first of all it must be an element of A. Secondly, as an element in $B \cup C$ it is either in B or in C. Suppose the former; then as an element both of A and of B, x must be in $A \cap B$. The second possibility, namely, $x \in C$, leads us

to $x \in A \cap C$. Thus in either eventuality $x \in (A \cap B) \cup (A \cap C)$, whence $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

The two opposite containing relations combine to give us the equality asserted in the proposition.

We continue the discussion of sets to extend the notion of union and of intersection to arbitrary collections of sets.

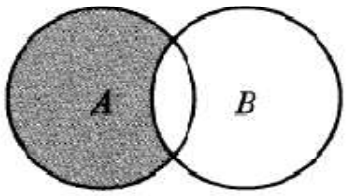
Given a set T we say that T serves as an *index set* for the family $\mathscr{F} = \{A_\alpha\}$ of sets if for every $\alpha \in T$ there exists a set of A_α in the family $\mathscr{F}$. The index set T can be any set, finite or infinite. Very often we use the set of non-negative integers as an index set, but, we repeat, T can be any (nonempty) set.

By the *union* of the sets A_α, where α is in T, we mean the set $\{x \mid x \in A_\alpha$ for at least one α in $T\}$. We shall denote it by $\bigcup_{\alpha \in T} A_\alpha$. By the *intersection* of the sets A_α, where α is in T, we mean the set $\{x \mid x \in A_\alpha$ for every $\alpha \in T\}$; we shall denote it by $\bigcap_{\alpha \in T} A_\alpha$. The sets A_α are *mutually disjoint* if for $\alpha \neq \beta$, $A_\alpha \cap A_\beta$ is the null set.

For instance, if S is the set of real numbers, and if T is the set of rational numbers, let, for $\alpha \in T$, $A_\alpha = \{x \in S \mid x \geq \alpha\}$. It is an easy exercise to see that $\bigcup_{\alpha \in T} A_\alpha = S$ whereas $\bigcap_{\alpha \in T} A_\alpha$ is the null set. The sets A_α are not mutually disjoint.

DEFINITION Given the two sets A, B then the *difference set*, $A - B$, is the set $\{x \in A \mid x \notin B\}$.

Returning to our little pictures, if A is the circle on the left, B that on the right, then $A - B$ is the shaded area.



Note that for any set B, the set A satisfies $A = (A \cap B) \cup (A - B)$. (Prove!) Note further that $B \cap (A - B)$ is the null set. A particular case of interest of the difference of two sets is when one of these is a subset of the other. In that case, when B is a subset of A, we call $A - B$ the *complement of B in A*.

We still want one more construct of two given sets A and B, their *Cartesian product* $A \times B$. This set $A \times B$ is defined as the set of all ordered pairs (a, b) where $a \in A$ and $b \in B$ and where we declare the pair (a_1, b_1) to be equal to (a_2, b_2) if and only if $a_1 = a_2$ and $b_1 = b_2$.

A few remarks about the Cartesian product. Given the two sets A and B we could construct the sets $A \times B$ and $B \times A$ from them. As sets these are distinct, yet we feel that they must be closely related. Given three sets A, B, C we can construct many Cartesian products from them: for instance, the set $A \times D$, where $D = B \times C$; the set $E \times C$, where $E = A \times B$; and also the set of all ordered triples (a, b, c) where $a \in A$, $b \in B$, and $c \in C$. These give us three distinct sets, yet here, also, we feel that these sets must be closely related. Of course, we can continue this process with more and more sets. To see the exact relation between them we shall have to wait until the next section, where we discuss one-to-one correspondences.

Given any index set T we could define the Cartesian product of the sets A_α as α varies over T; since we shall not need so general a product, we do not bother to define it.

Finally, we can consider the Cartesian product of a set A with itself, $A \times A$. Note that if the set A is a finite set having n elements, then the set $A \times A$ is also a finite set, but has n^2 elements. The set of elements (a, a) in $A \times A$ is called the *diagonal* of $A \times A$.

A subset R of $A \times A$ is said to define an *equivalence relation* on A if

1. $(a, a) \in R$ for all $a \in A$.
2. $(a, b) \in R$ implies $(b, a) \in R$.
3. $(a, b) \in R$ and $(b, c) \in R$ imply that $(a, c) \in R$.

Instead of speaking about subsets of $A \times A$ we can speak about a binary relation (one between two elements of A) on A itself, defining b to be related to a if $(a, b) \in R$. The properties 1, 2, 3 of the subset R immediately translate into the properties 1, 2, 3 of the definition below.

DEFINITION The binary relation $\sim$ on A is said to be an *equivalence relation* on A if for all a, b, c in A

1. $a \sim a$.
2. $a \sim b$ implies $b \sim a$.
3. $a \sim b$ and $b \sim c$ imply $a \sim c$.

The first of these properties is called *reflexivity*, the second, *symmetry*, and the third, *transitivity*.

The concept of an equivalence relation is an extremely important one and plays a central role in all of mathematics. We illustrate it with a few examples.

Example 1.1.1 Let S be any set and define $a \sim b$, for $a, b \in S$, if and only if $a = b$. This clearly defines an equivalence relation on S. In fact, an equivalence relation is a generalization of equality, measuring equality up to some property.

Example 1.1.2 Let S be the set of all integers. Given $a, b \in S$, define $a \sim b$ if $a - b$ is an even integer. We verify that this defines an equivalence relation of S.

1. Since $0 = a - a$ is even, $a \sim a$.
2. If $a \sim b$, that is, if $a - b$ is even, then $b - a = -(a - b)$ is also even, whence $b \sim a$.
3. If $a \sim b$ and $b \sim c$, then both $a - b$ and $b - c$ are even, whence $a - c = (a - b) + (b - c)$ is also even, proving that $a \sim c$.

Example 1.1.3 Let S be the set of all integers and let $n > 1$ be a fixed integer. Define for $a, b \in S$, $a \sim b$ if $a - b$ is a multiple of n. We leave it as an exercise to prove that this defines an equivalence relation on S.

Example 1.1.4 Let S be the set of all triangles in the plane. Two triangles are defined to be equivalent if they are similar (i.e., have corresponding angles equal). This defines an equivalence relation on S.

Example 1.1.5 Let S be the set of points in the plane. Two points a and b are defined to be equivalent if they are equidistant from the origin. A simple check verifies that this defines an equivalence relation on S.

There are many more equivalence relations; we shall encounter a few as we proceed in the book.

DEFINITION If A is a set and if $\sim$ is an equivalence relation on A, then the *equivalence class* of $a \in A$ is the set $\{x \in A \mid a \sim x\}$. We write it as $\mathrm{cl}(a)$.

