INTRODUCTION TO MODERN ALGEBRA

Marvin Marcus

INTRODUCTION TO MODERN ALGEBRA

Marvin Marcus
Institute for the Interdisciplinary
Applications of Algebra and Combinatorics
University of California, Santa Barbara
Santa Barbara, California

MARCEL DEKKER, INC. New York and Basel

Library of Congress Cataloging in Publication Data

Marcus, Marvin [Date]
Introduction to modern algebra.

(Monographs and textbooks in pure and applied mathematics ; v. 47)
1. Algebra, Abstract. I. Title.
QA162.M37 512'.02 78-2482
ISBN 0-8247-6479-X

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Current printing (last digit):
10 9 8 7 6 5 4 3 2 1

PRINTED IN THE UNITED STATES OF AMERICA

Preface

This book is written as a text for a basic one-year course in algebra at the advanced undergraduate or beginning graduate level. The presentation is oriented toward the applications of algebra to other branches of mathematics and to science in general. This point of view is reflected in the selection of constructive methods of proof, the choice of topics, and the space devoted to items related to current applications of algebra. Thus, modules over a principal ideal domain are studied via elementary operations on matrices. Considerable space is devoted to such topics as permutation groups and the Pòlya counting theory; polynomial theory; canonical forms for matrices; applications of linear algebra to differential equations; representations of groups.

Prerequisites for a course based on this book are minimal: standard one-quarter courses in the theory of equations and elementary matrix theory suffice. Altogether there are 390 exercises and these constitute an integral part of the book. Problems that require somewhat intricate arguments are accompanied by complete solutions. The exercises contain a number of important results and several definitions. Occasionally they are used to remove technical arguments from the mainstream of a proof within the section. Students should at least read the exercises. Frequently exercises appearing at the end of a section are mentioned within the section so that they can be easily assigned at the appropriate time.

We have not hesitated to reiterate definitions and results throughout the book. For example, conjugacy classes are discussed in the chapter on group theory and again in the chapter on group representation theory. Moreover, some arguments are repeated if they are separated from their last occurrence by a substantial amount of intervening material. Each section of the

book is followed by a Glossary which contains the page numbers for important definitions, "name" theorems, and special notations.

What follows is a description of the contents of each of the chapters. A diagram illustrating the interdependency of the various sections appears after the Preface.

Chapter 1, *Basic Structures*, introduces many of the basic ideas that occur later in the book. Section 1.1, *Sets and Functions*, contains the usual material on sets, functions, the de Morgan formulas, function composition, Cartesian products, equivalence relations, quotient sets, systems of distinct representatives, universal properties, partial and linear orderings, etc. Towards the end of the section, the Axiom of Choice is discussed in a heuristic way. The equivalence of Zorn's lemma and the Axiom of Choice is mentioned without going into much detail. Section 1.2, *Algebraic Structures*, introduces in order of increasing complexity some of the basic items developed in the remainder of the book. Thus groupoids, semigroups, monoids, groups, modules, vector spaces, algebras, and matrices appear here. In this section we also define permutation groups, free monoids, groupoid rings, polynomial rings, free power series, etc. The section contains an extensive list of elementary examples illustrating the definitions. The student can obtain considerable practice in the manipulation of these basic ideas in the exercise sections. Categories and morphisms appear in the exercises, but only peripherally.

Section 1.3, *Permutation Groups*, examines the details of permutation groups and their structure. The basic properties of permutations (including cycle structure and the simplicity of the alternating group of degree n for $n \geq 5$) appear here. Many of the basic ideas of group theory are illustrated in Section 1.3 in the context of permutation groups.

Chapter 2, *Groups*, is a rather thorough study of most of the major elementary theorems in group theory. Section 2.1, *Isomorphism Theorems*, carries the reader through the Jordan-Hölder theorem, properties of solvable groups, and composition series. Section 2.2, *Group Actions and the Sylow Theorems*, is devoted to a systematic study of the three major Sylow theorems. Since this section is highly combinatorial in nature, it seemed appropriate to include the Pòlya counting theorem and some interesting combinatorial applications.

Section 2.3, *Some Additional Topics*, contains a number of more advanced items in group theory, beginning with the Zassenhaus isomorphism lemma for groups. We then develop the Schreier refinement theorem for subnormal series of a (not necessarily finite) group. This section also includes the notion of a group with operators, admissible subgroups, and linear maps on vector spaces.

