ALGEBRA
...This great science which I have been calling "Characteristic," of which Algebra and Analysis are but small branches, ...is what gives words to languages, letters to words, digits to Arithmetic, notes to Music; this is what teaches us the secret of precise reasoning, requiring us to leave, as visible traces on paper, a volume for inspection at leisure: finally, this is what makes us reason, substituting characters in place of things, thereby unburdening the imagination.

Leibniz, De la Méthode de l'Universitalité (1674)
CONTENTS

Preface xi
Table of Principal Notation xv
Prerequisites xvii
Exercises xxii

PART I—GROUPS 1

Chapter 1. Monoids and Groups 2
Examples of Groups and Monoids 2
When Is a Monoid a Group? 4
Exercises 6

Chapter 2. How to Divide: Lagrange’s Theorem, Cosets, and an Application to Number Theory 8
Cosets 10
Fermat’s Little Theorem 12
Exercises 13
Euler’s Number 13

Chapter 3. Cauchy’s Theorem: How to Show a Number Is Greater Than 1 15
The Exponent 16
$S_n$: Our Main Example 17
Subgroups of $S_n$ 17
Cycles 18
The Product of Two Subgroups 19
The Classical Groups 20
| Exercises                                      | 21 |
| The Classical Groups                          | 22 |
| Chapter 4. Introduction to the Classification of Groups: |
| Homomorphisms, Isomorphisms, and Invariants   | 23 |
| Homomorphic Images                             | 26 |
| Exercises                                      | 27 |
| Chapter 5. Normal Subgroups—The Building Blocks of the Structure Theory | 29 |
| The Residue Group                              | 31 |
| Noether's Isomorphism Theorems                 | 32 |
| Conjugates in $S_n$                            | 34 |
| The Alternating Group                          | 35 |
| Exercises                                      | 37 |
| $S_n$ and $A_n$                                | 37 |
| Chapter 6. Classifying Groups—Cyclic Groups and Direct Products | 39 |
| Cyclic Groups                                  | 39 |
| Generators of a Group                          | 40 |
| Direct Products                                | 40 |
| Internal Direct Products                       | 42 |
| Exercises                                      | 44 |
| Chapter 7. Finite Abelian Groups               | 46 |
| Abelian $p$-Groups                             | 48 |
| Proof of the Fundamental Theorem for Finite Abelian Groups | 51 |
| The Classification of Finite Abelian Groups    | 51 |
| Exercises                                      | 53 |
| Finitely Generated Abelian Groups              | 54 |
| Chapter 8. Generators and Relations            | 56 |
| Description of Groups of Low Order             | 57 |
| Addendum: Erasing Relations                    | 61 |
| Exercises                                      | 63 |
| Explicit Generation of Groups by Arbitrary Subsets | 65 |
| Chapter 9. When Is a Group a Group? (Cayley's Theorem) | 66 |
| Generalized Cayley's Theorem                   | 68 |
| Group Representations                           | 70 |
| Exercises                                      | 71 |
## Chapter 10. Recounting: Conjugacy Classes and the Class Formula

The Center of a Group: 72

Groups Acting on Sets: A Recapitulation: 75

**Exercises**
- Double Cosets: 77
- Group Actions on Sets: 77

## Chapter 11. Sylow Subgroups: A New Invariant

Groups of Order Less Than 60: 82

Simple Groups: 83

**Exercises**
- Classification of Groups of Various Orders: 84

## Chapter 12. Solvable Groups: What Could Be Simpler?

Commutators: 86

Solvable Groups: 87

**Addendum:** Automorphisms of Groups: 89

**Exercises**
- Nilpotent Groups: 89
- The Special Linear Group $SL(n, F)$: 90
- The Projective Special Linear Group $PSL(n, F)$: 91
- Exercises for the Addendum: 92
- Semidirect Products, also cf. Example 8.8: 92
- The Wreath Product: 93

*Review Exercises for Part I* 95

## PART II—RINGS AND POLYNOMIALS

## Chapter 13. An Introduction to Rings

Domains and Skew Fields: 98

Left Ideals: 99

**Exercises**
- Rings of Matrices: 101
- Direct Products of Rings: 101

## Chapter 14. The Structure Theory of Rings

Ideals: 104

Noether's Isomorphism Theorems: 106

**Exercises**
- The Regular Representation: 107
- General Structure Theory: 107
Chapter 15. The Field of Fractions—A Study in Generalization

Intermediate Rings
Exercises
Subrings of $\mathbb{Q}$

Chapter 16. Polynomials and Euclidean Domains

The Ring of Polynomials
Euclidean Domains
Unique Factorization
Exercises
Formal Power Series
The Partition Number
Unique Factorization Domains

Chapter 17. Principal Ideal Domains: Induction without Numbers

Prime Ideals
Noetherian Rings
Exercises
Counterexamples
Consequences of Zorn's Lemma
UFD's
Noetherian Rings

Chapter 18. Roots of Polynomials

Finite Subgroups of Fields
Primitive Roots of 1
Exercises
The Structure of Euler($n$)

Chapter 19. (Optional) Applications: Famous Results from Number Theory

A Theorem of Fermat
Addendum: Fermat's Last Theorem
Exercises

Chapter 20. Irreducible Polynomials

Polynomials over UFD's
Eisenstein's Criterion
Exercises
Nagata's Theorem and its Applications
The Ring $\mathbb{Z}[x_1, \ldots, x_n]$, and the Generic Method

Review Exercises for Part II
PART III—FIELDS

Historical Background

Chapter 21. Field Extensions: Creating Roots of Polynomials

Algebraic Elements
Finite Field Extensions
Exercises
  Countability and Transcendental Numbers
  Algebraic Extensions

Chapter 22. The Problems of Antiquity

Construction by Straight Edge and Compass
Algebraic Description of Constructibility
Solution of the Problems of Antiquity
Exercises
  Constructibility
  Constructing a Regular \( n \)-gon

Chapter 23. Adjoining Roots to Polynomials: Splitting Fields

Splitting Fields
Separable Polynomials and Separable Extensions
The Characteristic of a Field
Exercises
  The Roots of a Polynomial in Terms of the Coefficients
  Separability and the Characteristic
  Calculus through the Looking Glass

Chapter 24. Finite Fields

Reduction Modulo \( p \)
Exercises

Chapter 25. The Galois Correspondence

The Galois Group of Automorphisms of a Field Extension
The Galois Group and Intermediate Fields
Exercises
  The Galois Group of the Compositum
  More on Artin’s Lemma
### Chapter 26. Applications of the Galois Correspondence 200

- Finite Separable Field Extensions and the Normal Closure 200
- The Galois Group of a Polynomial 201
- Constructible $n$-gons 202
- Finite Fields 202
- The Fundamental Theorem of Algebra 203