In the examples just discussed, what are the equivalence classes? In Example 1.1.1, the equivalence class of a consists merely of a itself. In Example 1.1.2 the equivalence class of a consists of all the integers of the form $a + 2m$, where $m = 0, \pm 1, \pm 2, \ldots$; in this example there are only two distinct equivalence classes, namely, $\mathrm{cl}(0)$ and $\mathrm{cl}(1)$. In Example 1.1.3, the equivalence class of a consists of all integers of the form $a + kn$ where $k = 0, \pm 1, \pm 2, \ldots$; here there are n distinct equivalence classes, namely $\mathrm{cl}(0), \mathrm{cl}(1), \ldots, \mathrm{cl}(n - 1)$. In Example 1.1.5, the equivalence class of a consists of all the points in the plane which lie on the circle which has its center at the origin and passes through a.

Although we have made quite a few definitions, introduced some concepts, and have even established a simple little proposition, one could say in all fairness that up to this point we have not proved any result of real substance. We are now about to prove the first genuine result in the book. The proof of this theorem is not very difficult—actually it is quite easy—but nonetheless the result it embodies will be of great use to us.

THEOREM 1.1.1 *The distinct equivalence classes of an equivalence relation on A provide us with a decomposition of A as a union of mutually disjoint subsets. Conversely, given a decomposition of A as a union of mutually disjoint, nonempty subsets, we can define an equivalence relation on A for which these subsets are the distinct equivalence classes.*

Proof. Let the equivalence relation on A be denoted by $\sim$.

We first note that since for any $a \in A$, $a \sim a$, a must be in $\mathrm{cl}(a)$, whence the union of the $\mathrm{cl}(a)$'s is all of A. We now assert that given two equivalence classes they are either equal or disjoint. For, suppose that $\mathrm{cl}(a)$ and $\mathrm{cl}(b)$ are not disjoint; then there is an element $x \in \mathrm{cl}(a) \cap \mathrm{cl}(b)$. Since $x \in \mathrm{cl}(a)$, $a \sim x$; since $x \in \mathrm{cl}(b)$, $b \sim x$, whence by the symmetry of the relation, $x \sim b$. However, $a \sim x$ and $x \sim b$ by the transitivity of the relation forces $a \sim b$. Suppose, now that $y \in \mathrm{cl}(b)$; thus $b \sim y$. However, from $a \sim b$ and $b \sim y$, we deduce that $a \sim y$, that is, that $y \in \mathrm{cl}(a)$. Therefore, every element in $\mathrm{cl}(b)$ is in $\mathrm{cl}(a)$, which proves that $\mathrm{cl}(b) \subset \mathrm{cl}(a)$. The argument is clearly symmetric, whence we conclude that $\mathrm{cl}(a) \subset \mathrm{cl}(b)$. The two opposite containing relations imply that $\mathrm{cl}(a) = \mathrm{cl}(b)$.

We have thus shown that the distinct $\mathrm{cl}(a)$'s are mutually disjoint and that their union is A. This proves the first half of the theorem. Now for the other half!

Suppose that $A = \bigcup A_\alpha$ where the A_α are mutually disjoint, nonempty sets (α is in some index set T). How shall we use them to define an equivalence relation? The way is clear; given an element a in A it is in *exactly one* A_α. We define for $a, b \in A$, $a \sim b$ if a and b are in the same A_α. We leave it as an exercise to prove that this is an equivalence relation on A and that the distinct equivalence classes are the A_α's.

Problems

1. (a) If A is a subset of B and B is a subset of C, prove that A is a subset of C.
   (b) If $B \subset A$, prove that $A \cup B = A$, and conversely.
   (c) If $B \subset A$, prove that for any set C both $B \cup C \subset A \cup C$ and $B \cap C \subset A \cap C$.
2. (a) Prove that $A \cap B = B \cap A$ and $A \cup B = B \cup A$.
   (b) Prove that $(A \cap B) \cap C = A \cap (B \cap C)$.
3. Prove that $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.
4. For a subset C of S let C' denote the complement of C in S. For any two subsets A, B of S prove the *De Morgan rules*:
   (a) $(A \cap B)' = A' \cup B'$.
   (b) $(A \cup B)' = A' \cap B'$.
5. For a finite set C let $o(C)$ indicate the number of elements in C. If A and B are finite sets prove $o(A \cup B) = o(A) + o(B) - o(A \cap B)$.

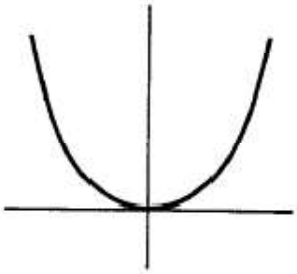
6. If A is a finite set having n elements, prove that A has exactly 2^n distinct subsets.
7. A survey shows that 63% of the American people like cheese whereas 76% like apples. What can you say about the percentage of the American people that like both cheese and apples? (The given statistics are not meant to be accurate.)
8. Given two sets A and B their *symmetric difference* is defined to be $(A - B) \cup (B - A)$. Prove that the symmetric difference of A and B equals $(A \cup B) - (A \cap B)$.
9. Let S be a set and let S^* be the set whose elements are the various subsets of S. In S^* we define an addition and multiplication as follows: If $A, B \in S^*$ (remember, this means that they are subsets of S):
   (1) $A + B = (A - B) \cup (B - A)$.
   (2) $A \cdot B = A \cap B$.
   Prove the following laws that govern these operations:
   (a) $(A + B) + C = A + (B + C)$.
   (b) $A \cdot (B + C) = A \cdot B + A \cdot C$.
   (c) $A \cdot A = A$.
   (d) $A + A =$ null set.
   (e) If $A + B = A + C$ then $B = C$.
   (The system just described is an example of a *Boolean algebra.*)
10. For the given set and relation below determine which define equivalence relations.
    (a) S is the set of all people in the world today, $a \sim b$ if a and b have an ancestor in common.
    (b) S is the set of all people in the world today, $a \sim b$ if a lives within 100 miles of b.
    (c) S is the set of all people in the world today, $a \sim b$ if a and b have the same father.
    (d) S is the set of real numbers, $a \sim b$ if $a = \pm b$.
    (e) S is the set of integers, $a \sim b$ if both $a > b$ and $b > a$.
    (f) S is the set of all straight lines in the plane, $a \sim b$ if a is parallel to b.
11. (a) Property 2 of an equivalence relation states that if $a \sim b$ then $b \sim a$; property 3 states that if $a \sim b$ and $b \sim c$ then $a \sim c$. What is wrong with the following proof that properties 2 and 3 imply property 1? Let $a \sim b$; then $b \sim a$, whence, by property 3 (using $a = c$), $a \sim a$.
    (b) Can you suggest an alternative of property 1 which will insure us that properties 2 and 3 do imply property 1?
12. In Example 1.1.3 of an equivalence relation given in the text, prove that the relation defined is an equivalence relation and that there are exactly n distinct equivalence classes, namely, $\mathrm{cl}(0), \mathrm{cl}(1), \ldots, \mathrm{cl}(n-1)$.
13. Complete the proof of the second half of Theorem 1.1.1.

