Universiteit Utrecht

Bachelor Thesis

Generalizations of the Casus Irreducibilis

Author:
Joost Franssen

Supervisor:
Prof. dr. G. L. M. Cornelissen

January 21, 2016

Introduction

The solution to the general cubic equation $ax^3 + bx^2 + cx + d = 0$, with $a \neq 0$, over the real numbers has been known for hundreds of years. Cardano, with attribution to Del Ferro and Tartaglia, published the solution in 1545, which is now usually known as 'Cardano's Formula'. Cardano already knew how to reduce the general cubic polynomial to a polynomial of the form $x^3 + px + q$, where $p, q \in \mathbb{R}$. This is also the first step of our approach. Instead of following Cardano's old methods, we use more modern techniques from Galois theory to derive a more general result: We shall derive Cardano's formula for cubic polynomials over any field that is not of characteristic 2 or 3. Since Cardano's formula does not work in those characteristics, we also derive two separate formulas for the roots of cubic polynomials using radicals for fields of characteristic 2 and 3.

Before we can get started on Cardano's formula in chapter 2, we shall first take a look at discriminants in chapter 1. The discriminant for quadratic polynomials is widely known and used. Indeed, for a polynomial $ax^2 + bx + c \in \mathbb{R}[x]$, most recognize $b^2 - 4ac$ as the discriminant. Moreover, we can derive some information about the roots from the sign of the discriminant. For instance, if the discriminant is positive, the two roots are real and distinct. However, what a discriminant is for any polynomial over some field, let alone how to compute it, is usually less familiar. We shall derive a general formula for discriminants using a special determinant known as the *resultant.* As we shall see, in the case of a polynomial x^3+px+q, the discriminant equals $-4p^3-27q^2$. For cubic polynomials the discriminant provides similar information as for quadratic ones. In particular, when we work over $\mathbb{R}$ and the discriminant is positive, then all roots are real and distinct.

Back to Cardano's formula: For a polynomial $f(x) := x^3+px+q$ over some field F, not of characteristic 2 or 3, Cardano's formula yields its roots as follows:

$$\frac{\omega^{i-1}}{3}\sqrt[3]{-\frac{27}{2}q+\frac{3}{2}\sqrt{-3\Delta}}+\frac{\omega^{1-i}}{3}\sqrt[3]{-\frac{27}{2}q-\frac{3}{2}\sqrt{-3\Delta}}, \qquad i = 1, 2, 3.$$

Here ω is a primitive third root of unity, i.e., an element unequal to 1 such that $\omega^3 = 1$. The discriminant of f is denoted by Δ.
Let us look at an example. Consider $g(x) := x^3 - 3x + 1 \in \mathbb{Q}[x]$, which is plotted in Figure 1 below.

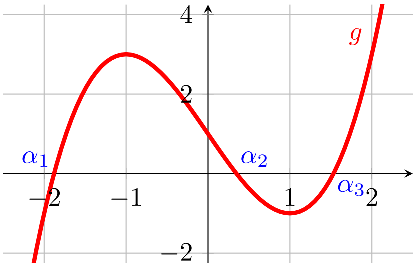

Figure 1: Plot of g.

In this case, the discriminant equals $-4(-3)^3 - 27 \cdot 1^2 = 81$. We see that the discriminant is positive, which means that all roots are real and distinct, which the graph confirms. Also, using Mathematica, we find approximations of the roots of g:

$$\alpha_1 \approx -1.87939, \quad \alpha_2 \approx 0.347296, \quad \alpha_3 \approx 1.53209.$$

We observe that the only rational roots could be ± 1 for the rational root test. Neither of these are roots of g, so g is irreducible over $\mathbb{Q}$. Thus in order to find the roots algebraically, we wish to apply Cardano's formula. However, we see that $\sqrt{-3 \cdot 81}$ becomes a complex number. Whence the radicand of the cube root, and therefore the cube root itself, is also a complex number in Cardano's formula. Since the roots of g are all real, two questions arise:

Question: Are the complex radicals necessary to express the roots of g using radicals?

Question: For what kinds of polynomials does such a situation occur?

The answer to the first question turns out to be *yes*. Moreover, we shall find in section 4.1 that computing the cube roots is just as hard as finding a root of g in the first place. Therefore we are stuck with the complex radicals. This situation, where the three real roots of a cubic polynomial over $\mathbb{Q}$ cannot be expressed with real radicals, is known as the *casus irreducibilis*. This, appropriately, is Latin for the 'irreducible case'. We examine this case in depth in section 4.1.
For the mathematicians from a few hundred years ago the casus irreducibilis had an important implication. Namely, it forced them to confront the complex numbers: Despite the fact that they worked with polynomials over the real numbers with real roots, the complex numbers still arose. As opposed to quadratic polynomials, where the complex roots can simply be ignored or deemed non-existent, the complex numbers were still a necessity for the real roots of cubic polynomials.

Before we rigorously investigate these questions in chapter 4, we take a useful detour in chapter 3 in order to obtain a more general answer to these questions. Namely, we shall not limit ourselves to polynomials over the real numbers, but over any field with the necessary properties of $\mathbb{R}$. We shall choose several useful properties from $\mathbb{R}$—such as the order relation $>$ and the fact that adjoining the square root of -1 to $\mathbb{R}$ yields an algebraically closed field $\mathbb{C}$—and find a common type of fields that satisfies these properties. The result will be *real closed fields*. One important result that makes this possible is the following:

Theorem (3.39 & 3.40)**:** Every ordered field has a real closed, algebraic extension that extends its order, known as a *real closure*. Moreover, these real closures are unique up to order-preserving isomorphism. □

The proof of the uniqueness we provide uses Zorn's lemma. Although others use Zorn's lemma to prove this statement, such as [13] and [8], our proof has been written entirely independently.
The analogy to $\mathbb{R}$ is now as follows. As we before considered polynomials over $\mathbb{Q}$, or any subfield F of $\mathbb{R}$, we can now consider any ordered field R. This ordered field is contained in a, for all intents and purposes, unique real closed field $\widetilde{R}$; in the same way as what $\mathbb{R}$ is to $\mathbb{Q}$ or F. One of the mentioned properties asserts that $\widetilde{R}(\boldsymbol{i})$, where $\boldsymbol{i}$ is a square root of -1, is algebraically closed. The elements of $\widetilde{R}$ we consider *formally real* and the ones of $\widetilde{R}(\boldsymbol{i})$ *formally complex*. This naturally extends the definitions of real and complex radicals to real closed fields.