Chapter 3, *Rings and Fields*, is the longest chapter in the book. Section

3.1, *Basic Facts*, covers ring characteristics, universal factorization properties of quotient rings, and the three ring isomorphism theorems.

Section 3.2, *Introduction to Polynomial Rings*, shows how an indeterminate can be constructed over an arbitrary ring. The ring extension theorem is proved and used here to imbed a ring in a ring with an indeterminate. This section also contains material on polynomials in several variables, including the basis theorem for symmetric polynomials. Ascending and descending chain conditions for ideals in a ring are discussed and the Hilbert basis theorem for Noetherian rings is proved. Quotient fields of integral domains, and more generally, rings of fractions with respect to ideals appear toward the end of the section.

Section 3.3, *Unique Factorization Domains*, starts with the usual material on polynomial division, the division algorithm and the remainder theorem. The division algorithm is proved over a noncommutative ring—it is required later in the study of elementary divisors over matrix rings. The basic fact that any principal ideal domain is a unique factorization domain is proved in Theorem 3.6. Example 6 shows how to calculate the greatest common divisor of two Gaussian integers using the Euclidean algorithm. Nilradicals, quotients of ideals, and the Jacobson radical all occur at the end of this section.

Section 3.4, *Polynomial Factorization*, begins in the standard way with Gauss' lemma and goes on to show that unique factorization is inherited by the polynomial ring over a unique factorization domain. Considerable space is devoted here to the practical problem of factoring polynomials. Theorem 4.7 shows how to construct a splitting field for a polynomial, and Theorem 4.13 exhibits the relationship between any two such splitting fields. The section concludes with a proof of the primitive element theorem for fields of characteristic zero. The exercises in this chapter contain a good deal of material, but detailed solutions are included for all but the most routine problems.

Section 3.5, *Polynomials and Resultants*, deals with the classical theory of polynomials. Sylvester's determinant, homogeneous polynomials, resultants, and discriminants appear here, and the fundamental question of when two polynomials have a common factor is investigated in some detail. This section concludes with a statement and proof of the Hilbert invariant theorem and a discussion of algebraic independence.

Section 3.6, *Applications to Geometric Constructions*, applies the preceding material on field theory to problems of ruler and compass construction of regular polygons and angle trisection.

Section 3.7, *Galois Theory*, is devoted to the proof of the fundamental theorem of Galois theory for fields of characteristic zero and its application

to the classical problem of the solvability of a general polynomial of degree n by radicals.

Chapter 4, *Modules and Linear Operators*, begins in Section 4.1, *The Hermite Normal Form*, with the derivation of a normal form under left equivalence of matrices over a principal ideal domain. This theorem is then applied to finitely generated modules, yielding many of the standard results in module theory as easy consequences. The Hermite normal form is also useful in showing how to compute generators for ideals in a matrix ring. This section also contains the basic theory of finite dimensional vector spaces, the Steinitz exchange theorem, and the theory of linear equations.

Section 4.2, *The Smith Normal Form*, shows how to compute canonical forms for matrices under right and left equivalence over a principal ideal domain. The Smith form is then used to analyze the structure of finitely generated modules as direct sums of free submodules. The fundamental theorem of abelian groups appears in Corollary 7. We then determine all low-order abelian groups in some of the examples and exercises. The cyclic primary decomposition of a module is carried out in the exercises, together with an analysis of finitely generated abelian groups given in terms of certain defining relations, i.e., group presentations.

Section 4.3, *The Structure of Linear Transformations*, develops the standard elementary divisor theory and matrix canonical forms over a field. Our approach is computational, and the canonical forms under similarity are derived in terms of the reduction of the characteristic matrix via equivalence over a polynomial ring. Most of the important normal forms for matrices under similarity occur here, e.g., the Frobenius normal form and the Jordan normal form. A considerable part of the section deals with the problem of computing the elementary divisors of a function of a matrix. These important results are used in other parts of mathematics, e.g., the theory of ordinary differential equations.

In the last section of the chapter, *Introduction to Multilinear Algebra*, we introduce symmetry classes of tensors and briefly study the tensor, Grassmann, and completely symmetric spaces. As an example of the use of an inner product in a symmetry class, we show how the famous van der Waerden conjecture concerning doubly stochastic matrices can be partially resolved.

The fifth and final chapter of the book, *Representations of Groups*, is essentially self-contained and could be used for a short course on group representation theory. The major part of the chapter is concerned with matrix representations of finite groups. This permits us to achieve deep penetration of the subject rather rapidly.