1.2 Mappings

We are about to introduce the concept of a mapping of one set into another. Without exaggeration this is probably the single most important and universal notion that runs through all of mathematics. It is hardly a new thing to any of us, for we have been considering mappings from the very earliest days of our mathematical training. When we were asked to plot the relation $y = x^2$ we were simply being asked to study the particular mapping which takes every real number onto its square.

Loosely speaking, a mapping from one set, S, into another, T, is a "rule" (whatever that may mean) that associates with each element in S a *unique* element t in T. We shall define a mapping somewhat more formally and precisely but the purpose of the definition is to allow us to think and speak in the above terms. We should think of them as rules or devices or mechanisms that transport us from one set to another.

Let us motivate a little the definition that we will make. The point of view we take is to consider the mapping to be defined by its "graph." We illustrate this with the familiar example $y = x^2$ defined on the real numbers S and taking its values also in S. For this set S, $S \times S$, the set of all pairs (a, b) can be viewed as the plane, the pair (a, b) corresponding to the point whose coordinates are a and b, respectively. In this plane we single out all those points whose coordinates are of the form (x, x^2) and call this set of points the graph of $y = x^2$. We even represent this set pictorially as



To find the "value" of the function or mapping at the point $x = a$, we look at the point in the graph whose first coordinate is a and read off the second coordinate as the value of the function at $x = a$.

This is, no more or less, the approach we take in the general setting to define a mapping from one set into another.

DEFINITION If S and T are nonempty sets, then a *mapping* from S to T is a subset, M, of $S \times T$ such that for *every* $s \in S$ there is a *unique* $t \in T$ such that the ordered pair (s, t) is in M.

This definition serves to make the concept of a mapping precise for us but we shall almost never use it in this form. Instead we do prefer to think of a

mapping as a rule which associates with any element s in S some element t in T, *the rule being, associate (or map) $s \in S$ with $t \in T$ if and only if $(s, t) \in M$.* We shall say that t is the *image* of s under the mapping.

Now for some notation for these things. Let σ be a mapping from S to T; we often denote this by writing $\sigma: S \to T$ or $S \xrightarrow{\sigma} T$. If t is the image of s under σ we shall sometimes write this as $\sigma: s \to t$; more often, we shall represent this fact by $t = s\sigma$. Note that we write the mapping σ on the *right*. There is no overall consistency in this usage; many people would write it as $t = \sigma(s)$. Algebraists often write mappings on the right; other mathematicians write them on the left. In fact, we shall not be absolutely consistent in this ourselves; when we shall want to emphasize the functional nature of σ we may very well write $t = \sigma(s)$.

Examples of Mappings

In all the examples the sets are assumed to be nonempty.

Example 1.2.1 Let S be any set; define $\iota: S \to S$ by $s = s\iota$ for any $s \in S$. This mapping ι is called the *identity mapping* of S.

Example 1.2.2 Let S and T be any sets and let t_0 be an element of T. Define $\tau: S \to T$ by $\tau: s \to t_0$ for every $s \in S$.

Example 1.2.3 Let S be the set of positive rational numbers and let $T = J \times J$ where J is the set of integers. Given a rational number s we can write it as $s = m/n$, where m and n have no common factor. Define $\tau: S \to T$ by $s\tau = (m, n)$.

Example 1.2.4 Let J be the set of integers and $S = \{(m, n) \in J \times J \mid n \neq 0\}$; let T be the set of rational numbers; define $\tau: S \to T$ by $(m, n)\tau = m/n$ for every (m, n) in S.

Example 1.2.5 Let J be the set of integers and $S = J \times J$. Define $\tau: S \to J$ by $(m, n)\tau = m + n$.

Note that in Example 1.2.5 the addition in J itself can be represented in terms of a mapping of $J \times J$ into J. Given an arbitrary set S we call a mapping of $S \times S$ into S a *binary operation* on S. Given such a mapping $\tau: S \times S \to S$ we could use it to define a "product" $*$ in S by declaring $a * b = c$ if $(a, b)\tau = c$.

Example 1.2.6 Let S and T be any sets; define $\tau: S \times T \to S$ by $(a, b)\tau = a$ for any $(a, b) \in S \times T$. This τ is called the *projection* of $S \times T$ on S. We could similarly define the projection of $S \times T$ on T.

Example 1.2.7 Let S be the set consisting of the elements x_1, x_2, x_3. Define $\tau: S \to S$ by $x_1\tau = x_2$, $x_2\tau = x_3$, $x_3\tau = x_1$.

Example 1.2.8 Let S be the set of integers and let T be the set consisting of the elements E and 0. Define $\tau: S \to T$ by declaring $n\tau = E$ if n is even and $n\tau = 0$ if n is odd.

If S is any set, let $\{x_1, \ldots, x_n\}$ be its subset consisting of the elements $x_1, x_2, \ldots, x_n$ of S. In particular, $\{x\}$ is the subset of S whose only element is x. Given S we can use it to construct a new set S^*, the set whose elements are the subsets of S. We call S^* the *set of subsets* of S. Thus for instance, if $S = \{x_1, x_2\}$ then S^* has exactly four elements, namely, $a_1 =$ null set, $a_2 =$ the subset, S, of S, $a_3 = \{x_1\}$, $a_4 = \{x_2\}$. The relation of S to S^*, in general, is a very interesting one; some of its properties are examined in the problems.

Example 1.2.9 Let S be a set, $T = S^*$; define $\tau: S \to T$ by $s\tau =$ complement of $\{s\}$ in $S = S - \{s\}$.

Example 1.2.10 Let S be a set with an equivalence relation, and let T be the set of equivalence classes in S (note that T is a subset of S^*). Define $\tau: S \to T$ by $s\tau = \mathrm{cl}(s)$.

We leave the examples to continue the general discussion. Given a mapping $\tau: S \to T$ we define for $t \in T$, the *inverse image* of t with respect to τ to be the set $\{s \in S \mid t = s\tau\}$. In Example 1.2.8, the inverse image of E is the subset of S consisting of the even integers. It may happen that for some t in T that its inverse image with respect to τ is empty; that is, t is not the image under τ of any element in S. In Example 1.2.3, the element (4, 2) is not the image of any element in S under the τ used; in Example 1.2.9, S, as an element in S^*, is not the image under the τ used of any element in S.

DEFINITION The mapping τ of S into T is said to be *onto* T if given $t \in T$ there exists an element $s \in S$ such that $t = s\tau$.

If we call the subset $S\tau = \{x \in T \mid x = s\tau \text{ for some } s \in S\}$ the *image* of S under τ, then τ is onto if the image of S under τ is all of T. Note that in Examples 1.2.1, 1.2.4–1.2.8, and 1.2.10 the mappings used are all onto.