*Exercises*

<table>
<thead>
<tr>
<th>Exercise</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Normal Closure</td>
<td>204</td>
</tr>
<tr>
<td>Separability Degree</td>
<td>205</td>
</tr>
<tr>
<td>Finite Fields</td>
<td>205</td>
</tr>
<tr>
<td>The Algebraic Closure</td>
<td>205</td>
</tr>
</tbody>
</table>

### Chapter 27. Solving Equations by Radicals 207

- Root Extensions 207
- Solvable Galois Groups 209
- Computing the Galois Group 211

*Exercises*

<table>
<thead>
<tr>
<th>Exercise</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prescribed Galois Groups</td>
<td>213</td>
</tr>
<tr>
<td>Root Towers</td>
<td>213</td>
</tr>
<tr>
<td>Finding the Number of Real Roots via the Discriminant</td>
<td>214</td>
</tr>
<tr>
<td>Galois Groups and Solvability</td>
<td>215</td>
</tr>
<tr>
<td>The Norm and Trace</td>
<td>216</td>
</tr>
</tbody>
</table>

*Review Exercises for Part III*

<table>
<thead>
<tr>
<th>Exercise</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prescribed Galois Groups</td>
<td>213</td>
</tr>
<tr>
<td>Root Towers</td>
<td>213</td>
</tr>
<tr>
<td>Finding the Number of Real Roots via the Discriminant</td>
<td>214</td>
</tr>
<tr>
<td>Galois Groups and Solvability</td>
<td>215</td>
</tr>
<tr>
<td>The Norm and Trace</td>
<td>216</td>
</tr>
</tbody>
</table>

### Appendix A. Transcendental Numbers: $e$ and $\pi$ 219

- Transcendence of $e$ 219
- Transcendence of $\pi$ 222

### Appendix B. Skew Field Theory 224

- The Quaternion Algebra 224
- Proof of Lagrange's Four Square Theorem 226
- Polynomials over Skew Fields 227
- Structure Theorems for Skew Fields 229

*Exercises*

<table>
<thead>
<tr>
<th>Exercise</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proof of Lagrange's Four Square Theorem</td>
<td>231</td>
</tr>
<tr>
<td>Frobenius' Theorem</td>
<td>232</td>
</tr>
</tbody>
</table>

*Index* 233
PREFACE

Algebra is a subject with which we become acquainted during most of our education, largely in connection with the solution of equations. Some of the most famous questions in mathematical history have involved equations with coefficients in \( \mathbb{Z} \), the set of integers. This course deals with their solutions. We shall see that the process of abstraction enables us to solve a variety of problems with economy of effort. This is the principle at the heart of abstract algebra, a subject that enables one to deduce sweeping conclusions from elementary premises. As such, this course can be used to initiate an intelligent student to the glorious world of mathematical discovery. At the same time, a course in abstract algebra, properly presented, could treat mathematics as an art as well as a science. In these notes I have tried to present underlying ideas, as well as the results they yield.

Abstract algebra received a major impetus toward the beginning of this century, when "intuition" in geometry began to lead to false assertions. It was seen that algebraic structure provides a firm foundation for a new subject called algebraic geometry, which enabled Zariski and others to put the latest developments of geometry on a solid footing. Indeed, it is not surprising that much of the structure theory of algebra was developed by Emmy Noether, the daughter of a geometer. Although algebraic geometry is outside the scope of this short book, we do attempt to lay the pedagogical foundations by introducing Noetherian rings and prime ideals.

These notes are far from comprehensive, and the serious student might continue with the far more thorough texts of Jacobson (Basic Algebra, Freeman 1985), Artin (Algebra, Prentice Hall 1991), and/or Cohn (Algebra I, Wiley 1974). Throughout, I have tried to stick to the main track of a one-year course, with the aim of touching on as many important theorems and
applications as possible. The observant reader will notice the heavy debt to
the classic text of Herstein (Topics in Algebra, Xerox, 1964), which in turn
owes a great deal to Van der Waerden’s pioneering work. The cursory treat­
ment here is indicated by the division of the material into chapters, each
of which is supposed to correspond to one of the 26 weeks of a year-long
algebra course. (Some of these chapters are extended a bit by appendices.)

References are normally within the same chapter, unless indicated other­
wise by a decimal point; for example, “Theorem 2.15” means “Theorem 15
of Chapter 2.”

My personal experience with this material is that often I am pressed
towards the end and usually end a one-year course with Chapter 24, leaving
most Galois theory for the next year. For those lecturers who would treat
Galois theory in the first year, this raises the question of how one could push
ahead in order to reach Chapters 25, 26, and 27. First of all, the addenda
to Chapters 3 and 12 are not needed elsewhere in the text. Another two
possible chapters that could be curtailed are Chapters 17 and 19. The
second half of Chapter 17 contains a second proof of the basic theorem
that Euclidean rings are Unique Factorization Domains (UFDs), by means
of translating all the relevant concepts to ideals and then replacing the
Euclidean degree function by considerations about ideals. Although one
sees more generally that all principal ideal domains (PIDs) are UFDs, this
generalization is nominal, since the motivating examples of PIDs are all
Euclidean; indeed it is quite difficult to come up with an example of a UFD
that is not Euclidean. Nevertheless, I would advocate a full presentation
of this material for two reasons: It serves as an introduction to Noetherian
rings and their techniques, one of the important advances in the early part
of the twentieth century; and the methods are more elegant, belonging
more intrinsically to the algebraic structure of rings, and indeed, ideals
have taken on an even more important role in ring theory than has unique
factorization.

The material of Chapter 19 might perhaps be more vulnerable to the
red pen, since its interest is partly historical. However, I feel that every so
often the lecturer should step back and let the students see what can be
reaped from the theories that they have labored so hard to master. For
many years number theory was considered the epitome of human inquiry,
attracting the attention of some of the greatest intellects in history, and
the theories described in this book were motivated largely by the material
in Chapter 19.

Concerning the two appendices at the end of these notes, I feel uneasy
relegating the transcendence of π to a corner unlikely to be reached in most
courses, but classroom experience indicates that this proof is perhaps the hardest of all for the students to digest, and perhaps the effort could be put to better purpose by explaining Galois theory more completely.

Appendix B is a luxury, but it ties up several loose ends; and it intrigues me to see how far one can get in noncommutative algebra (including most of the Skolem-Noether theorem) simply by pushing forward some of the basic techniques of elementary abstract algebra.

Although several of the exercises are more or less routine applications of the theorems, most exercises are extensions of the text, which usually require substantially more time to solve. A few exercises are intended to lead the reader to anticipate an upcoming topic; these exercises are usually difficult where they are presented, but become much easier later in the course.

Of course the traditional role of exercises in a course is to provide more-or-less routine applications of the main results, for the student's edification and also as possible material for examinations. These are provided as the Review Exercises at the end of the book.

A note about proofs: In an elementary course one takes considerable care to find the “best” proof. The usual criterion in a mathematics book is the length of the proof (the shorter the better). However, several standard proofs involve sleight of hand, pulling some computation out of the thin air, so, in an effort to explain the underlying ideas, I have turned to the criterion, “Which proof is easiest to remember?” For the reader's edification and amusement, several of the magic proofs have been put into the exercises.