The contents of a course in algebra vary considerably and seem to depend more on individual tastes and prejudices than do corresponding courses in analysis. The present book is no exception. However, a good deal

of the material included can be justified in terms of its applications to other parts of mathematics and science. We anticipate that a student who gains reasonable mastery of the contents will be ready for more advanced courses in algebra and the applications of algebra to a wide range of fields, e.g., computer science, control theory, algebraic coding theory, system theory, numerical linear algebra, quantum mechanics, and crystallography.

References

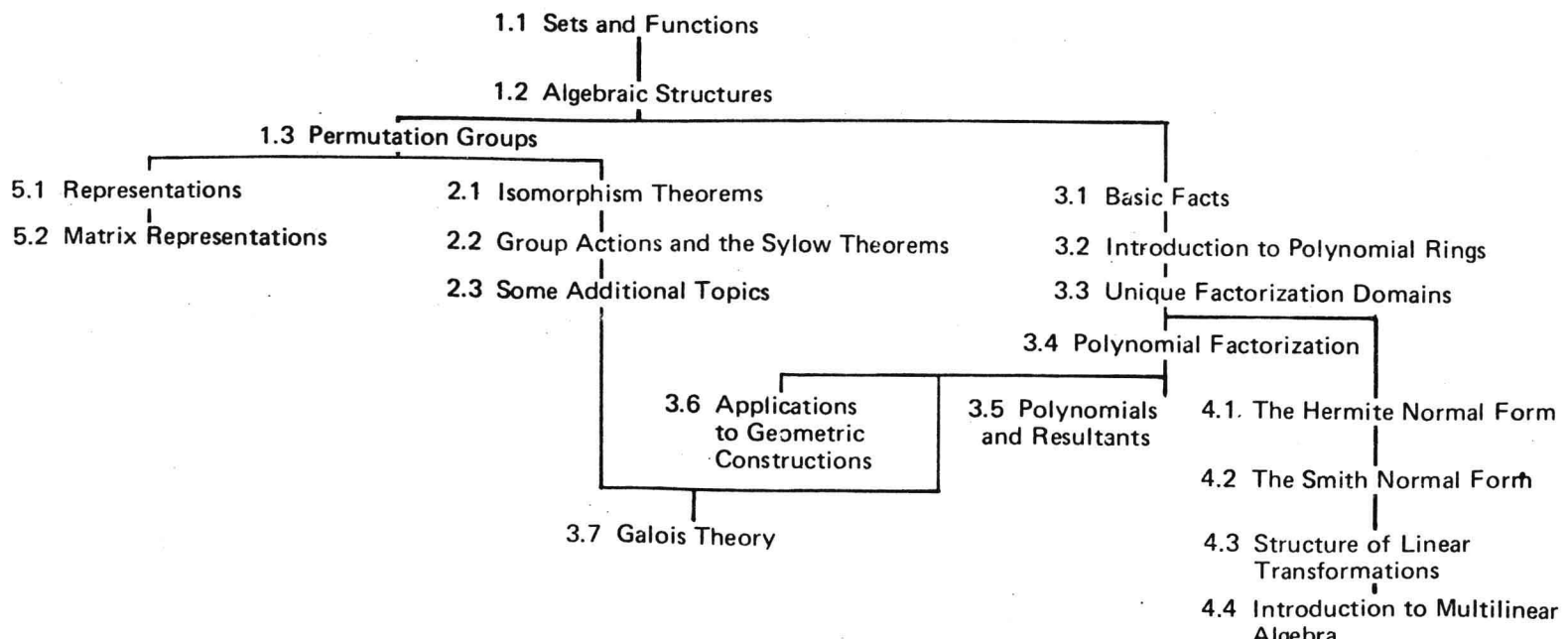
Each chapter is divided into sections. Thus Section 4.2 is Section 2 of Chapter 4. Definitions, theorems, and examples are numbered serially within a section. Thus Theorem 1.4 is the fourth theorem in Section 1. Any reference to a definition, theorem, or example within the chapter in which it appears does not identify the chapter. Any reference to a definition, theorem, or example occurring in another chapter includes the chapter and section number. The symbol ∎ is used to denote the end of a proof.

Acknowledgment

The author is pleased to acknowledge the assistance of Dr. Ivan Filippenko in reading the original manuscript and providing invaluable help in proofreading the printed copy.