This has equipped us with all the necessary tools to answer the main two questions thoroughly in chapter 4. The main result, which answers the second question, is Theorem 4.16. The formulation as well as the main idea of the proof of this theorem comes from Theorem 8.6.5 from [5]. However, in [5] the statement is merely stated and proved for subfields of the real numbers. Thus, although the proof works the same, we shall obtain a more general result thanks to the work we have done in chapter 3.

Acknowledgments

Firstly, I would like to express my gratitude to prof. dr. Cornelissen for his supervision and continual suggestions. Secondly, I wish to thank my dear friend Scott for his thorough proofreading of this thesis, which yielded various small corrections and improvements. Finally, I want to thank my parents for their continued support throughout.

Contents

Symbol Disambiguation

$\mathbb{N}$	The set of natural numbers: $\mathbb{N} := \{1, 2, 3, \ldots\}$.
$\mathbb{N}_0$	$\mathbb{N} \cup \{0\}$.
$\mathbb{Z}/n$	The integers modulo n.
$\mathbb{F}_p$	The finite field with p elements.
S_n	The symmetric group of degree n.
A_n	The alternating group of degree n.
R^*	The group of units (i.e., invertible elements with respect to multiplication) of a ring R.
v^{T}	The transpose of a vector or matrix.
$A \subseteq B$	A is a subset of B, including the possibility $A = B$.
$A \subset B$	$A \subseteq B$ and $A \neq B$, i.e., A is a proper subset of B.
$\bigcup \mathcal{C}$	$\bigcup_{C \in \mathcal{C}} C$, where a $\mathcal{C}$ is a set of sets.
$H \leq G$	H is a subgroup of G; H is less than or equal to G.
$H < G$	$H \leq G$ and $H \neq G$.
$\langle g \rangle$	The cyclic group or the ideal generated by g, depending on the context.
$\operatorname{Char} F$	The characteristic of a field F.
$[E : F]$	The degree of a field extension E of F.
$\boldsymbol{i}$	The imaginary unit $\sqrt{-1} \in \mathbb{C}$; A root of $x^2 + 1$.
$\overline{\alpha}$	The complex conjugate of $\alpha \in \mathbb{C}$; The formally complex conjugate of $\alpha \in R(\boldsymbol{i})$, where R is a formally real field (Definition 3.43).
$\wp_p$	The Artin-Schreier polynomial $x^p - x$ of degree p, where p is a prime number.
$\Delta(f)$	The discriminant of a polynomial f.
$\operatorname{Gal}(E/F)$	The Galois group of E over F.
E^G	The fixed field of G in E, where $G \leq \operatorname{Gal}(E/F)$ and E/F finite Galois.
Σ_F	The set of sums of squares of a field F.
Σ_F^*	$\Sigma_F \setminus \{0\}$.
$\overline{F}$	The algebraic closure of a field F.
$\widetilde{F}$	The real closure of a formally real field F.
$\langle K, L \rangle$	The compositum of K and L.

0 Preliminary Field and Galois Theory

This chapter mainly functions as a way to repeat definitions and known facts that we shall often use. We also use this to introduce some notation and terminology.

Definition 0.1: Let F be a field and let K and L be extensions of F. A homomorphism $\varphi : K \to L$ is called an *F-homomorphism* if φ is the identity map on F. A similar definition applies to *F-isomorphisms*, *F-monomorphisms*, and so on.
In particular, if $K = L$ and φ is an automorphism, then φ is called an *F-automorphism* if φ fixes F.

Definition 0.2: Let F be a field and let E be an algebraic extension of F. We call E a *normal extension* if each irreducible polynomial over F, which has a root in E, splits over E.

Definition 0.3: Let F be a field. We call an irreducible polynomial $p \in F[x]$ *separable* if the roots of p in its splitting field are distinct. An arbitrary polynomial $f \in F[x]$ is called *separable* if each irreducible factor of f in $F[x]$ is separable.
An algebraic extension E of F is a *separable extension* if for each $\alpha \in E$ the minimal polynomial of α over F is separable.

Definition 0.4: Let F be a field and E be a finite Galois extension, i.e., a finite separable, normal extension. The *Galois group of E over F* is the group of F-automorphisms on E with composition as group operation. We denote it by $\mathrm{Gal}(E/F)$.

Notation: We often write 'E/F' for 'E over F' when speaking of extensions.

***Theorem* 0.5** (Proposition 5.1.8 [5])**:** Let F be a field and $f \in F[x]$ an irreducible polynomial. Let E be a splitting field of f over F and let $\alpha, \beta \in E$ be roots of f. Then there exists an F-automorphism $\varphi : E \to E$ such that $\varphi(\alpha) = \beta$. □

Definition 0.6: Let E/F be a finite Galois extension and let $H \leq \mathrm{Gal}(E/F)$. The *fixed field* of H in E, denoted by E^H, is the field $\{a \in E \mid \varphi(a) = a \text{ for all } \varphi \in H\}$.