Another special type of mapping arises often and is important: the one-to-one mapping.

DEFINITION The mapping τ of S into T is said to be a *one-to-one mapping* if whenever $s_1 \neq s_2$, then $s_1\tau \neq s_2\tau$.

In terms of inverse images, the mapping τ is one-to-one if for any $t \in T$ the inverse image of t is either empty or is a set consisting of one element. In the examples discussed, the mappings in Examples 1.2.1, 1.2.3, 1.2.7, and 1.2.9 are all one-to-one.

When should we say that two mappings from S to T are equal? A natural definition for this is that they should have the same effect on every element of S; that is, the image of any element in S under each of these mappings should be the same. In a little more formal manner:

DEFINITION The two mappings σ and τ of S into T are said to be *equal* if $s\sigma = s\tau$ for every $s \in S$.

Consider the following situation: We have a mapping σ from S to T and another mapping τ from T to U. Can we compound these mappings to produce a mapping from S to U? The most natural and obvious way of doing this is to send a given element s, in S, in two stages into U, first by applying σ to s and then applying τ to the resulting element $s\sigma$ in T. This is the basis of the

DEFINITION If $\sigma: S \to T$ and $\tau: T \to U$ then the *composition* of σ and τ (also called their *product*) is the mapping $\sigma \circ \tau: S \to U$ defined by means of $s(\sigma \circ \tau) = (s\sigma)\tau$ for every $s \in S$.

Note that the order of events reads from left to right; $\sigma \circ \tau$ reads: first perform σ and then follow it up with τ. Here, too, the left-right business is not a uniform one. Mathematicians who write their mappings on the left would read $\sigma \circ \tau$ to mean first perform τ and then σ. Accordingly, in reading a given book in mathematics one must make absolutely sure as to what convention is being followed in writing the product of two mappings. We reiterate, *for us $\sigma \circ \tau$ will always mean: first apply σ and then τ.*

We illustrate the composition of σ and τ with a few examples.

Example 1.2.11 Let $S = \{x_1, x_2, x_3\}$ and let $T = S$. Let $\sigma: S \to S$ be defined by

$$x_1\sigma = x_2,$$
$$x_2\sigma = x_3,$$
$$x_3\sigma = x_1;$$

and $\tau: S \to S$ by

$$x_1\tau = x_1,$$
$$x_2\tau = x_3,$$
$$x_3\tau = x_2.$$

Thus

$$x_1(\sigma \circ \tau) = (x_1\sigma)\tau = x_2\tau = x_3,$$
$$x_2(\sigma \circ \tau) = (x_2\sigma)\tau = x_3\tau = x_2,$$
$$x_3(\sigma \circ \tau) = (x_3\sigma)\tau = x_1\tau = x_1.$$

At the same time we can compute $\tau \circ \sigma$, because in this case it also makes sense. Now

$$x_1(\tau \circ \sigma) = (x_1\tau)\sigma = (x_1\sigma) = x_2,$$
$$x_2(\tau \circ \sigma) = (x_2\tau)\sigma = x_3\sigma = x_1,$$
$$x_3(\tau \circ \sigma) = (x_3\tau)\sigma = x_2\sigma = x_3.$$

Note that $x_2 = x_1(\tau \circ \sigma)$, whereas $x_3 = x_1(\sigma \circ \tau)$ whence $\sigma \circ \tau \neq \tau \circ \sigma$.

Example 1.2.12 Let S be the set of integers, T the set $S \times S$, and suppose $\sigma: S \to T$ is defined by $m\sigma = (m - 1, 1)$. Let $U = S$ and suppose that $\tau: T \to U(= S)$ is defined by $(m, n)\tau = m + n$. Thus $\sigma \circ \tau: S \to S$ whereas $\tau \circ \sigma: T \to T$; even to speak about the equality of $\sigma \circ \tau$ and $\tau \circ \sigma$ would make no sense since they do not act on the same space. We now compute $\sigma \circ \tau$ as a mapping of S into itself and then $\tau \circ \sigma$ as one on T into itself.

Given $m \in S$, $m\sigma = (m - 1, 1)$ whence $m(\sigma \circ \tau) = (m\sigma)\tau = (m - 1, 1)\tau = (m - 1) + 1 = m$. Thus $\sigma \circ \tau$ is the identity mapping of S into itself. What about $\tau \circ \sigma$? Given $(m, n) \in T$, $(m, n)\tau = m + n$, whereby $(m, n)(\tau \circ \sigma) = ((m, n)\tau)\sigma = (m + n)\sigma = (m + n - 1, 1)$. Note that $\tau \circ \sigma$ is *not* the identity map of T into itself; it is not even an onto mapping of T.

Example 1.2.13 Let S be the set of real numbers, T the set of integers, and $U = \{E, 0\}$. Define $\sigma: S \to T$ by $s\sigma =$ largest integer less than or equal to s, and $\tau: T \to U$ defined by $n\tau = E$ if n is even, $n\tau = 0$ if n is odd. Note that in this case $\tau \circ \sigma$ cannot be defined. We compute $\sigma \circ \tau$ for two real numbers $s = \frac{8}{3}$ and $s = \pi$. Now since $\frac{8}{3} = 2 + \frac{2}{3}$, $(\frac{8}{3})\sigma = 2$, whence $(\frac{8}{3})(\sigma \circ \tau) = (\frac{8}{3}\sigma)\tau = (2)\tau = E$; $(\pi)\sigma = 3$, whence $\pi(\sigma \circ \tau) = (\pi\sigma)\tau = (3)\tau = 0$.

For mappings of sets, provided the requisite products make sense, a general *associative law* holds. This is the content of

LEMMA 1.2.1 (ASSOCIATIVE LAW) *If* $\sigma: S \to T$, $\tau: T \to U$, *and* $\mu: U \to V$, *then* $(\sigma \circ \tau) \circ \mu = \sigma \circ (\tau \circ \mu)$.

Proof. Note first that $\sigma \circ \tau$ makes sense and takes S into U, thus $(\sigma \circ \tau) \circ \mu$ also makes sense and takes S into V. Similarly $\sigma \circ (\tau \circ \mu)$ is meaningful and takes S into V. Thus we can speak about the equality, or lack of equality, of $(\sigma \circ \tau) \circ \mu$ and $\sigma \circ (\tau \circ \mu)$.

To prove the asserted equality we merely must show that for any $s \in S$, $s((\sigma \circ \tau) \circ \mu) = s(\sigma \circ (\tau \circ \mu))$. Now by the very definition of the composition

of maps, $s((\sigma \circ \tau) \circ \mu) = (s(\sigma \circ \tau))\mu = ((s\sigma)\tau)\mu$ whereas $s(\sigma \circ (\tau \circ \mu)) = (s\sigma)(\tau \circ \mu) = ((s\sigma)\tau)\mu$. Thus, the elements $s((\sigma \circ \tau) \circ \mu)$ and $s(\sigma \circ (\tau \circ \mu))$ are indeed equal. This proves the lemma.

