Field Extensions and Galois Theory

ENCYCLOPEDIA OF MATHEMATICS and Its Applications

GIAN-CARLO ROTA, Editor
Department of Mathematics
Massachusetts Institute of Technology
Cambridge, Massachusetts

GIAN-CARLO ROTA, *Editor*
ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS

Volume		Section
1	LUIS A. SANTALÓ **Integral Geometry and Geometric Probability,** 1976 (2nd printing, with revisions, 1979)	Probability
2	GEORGE E. ANDREWS **The Theory of Partitions,** 1976 (2nd printing, 1981)	Number Theory
3	ROBERT J. McELIECE **The Theory of Information and Coding** A Mathematical Framework for Communication, 1977 (2nd printing, with revisions, 1979)	Probability
4	WILLARD MILLER, Jr. **Symmetry and Separation of Variables,** 1977	Special Functions
5	DAVID RUELLE **Thermodynamic Formalism** The Mathematical Structures of Classical Equilibrium Statistical Mechanics, 1978	Statistical Mechanics
6	HENRYK MINC **Permanents,** 1978	Linear Algebra
7	FRED S. ROBERTS **Measurement Theory** with Applications to Decisionmaking, Utility, and the Social Sciences, 1979	Mathematics and the Social Sciences
8	L. C. BIEDENHARN and J. D. LOUCK **Angular Momentum in Quantum Physics:** Theory and Application, 1981	Mathematics of Physics
9	L. C. BIEDENHARN and J. D. LOUCK **The Racah-Wigner Algebra in Quantum Theory**, 1981	Mathematics of Physics

GIAN-CARLO ROTA, *Editor*
ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS

Volume		Section
10	JOHN D. DOLLARD and CHARLES N. FRIEDMAN **Product Integration** with Application to Differential Equations, 1979	Analysis
11	WILLIAM B. JONES and W. J. THRON **Continued Fractions:** Analytic Theory and Applications, 1980	Analysis
12	NATHANIEL F. G. MARTIN and JAMES W. ENGLAND **Mathematical Theory of Entropy,** 1981	Real Variable
13	GEORGE A. BAKER, Jr. and PETER R. GRAVES-MORRIS **Padé Approximants, Part I Basic Theory,** 1981	Mathematics of Physics
14	GEORGE A. BAKER, Jr. and PETER R. GRAVES-MORRIS **Padé Approximants, Part II: Extensions and Applications,** 1981	Mathematics of Physics
15	E. C. BELTRAMETTI and G. CASSINELLI **The Logic of Quantum Mechanics,** 1981	Mathematics of Physics
16	G. D. JAMES and A. KERBER **The Representation Theory of the Symmetric Group,** 1981	Algebra
17	M. LOTHAIRE **Combinatorics on Words,** 1982	Algebra

GIAN-CARLO ROTA, *Editor*
ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS

Volume		Section
18	H. O. FATTORINI **The Cauchy Problem,** 1983	Analysis
19	G. G. LORENTZ, K. JETTER, and S. D. RIEMENSCHNEIDER **Birkhoff Interpolation,** 1983	Interpolation and Approximation
20	RUDOLF LIDL and HARALD NIEDERREITER **Finite Fields,** 1983	Algebra
21	WILLIAM T. TUTTE **Graph Theory,** 1984	Combinatorics
22	JULIO R. BASTIDA **Field Extensions and Galois Theory,** 1984	Algebra
23	JOHN R. CANNON **The One-Dimensional Heat Equation,** 1984	Analysis

Other volumes in preparation

GIAN-CARLO ROTA, *Editor*

ENCYCLOPEDIA OF MATHEMATICS AND ITS APPLICATIONS

Volume 22

Section: Algebra
P. M. Cohn and Roger Lyndon, *Section Editors*

Field Extensions and Galois Theory

Julio R. Bastida
Department of Mathematics
Florida Atlantic University
Boca Raton, Florida

With a Foreword by

Roger Lyndon
The University of Michigan
Ann Arbor, Michigan

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,
São Paulo, Delhi, Dubai, Tokyo, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521302425

First published 1984

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data

Bastida, Julio R.
Field extensions and Galois theory.

(Encyclopedia of mathematics and its applications;
v. 22)
Bibliography: p.
Includes index.
1. Field extensions (Mathematics). 2. Galois theory.
I. Title. II. Series.
QA247.B37 1984 512′.32 83-7160

ISBN 978-0-521-30242-5 Hardback
ISBN 978-0-521-17396-4 Paperback

A mi hijo,
Ricardo Antonio

Contents

Editor's Statement

A large body of mathematics consists of facts that can be presented and described much like any other natural phenomenon. These facts, at times explicitly brought out as theorems, at other times concealed within a proof, make up most of the applications of mathematics, and are the most likely to survive change of style and of interest.

This ENCYCLOPEDIA will attempt to present the factual body of all mathematics. Clarity of exposition, accessibility to the non-specialist, and a thorough bibliography are required of each author. Volumes will appear in no particular order, but will be organized into sections, each one comprising a recognizable branch of present-day mathematics. Numbers of volumes and sections will be reconsidered as times and needs change.

It is hoped that this enterprise will make mathematics more widely used where it is needed, and more accessible in fields in which it can be applied but where it has not yet penetrated because of insufficient information.

GIAN-CARLO ROTA

Foreword

Galois theory is often cited as the beginning of modern "abstract" algebra. The ancient problem of the algebraic solution of polynomial equations culminated, through the work of Ruffini, Abel, and others, in the ideas of Galois, who set forth systematically the connection between polynomial equations and their associated groups. This was the beginning of the systematic study of group theory, nurtured by Cauchy and Jordan to its flowering at the end of the last century. It can also be viewed as the beginning of algebraic number theory (although here other forces were also clearly at work), developed later in the century by Dedekind, Kronecker, Kummer, and others. It is primarily this number-theoretic line of development that is pursued in this book, where the emphasis is on fields, and only secondarily on their groups.