One dilemma in a course in algebra is deciding where to introduce matrices. Although various subsets of matrices provide some of the most important examples of groups, the full set of \( n \times n \) matrices over a field has the structure of a “ring,” which usually is not defined until after groups have been studied for several months. (Indeed, we shall define a group in the first chapter.) Thus, strictly speaking, groups arising from matrices should not be introduced until far into the course; however, I did not want to wait to bring in such an important example. As a compromise, I have assumed that the reader is familiar with the definition and basic properties of matrices.

I would like to express my gratitude to my colleagues Steve Shnider and Shalom Feigelstock for suggestions on improving the exposition, to Boaz Saban and Miriam Rosset for spotting errors in the draft version, and to my helpmate Rachel Rowen for sharing her expertise in computer typesetting.
TABLE OF PRINCIPAL NOTATION

NOTE: Similar notation might depend on the context, e.g. group or ring.

<table>
<thead>
<tr>
<th>Prerequisites</th>
<th>( \varphi(n) )</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A \subseteq B; \ A \subseteq B )</td>
<td>( \text{GL}(n, F) )</td>
<td>4</td>
</tr>
<tr>
<td>( \emptyset )</td>
<td>( A(S) )</td>
<td>4</td>
</tr>
<tr>
<td>( N )</td>
<td>( S_n )</td>
<td>4</td>
</tr>
<tr>
<td>( A \cap B; \ A \cup B )</td>
<td>( \ell_s() )</td>
<td>5</td>
</tr>
<tr>
<td>( \cap_{i \in I} A_i \cup_{i \in I} A_i )</td>
<td>( \text{gcd}(a, b) )</td>
<td>5</td>
</tr>
<tr>
<td>( f : A \to B )</td>
<td>( H \leq G; \ H &lt; G )</td>
<td>8</td>
</tr>
<tr>
<td>( f(a); f(A) )</td>
<td>( (a) )</td>
<td>9</td>
</tr>
<tr>
<td>( 1:1 )</td>
<td>( o(a) )</td>
<td>10</td>
</tr>
<tr>
<td>( f^{-1} )</td>
<td>( Hg )</td>
<td>11</td>
</tr>
<tr>
<td>( A \times B )</td>
<td>( [G : H] )</td>
<td>11</td>
</tr>
<tr>
<td>( [s] )</td>
<td>( \mu(n) )</td>
<td>14</td>
</tr>
<tr>
<td>( A/\sim )</td>
<td>( \text{lcm} )</td>
<td>16</td>
</tr>
<tr>
<td>( Z )</td>
<td>( \exp(G) )</td>
<td>16</td>
</tr>
<tr>
<td>( \mathbb{Q}, \mathbb{R}, \mathbb{C} )</td>
<td>( \begin{pmatrix} 1 &amp; 2 &amp; \cdots &amp; n \ \pi_1 &amp; \pi_2 &amp; \cdots &amp; \pi_n \end{pmatrix} )</td>
<td>17</td>
</tr>
<tr>
<td>( \mathbb{Z} )</td>
<td>( \begin{pmatrix} i_1 &amp; i_2 &amp; \cdots &amp; i_t \end{pmatrix} )</td>
<td>18</td>
</tr>
<tr>
<td>( m\mathbb{Z} )</td>
<td>( HK )</td>
<td>19</td>
</tr>
<tr>
<td>( M_n(F) )</td>
<td>( \text{SL}(n, F) )</td>
<td>20</td>
</tr>
<tr>
<td>( e_{ij} )</td>
<td>( O(n, F), \text{SO}(n, F), \text{Sp}(n, F) )</td>
<td>21</td>
</tr>
<tr>
<td>( \ker \varphi )</td>
<td>( \text{ker} \varphi )</td>
<td>25</td>
</tr>
<tr>
<td>( (M, \cdot, e) )</td>
<td>( \varphi</td>
<td>_H )</td>
</tr>
<tr>
<td>( e )</td>
<td>( N &lt; G )</td>
<td>29</td>
</tr>
<tr>
<td>(</td>
<td>G</td>
<td>)</td>
</tr>
<tr>
<td>( a^{-1} )</td>
<td>( A_n )</td>
<td>36</td>
</tr>
<tr>
<td>( (\mathbb{Z}, +) )</td>
<td>( H \times K )</td>
<td>40</td>
</tr>
<tr>
<td>( \text{adj}(A) )</td>
<td>( A(M) )</td>
<td>46</td>
</tr>
<tr>
<td>( \det A )</td>
<td>( D_n )</td>
<td>59</td>
</tr>
<tr>
<td>( \text{Map}(S, S) )</td>
<td>( Q )</td>
<td>61</td>
</tr>
<tr>
<td>( \text{Unit}(M) )</td>
<td>( G(S) )</td>
<td>62</td>
</tr>
<tr>
<td>( \text{Euler}(n) )</td>
<td>( Z(G) )</td>
<td>73</td>
</tr>
</tbody>
</table>
### Table of Principal Notations

<table>
<thead>
<tr>
<th>Notation</th>
<th>Page</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C(a)$</td>
<td>73</td>
<td>$\mathbb{Z}[i]$</td>
<td>139</td>
</tr>
<tr>
<td>$G_s$</td>
<td>75</td>
<td>$N(y)$</td>
<td>139</td>
</tr>
<tr>
<td>$G \cdot s$</td>
<td>75</td>
<td>$c(f)$</td>
<td>146</td>
</tr>
<tr>
<td>$HgK$</td>
<td>77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N(A)$</td>
<td>79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$m_p$</td>
<td>80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{Aut}(G)$</td>
<td>89</td>
<td>$(f)$</td>
<td>159</td>
</tr>
<tr>
<td>$\text{Inn Aut}(G)$</td>
<td>89</td>
<td>$f_a$</td>
<td>159</td>
</tr>
<tr>
<td>$H \text{ char } G$</td>
<td>89</td>
<td>$[V : F]$</td>
<td>160</td>
</tr>
<tr>
<td>$Z_i(G)$</td>
<td>90</td>
<td>$F[a]$</td>
<td>161</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$F[a,b]$</td>
<td>163</td>
</tr>
</tbody>
</table>