We should like to show that if two mappings σ and τ are properly conditioned the very same conditions carry over to $\sigma \circ \tau$.

LEMMA 1.2.2 *Let $\sigma: S \to T$ and $\tau: T \to U$; then*

1. *$\sigma \circ \tau$ is onto if each of σ and τ is onto.*
2. *$\sigma \circ \tau$ is one-to-one if each of σ and τ is one-to-one.*

Proof. We prove only part 2, leaving the proof of part 1 as an exercise.

Suppose that $s_1, s_2 \in S$ and that $s_1 \neq s_2$. By the one-to-one nature of σ, $s_1\sigma \neq s_2\sigma$. Since τ is one-to-one and $s_1\sigma$ and $s_2\sigma$ are distinct elements of T, $(s_1\sigma)\tau \neq (s_2\sigma)\tau$ whence $s_1(\sigma \circ \tau) = (s_1\sigma)\tau \neq (s_2\sigma)\tau = s_2(\sigma \circ \tau)$, proving that $\sigma \circ \tau$ is indeed one-to-one, and establishing the lemma.

Suppose that σ is a one-to-one mapping of S *onto* T; we call σ a *one-to-one correspondence* between S and T. Given any $t \in T$, by the "onto-ness" of σ there exists an element $s \in S$ such that $t = s\sigma$; by the "one-to-oneness" of σ this s is unique. We define the mapping $\sigma^{-1}: T \to S$ by $s = t\sigma^{-1}$ if and only if $t = s\sigma$. The mapping σ^{-1} is called the *inverse* of σ. Let us compute $\sigma \circ \sigma^{-1}$ which maps S into itself. Given $s \in S$, let $t = s\sigma$, whence by definition $s = t\sigma^{-1}$; thus $s(\sigma \circ \sigma^{-1}) = (s\sigma)\sigma^{-1} = t\sigma^{-1} = s$. We have shown that $\sigma \circ \sigma^{-1}$ is the identity mapping of S onto itself. A similar computation reveals that $\sigma^{-1} \circ \sigma$ is the identity mapping of T onto itself.

Conversely, if $\sigma: S \to T$ is such that there exists a $\mu: T \to S$ with the property that $\sigma \circ \mu$ and $\mu \circ \sigma$ are the identity mappings on S and T, respectively, then we claim that σ is a one-to-one correspondence between S and T. First observe that σ is onto for, given $t \in T$, $t = t(\mu \circ \sigma) = (t\mu)\sigma$ (since $\mu \circ \sigma$ is the identity on T) and so t is the image under σ of the element $t\mu$ in S. Next observe that σ is one-to-one, for if $s_1\sigma = s_2\sigma$, using that $\sigma \circ \mu$ is the identity on S, we have $s_1 = s_1(\sigma \circ \mu) = (s_1\sigma)\mu = (s_2\sigma)\,\mu = s_2(\sigma \circ \mu) = s_2$. We have now proved

LEMMA 1.2.3 *The mapping $\sigma: S \to T$ is a one-to-one correspondence between S and T if and only if there exists a mapping $\mu: T \to S$ such that $\sigma \circ \mu$ and $\mu \circ \sigma$ are the identity mappings on S and T, respectively.*

DEFINITION If S is a nonempty set then $A(S)$ is the *set of all one-to-one mappings* of S onto itself.

Aside from its own intrinsic interest $A(S)$ plays a central and universal type of role in considering the mathematical system known as a group

(Chapter 2). For this reason we state the next theorem concerning its nature. All the constituent parts of the theorem have already been proved in the various lemmas, so we state the theorem without proof.

THEOREM 1.2.1 *If σ, τ, μ are elements of $A(S)$, then*

1. $\sigma \circ \tau$ *is in* $A(S)$.
2. $(\sigma \circ \tau) \circ \mu = \sigma \circ (\tau \circ \mu)$.
3. *There exists an element* ι *(the identity map) in* $A(S)$ *such that* $\sigma \circ \iota = \iota \circ \sigma = \sigma$.
4. *There exists an element* $\sigma^{-1} \in A(S)$ *such that* $\sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = \iota$.

We close the section with a remark about $A(S)$. Suppose that S has more than two elements; let x_1, x_2, x_3 be three distinct elements in S; define the mapping $\sigma: S \to S$ by $x_1\sigma = x_2$, $x_2\sigma = x_3$, $x_3\sigma = x_1$, $s\sigma = s$ for any $s \in S$ different from x_1, x_2, x_3. Define the mapping $\tau: S \to S$ by $x_2\tau = x_3$, $x_3\tau = x_2$, and $s\tau = s$ for any $s \in S$ different from x_2, x_3. Clearly both σ and τ are in $A(S)$. A simple computation shows that $x_1(\sigma \circ \tau) = x_3$ but that $x_1(\tau \circ \sigma) = x_2 \neq x_3$. Thus $\sigma \circ \tau \neq \tau \circ \sigma$. This is

LEMMA 1.2.4 *If S has more that two elements we can find two elements σ, τ in $A(S)$ such that $\sigma \circ \tau \neq \tau \circ \sigma$.*

Problems

1. In the following, where $\sigma: S \to T$, determine whether the σ is onto and/or one-to-one and determine the inverse image of any $t \in T$ under σ.
   (a) S = set of real numbers, T = set of nonnegative real numbers, $s\sigma = s^2$.
   (b) S = set of nonnegative real numbers, T = set of nonnegative real numbers, $s\sigma = s^2$.
   (c) S = set of integers, T = set of integers, $s\sigma = s^2$.
   (d) S = set of integers, T = set of integers, $s\sigma = 2s$.
2. If S and T are nonempty sets, prove that there exists a one-to-one correspondence between $S \times T$ and $T \times S$.
3. If S, T, U are nonempty sets, prove that there exists a one-to-one correspondence between
   (a) $(S \times T) \times U$ and $S \times (T \times U)$.
   (b) Either set in part (a) and the set of ordered triples (s, t, u) where $s \in S$, $t \in T$, $u \in U$.
4. (a) If there is a one-to-one correspondence between S and T, prove that there exists a one-to-one correspondence between T and S.

(b) If there is a one-to-one correspondence between S and T and between T and U, prove that there is a one-to-one correspondence between S and U.

5. If ι is the identity mapping on S, prove that for any $\sigma \in A(S)$, $\sigma \circ \iota = \iota \circ \sigma = \sigma$.

*6. If S is any set, prove that it is *impossible* to find a mapping of S *onto* S^*.

7. If the set S has n elements, prove that $A(S)$ has $n!$ (n factorial) elements.

8. If the set S has a finite number of elements, prove the following:
   (a) If σ maps S onto S, then σ is one-to-one.
   (b) If σ is a one-to-one mapping of S onto itself, then σ is onto.
   (c) Prove, by example, that both part (a) and part (b) are false if S does not have a finite number of elements.