In addition to these two specific outgrowths of Galois's ideas, there came something much broader, perhaps the essence of Galois theory: the systematically developed connection between two seemingly unrelated subjects, here the theory of fields and that of groups. More specifically, but in the same line, is the idea of studying a mathematical object by its group of automorphisms, an idea emphasized especially in Klein's Erlanger Program, which has since been accepted as a powerful tool in a great variety of mathematical disciplines.

Apart from the historical importance of the Galois theory of fields, its intrinsic interest and beauty, and its more or less direct applications to

number theory, these many generalizations and their important applications give further compelling reasons for seeking an understanding of the theory in its classical form, as presented in this volume. The Galois theory of field extensions combines the esthetic appeal of a theory of nearly perfect beauty with the technical development and difficulty that reveal the depth of the theory and that make possible its great usefulness, primarily in algebraic number theory and related parts of algebraic geometry.

In this book Professor Bastida has set forth this classical theory, of field extensions and their Galois groups, with meticulous care and clarity. The treatment is self-contained, at a level accessible to a sufficiently well-motivated beginning graduate student, starting with the most elementary facts about fields and polynomials and proceeding painstakingly, never omitting precise definitions and illustrative examples and problems. The qualified reader will be able to progress rapidly, while securing a firm grasp of the fundamental concepts and of the important phenomena that arise in the theory of fields. Ultimately, the study of this book will provide an intuitively clear and logically exact familiarity with the basic facts of a comprehensive area in the theory of fields. The author has judiciously stopped short (except in exercises) of developing specialized topics important to the various applications of the theory, but we believe he has realized his aim of providing the reader with a sound foundation from which to embark on the study of these more specialized subjects.

This book, then, should serve first as an easily accessible and fully detailed exposition of the classical Galois theory of field extensions in its simplest and purest form; and second, as a solid foundation for and introduction to the study of more advanced topics involving the same concepts, especially in algebraic number theory and algebraic geometry.

We believe that Professor Bastida has offered the reader, for a minimum of effort, a direct path into an enchantingly beautiful and exceptionally useful subject.

ROGER LYNDON

Preface

Since its inception at the beginning of the nineteenth century, the theory of field extensions has been a very active area of algebra. Its vitality stems not only from the interesting problems generated by the theory itself, but also from its connections with number theory and algebraic geometry. In writing this book, our principal objective has been to make the general theory of field extensions accessible to any reader with a modest background in groups, rings, and vector spaces.

The book is divided into four chapters. In order to give a precise idea of the background that the reader is expected to possess, we have preceded the text by a section on prerequisites. Except for the initial remarks, in which we indicate the restrictions that will be imposed on the rings considered throughout our presentation, the reader should not be concerned with the contents of this section until explicit reference is made to them. The first chapter is devoted to the general facts on fields and polynomials required in the study of field extensions. Although most of these facts can be found in one or another of the references given in the section on prerequisites, we have attempted to facilitate the reader's task by having them collected and stated in a manner suitably adapted to our purposes.

The theory of field extensions is presented in the subsequent three chapters, which deal, respectively, with algebraic extensions, Galois theory,

and transcendental extensions. The chapter on algebraic extensions is of basic importance for the entire theory, and has to be thoroughly understood before proceeding further. The last two chapters, on the other hand, can be read independently of each other.

Chapters are divided into sections, and each section ends with a set of problems. The problems include routine exercises, suggest alternative proofs of various results, or develop topics not discussed in the text. We have refrained from identifying the more difficult, and as a rule, no hints are given for the solutions. A result stated in a problem is not used in the text, but it may be required for the solution of a later problem.

The choice of material was dictated by the dual objective of providing thorough coverage of each topic treated and of keeping the length of the book within reasonable bounds. We decided to include in the text the results that constitute the core of the general theory of field extensions. Those parts of the theory sufficiently developed to merit a book of their own have been left out entirely, and several specialized topics of considerable interest have been relegated to the problems. We have not attempted to discuss any serious applications of our subject to number theory or algebraic geometry, since doing this would have required the introduction of additional background material. However, as the reader cannot fail to notice, connections with number theory manifest themselves occasionally in the presentation.

We have included bibliographical notes at the end of each chapter. These will provide the reader with references to the works in which important contributions were first published, with easily available references on topics presented as problems and on alternative treatments of topics covered in the text, and with suggestions for further reading.

The reference list at the end of the book comprises mainly the works cited in the text and notes. The vast literature on field extensions and Galois theory and on their applications to number theory and algebraic geometry cannot be surveyed, even superficially, within the confines of a few pages. To get a good idea of the present state of the literature, the reader may consult the pertinent sections of *Mathematical Reviews*, the review journal of the American Mathematical Society.

It is with the deepest gratitude and respect that we acknowledge the help given to us by Professor Harley Flanders, without which this book could not have been written. He read the manuscript and made very substantive suggestions on both content and style; offered us unrestricted access to his notes on field extensions; discussed proofs, examples, and problems with us; and never betrayed the slightest impatience in dealing with us during the four-year period that we worked on this book.

We would also like to express our sincere appreciation to Professor Gian-Carlo Rota, for his kind invitation to write a volume for the *Encyclo-*

pedia; to Professors Paul M. Cohn and Roger C. Lyndon, for their valuable suggestions; to Professors Tomás P. Schonbek and Scott H. Demsky, for their help with the bibliographical material; to my students Lynn Garrett and Jaleh Owliaei, for their comments; to Ruth Ebel and especially Rita Pelava, for their efficient typing; and to my colleagues at Florida Atlantic University, for their constant encouragement.