Marvin Marcus

DEPENDENCE OF SECTIONS



Contents

1

Basic Structures

1.1 Sets and Functions

We shall assume that the reader is familiar with the notion of a *set* or collection of objects. The purpose of this section is to set forth the notation and language used throughout this book.

If S is a set and x is a *member* or an *element* of S, we write

$$x \in S;$$

if x is not an element of S, we write

$$x \notin S.$$

If X is a set consisting of all elements x for which a certain proposition $p(x)$ is true, we write

$$X = \{x \mid p(x)\}.$$

Thus, for example,

$$\{x \mid x \text{ is an integer and } \tfrac{1}{2} \leq x < 5\}$$

is the set consisting of the integers 1, 2, 3, and 4. It is often feasible to explicitly write out the elements of a set, e.g.,

$$X = \{2,4,6\} \tag{1}$$

means that X consists of the numbers 2, 4, and 6. The curly bracket notation in (1) is usually reserved for finite sets, but sometimes infinite sets can be written this way by use of the ubiquitous "triple dot" notation, e.g.,

$$N = \{1, 2, 3, \ldots\}$$

is the set of positive integers.

If every element of the set X is in the set Y, we write

$$X \subset Y,$$

or

$$Y \supset X,$$

and call X a *subset* of Y. If $X \subset Y$ but $X \neq Y$, then X is a *proper* subset of Y. The *empty set* or *null set*, denoted by $\varnothing$, is the set with no elements; clearly,

$$\varnothing \subset X$$

for any X. The *power set* of a set X is the set of all subsets of X. It is denoted by $P(X)$:

$$P(X) = \{Y \mid Y \subset X\}.$$

If X contains only finitely many elements, we denote the number of elements in X by $|X|$. It is an easy exercise to verify that for a finite set X,

$$|P(X)| = 2^{|X|} \tag{2}$$

(See Exercise 1). For example, if X is the set (1), then the elements of $P(X)$ are the eight subsets

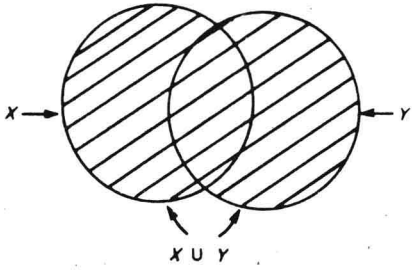
$$\varnothing, \quad \{2\}, \quad \{4\}, \quad \{6\}, \quad \{2,4\}, \quad \{2,6\}, \quad \{4,6\}, \quad X.$$

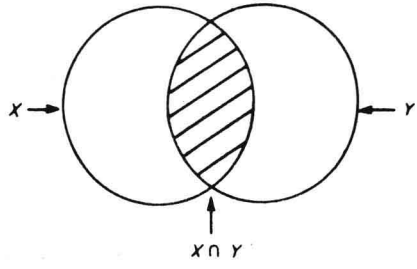
The *union* of two sets X and Y is the set of elements in either X or Y and is denoted by

$$X \cup Y.$$

The *intersection* of X and Y is the totality of elements in both X and Y and is denoted by

$$X \cap Y.$$





We say that I is an *indexing* set or *labeling* set for a family of sets if to each

element of I there corresponds a well-defined set X_i in the family. The union and intersection of a family of sets indexed by I are written, respectively,

$$\bigcup_{i \in I} X_i$$

and

$$\bigcap_{i \in I} X_i.$$

Thus, $x \in \bigcup_{i \in I} X_i$ means that $x \in X_i$ for some $i \in I$, whereas $x \in \bigcap_{i \in I} X_i$ means that $x \in X_i$ for every $i \in I$. For example, if $I = N$ and X_i is the closed interval on the real line consisting of all x such that $1/i \leq x \leq 1$, then

$$\bigcup_{i \in N} X_i = \{x \mid 0 < x \leq 1\}$$

and

$$\bigcap_{i \in N} X_i = \{1\}.$$

If $\{X_i \mid i \in I\}$ is a family of sets and $X_i \cap X_j = \varnothing$ whenever $i \neq j$, we say that the sets in the family are *pairwise disjoint*. If

$$X = \bigcup_{i \in I} X_i$$

and the sets X_i are pairwise disjoint, then $\{X_i \mid i \in I\}$ is a *partition* of X.