***Theorem* 0.7** (Fundamental Theorem of Galois Theory, a.k.a. Galois Correspondence; 4.10.1 [6])**:** Let E/F be a finite Galois extension of a field F. There is a bijection between the set of intermediate fields—i.e., a field K such that $F \subseteq K \subseteq E$—and the set of subgroups of the Galois group of E/F.
The map $K \mapsto \mathrm{Gal}(E/K)$—that sends an intermediate field K to the Galois group of E/K—and the map $H \mapsto E^H$—that sends a subgroup H of $\mathrm{Gal}(E/F)$ to its fixed field—are inverses of each other.
Moreover, if K and L are intermediate fields with $K \subseteq L$, then $\mathrm{Gal}(E/L) \leq \mathrm{Gal}(E/K)$. Conversely, if H and G are subgroups of $\mathrm{Gal}(E/F)$ with $H \leq G$, then $E^G \subseteq E^H$. Thus the correspondence reverses inclusions.
Finally, $H \leq \mathrm{Gal}(E/F)$ is a normal subgroup if and only if E^H/F is a normal extension. □

Definition 0.8: Let F be a field. An *n-th root of unity* is a root of the polynomial $x^n - 1$, where $n \in \mathbb{N}$, in some extension field of F. A *primitive n-th root of unity ζ* is an n-th root of unity such that for each $k \in \mathbb{N}$, $1 \leq k < n$, $\zeta^k \neq 1$.

***Theorem* 0.9** (4.11.1 [6])**:** Let F be a field, $n \in \mathbb{N}$ and E a splitting field of $x^n - 1$ over F. If $\mathrm{Char}\, F \nmid n$ (including $\mathrm{Char}\, F = 0$), then the set μ_n of n-th roots of unity is a cyclic subgroup of E^*. □

Note that, for such a field F, the primitive n-th roots of unity are precisely the generators of μ_n.

***Theorem* 0.10** (4.10.2 [6])**:** Let F be a field and E an extension field over F. If E is the splitting field of a separable polynomial over F, then E/F is a finite Galois extension. □

***Theorem* 0.11** (4.8.2 [6])**:** Let F be a field. The following hold:

(i) If $\mathrm{Char}\, F = 0$, then every polynomial over F is separable.

(ii) If $\mathrm{Char}\, F = p$, with $p > 0$, then an irreducible polynomial $f \in F[x]$ is *in*separable if and only if there exists a polynomial $g \in F[x]$ such that $f(x) = g(x^p)$. □

***Theorem* 0.12** (Lemma from [6, p. 302])**:** Let F be a field and $f \in F[x]$ a separable polynomial of degree $n \in \mathbb{N}$. Then the Galois group of the splitting field E of f over F is isomorphic to a subgroup of the symmetric group S_n. □

It is worth noting that the proof of [6] uses the fact that every element of the Galois group of E/F induces a permutation on the roots of f. This is something we often use when applying this theorem.

***Theorem* 0.13** (4.9.1 [6])**:** Let E/F be a finite Galois extension. Then $|\mathrm{Gal}(E/F)| = [E : F]$. □

Definition 0.14: Let F be a field. A *radical* over F is an element α in some extension field of F, which satisfies one of the following conditions:

(i) α is a root of a polynomial $x^n - a \in F[x]$ with $\mathrm{Char}\, F \nmid n$;

(ii) α is a root of a polynomial $x^p - x - a \in F[x]$ with $p = \mathrm{Char}\, F$.

Radicals of the form (i) are called *n-th roots over F* (or *square* or *cube* when $n = 2$ or $n = 3$, respectively) and may be denoted by $\sqrt[n]{a}$. Radicals of the form (ii) are called *p-th Artin-Schreier roots over F* and are denoted by $\wp_p^{-1}(a)$. The polynomial $x^p - x$ is called the *Artin-Schreier polynomial* (of degree p) and denoted by $\wp_p$. Hence this terminology and notation.

Remark 0.15: In case of (i), by Theorem 0.9, there exists a primitive n-th root of unity ζ in some extension of F. If $\sqrt[n]{a}$ is any root of $x^n - a$, then all the roots are $\sqrt[n]{a}$, $\zeta \sqrt[n]{a}$, $\zeta^2 \sqrt[n]{a}$, $\ldots$, $\zeta^{n-1} \sqrt[n]{a}$. See also Remark 2.2 on page 13 for a convention of choosing these roots.
For (ii), we can use Fermat's Little Theorem and the binomial expansion to see that if $\wp_p^{-1}(a)$ is a root of $x^p - x - a$, then all the roots are $\wp_p^{-1}(a)$, $\wp_p^{-1}(a) + 1$, $\wp_p^{-1}(a) + 2$, $\ldots$, $\wp_p^{-1}(a) + p - 1$. Indeed, let $n \in \mathbb{Z}/p$. We have

$$\begin{aligned} \wp_p(\wp_p^{-1}(a) + n) &= (\wp_p^{-1}(a) + n)^p - (\wp_p^{-1}(a) + n) \\ &= \wp_p^{-1}(a)^p + n^p - \wp_p^{-1}(a) - n \\ &= \wp_p^{-1}(a)^p - \wp_p^{-1}(a) = a. \end{aligned}$$

Thus $\wp_p^{-1}(a) + n$ is a root of $x^p - x - a$.

Remark 0.16: The reason that roots of the form (i) are generally not considered when $\mathrm{Char}\, F \mid n$, is that in this case n-th roots behave 'badly': If $n = pm$, $\mathrm{Char}\, F = p$, and α is a root of $x^n - a$, then $x^n - a = (x^m - \alpha^m)^p$. Consequently, α^m has multiplicity p and there are only up to m distinct roots. In particular, when $m = 1$, $x^n - a$ only has one root with multiplicity p. In this case we replace these radicals with inverses of the Artin-Schreier polynomial of degree p. As Remark 0.15 showed, these types of radicals have a similar property to n-th roots in that all roots are known if one is known.

***Theorem* 0.17** (4.13.3 [6])**:** Let E/F be a finite Galois extension of degree n, with $\mathrm{Char}\, F \nmid n$, such that F contains a primitive n-th root of unity. If $\mathrm{Gal}(E/F)$ is a cyclic group, then $E = F(\xi)$, where ξ is an n-th root over F, i.e., a radical of the form (i) in Definition 0.14. □

***Theorem* 0.18** (4.13.6 [6])**:** Let $\mathrm{Char}\, F = p > 0$ and let E/F be a finite Galois extension of degree p. If $\mathrm{Gal}(E/F)$ is a cyclic group, then $E = F(\xi)$, where ξ is a p-th Artin-Schreier root over F, i.e., a radical of the form (ii) in Definition 0.14. □

1 Discriminants

In this chapter we shall define the *discriminant* of a polynomial over a field. We will prove several properties of discriminants and will see how these relate to the roots of polynomials. We will primarily look at its properties for polynomials with real coefficients. Finally, we shall present a somewhat complicated, but computable, formula for the discriminant in terms of the coefficients of the polynomial.