### Part II. Rings

<table>
<thead>
<tr>
<th>Notation</th>
<th>Page</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Ra$</td>
<td>100</td>
<td>$\text{Hom}_F(V,V)$</td>
<td>166</td>
</tr>
<tr>
<td>$L_1 + L_2$</td>
<td>100</td>
<td>$\hat{a}$</td>
<td>167</td>
</tr>
<tr>
<td>$A \triangleleft R$</td>
<td>104</td>
<td>$C(P;r)$</td>
<td>167</td>
</tr>
<tr>
<td>$R/A$</td>
<td>105</td>
<td>$n_{E/K,L}$</td>
<td>180</td>
</tr>
<tr>
<td>$\mathbb{Z}$</td>
<td>109</td>
<td>$\text{char}(F)$</td>
<td>182</td>
</tr>
<tr>
<td>$\hat{R}[S^{-1}], S^{-1}R$</td>
<td>110</td>
<td>$\text{GF}(n)$</td>
<td>189</td>
</tr>
<tr>
<td>$AS^{-1}$</td>
<td>111</td>
<td>$\text{Gal}(K/F)$</td>
<td>191</td>
</tr>
<tr>
<td>$\mathbb{Z}[\frac{1}{s}]$</td>
<td>112</td>
<td>$K^S$</td>
<td>192</td>
</tr>
<tr>
<td>$\mathbb{Z}_p$</td>
<td>112</td>
<td>$N_{K/F}(a)$</td>
<td>216</td>
</tr>
<tr>
<td>$R[x]$</td>
<td>114</td>
<td>$T_{K/F}(a)$</td>
<td>216</td>
</tr>
<tr>
<td>$a</td>
<td>b$</td>
<td>117</td>
<td>$\mathbb{H}$</td>
</tr>
<tr>
<td>$R[[x]]$</td>
<td>120</td>
<td>$\text{Cent}(R)$</td>
<td>224</td>
</tr>
<tr>
<td>$\exp(f)$</td>
<td>121</td>
<td>$N(y + zj)$</td>
<td>226</td>
</tr>
<tr>
<td>$\log(f)$</td>
<td>121</td>
<td>$C_R(S)$</td>
<td>229</td>
</tr>
</tbody>
</table>

### Part III. Fields
Mathematics evolves around the study of various sets, so let us review briefly some foundations about sets. We write $A \subseteq B$ to denote that the set $A$ is a subset of $B$, i.e., every element of $A$ is contained in $B$; if, furthermore, $A \neq B$, we may write $A \subset B$. The empty set is denoted as $\emptyset$. A set $S$ can be described either by a list of its elements or by means of a larger set. For example, perhaps the most fundamental set in algebra is $\mathbb{N} = \{0, 1, 2, \ldots \}$ (the set of natural numbers); the set of even natural numbers can be designated either as $\{0, 2, 4, \ldots \}$ or $\{n \in \mathbb{N} : n \text{ is even} \}$ or simply as $\{2n : n \in \mathbb{N} \}$ (which will be condensed even further to $2\mathbb{N}$).

Given an arbitrary set $S$ and $A, B \subseteq S$, we write $A \cap B$ for $\{s \in S : s \in A \text{ and } s \in B\}$; we say $A$ and $B$ are disjoint if $A \cap B = \emptyset$. Likewise we write $A \cup B$ for $\{s \in S : s \in A \text{ or } s \in B\}$. One can think of $A \cap B$ as the largest set contained in both $A$ and $B$, in the sense that any set contained in both $A$ and $B$ must also be contained in $A \cap B$. Similarly $A \cup B$ is the smallest subset of $S$ containing both $A$ and $B$, in the sense that any subset of $S$ containing both $A$ and $B$ must contain $A \cup B$. We shall use “the largest” and “the smallest” in this sense, throughout the text. (On the other hand, “maximal” is used in the slightly weaker sense that there is nothing larger. For example, if we are given the subsets $A_1 = \{0, 1, 3\}$, $A_2 = \{0, 2, 3\}$, and $A_3 = \{0, 3\}$ of $\mathbb{N}$, we would say $A_1$ and $A_2$ are maximal among these three subsets, although none of these is the largest.)

$\cup$ and $\cap$ satisfy the familiar axioms of associativity and distributivity for unions and intersections of subsets $A, B, C$ of a set $S$:
Similarly, given an arbitrary collection of subsets \( A_i : i \in I \) of \( A \), we define their intersection \( \bigcap_{i \in I} A_i \) or union \( \bigcup_{i \in I} A_i \).

A function \( f \) from a set \( A \) to a set \( B \) is denoted as \( f : A \to B \), and is also called a map. (We stipulate that the image \( f(a) \) is defined for each \( a \) in \( A \).) Maps between sets turn out to be as important as the sets themselves. Given a map \( f : A \to B \) we write \( f(A) \) for \( \{ f(a) : a \in A \} \subseteq B \); we say \( f \) is onto if \( f(A) = B \). On the other hand, we say \( f \) is one-to-one, written 1:1, if for any elements \( a_1 \neq a_2 \) in \( A \) we have \( f(a_1) \neq f(a_2) \) in \( B \). (It is usually easier to verify the equivalent formulation that \( f(a_1) = f(a_2) \) implies \( a_1 = a_2 \).) A map which is 1:1 and onto is called a 1:1 correspondence, or bijection.

Given a map \( f : A \to B \) and \( b \in B \), we define the inverse image \( f^{-1}(b) \) as \( \{ a \in A : f(a) = b \} \). Analogously for \( S \subseteq B \), we define \( f^{-1}(S) \) as \( \{ a \in A : f(a) \in S \} \). Of course \( f^{-1} \) need not be a map, but if \( f \) is a bijection, then \( f^{-1} \) is a map and indeed is also a bijection. For maps \( f : A \to B \) and \( g : B \to C \), the composite \( h = g \circ f : A \to C \) is defined by \( h(a) = g(f(a)) \).

The Cartesian product \( A \times B \) of sets \( A, B \) is the set of ordered pairs \( \{(a, b) : a \in A, b \in B \} \). A binary relation \( \sim \) on a set \( A \) is defined formally as a subset \( R \) of \( A \times A \); usually we simply write \( a \sim b \) if \( (a, b) \in R \). For example the relation \( "a < b" \) is defined on \( \mathbb{N} \) as \( \{(0, 1), (0, 2), (1, 2), \ldots \} \); the relation "equality" in \( A \) is defined as \( \{(a, a) : a \in A \} \). Generalizing "equality," one says a relation \( \sim \) is an equivalence if it satisfies the following properties:

1. ("reflexivity") \( a \sim a \) for all \( a \) in \( A \);
2. ("symmetry") \( a \sim b \) implies \( b \sim a \);
3. ("transitivity") \( a \sim b \) and \( b \sim c \) imply \( a \sim c \).

For example parallelism of lines in the Euclidean plane is an equivalence relation (provided one says a line is parallel to itself.)

Given an equivalence \( \sim \) on \( A \), we define the "equivalence class" \([s]\) of any element \( s \) of \( A \) to be \( \{ a \in A : a \sim s \} \). Note that any element belongs to its own equivalence class, by reflexivity, so \( A \) is the union of various equivalence classes. On the other hand, two equivalence classes either coincide or are
disjoint; hence, $A$ is the disjoint union of its equivalence classes. The set of equivalence classes of $A$ is denoted as $A/\sim$.

Conversely, we define a “partition” of $A$ to be a set of disjoint subsets whose union is $A$; any partition defines the equivalence relation defined by stipulating that $a \sim b$ iff they lie in the same subset in the partition. (By "iff" we mean "if and only if.")