9. Prove that the converse to both parts of Lemma 1.2.2 are false; namely,
   (a) If $\sigma \circ \tau$ is onto, it need not be that both σ and τ are onto.
   (b) If $\sigma \circ \tau$ is one-to-one, it need not be that both σ and τ are one-to-one.

10. Prove that there is a one-to-one correspondence between the set of integers and the set of rational numbers.

11. If $\sigma: S \to T$ and if A is a subset of S, the *restriction of* σ *to* A, σ_A, is defined by $a\sigma_A = a\sigma$ for any $a \in A$. Prove
   (a) σ_A defines a mapping of A into T.
   (b) σ_A is one-to-one if σ is.
   (c) σ_A may very well be one-to-one even if σ is not.

12. If $\sigma: S \to S$ and A is a subset of S such that $A\sigma \subset A$, prove that $(\sigma \circ \sigma)_A = \sigma_A \circ \sigma_A$.

13. A set S is said to be *infinite* if there is a one-to-one correspondence between S and a proper subset of S. Prove
   (a) The set of integers is infinite.
   (b) The set of real numbers is infinite.
   (c) If a set S has a subset A which is infinite, then S must be infinite.
   (*Note:* By the result of Problem 8, a set finite in the usual sense is not infinite.)

†14. If S is infinite and can be brought into one-to-one correspondence with the set of integers, prove that there is one-to-one correspondence between S and $S \times S$.

†15. Given two sets S and T we declare $S < T$ (S is smaller than T) if there is a mapping of T *onto* S but *no* mapping of S *onto* T. Prove that if $S < T$ and $T < U$ then $S < U$.

16. If S and T are finite sets having m and n elements, respectively, prove that if $m < n$ then $S < T$.

1.3 The Integers

We close this chapter with a brief discussion of the set of integers. We shall make no attempt to construct them axiomatically, assuming instead that we already have the set of integers and that we know many of the elementary facts about them. In this number we include the principle of mathematical induction (which will be used freely throughout the book) and the fact that a nonempty set of positive integers always contains a smallest element. As to notation, the familiar symbols: $a > b$, $a \leq b$, $|a|$, etc., will occur with their usual meaning. To avoid repeating that something is an integer, we make the assumption *that all symbols, in this section, written as lowercase Latin letters will be integers.*

Given a and b, with $b \neq 0$, we can divide a by b to get a nonnegative remainder r which is smaller in size than b; that is, we can find m and r such that $a = mb + r$ where $0 \leq r < |b|$. This fact is known as the *Euclidean algorithm* and we assume familiarity with it.

We say that $b \neq 0$ *divides* a if $a = mb$ for some m. We denote that b divides a by $b \mid a$, and that b does not divide a by $b \nmid a$. Note that if $a \mid 1$ then $a = \pm 1$, that when both $a \mid b$ and $b \mid a$, then $a = \pm b$, and that any b divides 0. If $b \mid a$, we call b a *divisor* of a. Note that if b is a divisor of g and of h, then it is a divisor of $mg + nh$ for arbitrary integers m and n. We leave the verification of these remarks as exercises.

DEFINITION The positive integer c is said to be the *greatest common divisor* of a and b if

1. c is a divisor of a and of b.
2. Any divisor of a and b is a divisor of c.

We shall use the notation (a, b) for the greatest common divisor of a and b. Since we insist that the greatest common divisor be positive, $(a, b) = (a, -b) = (-a, b) = (-a, -b)$. For instance, $(60, 24) = (60, -24) = 12$. Another comment: The mere fact that we have defined what is to be meant by the greatest common divisor does not guarantee that it exists. This will have to be proved. However, we can say that if it exists then it is unique, for, if we had c_1 and c_2 satisfying both conditions of the definition above, then $c_1 \mid c_2$ and $c_2 \mid c_1$, whence we would have $c_1 = \pm c_2$; the insistence on positivity would then force $c_1 = c_2$. Our first business at hand then is to dispose of the existence of (a, b). In doing so, in the next lemma, we actually prove a little more, namely that (a, b) must have a particular form.

LEMMA 1.3.1 *If a and b are integers, not both 0, then (a, b) exists; moreover, we can find integers m_0 and n_0 such that $(a, b) = m_0 a + n_0 b$.*

Proof. Let $\mathcal{M}$ be the set of all integers of the form $ma + nb$, where m and n range freely over the set of integers. Since one of a or b is not 0, there are nonzero integers in $\mathcal{M}$. Because $x = ma + nb$ is in $\mathcal{M}$, $-x = (-m)a + (-n)b$ is also in $\mathcal{M}$; therefore, $\mathcal{M}$ always has in it some positive integers. But then there is a smallest positive integer, c, in $\mathcal{M}$; being in $\mathcal{M}$, c has the form $c = m_0a + n_0b$. We claim that $c = (a, b)$.

Note first that if $d \mid a$ and $d \mid b$, the $d \mid (m_0a + n_0b)$, whence $d \mid c$. We now must show that $c \mid a$ and $c \mid b$. Given any element $x = ma + nb$ in $\mathcal{M}$, then by the Euclidean algorithm, $x = tc + r$ where $0 \leq r < c$. Writing this out explicitly, $ma + nb = t(m_0a + n_0b) + r$, whence $r = (m - tm_0)a + (n - tn_0)b$ and so must be in $\mathcal{M}$. Since $0 \leq r$ and $r < c$, by the choice of c, $r = 0$. Thus $x = tc$; we have proved that $c \mid x$ for any $x \in \mathcal{M}$. But $a = 1a + 0b \in \mathcal{M}$ and $b = 0a + 1b \in \mathcal{M}$, whence $c \mid a$ and $c \mid b$.

We have shown that c satisfies the requisite properties to be (a, b) and so we have proved the lemma.

DEFINITION The integers a and b are *relatively prime* if $(a, b) = 1$.

As an immediate consequence of Lemma 1.3.1, we have the

COROLLARY *If a and b are relatively prime, we can find integers m and n such that $ma + nb = 1$.*

We introduce another familiar notion, that of prime number. By this we shall mean an integer which has no nontrivial factorization. For technical reasons, we exclude 1 from the set of prime numbers. The sequence 2, 3, 5, 7, 11, ... are all prime numbers; equally, $-2, -3, -5, \ldots$ are prime numbers. Since, in factoring, the negative introduces no essential differences, for us prime numbers will always be positive.

DEFINITION The integer $p > 1$ is a *prime number* if its only divisors are $\pm 1, \pm p$.