1.1 Introduction

The word 'discriminant' presumably comes from the Latin *discriminare*, which means 'to distinguish'. This is indeed what the discriminant allows us to do: For polynomials with coefficients in some field F, the discriminant distinguishes between whether or not there are multiple roots. Moreover, for polynomials of degree 2 and 3 over $\mathbb{R}$, it also distinguishes between whether or not there are non-real complex roots, as we shall see later. We define the discriminant as follows:

Definition 1.1: Let $f \in F[x]$ be a polynomial of degree $n \geq 2$, with F a field. Let $a_n \in F$ be the leading coefficient of f and let $\alpha_1, \dots, \alpha_n$ be its roots, not necessarily distinct, in an extension field of F. We define the *discriminant* $\Delta(f)$ of f over F by

$$\Delta(f) \coloneqq a_n^{2n-2} \prod_{1 \leq i < j \leq n} (\alpha_i - \alpha_j)^2.$$

The factor a_n^{2n-2} seems somewhat arbitrary, but we add this to keep the definition consistent with historical formulas for the discriminant (see also Example 1.14). As mentioned above, the discriminant of a polynomial can be computed without knowing its roots (see Theorem 1.17). Hence it makes sense to use the discriminant to infer information about the nature of the roots. This computation will in particular imply that $\Delta(f) \in F$. We can also prove this directly using Galois theory.

***Proposition* 1.2:** Let $f \in F[x]$ be a polynomial of degree $n \geq 2$, F a field. Then $\Delta(f) \in F$.

Proof. Because $\Delta(f) \in F$ if and only if $\Delta(f)/a_n^{2n-2} \in F$, where a_n is the leading coefficient of f, it suffices to consider the case where f is monic.
If $\Delta(f) = 0$, then clearly $\Delta(f) \in F$. Suppose now $\Delta(f) \neq 0$. Let $\alpha_1, \dots, \alpha_n$ be the roots of f. The product in the discriminant consists of $(n-1) + (n-2) + \cdots + 2 + 1 = \frac{1}{2}n(n-1)$ factors. Each factor $(\alpha_i - \alpha_j)^2$ can be written as $-(\alpha_i - \alpha_j)(\alpha_j - \alpha_i)$. Thus we see that we can rewrite the discriminant as

$$\Delta(f) = (-1)^{\frac{1}{2}n(n-1)} \prod_{i \neq j} (\alpha_i - \alpha_j). \tag{1.1}$$

Since $\Delta(f) \neq 0$, all factors of the product are nonzero, meaning that $\alpha_i \neq \alpha_j$ for $i \neq j$. Thus f has no multiple roots, which means that f is separable. Let E be the splitting field of f over F. By Theorem 0.10, E is a finite Galois extension of F. Let $G \coloneqq \mathrm{Gal}(E/F)$ be the Galois group of E over F. By the fundamental theorem of Galois theory, the fixed field E^G of G in E is equal to F. Therefore, $\Delta(f) \in F$ if and only if $\Delta(f)$ is fixed by G. Theorem 0.12 says that G is isomorphic to a subgroup of S_n. Hence any element φ of G is determined by a permutation $\sigma \in S_n$ on the indices of the roots of f. That is to say, $\varphi(\alpha_i) = \alpha_{\sigma(i)}$ for each i. From (1.1) we see that any permutation in S_n does not change $\Delta(f)$. Whence $\Delta(f)$ is also invariant under all F-automorphisms on E, which are the elements of G. Thus $\Delta(f) \in E^G = F$. □

1.2 Relation to the Roots

By definition $\mathbb{C} = \mathbb{R}(\boldsymbol{i})$, where $\boldsymbol{i}$ is a root of $x^2 + 1 \in \mathbb{R}[x]$. The latter polynomial clearly is irreducible and separable over $\mathbb{R}$, which means that $\mathbb{C}/\mathbb{R}$ is a Galois extension of degree 2. Then $\mathrm{Gal}(\mathbb{C}/\mathbb{R})$ has two elements by Theorem 0.13, namely the identity and the map that sends $\boldsymbol{i}$ to $-\boldsymbol{i}$. The latter is called *complex conjugation* and is denoted by $\overline{\cdot}$, e.g., $\overline{\boldsymbol{i}} = -\boldsymbol{i}$. (This has a generalization: see Lemma 3.24.) This has a nice consequence: If f is a polynomial over $\mathbb{R}$ with a root α (which lies in $\mathbb{C}$ by the fundamental theorem of algebra; see also Theorem 3.30), then the complex conjugate $\overline{\alpha}$ is also a root of f. Indeed, the previous argument shows that complex conjugation is an $\mathbb{R}$-automorphism, hence $\overline{f(\alpha)} = f(\overline{\alpha})$. From

this we conclude that non-real complex roots of f come in pairs. We will usually refer to a complex number and its conjugate as a *complex pair.*
We can now prove an interesting relation between the sign of the discriminant and the number of non-real complex numbers. Note that, by Proposition 1.2, the discriminant is a real number if we are talking about polynomials over $\mathbb{R}$. So we may speak about its sign.

***Theorem* 1.3:** Let $f \in \mathbb{R}[x]$ be a polynomial of degree $n \geq 2$ with distinct roots $\alpha_1, \dots, \alpha_n$. Let r be half of the number of complex roots of f (note that $r \in \mathbb{N}_0$ by the foregoing). Then $\operatorname{sgn}(\Delta(f)) = (-1)^r$, where sgn is the signum function.