In algebra one often introduces new algebraic structures by means of equivalence classes. The reader may have encountered the following construction of the integers $\mathbb{Z}$, from $\mathbb{N}$: Define the equivalence $\sim$ on $\mathbb{N} \times \mathbb{N}$ by

$$(a_1, a_2) \sim (b_1, b_2) \iff a_1 + b_2 = a_2 + b_1.$$  

For example, $(4, 7) \sim (5, 8)$. We define $\mathbb{Z}$ formally as the set $\mathbb{N} \times \mathbb{N}/\sim$. Intuitively the equivalence class of $(a_1, a_2)$ has been identified with the integer $a_1 - a_2$, since (a posteriori)

$$(a_1, a_2) \sim (b_1, b_2) \iff a_1 - a_2 = b_1 - b_2.$$  

One can define algebraic operations on the equivalence classes, by means of representatives from the classes. For example one defines addition on $\mathbb{Z}$ by

$$[(a_1, a_2)] + [(b_1, b_2)] = [(a_1 + b_1, a_2 + b_2)].$$

The equality $-3 + 4 = 1$ can be expressed as $[(1, 4)] + [(6, 2)] = [(7, 6)]$. This poses a new difficulty: one has to show that the outcome is independent of the particular representatives we chose in the equivalence classes. For example, if we used $(2,5)$ instead of $(1,4)$ we would wind up with $(8,7)$ instead of $(7,6)$, which is all right since they are equivalent. This condition is called "well-defined" and is one of the nuisances that we must contend with in many constructions.

Another example of construction by means of conjugate classes is given following Definition 2, below.

We shall assume such familiar properties of $\mathbb{N}$ as the unique factorization of any whole number $> 1$ into primes. In fact this follows from the theory to be developed in Chapter 16, but it is convenient to use these properties earlier, to develop examples in group theory. The reader is assumed to be familiar with the method of proof by mathematical induction, used in proving the bulk of our theorems, cf. Exercise 1.

One fundamental property of $\mathbb{N}$ is that every nonempty subset $S$ of $\mathbb{N}$ contains a unique smallest element. This can be proved by mathematical induction, and actually provides an alternate version of induction that will be used quite frequently in the theory of finite groups.
Much of mathematics involves the sets $\mathbb{Q}$ (the rational numbers), $\mathbb{R}$ (the real numbers), and $\mathbb{C}$ (the complex numbers), so we are led to study the properties these sets have in common. One can check easily that in each of these sets the following axioms hold with respect to addition and multiplication, for any elements $a, b,$ and $c$:

(F1) $(a + b) + c = a + (b + c)$
(F2) $a + 0 = 0 + a = a$
(F3) $a + b = b + a$
(F4) $a + (-a) = (-a) + a = 0$
(F5) $(ab)c = a(bc)$
(F6) $a1 = 1a = a$
(F7) $ab = ba$
(F8) $a\frac{1}{a} = \frac{1}{a}a = 1$ whenever $a \neq 0$
(F9) $a(b + c) = ab + ac$
(F9') $(b + c)a = ba + ca$

(Of course (F9') is superfluous here in view of (F7).) We are now in position to cross the threshold of abstract algebra — why not define an abstract entity satisfying (F1) through (F9)? Then we could examine its properties and apply the ensuing results to $\mathbb{Q}, \mathbb{R}$, and $\mathbb{C}$, thereby saving ourselves the trouble of studying each set separately. In order to do this we need more than just the set; we need to define "operations" corresponding to addition and multiplication. Accordingly, one defines a binary operation on a set $S$ to be a map $S \times S \rightarrow S$, i.e., it takes an ordered pair of elements of $S$ and assigns an element of $S$ as the answer. For example, the binary operation $+$ on $\mathbb{Q}$ takes $(4, -7.5)$ to their sum, $-3.5$.

Please note that each binary operation is assumed to be defined for every pair of elements $(s_1, s_2)$ of $S \times S$; in some treatments this property is isolated as a separate axiom (called closure), but here it is subsumed in the definition.

Definition 1. A field is a set $F$, together with binary operations (denoted as $+$ and $\cdot$ and called addition and multiplication) and designated elements $0, 1 \in F$, such that:

(i) properties (F1), (F2), (F3), (F5), (F6), (F7), and (F9) hold for all $a, b, c$ in $F$;
(ii) For any element $a$ in $F$ there is a unique element denoted $(-a)$ satisfying (F4);
(iii) For any $a \neq 0$ in $F$ there is a unique element denoted $\frac{1}{a}$, or $a^{-1}$, satisfying (F8).

Having defined a field we might look for more examples. The advantage
of arranging the properties as in Definition 1 is that the properties in (i) pass at once to subsets. On the other hand, $\mathbb{Z}$, with the usual addition and multiplication, satisfies (i) and (ii) but not (iii), since $\frac{1}{2}$ is not in $\mathbb{Z}; \mathbb{N}$ satisfies (i) but neither (ii) nor (iii). Nevertheless, one can show that $\mathbb{Q}$ as a set is in 1:1 correspondence with $\mathbb{Z}$ (as well as with $\mathbb{N}$). This points to the principle that operations accompanying a set are quintessential to the algebraic theory; the aggregate of set, operations, and designated elements, is called the (algebraic) structure. Here is an example that is very important in number theory, as we shall see.

**Definition 2.** Given $m \geq 1$, define $\mathbb{Z}_m = \{0, 1, \ldots, m - 1\}$, provided with the addition and multiplication of "clock arithmetic," i.e., addition and multiplication taken modulo $m$, also written (mod $m$).

To understand $\mathbb{Z}_m$ more fully, we need a formal description. Define $m\mathbb{Z} = \{mn : n \in \mathbb{Z}\} = \{\ldots, -m, 0, m, 2m, \ldots\}$, the multiples of $m$. We define the equivalence $a \equiv b$ (mod $m$) (for $a, b \in \mathbb{Z}$), read "a is congruent to b modulo $m\)," iff $a - b \in m\mathbb{Z}$. Let us write $[a]$ for the equivalence class containing $a$; thus $[a] = [a + m] = [a - m] = \ldots$. In this way $\mathbb{Z}$ is partitioned into $m$ equivalence classes $[0], [1], \ldots, [m - 1]$. Now define addition and multiplication by $[a] + [b] = [a + b]$ and $[a][b] = [ab]$. One must check that this definition is well-defined, cf. Exercise 1, but then it is easy to transfer properties from $\mathbb{Z}$ to $\mathbb{Z}_m$, merely by writing brackets wherever appropriate.