Another way of putting this is to say that an integer p (larger than 1) is a prime number if and only if given any other integer n then either $(p, n) = 1$ or $p \mid n$. As we shall soon see, the prime numbers are the building blocks of the integers. But first we need the important observation,

LEMMA 1.3.2 *If a is relatively prime to b but $a \mid bc$, then $a \mid c$.*

Proof. Since a and b are relatively prime, by the corollary to Lemma 1.3.1, we can find integers m and n such that $ma + nb = 1$. Thus $mac + nbc = c$. Now $a \mid mac$ and, by assumption, $a \mid nbc$; consequently,

$a \mid (mac + nbc)$. Since $mac + nbc = c$, we conclude that $a \mid c$, which is precisely the assertion of the lemma.

Following immediately from the lemma and the definition of prime number is the important

COROLLARY *If a prime number divides the product of certain integers it must divide at least one of these integers.*

We leave the proof of the corollary to the reader.

We have asserted that the prime numbers serve as the building blocks for the set of integers. The precise statement of this is the *unique factorization theorem*:

THEOREM 1.3.1 *Any positive integer $a > 1$ can be factored in a unique way as $a = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_t^{\alpha_t}$, where $p_1 > p_2 > \cdots > p_t$ are prime numbers and where each $\alpha_i > 0$.*

Proof. The theorem as stated actually consists of two distinct subtheorems; the first asserts the possibility of factoring the given integer as a product of prime powers; the second assures us that this decomposition is unique. We shall prove the theorem itself by proving each of these subtheorems separately.

An immediate question presents itself: How shall we go about proving the theorem? A natural method of attack is to use mathematical induction. A short word about this; we shall use the following version of mathematical induction: If the proposition $P(m_0)$ is true and if the truth of $P(r)$ for all r such that $m_0 \leq r < k$ implies the truth of $P(k)$, then $P(n)$ is true for all $n \geq m_0$. This variant of induction can be shown to be a consequence of the basic property of the integers which asserts that any nonempty set of positive integers has a minimal element (see Problem 10).

We first prove that every integer $a > 1$ can be factored as a product of prime powers; our approach is via mathematical induction.

Certainly $m_0 = 2$, being a prime number, has a representation as a product of prime powers.

Suppose that any integer r, $2 \leq r < k$ can be factored as a product of prime powers. If k itself is a prime number, then it is a product of prime powers. If k is not a prime number, then $k = uv$, where $1 < u < k$ and $1 < v < k$. By the induction hypothesis, since both u and v are less than k, each of these can be factored as a product of prime powers. Thus $k = uv$ is also such a product. We have shown that the truth of the proposition for all integers r, $2 \leq r < k$, implies its truth for k. Consequently, by the basic induction principle, the proposition is true for all integers $n \geq m_0 = 2$; that is, every integer $n \geq 2$ is a product of prime powers.

Now for the uniqueness. Here, too, we shall use mathematical induction, and in the form used above. Suppose that

$$a = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r} = q_1^{\beta_1} q_2^{\beta_2} \cdots q_s^{\beta_s},$$

where $p_1 > p_2 > \cdots p_r$, $q_1 > q_2 > \cdots > q_s$ are prime numbers, and where each $\alpha_i > 0$ and each $\beta_i > 0$. Our object is to prove

1. $r = s$.
2. $p_1 = q_1, p_2 = q_2, \ldots, p_r = q_r$.
3. $\alpha_1 = \beta_1, \alpha_2 = \beta_2, \ldots, \alpha_r = \beta_r$.

For $a = 2$ this is clearly true. Proceeding by induction we suppose it to be true for all integers u, $2 \leq u < a$. Now, since

$$a = p_1^{\alpha_1} \cdots p_r^{\alpha_r} = q_1^{\beta_1} \cdots q_s^{\beta_s}$$

and since $\alpha_1 > 0$, $p_1 \mid a$, hence $p_1 \mid q_1^{\beta_1} \cdots q_s^{\beta_s}$. However, since p_1 is a prime number, by the corollary to Lemma 1.3.2, it follows easily that $p_1 = q_i$ for some i. Thus $q_1 \geq q_i = p_1$. Similarly, since $q_1 \mid a$ we get $q_1 = p_j$ for some j, whence $p_1 \geq p_j = q_1$. In short, we have shown that $p_1 = q_1$. Therefore $a = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_r^{\alpha_r} = p_1^{\beta_1} q_2^{\beta_2} \cdots q_s^{\beta_s}$. We claim that this forces $\alpha_1 = \beta_1$. (Prove!) But then

$$b = \frac{a}{p^{\alpha_1}} = p_2^{\alpha_2} \cdots p_r^{\alpha_r} = q_2^{\beta_2} \cdots q_s^{\beta_s}.$$

If $b = 1$, then $\alpha_2 = \cdots = \alpha_r = 0$ and $\beta_2 = \cdots = \beta_s = 0$; that is, $r = s = 1$, and we are done. If $b > 1$, then since $b < a$ we can apply our induction hypothesis to b to get

1. The number of distinct prime power factors (in b) on both sides is equal, that is, $r - 1 = s - 1$, hence $r = s$.
2. $\alpha_2 = \beta_2, \ldots, \alpha_r = \beta_r$.
3. $p_2 = q_2, \ldots, p_r = q_r$.

Together with the information we already have obtained, namely, $p_1 = q_1$ and $\alpha_1 = \beta_1$, this is precisely what we were trying to prove. Thus we see that the assumption of the uniqueness of factorization for the integers less than a implied the uniqueness of factorization for a. In consequence, the induction is completed and the assertion of unique factorization is established.

We change direction a little to study the important notion of congruence modulo a given integer. As we shall see later, the relation that we now introduce is a special case of a much more general one that can be defined in a much broader context.

DEFINITION Let $n > 0$ be a fixed integer. We define $a \equiv b \bmod n$ if $n \mid (a - b)$.

The relation is referred to as *congruence modulo n*, n is called the *modulus* of the relation, and we read $a \equiv b \bmod n$ as "a is congruent to b modulo n." Note, for example, that $73 \equiv 4 \bmod 23$, $21 \equiv -9 \bmod 10$, etc.

This congruence relation enjoys the following basic properties:

LEMMA 1.3.3

1. *The relation congruence modulo n defines an equivalence relation on the set of integers.*
2. *This equivalence relation has n distinct equivalence classes.*
3. *If* $a \equiv b \bmod n$ *and* $c \equiv d \bmod n$, *then* $a + c \equiv b + d \bmod n$ *and* $ac \equiv bd \bmod n$.
4. *If* $ab \equiv ac \bmod n$ *and a is relatively prime to n, then* $b \equiv c \bmod n$.

Proof. We first verify that the relation congruence modulo n is an equivalence relation. Since $n \mid 0$, we indeed have that $n \mid (a - a)$ whence $a \equiv a \bmod n$ for every a. Further, if $a \equiv b \bmod n$ then $n \mid (a - b)$, and so $n \mid (b - a) = -(a - b)$; thus $b \equiv a \bmod n$. Finally, if $a \equiv b \bmod n$ and $b \equiv c \bmod n$, then $n \mid (a - b)$ and $n \mid (b - c)$ whence $n \mid \{(a - b) + (b - c)\}$, that is, $n \mid (a - c)$. This, of course, implies that $a \equiv c \bmod n$.