Proof. First note that the factor $a_n^{2n-2} = (a_n^{n-1})^2$ is a square, hence does not influence the sign of the discriminant. Therefore, we shall determine the sign of $\Delta(f)$ by looking at the factors $(\alpha_i - \alpha_j)^2$, $i < j$, and considering the following (exhaustive) cases:

(i) If $\alpha_i - \alpha_j \in \mathbb{R}$, then $(\alpha_i - \alpha_j)^2 > 0$.

(ii) If $\overline{\alpha_i} = \alpha_j$, then $\alpha_i - \alpha_j = 2\boldsymbol{i}\operatorname{Im}(\alpha_i)$ and so $(\alpha_i - \alpha_j)^2 < 0$. Note that the conjugate $\overline{\alpha_i - \alpha_j}$ equals $\alpha_j - \alpha_i$ and does not appear in the product, as $j > i$.

(iii) If $\alpha_i - \alpha_j \notin \mathbb{R}$ and $\overline{\alpha_i} \neq \alpha_j$, then in the complex conjugate $\overline{\alpha_i - \alpha_j} = \overline{\alpha_i} - \overline{\alpha_j}$ appears at least one root different from both α_i and α_j; that is to say, $\overline{\alpha_i} \neq \alpha_i, \alpha_j$ or $\overline{\alpha_j} \neq \alpha_i, \alpha_j$. Since $(\overline{\alpha_i} - \overline{\alpha_j})^2 = (\overline{\alpha_j} - \overline{\alpha_i})^2$, this factor appears in the product as well. This means that the product has a factor

$$(\alpha_i - \alpha_j)^2(\overline{\alpha_i} - \overline{\alpha_j})^2 = (\alpha_i - \alpha_j)^2\overline{(\alpha_i - \alpha_j)^2} = \operatorname{Re}(\alpha_i - \alpha_j)^2 + \operatorname{Im}(\alpha_i - \alpha_j)^2 > 0.$$

We see that each pair of complex conjugates yields one negative factor in the product and that the other factors only contribute positive ones. There are r pairs of complex conjugates and hence r negative factors. Therefore, the sign of $\Delta(f)$ equals $(-1)^r$, as desired. □

From this theorem we can infer whether or not all roots of polynomials of degree 2 or 3 are real. This is shown in the corollaries below.

***Corollary* 1.4:** Let $f \in \mathbb{R}[x]$ be a polynomial of degree 2. Let α_1 and α_2 be the roots of f. Then the following hold:

- $\Delta(f) = 0$ if and only if $\alpha_1 = \alpha_2$;
- $\Delta(f) > 0$ if and only if $\alpha_1, \alpha_2 \in \mathbb{R}$ and $\alpha_1 \neq \alpha_2$;
- $\Delta(f) < 0$ if and only if $\alpha_1, \alpha_2 \in \mathbb{C} \setminus \mathbb{R}$ and $\alpha_1 = \overline{\alpha_2}$.

Proof. Clearly $(\alpha_1 - \alpha_2)^2 = 0$ if and only if $\alpha_1 = \alpha_2$.
If $\Delta(f) > 0$, the theorem tells us there is an even amount of complex pairs. Since there can be at most one, as f only has two roots, this amount must be zero. Hence $\alpha_1, \alpha_2 \in \mathbb{R}$. They are distinct, as $\Delta(f) \neq 0$. Conversely, if $\alpha_1, \alpha_2 \in \mathbb{R}$ and distinct, there are no complex pairs, so, by the theorem, $\operatorname{sgn}(\Delta(f)) = (-1)^0 = 1$.
If $\Delta(f) < 0$, by the theorem there is an odd number of complex pairs. There can only be one, so α_1 and α_2 form a complex pair, i.e., $\alpha_1, \alpha_2 \in \mathbb{C} \setminus \mathbb{R}$ and $\alpha_1 = \overline{\alpha_2}$. Conversely, α_1 and α_2 form the only complex pair. The theorem says that $\operatorname{sgn}(\Delta(f)) = (-1)^1 = -1$. □

***Corollary* 1.5:** Let $f \in \mathbb{R}[x]$ be a polynomial of degree 3. Let α_1, α_2 and α_3 be the roots of f. Then the following hold:

- $\Delta(f) = 0$ if and only if $\alpha_i = \alpha_j$ for some $i \neq j$, i.e., there is a multiple root;
- $\Delta(f) > 0$ if and only if $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ and are all distinct;
- $\Delta(f) < 0$ if and only if one root is real and the other two form a complex pair.

Proof. The product $(\alpha_1 - \alpha_2)^2(\alpha_1 - \alpha_3)^2(\alpha_2 - \alpha_3)^2$ equals zero if and only if one of its factors is zero. The latter clearly occurs precisely when two of the roots coincide.
Note that the amount of complex pairs is either zero or one, as f only has three roots.
If $\Delta(f) > 0$, then, by the theorem, there are zero complex pairs, and so $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$. They are distinct as $\Delta(f) \neq 0$. Conversely, if all roots are distinct and in $\mathbb{R}$, then by the theorem $\operatorname{sgn}(\Delta(f)) = (-1)^0 = 1$.
If $\Delta(f) < 0$, then the theorem implies there is one complex pair. This also means that the remaining root is real. Conversely, there is one complex pair. Therefore, by the theorem, $\operatorname{sgn}(\Delta(f)) = (-1)^1 = -1$. □

The fact that the complex roots come in pairs also has another consequence:

***Proposition* 1.6:** Every polynomial $f \in \mathbb{R}[x]$ of odd degree has a root in $\mathbb{R}$.

Proof. Let $n := \deg f$. Since n is odd, there are at most $\frac{n-1}{2}$ pairs of complex roots, not necessarily distinct. This means that there are at most $n-1$ complex roots (counting multiplicity), hence at least one of the roots of f must lie in $\mathbb{R}$. □

The statement of Proposition 1.6 is usually obtained using real analysis by arguing that $f(x)$ approaches $\pm\infty$ as $x \to -\infty$, while $f(x)$ goes to $\mp\infty$ as $x \to \infty$. Then, because f is continuous, the intermediate value theorem implies that f attains the value 0 somewhere, meaning that f has a root in $\mathbb{R}$. In the proof above we have actually obtained the same result algebraically.[1]