JULIO R. BASTIDA
Boca Raton, Florida

Historical Introduction

Problems of geometric construction appeared early in the history of mathematics. They were first considered by the Greek mathematicians of the fifth century B.C. Only two instruments—an unmarked ruler and a compass—were permitted in these constructions. Although many such constructions could be performed, others eluded the efforts of these mathematicians. Four famous problems from the period that remained unsolved for a long time are the following: doubling the cube, which consists of constructing a cube whose volume is twice that of a given cube; trisecting the angle; squaring the circle, which consists of constructing a square whose area is that of a given circle; and constructing regular polygons.

At the end of the eighteenth century, when it was observed that questions on geometric constructions can be translated into questions on fields, a breakthrough finally occurred. The 19-year-old Gauss [2: art. 365] proved in 1796 that the regular 17-sided polygon is constructible. A few years later, Gauss [2: art. 365, 366] stated necessary and sufficient conditions for the constructibility of the regular n-sided polygon. He gave a proof only of the sufficiency, and claimed to have a proof of the necessity; the latter was first given by Wantzel [1] in 1837. In his investigations, Gauss introduced and used a number of concepts that became of central importance in subsequent developments. A by-product of the works of Gauss and Wantzel on regular polygons was a proof that an arbitrary angle cannot be

trisected. The proof of the impossibility of doubling the cube is more elementary, but its discovery is difficult to trace. As to the remaining problem, it was realized that the proof of the impossibility of squaring the circle depended on knowing that the number π is transcendental; this missing ingredient was supplied in 1882 by Lindemann [1], who used analytic techniques to settle one of the more fascinating questions in this area of mathematics.

The general theory of fields evolved during the last half of the nineteenth century, when the algebraists made significant advances in the study of algebraic numbers and algebraic functions. The first systematic exposition of the theory of algebraic numbers was given in 1871 by Dedekind [4]; in this work, Dedekind introduced the basic notions on fields, but restricted the field elements to complex numbers. As regards transcendental numbers, the early contributions were made by analysts. The most notable of these contributions were that by Liouville [1] in 1851, devoted to the construction of classes of transcendental numbers, and those by Hermite [2] in 1873 and Lindemann [1] in 1882, in which proofs are given of the transcendence of the numbers e and π, respectively. But it was not until 1882 that transcendentals made their appearance in the theory of fields, when Kronecker [2] succeeded in using the adjunction of indeterminates as the basis for a formulation of the theory of algebraic numbers. It was also in 1882 that fields of algebraic functions of complex variables were introduced by Dedekind and Weber [1] in order to lay the foundations of the arithmetical theory of algebraic functions. This work, in which a purely algebraic treatment of Riemann surfaces is given, marks the beginning of what was to become a very fruitful interplay between commutative algebra and algebraic geometry. It was next discovered in 1887 by Kronecker [3] that every algebraic number field can be obtained as the quotient of the polynomial domain $\mathbf{Q}[X]$ by the principal ideal generated by an irreducible polynomial, showing in effect that the theory of algebraic numbers does not require the use of complex numbers. Finally, the abstract definition of a field as we know it today was given in 1893 by Weber [1] in an article on the foundations of Galois theory. Weber also observed in this work that Kronecker's construction can be applied to arbitrary fields, and in particular to every field of integers modulo a prime; and that as a result, we recover the theory of higher congruences previously developed by Galois [2], Serret [1: 343–370], and Dedekind [2].

The final step toward the axiomatic foundations of the theory of fields was taken by Steinitz [1] in 1910. Spurred on by both the earlier contributions and the discovery by Hensel [1] of the p-adic fields, Steinitz set out to derive the consequences of Weber's axioms. His work, in which field extensions were first studied in full generality and in which normality, separability, and pure inseparability were introduced in order to give a detailed analysis of the structure of algebraic extensions, became the corner-

stone in the development of abstract algebra. In the words of Artin and Schreier [1]: "E. Steinitz hat durch seine 'Algebraische Theorie der Körper' weite Gebiete der Algebra einer abstrakten Behandlungsweise erschlossen; seiner bahnbrechenden Untersuchung ist zum grossen Teil die starke Entwicklung zu danken, die seither die moderne Algebra genommen hat". It is in the closing pages of Steinitz's article that the theory of transcendental extensions was first presented. However, before this theory could be brought to its present state, two significant additions were yet to be made, both partially motivated by questions in algebraic geometry. In 1939, MacLane [1] introduced the notion of separability for transcendental extensions. This was then followed in 1946 by the treatise on the foundations of algebraic geometry by Weil [1], in which the abstract notion of derivation is introduced in the study of separability.

Galois theory is generally regarded as one of the central and most beautiful parts of algebra. Its creation marked the culmination of investigations by generations of mathematicians into one of the oldest problems in algebra, the solvability of polynomial equations by radicals. The familiar formula for the roots of the quadratic equation was essentially known to the Babylonian mathematicians of the twentieth century B.C. No significant progress was made on polynomial equations of higher degree until the sixteenth century, when del Ferro and Ferrari discovered the formulas for the cubic and quartic equations, respectively. These results were first published by Cardano [1] in 1545; it is probably for this reason that Cardano's name has been traditionally associated with the formulas for the cubic equation.

These formulas express the roots of the equations in terms of the coefficients, using exclusively the field operations and the extraction of roots. Attempts to find such formulas for polynomial equations of higher degree were unsuccessful; and partly as a consequence of the work of Lagrange [2; 3] in 1770–1772, the algebraists of the period came to believe that it was impossible to derive them. This was proved to be the case at the beginning of the nineteenth century. Several proofs were published by Ruffini [1] between 1799 and 1813, but they were incomplete. The first satisfactory proof was given by Abel [2] in 1826, three years before his tragic death before the age of 27; between 1826 and 1829 he obtained further results on the solvability of polynomial equations by radicals, which were published in Abel [3; 1: II, 217–243, 269–270, 271–279].

The contributions of Ruffini and Abel were followed by the decisive results of Galois [1: 25–61] in 1832. Galois proved that the solvability of a polynomial equation by radicals is equivalent to a special property of a group naturally associated with the equation. Galois made this discovery before the age of 20, at a time when abstract algebra virtually did not exist!