In this manner we see $\mathbb{Z}_m$ satisfies (i) and (ii) of Definition 1 (taking $-[a]$ to be $[-a] = [m - a]$), but again (iii) may fail. Indeed the reader can easily check that $[2]^{-1}$ does not exist in $\mathbb{Z}_4$, leading us to the question: "For what $m$ is $\mathbb{Z}_m$ a field?" We shall soon see (Corollary 1.13) that $\mathbb{Z}_m$ is a field iff $m$ is a prime number; for example $[2]^{-1} = [6]$ in $\mathbb{Z}_{11}$. Of course the field $\mathbb{Z}_p$ (for $p$ prime) has $p$ elements, and the existence of finite fields will be important in our study of groups. Although $\mathbb{Z}_p$ suffices for many applications, the reader should be apprised that there is a field having precisely $n$ elements, if and only if $n$ is a power of a prime number; this and other facts about finite fields will be proved much later, in Chapters 24 and 26.

As mentioned earlier, the reader is assumed to be familiar with matrices over $\mathbb{R}$. Analogously, the set of $n \times n$ matrices with entries in an arbitrary field $F$ also is endowed with the analogous matrix addition and matrix multiplication and is denoted as $M_n(F)$. Let us define the $n \times n$ matrix unit $e_{ij}$ of $M_n(F)$ to be the matrix with 1 in the $i - j$ position and 0 everywhere else. Then the following properties are satisfied:
Prerequisites

(i) \( e_{11} + \cdots + e_{nn} = 1; \)
(ii) \( e_{ij}e_{jk} = e_{ik}; \)
(iii) \( e_{ij}e_{uv} = 0 \) if \( j \neq u. \)

Any element of \( M_n(R) \) can be written uniquely in the form \( \sum_{i,j=1}^{n} r_{ij}e_{ij} \)
for \( r_{ij} \) in \( R. \) Matric units are a significant aid for computing with matrices.

Exercises

1. Prove that addition in \( \mathbb{Z} \) defined formally in the text actually is
well-defined. Similarly, define multiplication in \( \mathbb{Z}, \) and verify (i),(ii)
of Definition 1, by induction. (This exercise becomes rather boring
after a while, once one gets the hang of what is going on.)
2. Prove in any field \( F \) that \( a, b \neq 0 \) implies \( ab \neq 0. \)
3. Define \( \mathbb{Q}(\sqrt{1}) \) to be \( \{a + b\sqrt{1} : a, b \in \mathbb{Q}\}, \) which can be viewed as
a subset of \( \mathbb{C}. \) Prove \( \mathbb{Q}(\sqrt{1}) \) is a field. (Hint: the only challenging
part is the multiplicative inverse.)
4. Define \( \mathbb{Q}(\sqrt{2}) \) to be \( \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}, \) viewed as a subset of \( \mathbb{R}. \)
Prove this is a field, under the natural addition and multiplication:

\[
(a + b\sqrt{2}) + (c + d\sqrt{2}) = (a + c) + (b + d)\sqrt{2}
\]
\[
(a + b\sqrt{2})(c + d\sqrt{2}) = (ac + 2bd) + (ad + bc)\sqrt{2}
\]

Although \( \mathbb{Q}(\sqrt{2}) \) is clearly in 1:1 correspondence with \( \mathbb{Q}(\sqrt{-1}), \)
their algebraic structures as fields are different. Explain.
PART I — GROUPS

In this course, abstract algebra focuses on sets endowed with "algebraic structure," and axioms describing how the elements behave with respect to the given operations. The operations of basic concern to us are multiplication and/or addition, in various contexts. In this spirit, we should start with some observations about Definition 0.1:

1. Property (iii) (axiom (F8)) relies only on multiplication, not addition;
2. Properties (F5) through (F7) are the multiplicative analogs of (F1) through (F3), and (F8) is the analog of (F4), except that we delete \{0\}.

Thus it makes sense to isolate properties (F1)–(F3) with an eye to include (F4) shortly thereafter; then using the multiplicative analog (deleting \{0\}) we would understand (F5) through (F8). This will lead us first to the definition of "monoid" and then to "group." Abstract groups appear in almost every branch of mathematics and physics, as well as in other sciences and even in many aspects of day-to-day life (such as telling time). Our object is to develop enough of the theory of groups to enable us to answer basic questions concerning their structure and to familiarize the reader with certain groups that he or she is likely to encounter repeatedly in the future. Various interesting classes of groups are easier to study than groups in general. We shall obtain rather decisive theorems concerning Abelian groups; in Part III we shall need "solvable groups," a useful generalization of Abelian groups that is studied in Chapter 12.
Chapter 1. Monoids and Groups

Definition 1. A monoid \((M, \cdot, e)\) is a set \(M\) with a binary operation \(\cdot\) and a neutral element \(e\) (also called the identity) satisfying the following properties for all \(a, b, c\) in \(M\):

\[(M1)\] (associativity) \((ab)c = a(bc)\);
\[(M2)\] \(ae = ea = a\).

(Note that, as is customary in multiplicative notation, we write \(ab\) instead of \(a \cdot b\). On the other hand, often \(\cdot\) will be +, which is never suppressed in the notation.)

Note that the neutral element is uniquely determined by the operation; indeed if \(e\) and \(e'\) are neutral elements then \(e' = e'e = e\). Thus we usually delete "\(e\)" from the notation; for example \((\mathbb{Z}, +)\) can only mean \((\mathbb{Z}, +, 0)\). We shall often delete the operation also, and write \(M\) for the monoid \((M, \cdot, e)\).

We say elements \(a\) and \(b\) in \(M\) commute if \(ab = ba\). The monoid \(M\) is commutative if \(ab = ba\) for all \(a, b\) in \(M\). When the set \(M\) is finite, we say \(M\) is finite and write \(|M|\) for the number of elements of \(M\), called the order of \(M\). The particular operation plays a crucial role in the structure of the monoid. For example \((\mathbb{Z}_2, +, 0)\) and \((\mathbb{Z}_2, \cdot, 1)\) are finite monoids each of order 2, but whose structures are not analogous. (In the first case the neutral element is 0 and satisfies 0 = 1 + 1, but in the second case the neutral element 1 satisfies 1 \(\neq 0\).)

Note. Associativity enables us to write products without parentheses, without ambiguity, cf. Exercise 1. Let us now turn to the key notion, bringing in (F8).

Definition 2. An element \(a\) of \(M\) is left invertible (resp. right invertible) if there is \(b\) in \(M\) such that \(ba = e\) (resp. \(ab = e\)); \(a\) is invertible if \(a\) is left and right invertible. A group is a monoid in which every element is invertible.

Suppose \(a\) is invertible. Then there are elements \(b, b'\) such that \(ba = e = ab'\). We shall now see that \(b = b'\). Indeed \(b = b(ab') = (ba)b' = b'\). Thus \(a\) has a unique element \(b\) such that \(ba = e = ab\), which is called the inverse of \(a\), denoted \(a^{-1}\).

A commutative group is called Abelian, after the Norwegian mathematician Abel.

Examples of Groups and Monoids
Let us start with some basic examples.

(1) If \(F\) is a field then \((F, +)\) and \((F \setminus \{0\}, \cdot)\) are Abelian groups, seen
by axioms (F1) through (F4) and (F5) through (F8) respectively, given in the prerequisites. (See exercise 0.2 to clean up a sticky point.)