1.3 The Resultant

Here we shall define a useful tool, with which we can later derive the desired formula for the discriminant. We will more or less be following Lang's approach (see Chapter IV, §8 of [14]). We jump right into the definition:

Definition 1.7: Let F be a field. Let $x_0, \ldots, x_n, y_0, \ldots, y_m$ be algebraically independent over F, where $n, m \in \mathbb{N}$. We define the *resultant* $\operatorname{Res} \in F[x_0, \ldots, x_n, y_0, \ldots, y_m]$ by

$$\operatorname{Res}(\vec{x},\vec{y}) := \underbrace{\left|\begin{matrix} x_n & x_{n-1} & \cdots & x_0 & & & \\ & x_n & x_{n-1} & \cdots & x_0 & & \\ & & \ddots & \ddots & \ddots & \ddots & \\ & & & x_n & x_{n-1} & \cdots & x_0 \\ y_m & y_{m-1} & \cdots & y_0 & & & \\ & y_m & y_{m-1} & \cdots & y_0 & & \\ & & \ddots & \ddots & \ddots & \ddots & \\ & & & y_m & y_{m-1} & \cdots & y_0 \end{matrix}\right|}_{n+m} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} m \\ \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} n \end{matrix}$$

where the empty spaces are zeros and $\vec{x} = (x_0, \ldots, x_n)$ and $\vec{y} = (y_0, \ldots, y_m)$. If $f, g \in F[z]$, z a variable, and $f(z) = a_n z^n + \cdots + a_1 z + a_0$ and $g(z) = b_m z^m + \cdots + b_1 z + b_0$, then we define $\operatorname{Res}(f,g) := \operatorname{Res}(a_0, \ldots, a_n, b_0, \ldots, b_m)$.

We will now establish a range of properties of the resultant.
Let z be a variable. The first property, which we can deduce directly from the properties of determinants, is homogeneity in $\vec{x}$ and $\vec{y}$ of degree m and n, respectively. To see this, consider $\operatorname{Res}(z\vec{x},\vec{y})$. Then z appears precisely in the first m rows in all non-zero entries of the matrix above. By factoring out one z from each row, we obtain m z's in total. Now all z's have been removed, so $\operatorname{Res}(z\vec{x},\vec{y}) = z^m \operatorname{Res}(\vec{x},\vec{y})$. Similarly, factoring out z from the bottom n rows yields $\operatorname{Res}(\vec{x},z\vec{y}) = z^n \operatorname{Res}(\vec{x},\vec{y})$.
Now define the polynomials $f(z) := x_n z^n + \cdots + x_1 z + x_0$ and $g(z) := y_m z^m + \cdots + y_1 z + y_0$ over $F[\vec{x},\vec{y}]$. We show that $\operatorname{Res}(f,g)$ can be expressed as a linear combination of f and g in $F[\vec{x},\vec{y}][z]$. Let $M(\vec{x},\vec{y})$ be the matrix from Definition 1.7 (thus $\det M(\vec{x},\vec{y}) = \operatorname{Res}(\vec{x},\vec{y})$). Applying this matrix to the vector $\vec{z} := (z^{n+m-1}, \ldots, z, 1)^{\mathsf{T}}$ yields the following:

$$\begin{pmatrix} x_n z^{n+m-1} + x_{n-1} z^{n-1+m-1} + \cdots + x_1 z^{1+m-1} + x_0 z^{m-1} \\ x_n z^{n+m-2} + x_{n-1} z^{n-1+m-2} + \cdots + x_1 z^{1+m-2} + x_0 z^{m-2} \\ \vdots \\ x_n z^n + x_{n-1} z^{n-1} + \cdots + x_1 z + x_0 \\ y_m z^{m+n-1} + y_{m-1} z^{m-1+n-1} + \cdots + y_1 z^{1+n-1} + y_0 z^{n-1} \\ y_m z^{m+n-2} + y_{m-1} z^{m-1+n-2} + \cdots + y_1 z^{1+n-2} + y_0 z^{n-2} \\ \vdots \\ y_m z^m + y_{m-1} z^{m-1} + \cdots + y_1 z + y_0 \end{pmatrix} = \begin{pmatrix} z^{m-1} f(z) \\ z^{m-2} f(z) \\ \vdots \\ f(z) \\ z^{n-1} g(z) \\ z^{n-2} g(z) \\ \vdots \\ g(z) \end{pmatrix}.$$

[1] Observe, however, that we did assume that $\mathbb{C}$ is algebraically closed. We prove this in Theorem 3.30, where we unavoidably use some analysis.

Let $w(z)$ be the vector on the right. Then, more compactly, we have the equality $M(\vec{x},\vec{y})\vec{z} = w(z)$. Write $\tilde{M}(\vec{x},\vec{y},z)$ for the matrix $M(\vec{x},\vec{y})$, whose final column is replaced by $w(z)$. Note that the last component of $\vec{z}$ is 1. Also note that the determinant of $M(\vec{x},\vec{y})$ is non-zero, as the diagonal yields the non-zero term $x_n^m y_0^n$, which does not occur any other way in the determinant. Therefore, by Cramer's rule, we have

$$1 = \frac{\det \tilde{M}(\vec{x},\vec{y},z)}{\det M(\vec{x},\vec{y})}.$$

Since $\operatorname{Res}(\vec{x},\vec{y}) = \det M(\vec{x},\vec{y})$, we see that also $\operatorname{Res}(\vec{x},\vec{y}) = \det \tilde{M}(\vec{x},\vec{y},z)$. By definition, the determinant is a summation of products, which consist of entries of the matrix such that from each column exactly one entry is taken, while no two entries come from the same row. Since each entry in the final column of $\tilde{M}(\vec{x},\vec{y},z)$ contains a power of z and either $f(z)$ or $g(z)$, it follows that each term in the determinant of $\tilde{M}(\vec{x},\vec{y},z)$ contains either $f(z)$ or $g(z)$ (along with a power of z). By factoring out $f(z)$ and $g(z)$, we see that there exist polynomials $u, v \in F[\vec{x},\vec{y}][z]$ such that

$$\operatorname{Res}(\vec{x},\vec{y}) = u(z)f(z) + v(z)g(z). \tag{1.2}$$

Note that $\operatorname{Res}(\vec{x},\vec{y}) \in F[\vec{x},\vec{y}]$, so z completely vanishes.