Although Galois's result on the solvability of polynomial equations by radicals settled a problem that had eluded the efforts of some of the

greatest mathematicians of earlier generations, later developments have shown that the ideas introduced by Galois in his solution surpass by far the importance of the problem that he originally set out to solve. First, Galois defined and used the group-theoretical properties of normality, simplicity, and solvability, which play a significant role in the theory of groups. Moreover, he solved a problem of fields by translating it into a more tractable problem on groups; in so doing, he probably made the earliest application of a method that has become pervasive in algebra, namely, that of studying a mathematical object by suitably relating it to a mathematical object with a simpler structure. Nor is it an exaggeration to say that Galois theory is a prerequisite for much current research in number theory and algebraic geometry.

The story of Galois's life is a topic of considerable controversy. A gifted mathematician who is killed in a duel at the age of 20 presents unlimited opportunities for the creation of a myth. Unfortunately, this is precisely what several well-known authors have done in their writings on Galois. By means of intentional or unintentional omissions and distortions, legends have been created in which Galois is portrayed as a struggling genius unappreciated not only by the general public, but also by some of the leading mathematicians of his time. The recent article by Rothman [1] offers a lively account of such theories, as well as a careful attempt to unravel them.

Galois's ideas were expressed originally within the context of the theory of equations: To each polynomial equation is assigned a group of permutations of its roots. The progress made toward the axiomatic foundations of algebra in the last part of the nineteenth century had a considerable impact on Galois theory. Dedekind [4] observed that a more natural setting for Galois theory is obtained by regarding the groups associated with polynomial equations as groups of automorphisms of the corresponding splitting fields. Furthermore, he pioneered the systematic use of linear algebra in Galois theory. Since the abstract theory of field extensions was not developed until the first decade of the present century, Dedekind had to restrict his considerations to special types of fields. That his formulation of Galois theory remains meaningful for arbitrary fields was shown subsequently by the works of Weber [1] in 1893, of Steinitz [1] in 1910, and of Artin [3] in 1942. It is to these algebraists, and especially to Artin, that we owe what is now considered to be the definitive exposition of the Galois theory of finite groups of field automorphisms. A further contribution that must be mentioned is the generalization of the principal results of this theory to a special type of infinite groups of field automorphisms, discovered by Krull [1] in 1928.

Prerequisites

We shall assume that the reader possesses a certain familiarity with the rudiments of abstract algebra. More specifically, in addition to the basic properties of integers, sets, and mappings, the reader is expected to know the elementary parts of the theory of groups and the theory of rings, and to possess a reasonable background in linear algebra. Suggested references on these prerequisites are the following.

1. Adamson, I. T. *Elementary Rings and Modules*. New York: Harper & Row, 1972.
2. Godement, R. *Cours d'Algèbre*. Paris: Hermann, 1963. (English translation: *Algebra*. New York: Houghton Mifflin, 1968.)
3. Halmos, P. R. *Naive Set Theory*. New York: Springer-Verlag, 1974.
4. Hoffman, K., and Kunze, R. *Linear Algebra*. Englewood Cliffs, NJ: Prentice-Hall, 1971.
5. Ledermann, W. *Introduction to Group Theory*. Edinburgh: Oliver & Boyd, 1973.
6. Rotman, J. J. *The Theory of Groups, an Introduction*. Boston: Allyn & Bacon, 1973.

This list is not intended as an exhaustive bibliography on the basic concepts of algebra. We have simply selected six easily accessible books that, for our purposes, are particularly suitable as references. The books [1] and [2] seem the most convenient: In the first place, we shall adhere almost

completely to the terminology and notation used in these books; furthermore, taken together, these cover all the required background on rings, ideals, polynomials, modules, and vector spaces. The few facts on ordering and cardinal numbers occasionally used here are contained in the book [3]; and each of the books [5] and [6] contains all the background on groups needed in our presentation of Galois theory. Finally, the book [4] can be used as an alternative reference on linear algebra.

It should be noted that many books on abstract algebra, in chapters dealing separately with sets, groups, rings, and linear algebra, contain all or more of the prerequisites just described. Some of these are listed in the bibliography at the end of this book. (There is one section of the present book that requires additional prerequisites. This is section 3.12, which is devoted to infinite Galois theory, and in which some facts on topological groups are used. This section, however, is intended for readers interested in modern number theory; such readers would have to be well-versed in the theory of topological groups, and so it would be superfluous to give references on this subject.)

We now proceed to state in precise terms the conventions that will be adopted, and to explain the terminology and notation that will be used. Since there is no total agreement on these matters in the literature, the reader should make sure that we are using the same language.

Three types of algebraic structure are considered in our presentation. The first is defined by one operation, the second by two operations, and the third by one operation and one action. The term *operation* is being used here with the same meaning as "law of composition", "internal law of composition", and "binary operation", all of which are standard in the literature; and the term *action* is being used with the same meaning as "external law of composition", which is also of common usage.

We shall be concerned exclusively with operations that are associative and admit a neutral element. Moreover, for the most part, we shall use the multiplicative and additive notations. In the former case, the neutral element is called the **unit element** and is denoted by 1; and in the latter, it is called the **zero element** and is denoted by 0.

In the case of groups, subgroups, and group-homomorphisms, we shall usually follow [5] and [6]. In particular, the operation of a group will be written multiplicatively; the only exception to this occurs when reference is being made to the additive group of a ring, where the context always makes the intended meaning clear.

On the other hand, it will not be necessary for us to use the concept of ring in its full generality. First, our rings, subrings, and ring-homomorphisms will be restricted as in [2]: Rings possess a unit element; ring and subring have the same unit element; and ring-homomorphisms send unit element to unit element. Also, the nature of our subject dictates that we restrict our consideration to commutative rings in which the zero and unit

elements are distinct. Whenever we speak of rings, subrings, and ring-homomorphisms, it will be tacitly understood that all these restrictions apply.