(2) \((\mathbb{Z}, +)\) is an Abelian group, but \((\mathbb{Z} \setminus \{0\}, \cdot)\) is a commutative monoid that is not a group.

(3) \((\mathbb{Z}_m, +)\) is an Abelian group of order \(m\); \((\mathbb{Z}_m, \cdot)\) is a commutative monoid.

(4) (See prerequisites.) For any field \(F\) and any \(n\), write \(M_n(F)\) for the set of \(n \times n\) matrices with entries in an arbitrary field \(F\), endowed with the usual matrix addition and multiplication. \((M_n(F), +)\) is an Abelian group. \((M_n(F), \cdot, 1)\) is a monoid (but not a group) that is not commutative for \(n > 1\), since the matrices \(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\) and \(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\) do not commute. (These matrices are not invertible, since each has rank 1. Two invertible \(2 \times 2\) matrices that do not commute are \(\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\) and \(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\).) One can define adjoints and determinants in the usual way; then \(A \text{adj}(A) = \det A\) for any \(A\) in \(M_n(F)\), so \(A^{-1} = \text{adj}(A)/\det A\) exists whenever \(\det A \neq 0\).

Aside. This argument also shows \(AB = 1\) implies \(BA = 1\), for any two matrices \(A, B\), cf. Exercise 8.

(5) Suppose \(S\) is a set. Then \(\{\text{functions } f : S \to S\}\) form a monoid which we denote as \(\text{Map}(S, S)\), whose operation is composition; the neutral element is the identity map \(1_S\) defined by \(1_S(s) = s\) for all \(s\) in \(S\).

(6) The trivial group \(\{e\}\) has multiplication given by \(ee = e\).

Note. Although the operation of a group often is +, it is customary to choose multiplicative notation when studying groups in general, since "+" biases us toward commutativity.

One can extract a group from an arbitrary monoid. To understand this procedure let us first examine invertibility.

Remark 3. Suppose \(a, b\) are invertible. From the equation \(a^{-1}a = e = aa^{-1}\) we also see \(a = (a^{-1})^{-1}\). Furthermore \(abb^{-1}a^{-1} = e = b^{-1}a^{-1}ab\), implying \((ab)^{-1} = b^{-1}a^{-1}\).

Given a monoid \(M\), write \(\text{Unit}(M)\) for \(\{\text{invertible elements of } M\}\).

**Proposition 4.** \(\text{Unit}(M)\) is a group.

**Proof.** Associativity in \(\text{Unit}(M)\) follows at once from associativity in \(M\). But \(e \in \text{Unit}(M)\), so Remark 3 implies \(\text{Unit}(M)\) is closed under multiplication and is thus a monoid. Moreover if \(a \in \text{Unit}(M)\), then \(a^{-1} \in \text{Unit}(M)\) so \(\text{Unit}(M)\) is a group. □
Example 5. Applying Proposition 4 to Examples (1) through (5) yield the following useful groups:

1. $\text{Unit}(F, \cdot) = F \setminus \{0\}$ for any field $F$.
2. $\text{Unit}(\mathbb{Z}, \cdot) = \{\pm 1\}$.
3. $\text{Unit}(\mathbb{Z}_m, \cdot) = \{a : 1 \leq a < m, \text{and } a \text{ is invertible } \text{mod } m\}$. This is called the Euler group, denoted $\text{Euler}(m)$. For example, $\text{Euler}(6) = \{1,5\}$; $\text{Euler}(7) = \{1,2,3,4,5,6\}$, and $\text{Euler}(8) = \{1,3,5,7\}$. The order of $\text{Euler}(m)$ is called the Euler number $\varphi(m)$. Thus $\varphi(3) = \varphi(4) = \varphi(6) = 2$, and $\varphi(5) = \varphi(8) = 4$.
4. $\text{Unit}(M_n(F), \cdot)$ is the group of regular $n \times n$ matrices over $F$, called the general linear group and denoted $\text{GL}(n,F)$.

Let us pause for a moment, to assert that $\text{GL}(n,F)$ is perhaps the most important group in mathematics, since it can be interpreted as the group of invertible linear transformations of an $n$-dimensional vector space over the field $F$; as such, it has fundamental significance in geometry and in physics. (This identification also enables us to prove various properties of $\text{GL}(n,F)$, cf. Exercise 6.)

Other “geometrical” groups can be defined similarly in terms of various kinds of linear transformations, cf. Chapter 2 and Exercises 12.9ff, 12.15ff. Nevertheless, the focal role in these notes is played by the next example.

5. $\text{Unit}(\text{Map}(S, S))$ is denoted as $A(S)$, the 1:1 onto maps from $S$ to $S$. In the special case $S = \{1, \ldots, n\}$ we denote $A(S)$ as $S_n$, the group of permutations of $n$ symbols. $S_n$, often called the symmetric group, plays a special role in finite group theory and will be used throughout as our main example. The reader is urged to trace its development via the index.

When Is a Monoid a Group?

We want to explore the fundamental question of when a given monoid is already a group. One basic property of groups will become focal.

Remark 6. If $ab = ac$ with $a$ left invertible, then $b = c$. (Indeed, multiply by the left inverse of $a$.)

Accordingly we call a monoid (left) cancellative if it satisfies the property

$$ab = ac \text{ implies } b = c$$

for any elements $a, b, c$. An example of a cancellative monoid which is not a group is $(\mathbb{Z} \setminus \{0\}, \cdot)$. On the other hand we have

Theorem 7. Any finite cancellative monoid $M$ is a group.
Before proving Theorem 7, let us accumulate some facts.

**Lemma 8.** If every element of a monoid $M$ is right invertible, then $M$ is a group.

*Proof.* We need to show any element $a$ of $M$ is invertible. By hypothesis there is $b$ such that $ab = e$; likewise there is $c$ such that $bc = e$. But then $b$ is invertible, so $a = b^{-1}$, as noted after Definition 2, implying $b = a^{-1}$. □

**Fact 9.** Suppose $S$ is a finite set. A function $f : S \to S$ is 1:1 iff $f$ is onto.

This fact is called the "pigeonhole principle"; for if a letter carrier is to distribute 17 letters into 17 boxes, clearly each box receives a letter iff no box receives at least two letters. Of course the pigeonhole principle fails for infinite sets, as illustrated by the map $f : \mathbb{N} \to \mathbb{N}$ given by $f(n) = n + 1$ for all $n$ in $\mathbb{N}$ (which is 1:1 but not onto). The pigeonhole principle relates to monoids and cancellation via the following key observation.

**Proposition 10.** Suppose $M$ is a monoid. Given $s \in M$ define the left multiplication map $\ell_s : M \to M$ by $\ell_s(a) = sa$ for all $a$ in $M$.

(i) $\ell_s$ is onto iff $s$ is right invertible;

(ii) $\ell_s$ is 1:1 iff $sb \neq sc$ for all $b, c$ in $M$.