Remark 1.8: The coefficients $\vec{x}$ and $\vec{y}$ of f and g have thus far been algebraically independent. However, when we wish to apply these results to polynomials over some field F, their coefficients need not be algebraically independent anymore. The reason these results still apply is that we *substitute* the indeterminates with elements of F; this is the same process as the usual *evaluation* of polynomials. More formally, if $\vec{a}$ represents elements $a_0, \ldots, a_n \in F$, then we have the F-homomorphism $\varepsilon_{\vec{a}} : F[\vec{x}] \to F$ defined by $p \mapsto p(\vec{a})$. The fact that this is a homomorphism follows directly from addition and multiplication of polynomials. In particular, $\varepsilon_{\vec{a}}$ sends each x_i to a_i.
To exemplify this, let $p, q \in F[z]$ be polynomials of degree n and m, respectively, with coefficients $\vec{a}$ and $\vec{b}$. Regard f, g, u, v from equation (1.2) as polynomials in $F[z][\vec{x},\vec{y}]$. We see that $\operatorname{Res}(p,q) = \varepsilon_{\vec{a},\vec{b}}(\operatorname{Res})$, $p = \varepsilon_{\vec{a},\vec{b}}(f)$ and $q = \varepsilon_{\vec{a},\vec{b}}(g)$. We also obtain the polynomials $\tilde{u} := \varepsilon_{\vec{a},\vec{b}}(u)$ and $\tilde{v} := \varepsilon_{\vec{a},\vec{b}}(v)$ over F. This now results in equation (1.2) in F:

$$\operatorname{Res}(p,q) = \tilde{u}(z)p(z) + \tilde{v}(z)q(z).$$

We shall not explicitly use the homomorphism $\varepsilon_{\vec{a},\vec{b}}$, but just use this property of substituting values for indeterminates.

Applying this, equation (1.2) yields the following result:

***Proposition* 1.9:** Let $p, q \in F[z]$ be polynomials that have at least one root in common in some extension field of F. Then $\operatorname{Res}(p,q) = 0$.

Proof. From (1.2) we see that there exist polynomials $u, v \in F[z]$—where we substituted the coefficients of p and q for $\vec{x}$ and $\vec{y}$, respectively—such that $\operatorname{Res}(p,q) = u(z)p(z) + v(z)q(z)$. Let α be in some extension field of F such that $p(\alpha) = q(\alpha) = 0$. Then it follows that

$$\operatorname{Res}(p,q) = u(\alpha)p(\alpha) + v(\alpha)q(\alpha) = u(\alpha)\cdot 0 + v(\alpha)\cdot 0 = 0,$$

as desired. □

This proposition actually has a converse, which we shall now work toward. We will need the following two lemmas.

***Lemma* 1.10:** Let R be an integral domain and let $f \in R[x,y]$ be such that $f(x,x) = 0$. Then $y - x$ is a factor of f.

Proof. Regard f as a polynomial of y, so that $f(y) = \sum_{i=0}^n f_i(x)y^i$ with $f_i \in R[x]$ and $n = \deg f$ in $R[x][y]$. Since $f(x) = f(x,x) = 0$, we see that

$$f(y) = f(y) - f(x) = \sum_{i=0}^{n} f_i(x)(y^i - x^i).$$

As $y^i - x^i = (y-x)(y^{i-1} + xy^{i-2} + \cdots + x^{i-2}y + x^{i-1})$, it follows that $y - x$ divides $f(y)$. Thus $y - x$ is a factor of f. □

Before we continue with the second lemma, we first introduce some notation:

Notation: Denote by $\vec{\alpha}_i^{\beta_j}$ the vector $\vec{\alpha}$ with the i-th component replaced by β_j. We extend this notation so that $\vec{\alpha}_i$ means that the i-th entry is deleted entirely.

***Lemma* 1.11:** Let F be a field and let $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m$ be algebraically independent over F. Then $\alpha_i - \beta_j$ is prime in $F[\vec{\alpha}, \vec{\beta}]$ for each i and j. Moreover, if $(i, j) \neq (k, l)$, then $\alpha_i - \beta_j \neq \alpha_k - \beta_l$.

Proof. Suppose there are $a, b \in F[\vec{\alpha}, \vec{\beta}]$ such that $\alpha_i - \beta_j = a(\vec{\alpha}, \vec{\beta})b(\vec{\alpha}, \vec{\beta})$. If we regard a and b as polynomials in $F[\vec{\alpha}_i, \vec{\beta}][\alpha_i]$, then we see that $\deg(ab) = 1$, as $\deg(\alpha_i - \beta_j) = 1$ in $F[\vec{\alpha}_i, \vec{\beta}][\alpha_i]$. Since $a, b \neq 0$ and $F[\vec{\alpha}, \vec{\beta}]$ is an integral domain, we have $\deg a + \deg b = 1$. Hence, without loss of generality, $\deg a = 0$. This means that $a \in F[\vec{\alpha}_i, \vec{\beta}]$. It follows that $\deg b = 1$ and so there are elements $c, d \in F[\vec{\alpha}_i, \vec{\beta}]$ such that $b(\alpha_i) = c\alpha_i + d$. Since $ab(\alpha_i) = \alpha_i - \beta_j$, it follows that $ac = 1$. Therefore, a is a unit. Whence $\alpha_i - \beta_j$ does not have a proper devisor and thus is prime.
To see that $\alpha_i - \beta_j$ and $\alpha_k - \beta_l$ give rise to different primes for $(i, j) \neq (k, l)$, suppose $\alpha_i - \beta_j = \alpha_k - \beta_l$. Note that $i = k$ if and only if $j = l$, so $\alpha_i \neq \alpha_k$ and $\beta_j \neq \beta_l$. Consider the non-zero polynomial $p(z_1, z_2, z_3, z_4) := z_1 - z_2 - z_3 + z_4 \in F[z_1, z_2, z_3, z_4]$. Then $(\alpha_i, \beta_j, \alpha_k, \beta_l)$ is a root of p, contradicting their algebraic independence. Hence $\alpha_i - \beta_j \neq \alpha_k - \beta_l$. □

We shall now derive a formula for the resultant in terms of the roots of the polynomials. We do this by considering the resultant as a function of the roots, for which we need to deem the roots as indeterminates.