Finally, in the case of modules and vector spaces, we shall follow [1], [2], and [4]. As usual, the operation and action of a module are referred to as its **vector addition** and **scalar multiplication**, respectively. In view of the conventions just adopted, it will not be necessary to distinguish between left and right modules. We shall speak of A-modules, A-submodules, and A-linear mappings whenever we wish to indicate that the ring of scalars is A. The general concept of module will play only an ancillary role in this book, since we shall be concerned primarily with vector spaces; if the field of scalars is A, we shall speak of **A-spaces** instead of vector spaces over A. It is hoped that this departure from standard terminology will not cause misunderstandings.

So far, for each of the prerequisites, we have made reference to certain books whose terminology and notation we shall generally follow. We shall now indicate the few instances where deviations occur.

A relation is said to **order** a set when it is reflexive, antisymmetric, and transitive on the elements of the set. By an **ordered set** we shall understand a set provided with a relation that orders it.

Let E be an ordered set. If $x, y \in E$, we write $x \leq y$ or $y \geq x$ to express that the pair (x, y) is in the given relation ordering E; and we write $x < y$ or $y > x$ to express that (x, y) is in this relation and $x \neq y$. If $(x_i)_{i \in I}$ is a family of elements of E, to say that $(x_i)_{i \in I}$ is **filtered** means that for all $i, j \in I$, there exists a $k \in I$ for which $x_k \geq x_i$ and $x_k \geq x_j$; and to say that $(x_i)_{i \in I}$ is a **chain** means that for all $i, j \in I$, we have $x_i \leq x_j$ or $x_i \geq x_j$. If $S \subseteq E$, then S is said to be **filtered** when the family $(x)_{x \in S}$ is filtered; and similarly, S is said to be a **chain** when $(x)_{x \in S}$ is a chain. If $S \subseteq E$ and $b \in E$, then b is an **upper bound for** S when $b \geq x$ for every $x \in S$.

If E is an ordered set, there can be in E at most one upper bound for E; when it exists, it is said to be the **largest element of** E. A **maximal element of** E is an $x \in E$ such that $x < y$ for no $y \in E$. Note that if the largest element of E exists, it is the only maximal element of E; but when E does not admit a largest element, it may admit more than one maximal element.

The preceding considerations on ordered sets apply, in particular, to sets of sets. Whenever we speak of a set of sets as being ordered by the inclusion relation, it will be understood that the relation in question is $\subseteq$. It is clear, therefore, what is meant when we speak of a **filtered family of sets**, a **filtered set of sets**, a **chain of sets**, the **largest element of a set of sets**, and a **maximal element of a set of sets**.

It should be noted, on the other hand, that every set of sets is also ordered by the opposite inclusion relation $\supseteq$. This, however, will be applied in only two instances: when we speak of the **smallest element of a set of sets** and of a **minimal element of a set of sets**.

To conclude these remarks on ordered sets, we shall state the result called **Zorn's lemma**. By an **inductive set** we shall understand an ordered set in which every nonempty chain admits an upper bound. The result in question asserts the following:

Every nonempty inductive set admits a maximal element.

This is a powerful set-theoretical tool that we shall use to derive important properties of algebraically closed fields and to establish the extendibility of certain mappings. It is not an "intuitive" statement, and does not yield "constructive" proofs. It is known to be equivalent to the "more intuitive" **axiom of choice** in the theory of sets, which asserts that the cartesian product of every nonempty family of nonempty sets is nonempty. The reader interested in a detailed study of these questions may wish to consult the book [3]. We shall simply accept Zorn's lemma as a valid result, and apply it without further comment.

A group consisting of a single element will be called **trivial**. If G is a group and H is a subgroup of G, a **left transversal of H in G** is a subset of G having exactly one element in common with each left coset of H in G; a **right transversal of H in G** is defined similarly, using right cosets.

Let A be a ring. There exists a unique homomorphism from the ring $\mathbf{Z}$ of integers to A; this is the mapping $n \to n1$ from $\mathbf{Z}$ to A. It is customary to denote by the same symbol n the value of this homomorphism at an integer n; this is only a notational convenience, and it should be noted that if m and n are distinct integers, the equality $m = n$ may be valid in A. The image of this homomorphism is called the **image of $\mathbf{Z}$ in A**; it is the smallest element of the set of all subrings of A.

If A is a ring, the **invertible elements of A** are the multiplicatively invertible elements of A. The set of all invertible elements of A is multiplicatively stable, and, provided with the operation defined by restriction of the multiplication of A, is a group. This group is denoted by A^*; its neutral element is 1, the unit element of A. The subgroups of A^* are called the **multiplicative groups in A**. The elements of finite order in A^* are the **roots of unity in A**; and if n is a positive integer, an nth **root of unity in A** is an $\alpha \in A$ for which $\alpha^n = 1$, that is, a root of unity in A with order dividing n.

An ideal in a ring is **null** when it consists of a single element; **prime** when it is a proper ideal and its complement in the ring is multiplicatively stable; and **maximal** when it is a maximal element of the set of all proper ideals.

We shall speak of **domains** instead of integral domains, and of **factorial domains** instead of unique factorization domains. By a **system of representatives of irreducible elements** in a factorial domain we shall understand a set of irreducible elements having exactly one element in common with the set of all associates of each irreducible element.

Polynomials play an essential role in our subject. The letters X, Y, Z —with or without subscripts—will be reserved for the variables in our rings of polynomials. Polynomials in infinitely many variables will be required only occasionally in this book (and in the only important instance, alternatives are indicated); the reader who is not familiar with this more general type of polynomial should read 0.0.5 below, where it is explained how to construct rings of polynomials in infinitely many variables.

An injective group-homomorphism or ring-homomorphism will be called a **monomorphism** or an **embedding**. Given two groups or two rings A and B, to say that A is **embeddable in** B will mean that there exists a monomorphism from A to B. This terminology is particularly convenient when dealing with fields, since it serves as a constant reminder of the fact that every homomorphism from a field to a ring is injective.