Let the equivalence class, under this relation, of a be denoted by $[a]$; we call it the *congruence class* (mod n) of a. Given any integer a, by the Euclidean algorithm, $a = kn + r$ where $0 \leq r < n$. But then, $a \in [r]$ and so $[a] = [r]$. Thus there are at most n distinct congruence classes; namely, $[0], [1], \ldots, [n - 1]$. However, these are distinct, for if $[i] = [j]$ with, say, $0 \leq i < j < n$, then $n \mid (j - i)$ where $j - i$ is a positive integer less than n, which is obviously impossible. Consequently, there are exactly the n distinct congruence classes $[0], [1], \ldots, [n - 1]$. We have now proved assertions 1 and 2 of the lemma.

We now prove part 3. Suppose that $a \equiv b \bmod n$ and $c \equiv d \bmod n$; therefore, $n \mid (a - b)$ and $n \mid (c - d)$ whence $n \mid \{(a - d) + (c - d)\}$, and so $n \mid \{(a + c) - (b + d)\}$. But then $a + c \equiv b + d \bmod n$. In addition, $n \mid \{(a - b)c + (c - d)b\} = ac - bd$, whence $ac \equiv bd \bmod n$.

Finally, notice that if $ab \equiv ac \bmod n$ and if a is relatively prime to n, then the fact that $n \mid a(b - c)$, by Lemma 1.3.2, implies that $n \mid (b - c)$ and so $b \equiv c \bmod n$.

If a is not relatively prime to n, the result of part 4 may be false; for instance, $2.3 \equiv 4.3 \bmod 6$, yet $2 \not\equiv 4 \bmod 6$.

Lemma 1.3.3 opens certain interesting possibilities for us. Let J_n be the

set of the congruence classes mod n; that is, $J_n = \{[0], [1], \ldots, [n-1]\}$. Given two elements, $[i]$ and $[j]$ in J_n, let us define

$$[i] + [j] = [i + j]; \tag{a}$$

$$[i][j] = [ij]. \tag{b}$$

We assert that the lemma assures us that this "addition" and "multiplication" are *well defined*; that is, if $[i] = [i']$ and $[j] = [j']$, then $[i] + [j] = [i + j] = [i' + j'] = [i'] + [j']$ and that $[i][j] = [i'][j']$. (Verify!) These operations in J_n have the following interesting properties (whose proofs we leave as exercises): for any $[i]$, $[j]$, $[k]$ in J_n,

1. $[i] + [j] = [j] + [i]$ } commutative laws.
2. $[i][j] = [j][i]$ }
3. $([i] + [j]) + [k] = [i] + ([j] + [k])$ } associative laws.
4. $([i][j])[k] = [i]([j][k])$ }
5. $[i]([j] + [k]) = [i][j] + [i][k]$ distributive law.
6. $[0] + [i] = [i]$.
7. $[1][i] = [i]$.

One more remark: if $n = p$ is a prime number and if $[a] \neq [0]$ is in J_p, then there is an element $[b]$ in J_p such that $[a][b] = [1]$.

The set J_n plays an important role in algebra and number theory. It is called the set of *integers* mod n; before we proceed much further we will have become well acquainted with it.

Problems

1. If $a \mid b$ and $b \mid a$, show that $a = \pm b$.
2. If b is a divisor of g and of h, show it is a divisor of $mg + nh$.
3. If a and b are integers, the *least common multiple* of a and b, written as $[a, b]$, is defined as that positive integer d such that
   (a) $a \mid d$ and $b \mid d$.
   (b) Whenever $a \mid x$ and $b \mid x$ then $d \mid x$.
   Prove that $[a, b]$ exists and that $[a, b] = ab/(a, b)$, if $a > 0$, $b > 0$.
4. If $a \mid x$ and $b \mid x$ and $(a, b) = 1$ prove that $(ab) \mid x$.
5. If $a = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ and $b = p_1^{\beta_1} \cdots p_k^{\beta_k}$ where the p_i are distinct prime numbers and where each $\alpha_i \geq 0$, $\beta_i \geq 0$, prove
   (a) $(a, b) = p_1^{\delta_1} \cdots p_k^{\delta_k}$ where δ_i = minimum of α_i and β_i for each i.
   (b) $[a, b] = p_1^{\gamma_1} \cdots p_k^{\gamma_k}$ where γ_i = maximum of α_i and β_i for each i.

6. Given a, b, on applying the Euclidean algorithm successively we have

$$\begin{aligned} a &= q_0 b + r_1, & 0 \le r_1 < |b|, \\ b &= q_1 r_1 + r_2, & 0 \le r_2 < r_1, \\ r_1 &= q_2 r_2 + r_3, & 0 \le r_3 < r_2, \\ &\cdot \\ &\cdot \\ &\cdot \\ r_k &= q_{k+1} r_{k+1} + r_{k+2}, & 0 \le r_{k+2} < r_{k+1}. \end{aligned}$$

Since the integers r_k are decreasing and are all nonnegative, there is a first integer n such that $r_{n+1} = 0$. Prove that $r_n = (a, b)$. (We consider, here, $r_0 = |b|$.)

7. Use the method in Problem 6 to calculate
(a) (1128, 33). (b) (6540, 1206).

8. To check that n is a prime number, prove that it is sufficient to show that it is not divisible by any prime number p, such that $p \le \sqrt{n}$.

9. Show that $n > 1$ is a prime number if and only if for any a either $(a, n) = 1$ or $n \mid a$.

10. Assuming that any nonempty set of positive integers has a minimal element, prove
(a) If the proposition P is such that
(1) $P(m_0)$ is true,
(2) the truth of $P(m-1)$ implies the truth of $P(m)$,
then $P(n)$ is true for all $n \ge m_0$.
(b) If the proposition P is such that
(1) $P(m_0)$ is true,
(2) $P(m)$ is true whenever $P(a)$ is true for all a such that $m_0 \le a < m$,
then $P(n)$ is true for all $n \ge m_0$.

11. Prove that the addition and multiplication used in J_n are well defined.

12. Prove the properties 1–7 for the addition and multiplication in J_n.

13. If $(a, n) = 1$, prove that one can find $[b] \in J_n$ such that $[a][b] = [1]$ in J_n.

*14. If p is a prime number, prove that for any integer a, $a^p \equiv a \bmod p$.

15. If $(m, n) = 1$, given a and b, prove that there exists an x such that $x \equiv a \bmod m$ and $x \equiv b \bmod n$.

16. Prove the corollary to Lemma 1.3.2.

17. Prove that n is a prime number if and only if in J_n, $[a][b] = [0]$ implies that $[a] = [b] = [0]$.