If X and Y are sets, then the set of elements in Y but not in X is the *complement* of X *relative to* Y, denoted by

$$Y - X.$$

If $X \subset Y$, we write

$$X^c$$

instead of $Y - X$ when Y is understood. The *De Morgan formulas* connect the union, intersection, and complements of a family of subsets of Y:

$$\left(\bigcup_{i \in I} X_i\right)^c = \bigcap_{i \in I} X_i^c \tag{3}$$

and

$$\left(\bigcap_{i \in I} X_i\right)^c = \bigcup_{i \in I} X_i^c \tag{4}$$

(see Exercise 2).

If X and Y are sets, then a *function* (or *mapping* or *map*) from X to Y is a well-defined rule that associates with each element $x \in X$ an element $f(x) \in Y$. The set of all maps from X to Y is denoted by Y^X. We write

$$f: X \to Y$$

or in diagram form

$$X \xrightarrow{f} Y$$

to indicate that f is a function from X to Y.

The element $f(x) \in Y$ is the *value of* f *at* x, or the *image of* x *under* f;

the set X is called the *domain* of f, written dmn f; the set Y is called the *codomain* of f; the set

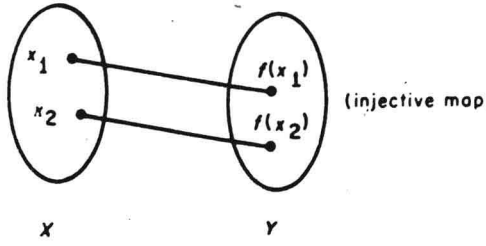
$$f(X) = \{y \in Y \mid y = f(x) \text{ for some } x \in X\}$$

is called the *image* or *range* of f, written im f.

If $Z \subset Y$, then $f^{-1}(Z)$, called the *inverse image* of Z, is the set

$$f^{-1}(Z) = \{x \in X \mid f(x) \in Z\}.$$

A map $f: X \to Y$ is *injective* (1–1) or an *injection* if $f(x_1) = f(x_2)$ implies $x_1 = x_2$; it is *surjective* (onto) or a *surjection* if $f(X) = Y$; it is *bijective* (1–1, onto) or a *bijection* if f is injective and surjective. Other words are *monomorphic* (injective); *epimorphic* (surjective); a *matching* (bijective) of X and Y.



For example, the map $f: N \to N$ defined by the formula

$$f(n) = 2n$$

is an injection but certainly not a surjection. However, if by $2N$ we mean the set of positive even integers, then $f: N \to 2N$ is a bijection. If $g: N \to \{1, -1\}$ is defined by $g(n) = (-1)^n$, then g is an epimorphism or a surjection, and

$$g^{-1}(\{1\}) = 2N,$$
$$g^{-1}(\{-1\}) = N - 2N.$$

Two functions $f: X \to Y$ and $g: X \to Y$ are *equal* if and only if (hereafter abbreviated "iff")

$$f(x) = g(x)$$

for all $x \in X$.

The *composite* of two functions $f: X \to Y$ and $g: Y \to Z$ is the map $h: X \to Z$ whose value at any $x \in X$ is $h(x) = g(f(x))$. The composite of f and g is written gf or $g \cdot f$. The composition of maps is *associative*: If $X \xrightarrow{f} Y$, $Y \xrightarrow{g} Z$, and $Z \xrightarrow{h} W$, then $h \cdot (g \cdot f) = (h \cdot g) \cdot f$ (verify!). Thus the triple composite may simply be denoted by $h \cdot g \cdot f$.

If $\{Y_i \mid i \in I\}$ is a family of sets, then the *cartesian product* of the

members of the family is the set of all functions $f\colon I \to \bigcup_{i\in I} Y_i$ such that $f(i) \in Y_i$ for each $i \in I$. The cartesian product is denoted by

$$\times_{i\in I} Y_i = \{f \mid f\colon I \to \bigcup_{i\in I} Y_i \text{ and } f(i) \in Y_i \text{ for each } i \in I\}.$$

If $\{Y_1, \ldots, Y_n\}$ is a family of n sets, their cartesian product is also written

$$Y_1 \times \cdots \times Y_n$$

and can be thought of as the totality of ordered n-tuples

$$f = (y_1, \ldots, y_n),$$

$y_i \in Y_i$, $i = 1, \ldots, n$; that is, $f(i) = y_i$, $i = 1, \ldots, n$. Two n-tuples $(y_1, \ldots, y_n)$ and $(z_1, \ldots, z_n)$ are equal iff $y_i = z_i$, $i = 1, \ldots, n$.