A module or vector space consisting of a single element will be called **null**. If A is a field and E is an A-space, the symbol $[E: A]$ will denote the dimension of E over A. Incidentally, the reader in need of a rapid review of the theory of dimension for general vector spaces may wish to learn Steinitz's axiomatic approach; this is given in section 4.1 and requires set-theoretical prerequisites exclusively, so that it can be read without reference to any other section.

If A is a ring and I is a set, the symbol $A^{(I)}$ will be used to denote the **free A-module based on I**. In order to define this module, we recall that if $(P_i)_{i\in I}$ is a family of statements, we say that P_i holds **for almost every** $i \in I$ when the set of all $i \in I$ for which P_i does not hold is finite. This being so, the elements of $A^{(I)}$ are the families $(\lambda_i)_{i\in I}$ of elements of A such that $\lambda_i = 0$ for almost every $i \in I$; and the vector addition and scalar multiplication of $A^{(I)}$ are defined "coordinate-wise":

$$(\lambda_i)_{i\in I} + (\mu_i)_{i\in I} = (\lambda_i + \mu_i)_{i\in I} \quad \text{and} \quad \alpha(\lambda_i)_{i\in I} = (\alpha\lambda_i)_{i\in I}.$$

For each $i \in I$, let ε_i denote the element of $A^{(I)}$ with 1 as its ith coordinate and with 0 as its jth coordinate for every $j \in I - \{i\}$. Then $(\varepsilon_i)_{i\in I}$ is a base of $A^{(I)}$, and so $A^{(I)}$ is indeed a free A-module; we refer to $(\varepsilon_i)_{i\in I}$ as the **standard base of $A^{(I)}$**.

If A is a ring and n is a positive integer, then the free A-module based on $\{1,2,\ldots,n\}$ is none other than the familiar A-module $A^{(n)}$ of "vectors" with n coordinates in A.

If a ring A is a subring of a ring B, then B can be regarded as an A-module in a natural way: The vector addition is the addition of B, and the scalar multiplication is the action of A on B defined by restriction of the multiplication of B. Whenever a ring is viewed as a module over a subring, it will be understood that the linear structure under consideration is defined in this manner.

If a ring A is a common subring of rings B and C, it is customary to define an **A-homomorphism from B to C** as a homomorphism from B to C

such that $\alpha \to \alpha$ for every $\alpha \in A$. It is readily seen that a homomorphism from B to C has this property if and only if it is an A-linear mapping from the A-module B to the A-module C.

Two notions of an independent family of elements arise in the present context. Let a ring A be a subring of a ring B, and let $(\beta_1, \beta_2, \ldots, \beta_n)$ be a finite sequence of elements of B—it is sufficient to consider only finite sequences, since the properties in question are of "finite character". First, it is meaningful to speak of $(\beta_1, \beta_2, \ldots, \beta_n)$ as a linearly independent sequence of elements of the A-module B; we shall express this by saying that $(\beta_1, \beta_2, \ldots, \beta_n)$ is **linearly independent over** A. On the other hand, recall that we have the A-homomorphism $f(X_1, X_2, \ldots, X_n) \to f(\beta_1, \beta_2, \ldots, \beta_n)$ from $A[X_1, X_2, \ldots, X_n]$ to B, and that its kernel is said to be the **ideal of algebraic relations of** $(\beta_1, \beta_2, \ldots, \beta_n)$ **over** A; then $(\beta_1, \beta_2, \ldots, \beta_n)$ is said to be **algebraically independent over** A when this ideal is null. Of course, these two notions are not identical; generally speaking, algebraic independence implies linear independence, but not conversely.

For each of the algebraic objects considered previously, there is a notion of quotient object. In the case of a group, it is defined by a normal subgroup; in that of a ring, by a proper ideal; and in that of a module, by a submodule. The elements of the quotient object are the cosets of the elements in the original object relative to the defining normal subgroup, proper ideal, or submodule. The mapping from the original object to the quotient object that to each element assigns its coset is called the **natural projection**; it is a surjective homomorphism in the case of groups and rings, and a surjective linear mapping in the case of modules.

To close the present discussion on prerequisites, we shall derive some isolated special results that will be used in this book and for which we were unable to give suitable references.

0.0.1. Given a set A, a ring R, and a bijection u from A to R, there is a unique ring structure on A relative to which u is an isomorphism. The addition and multiplication defining this ring structure are the operations

$$(\alpha, \beta) \to u^{-1}(u(\alpha) + u(\beta)) \quad \text{and} \quad (\alpha, \beta) \to u^{-1}(u(\alpha)u(\beta))$$

on A.

It is clear that these are the only possible choices for the required operations: If A is provided with a ring structure in such a way that u is an isomorphism, then for all $\alpha, \beta \in A$ we have

$$u(\alpha + \beta) = u(\alpha) + u(\beta) \quad \text{and} \quad u(\alpha\beta) = u(\alpha)u(\beta),$$

and hence

$$\alpha + \beta = u^{-1}(u(\alpha) + u(\beta)) \quad \text{and} \quad \alpha\beta = u^{-1}(u(\alpha)u(\beta)).$$

This being said, we can now state and prove the following result on ring-monomorphisms.

Let A and R be rings, and let u be a monomorphism from A to R. Then there exist a ring B and an isomorphism v from B to R such that A is a subring of B and such that v extends u.

Proof. First, choose a set S such that S and $R - \mathrm{Im}(u)$ are equipotent and such that A and S are disjoint. Then choose a bijection t from S to $R - \mathrm{Im}(u)$, and put $B = A \cup S$.