*Proof.* (i) ($\Rightarrow$) $e = \ell_s(s')$ for some $s'$ in $M$; thus $ss' = e$.

($\Leftarrow$) If $ss' = e$ then $a = ss'a = \ell_s(s'a)$ for any $a$ in $M$.

(ii) Self-evident. □

*Proof of Theorem 7.* By Proposition 10(ii) we see for each $s$ in $M$ that $\ell_s$ is 1:1 and thus onto by the pigeonhole principle. Hence each $s$ in $M$ is right invertible, by Proposition 10(i), so, by Lemma 8, $M$ is a group. □

**Remark 11.** The proof of Theorem 7 shows that, for any element $g$ of a group $G$, the left multiplication map $\ell_g : G \to G$ is 1:1 and onto. We shall return to this fact later.

Let us apply theorem 7 to divisibility and the Euler group.

**Remark 12.** Writing $\gcd(a, b)$ to denote the greatest common divisor of $a$ and $b$, we claim $\text{Euler}(m) = \{a \in \mathbb{N} : 1 \leq a < m \text{ and } \gcd(a, m) = 1\}$. Indeed, using unique factorization in $\mathbb{N}$, it is easy to identify the right-hand side with a cancellative monoid contained in $\mathbb{Z}_m$. (If $\gcd(a_1, m) = 1$ and $\gcd(a_2, m) = 1$, then $\gcd(a_1a_2, m) = 1$; also if $ab = ac$ in $\mathbb{Z}_m$, then $ab \equiv ac \pmod{m}$, implying $m$ divides $ab - ac = a(b - c)$, so $m$ divides $b - c$.) Hence the right-hand side is a group so by definition is contained in $\text{Euler}(m)$.

On the other hand any invertible element $a$ of $(\mathbb{Z}_m, \cdot)$ is relatively prime to $m$, for if $d \in \mathbb{N}$ divides both $a$ and $m$, then $a0 = 0 = \frac{a}{d}m = a \frac{m}{d}$. 


implying by remark 6 that \( \frac{m}{d} \equiv 0 \pmod{m} \), so \( d = 1 \). Thus we have proved equality. □

**Corollary 13.** A number is invertible \( \pmod{m} \) iff it is relatively prime to \( m \). In particular Euler(p) = \{1, \ldots, p - 1\} iff \( p \) is prime.

**Corollary 14.** \( \mathbb{Z}_p \) is a field, for any prime number \( p \).

**Corollary 15.** \( \{i, i + t, \ldots, i + (p - 1)t\} \) are all distinct \( \pmod{p} \), for \( p \) prime, whenever \( t \not\equiv 0 \pmod{p} \).

**Proof.** One may assume \( i = 0 \); but then the first term is 0, and the others \( \{t, \ldots, (p - 1)t\} \) are distinct and nonzero \( \pmod{p} \), seen by canceling \( t \) in Euler(p). □

**Corollary 16.** If \( \gcd(m, n) = 1 \), then there are \( a, b \) in \( \mathbb{Z} \) with \( am + bn = 1 \).

**Proof.** Let \( b = n^{-1} \) in Euler(m). Then \( m \) divides \( 1 - bn \), so \( 1 - bn = am \) for some \( a \) in \( \mathbb{Z} \). □

**Exercises**

1. Define \( a_1a_2a_3 = (a_1a_2)a_3 \) and, continuing in this way, define

\[
a_1 \ldots a_n = (a_1 \ldots a_{n-1})a_n.
\]

Prove that \((a_1 \ldots a_m)(a_{m+1} \ldots a_n) = a_1 \ldots a_n\) for any \( n \) and any \( m < n \). (Hint: Induction on \( m \).)

2. Define \( A + B = \{a + b : a \in A, b \in B\} \) for subsets \( A, B \) of \( \mathbb{Z} \). Compute \( 3\mathbb{Z} + 4\mathbb{Z}, 6\mathbb{Z} + 10\mathbb{Z} \), and in general \( a\mathbb{Z} + b\mathbb{Z} \).

3. Show if \( a_1 \equiv b_1 \) and \( a_2 \equiv b_2 \pmod{m} \) then \( a_1 + a_2 \equiv b_1 + b_2 \) and \( a_1a_2 \equiv b_1b_2 \). Conclude that addition and multiplication in \( \mathbb{Z}_m \) (as defined via equivalence classes) is well-defined.

4. If \( a^2 \equiv 1 \pmod{p} \) for \( p \) prime, then \( a \equiv \pm 1 \pmod{p} \). (Hint: Show that \( p \) divides \( a^2 - 1 = (a + 1)(a - 1) \).) What happens for \( p \) not prime, e.g., \( p = 15 \)? \( p = 21 \)?

5. Write down the Euler groups Euler(n) for all \( n \leq 10 \). For which values of \( n \) does Euler(n) have order 2? Show that these groups are the “same” in some sense (to be made precise in Chapter 4).

6. Suppose \( F \) is a finite field of \( q \) elements (for example, if \( q \) is prime one could have \( F = \mathbb{Z}_q \)). Then the group GL(n, F) has order

\[
(q^n - 1)(q^n - q) \ldots (q^n - q^{n-1})
\]

\[
=q^{n(n-1)/2}(q^n - 1)(q^{n-1} - 1) \ldots (q - 1).
\]
(Hint: One can define an arbitrary invertible linear transformation by sending a given base to any other base, so \(|\text{GL}(n, F)|\) is the number of different possible bases of the vector space \(F^n\) over \(F\) (where the order of the base elements is significant). The first vector in the base can be any nonzero vector and thus chosen \((q^n - 1)\) ways; the next vector can be chosen \((q^n - q)\) ways, and so forth.)

7. How many concrete examples of finite groups can you produce at this stage? What is the non-abelian group of smallest order?

8. Consider the property \(ab = 1\) implies \(ba = 1\). Show this property holds in \(M_n(F)\) (\(F\) a field) and in any finite monoid. However it fails in Map\((S, S)\) when \(S\) is an infinite set. (Hint: Define \(f: \mathbb{N} \to \mathbb{N}\) by \(f(n) = n + 1\).)

9. A **semigroup** is a set with a binary operation satisfying M1 of definition 1, but not necessarily M2. Any semigroup \(S\) can be made into a monoid, simply by adjoining a formal identity \(e\) and defining \(ea = ae = a\) for all \(a\) in \(S\). Nevertheless, semigroups (without 1) arise in several contexts, where the elements are not necessarily invertible. (For example, let \(S\) be the set of functions \(\mathbb{R} \to \mathbb{R}\) that are continuous in the neighborhood of a point.)

10. Inspired by the example in Exercise 9, call a semigroup \(S\) **inverse** if for each \(a\) in \(S\) there is \(b\) in \(S\) such that \(aba = a\) and \(bab = b\). Then \(b\) is unique and is called the **inverse** of \(a\). Prove the analog of Remark 3 for inverse semigroups. Inverse semigroups have commanded fair attention in research.

11. An inverse semigroup that is cancellable is a group.