Suppose $I = [0,1]$, i.e., I is the closed interval on the real line consisting of all x for which $0 \leq x \leq 1$. For each $i \in I$ let $Y_i = I$. We assert that the cartesian product

$$\times_{i\in I} Y_i$$

is in fact the set of all maps from I to I, i.e.,

$$I^I = \times_{i\in I} Y_i.$$

For, any $f \in \times_{i\in I} Y_i$ is a function whose value at each $i \in I$ is an element of $Y_i = I$.

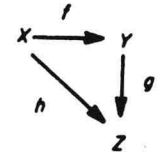
The special map $\iota_X\colon X \to X$, called the *identity map*, is defined by

$$\iota_X(x) = x$$

for each $x \in X$.

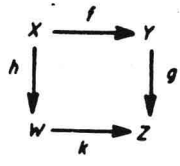
If $Z \subset X$ and $f\colon X \to Y$, then $f|Z$ is the function whose domain is Z and whose value for each $z \in Z$ is $f(z)$; $f|Z$ is called the *restriction* of f to Z, and f is called an *extension* of $f|Z$. If $Z \subset X$, then the map $\iota_X | Z$ is called the *canonical injection* of Z into X.

Compositions of maps are often depicted by *mapping diagrams*; for example,



(5)

means that $f\colon X \to Y$, $g\colon Y \to Z$, $h\colon X \to Z$, and $h = g \cdot f$. Another example,



(6)

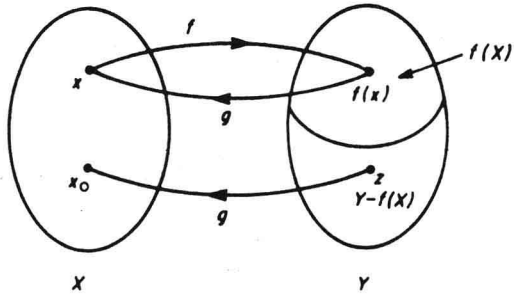
indicates that $g \cdot f = k \cdot h$. Diagrams showing the equality of compositions of sequences of maps, such as (5) and (6), are called *commutative diagrams.*

If $f\colon X \to Y$, $g\colon Y \to X$, and $gf = \iota_X$, then g is a *left inverse* of f; if $fg = \iota_Y$, then g is a *right inverse* of f. If g is a left and right inverse of f, then it is an *inverse* of f. (See Exercise 3.)

Theorem 1.1 *Assume $f\colon X \to Y$. Then*

(i) *f is injective iff it has a left inverse.*
(ii) *f is surjective iff it has a right inverse.*
(iii) *If f has a left inverse g and a right inverse h, then $g = h$.*
(iv) *f is bijective iff it has an inverse.*
(v) *If f has an inverse it is unique and is denoted by f^{-1}.*
(vi) *If f has an inverse, then $(f^{-1})^{-1} = f$.*

Proof: (i) If f has a left inverse $g\colon Y \to X$, then $f(x_1) = f(x_2)$ implies that $g(f(x_1)) = g(f(x_2))$ and hence that $x_1 = \iota_X(x_1) = (g \cdot f)(x_1) = g(f(x_1)) = g(f(x_2)) = (g \cdot f)(x_2) = \iota_X(x_2) = x_2$. Hence f is injective. Conversely, if f is injective, then for each $y \in f(X)$ there is exactly one element in X, call it $x_y \in X$, such that $f(x_y) = y$: define $g|f(X)$ by $g(y) = x_y$. For any other $z \in Y$, let $g(z) = x_0$, some fixed element in X. Obviously $(g \cdot f)(x) = g(f(x)) = x = \iota_X(x)$ for all $x \in X$, so g is a left inverse of f.



(ii) If $f\colon X \to Y$ is surjective, then $f(X) = Y$. Let $g\colon Y \to X$ be defined as follows: For each $y \in Y$ choose an $x_y \in f^{-1}(\{y\})$ and let $g(y) = x_y$. Then $(f \cdot g)(y) = f(g(y)) = f(x_y) = y = \iota_Y(y)$, i.e., $f \cdot g = \iota_Y$. Hence f has a right inverse. Conversely, if $g\colon Y \to X$ is a right inverse of f and $y \in Y$, then $y = \iota_Y(y) = (f \cdot g)(y) = f(g(y))$ and hence $y \in \operatorname{im} f$. Thus f is surjective.