Denote by v the mapping from B to R, such that

$$\alpha \to u(\alpha) \quad \text{for } \alpha \in A \quad \text{and} \quad \alpha \to t(\alpha) \quad \text{for } \alpha \in S.$$

Then v is a bijection extending u, and the desired conclusion follows by providing B with the ring structure relative to which v is an isomorphism. □

0.0.2. The following useful result relates the linear structures defined by a chain of rings.

Let A be a ring, let B be a ring having A as a subring, and let C be a ring having B as a subring.

(i) *If $(\beta_i)_{i \in I}$ and $(\gamma_j)_{j \in J}$ are, respectively, generating systems of the A-module B and of the B-module C, then $(\beta_i \gamma_j)_{(i,j) \in I \times J}$ is a generating system of the A-module C.*

(ii) *If $(\beta_i)_{i \in I}$ and $(\gamma_j)_{j \in J}$ are, respectively, families of elements of B and C that are linearly independent over A and B, then $(\beta_i \gamma_j)_{(i,j) \in I \times J}$ is linearly independent over A.*

(iii) *If $(\beta_i)_{i \in I}$ and $(\gamma_j)_{j \in J}$ are, respectively, bases of the A-module B and of the B-module C, then $(\beta_i \gamma_j)_{(i,j) \in I \times J}$ is a base of the A-module C.*

Proof. It suffices to prove (i) and (ii), since these clearly imply (iii).

First, let $(\beta_i)_{i \in I}$ and $(\gamma_j)_{j \in J}$ be as in (i), and let $\xi \in C$. We can write $\xi = \sum_{j \in J} \mu_j \gamma_j$, where $\mu_j \in B$ for every $j \in J$ and $\mu_j = 0$ for almost every $j \in J$. Then for every $j \in J$, we can write $\mu_j = \sum_{i \in I} \alpha_{ij} \beta_i$, where $\alpha_{ij} \in A$ for every $i \in I$ and $\alpha_{ij} = 0$ for almost every $i \in I$; and furthermore, when $\mu_j = 0$, we can take $\alpha_{ij} = 0$ for every $i \in I$. It then follows that $\alpha_{ij} = 0$ for almost every $(i, j) \in I \times J$ and

$$\xi = \sum_{j \in J} \mu_j \gamma_j = \sum_{(i,j) \in I \times J} \alpha_{ij} \beta_i \gamma_j,$$

which shows that ξ is a linear combination of $(\beta_i \gamma_j)_{(i,j) \in I \times J}$ with coefficients in A.

To conclude, let $(\beta_i)_{i \in I}$ and $(\gamma_j)_{j \in J}$ be as in (ii). Assume that

$$\sum_{(i,j) \in I \times J} \alpha_{ij} \beta_i \gamma_j = 0,$$

where $\alpha_{ij} \in A$ for every $(i, j) \in I \times J$ and $\alpha_{ij} = 0$ for almost every $(i, j) \in I \times J$. Put $\mu_j = \sum_{i \in I} \alpha_{ij} \beta_i$ for every $j \in J$; then $\mu_j \in B$ for every $j \in J$ and

$\mu_j = 0$ for almost every $j \in J$, and

$$\sum_{j \in J} \mu_j \gamma_j = \sum_{(i,j) \in I \times J} \alpha_{ij} \beta_i \gamma_j = 0.$$

The linear independence of $(\gamma_j)_{j \in J}$ over B now implies that

$$\sum_{i \in I} \alpha_{ij} \beta_i = \mu_j = 0$$

for every $j \in J$; and the linear independence of $(\beta_i)_{i \in I}$ over A implies then that $\alpha_{ij} = 0$ for every $(i, j) \in I \times J$. Thus, $(\beta_i \gamma_j)_{(i,j) \in I \times J}$ is linearly independent over A. □

0.0.3. The following proposition, which will be used in the study of algebraic extensions, gives a sufficient condition for a domain to be a field.

Let A be a domain. If there exists a subfield K of A such that A is a finite-dimensional K-space, then A is a field.

Proof. Indeed, let $\alpha \in A$ and $\alpha \neq 0$. If K is a subfield of A with the indicated property, then the mapping $\xi \to \alpha\xi$ from A to A is K-linear. Since A is a domain, this mapping is injective, and the assumed finite dimensionality of A as a K-space implies that it is bijective. In particular, we have $\alpha\beta = 1$ for some $\beta \in A$, and so α is invertible in A. □

0.0.4. As a typical illustration of how Zorn's lemma is used in algebra, we shall now derive a result on the existence of maximal ideals. This will be used to prove that every field admits an algebraic closure.

Every proper ideal in a ring is contained in a maximal ideal.

Proof. Let A be a ring, and let $\mathcal{J}$ be a proper ideal of A. Denote by Ω the set of all proper ideals of A containing $\mathcal{J}$, and order Ω by inclusion. Then the maximal ideals of A containing $\mathcal{J}$ are precisely the maximal elements of Ω; therefore, by Zorn's lemma, it suffices to verify that Ω is nonempty and inductive.

To do this, note first that $\mathcal{J} \in \Omega$, so that Ω is nonempty. Now let Ψ be a nonempty chain in Ω, and put $\mathcal{M} = \cup_{\mathcal{K} \in \Psi} \mathcal{K}$. Since Ψ is filtered, it is clear that $\mathcal{M}$ is an ideal of A; and since $1 \notin \mathcal{K}$ for every $\mathcal{K} \in \Psi$, we have $1 \notin \mathcal{M}$. Therefore $\mathcal{M}$ is a proper ideal of A such that $\mathcal{M} \supseteq \mathcal{K}$ for every $\mathcal{K} \in \Psi$, which implies that $\mathcal{M}$ is an upper bound for Ψ in Ω. We conclude that Ω is inductive. □

Applying this to the null ideal, we see that every ring possesses maximal ideals. Also, since an element in a ring is noninvertible if and only if it generates a proper ideal, it follows that the invertible elements in a ring are the elements belonging to no maximal ideal.

0.0.5. It is sometimes required to know how given algebraic structures of the same type can be combined in order to obtain an algebraic

structure suitably related to the given ones. To prove that every field admits an algebraic closure, we shall use the following special case of a general proposition in universal algebra.

Let $(A_i)_{i\in I}$ be a nonempty family of rings such that for all $i, j \in I$, there exists a $k \in I$ such that A_k contains A_i and A_j as subrings. Then $\cup_{i\in I} A_i$ can be uniquely provided with a ring structure in such a way that it contains A_i as a subring for every $i \in I$.

Proof. Let $A = \cup_{i\in I} A_i$. To define the two required operations on A, let $\alpha, \beta \in A$. The hypothesis implies, first, that there exists an $i \in I$ for which $\alpha, \beta \in A_i$; it also implies that if $i, j \in I$ and $\alpha, \beta \in A_i \cap A_j$, then A_i and A_j are subrings of one and the same ring, and so the symbols $\alpha + \beta$ and $\alpha\beta$ have the same meaning in A_i and in A_j.

It is meaningful, therefore, to speak of the addition and of the multiplication on A that to each $(\alpha, \beta) \in A \times A$ assign, respectively, the elements $\alpha + \beta$ and $\alpha\beta$ of A_i whenever $i \in I$ and $\alpha, \beta \in A_i$. Note also that for all $i, j \in I$, the rings A_i and A_j have the same unit element. The details involved in verifying that the addition and the multiplication on A just described satisfy all the required conditions are now seen to be straightforward, and can be omitted. □

The proposition just proved can be used in order to define rings of polynomials in infinitely many variables.

To see this, consider a ring A and an infinite set I. Denote by Ω the set of all finite subsets of I; and for every $S \in \Omega$, write $A_S = A[X_i]_{i\in S}$. It is clear that if $S, T \in \Omega$, then $S \cup T \in \Omega$ and $A_{S\cup T}$ contains A_S and A_T as subrings. The ring $A[X_i]_{i\in I}$ is defined as the ring obtained when $\cup_{S\in\Omega} A_S$ is provided with the ring structure described in the preceding proof.

The few general properties of rings of polynomials in infinitely many variables that will be used in this book are then seen to be immediate consequences of the corresponding properties of rings of polynomials in finitely many variables.

0.0.6. Recall that if F is a field and n is a positive integer, the symbol $GL_n(F)$ denotes the multiplicative group of nonsingular $n \times n$ matrices with entries in F. In the computations of certain Galois groups, it will be necessary to know the order of $GL_2(F)$ when F is a finite field. As we now show, this order can be determined by an elementary counting argument.

If F is a finite field, then $GL_2(F)$ is a group of order $(\mathrm{Card}(F)^2 - \mathrm{Card}(F))(\mathrm{Card}(F)^2 - 1)$.

Proof. Let us write $q = \mathrm{Card}(F)$.

The elements of $GL_2(F)$ are the 2×2 matrices whose rows are linearly independent vectors in $F^{(2)}$. Therefore, the first row of such a

matrix can be any nonzero vector in $F^{(2)}$; and the second row can be any vector in $F^{(2)}$ that is not a scalar multiple of the first row. Since there are q^2-1 choices for the first row, and since for each of these there are q^2-q for the second row, we conclude that the number of elements of $GL_2(F)$ is $(q^2-1)(q^2-q)$, as claimed. □

0.0.7. Here we shall collect a few facts on the elementary symmetric polynomials that will be used in some of the classical illustrations of Galois theory.

Given a positive integer n and an integer k such that $0 \le k \le n$, let $[n; k]$ denote the set of all subsets of $\{1,2,\ldots,n\}$ with cardinality k.

If A is a ring and n is a positive integer, the **elementary symmetric polynomials in** $A[X_1, X_2,\ldots,X_n]$ are the polynomials $e_0, e_2,\ldots,e_n$ defined by

$$e_k(X_1, X_2,\ldots,X_n) = \sum_{S\in[n;k]} \prod_{i\in S} X_i \quad \text{for } 0 \le k \le n.$$

Thus, we have

$$\begin{aligned}
e_0(X_1, X_2,\ldots,X_n) &= 1,\\
e_1(X_1, X_2,\ldots,X_n) &= \sum_{i=1}^{n} X_i,\\
e_2(X_1, X_2,\ldots,X_n) &= \sum_{1\le i<j\le n} X_iX_j,\\
e_3(X_1, X_2,\ldots,X_n) &= \sum_{1\le i<j<k\le n} X_iX_jX_k,\\
&\vdots \qquad\qquad \vdots \qquad\qquad \vdots\\
e_n(X_1, X_2,\ldots,X_n) &= \prod_{i=1}^{n} X_i.
\end{aligned}$$

We begin with an auxiliary result showing that the elementary symmetric polynomials satisfy simple recurrence relations. These will be used in order to derive two important facts about the elementary symmetric polynomials.

Let A be a ring, let n be an integer such that $n>1$, and let $\bar{e}_0, \bar{e}_1,\ldots,\bar{e}_{n-1}$ and $e_0, e_1,\ldots,e_n$ denote, respectively, the elementary symmetric polynomials in $A[X_1, X_2,\ldots,X_{n-1}]$ and $A[X_1, X_2,\ldots,X_n]$. Then

$$e_n(X_1, X_2,\ldots,X_n) = X_n\bar{e}_{n-1}(X_1, X_2,\ldots,X_{n-1}),$$

and

$$e_k(X_1, X_2,\ldots,X_n) = \bar{e}_k(X_1, X_2,\ldots,X_{n-1}) + X_n\bar{e}_{k-1}(X_1, X_2,\ldots,X_{n-1})$$

for $1 \le k \le n-1$.

Proof. The first equality is evident. Suppose now that $1 \le k \le n-1$, and let Ω denote the set of all sets of the form $S\cup\{n\}$ with $S \in [n-1; k-1]$. Since

$$[n; k] = [n-1; k]\cup\Omega \quad \text{and} \quad [n-1; k]\cap\Omega = \varnothing,$$