***Theorem* 1.12:** Let $\alpha_1, \ldots, \alpha_n, x_n, \beta_1, \ldots, \beta_m, y_m$ be algebraically independent over a field F. Let $\vec{\alpha} := (\alpha_1, \ldots, \alpha_n)$ and $\vec{\beta} := (\beta_1, \ldots, \beta_m)$. Define the polynomials $f_{\vec{\alpha}}(z) := x_n \prod_{i=1}^n (z - \alpha_i) =: x_n z^n + \cdots + x_1 z + x_0$ and $g_{\vec{\beta}}(z) := y_m \prod_{j=1}^m (z - \beta_j) =: y_m z^m + \cdots + y_1 z + y_0$. Then

$$\operatorname{Res}(f_{\vec{\alpha}}, g_{\vec{\beta}}) = x_n^m y_m^n \prod_{i=1}^n \prod_{j=1}^m (\alpha_i - \beta_j) =: P(\vec{\alpha}, x_n, \vec{\beta}, y_m).$$

Proof. By simply expanding the products above, we see that the coefficients x_i $(i < n)$ and y_j $(j < m)$ are polynomials in $F[\vec{\alpha}, x_n]$ and $F[\vec{\beta}, y_m]$, respectively. From this will follow that $x_0, \ldots, x_n, y_0, \ldots, y_m$ are also algebraically independent. Indeed, suppose there is a non-trivial polynomial h such that $h(\vec{x}, \vec{y}) = 0$. Then we get a non-trivial polynomial $\tilde{h}$ with $(\vec{\alpha}, x_n, \vec{\beta}, y_m)$ as a root:

$$\begin{aligned}\tilde{h}(\vec{\alpha}, x_n, \vec{\beta}, y_m) &:= h(x_0(\vec{\alpha}, x_n), \ldots, x_{n-1}(\vec{\alpha}, x_n), x_n, y_0(\vec{\beta}, y_m), \ldots, y_{m-1}(\vec{\beta}, y_m), y_m) \\ &= h(\vec{x}, \vec{y}) = 0,\end{aligned}$$

which contradicts the algebraic independence of $\vec{\alpha}, x_n, \vec{\beta}, y_m$.
We can now deem $\operatorname{Res}(\vec{x}, \vec{y})$ as a polynomial in terms of $\vec{\alpha}, x_n, \vec{\beta}, y_m$. Moreover, since every coefficient x_i has a factor x_n, and every y_i has a factor y_m, by the homogeneity property, we get $\operatorname{Res}(\vec{x}, \vec{y}) = x_n^m y_m^n \operatorname{Res}(\tilde{f}_{\vec{\alpha}}, \tilde{g}_{\vec{\beta}})$, where $\tilde{f}_{\vec{\alpha}} := f_{\vec{\alpha}}/x_n$ and $\tilde{g}_{\vec{\beta}} := g_{\vec{\beta}}/y_m$.
By the foregoing, we can set $R(\vec{\alpha}, \vec{\beta}) := \operatorname{Res}(\tilde{f}_{\vec{\alpha}}, \tilde{g}_{\vec{\beta}})$, as the latter is a polynomial in $F[\vec{\alpha}, \vec{\beta}]$. Observe that for every i and j we have $\tilde{f}_{\vec{\alpha}_i^{\beta_j}}(\beta_j) = \tilde{g}_{\vec{\beta}}(\beta_j) = 0$, hence $R(\vec{\alpha}_i^{\beta_j}, \vec{\beta}) = \operatorname{Res}(\tilde{f}_{\vec{\alpha}_i^{\beta_j}}, \tilde{g}_{\vec{\beta}}) = 0$, by Proposition 1.9. Now view R as an element of $F[\vec{\alpha}_i, \vec{\beta}_j][\beta_j, \alpha_i]$, so that $R(\beta_j, \beta_j) = R(\vec{\alpha}_i^{\beta_j}, \vec{\beta}) = 0$. Since $F[\vec{\alpha}_i, \vec{\beta}_j]$ is an integral domain, we see that Lemma 1.10 implies that $\alpha_i - \beta_j$ is a factor of R for every i and j. Since these factors are all distinct primes by Lemma 1.11, it follows that their product $\prod_{i=1}^n \prod_{j=1}^m (\alpha_i - \beta_j)$ is also a factor of $R(\vec{\alpha}, \vec{\beta})$.
Now observe that

$$P(\vec{\alpha}, x_n, \vec{\beta}, y_m) = x_n^m y_m^n \prod_{i=1}^n \prod_{j=1}^m (\alpha_i - \beta_j) = x_n^m y_m^n \prod_{i=1}^n \tilde{g}_{\vec{\beta}}(\alpha_i) = x_n^m \prod_{i=1}^n g_{\vec{\beta}}(\alpha_i). \tag{1.3}$$

So P can be seen as a polynomial in $\vec{\alpha}, x_n, \vec{y}$. We see that P is then homogeneous in $\vec{y}$ of degree n. On the other hand we have

$$P(\vec{\alpha}, x_n, \vec{\beta}, y_m) = x_n^m y_m^n \prod_{j=1}^m \prod_{i=1}^n (-1)(\beta_j - \alpha_i) = (-1)^{nm} x_n^m y_m^n \prod_{j=1}^m \tilde{f}_{\vec{\alpha}}(\beta_j) = (-1)^{nm} y_m^n \prod_{j=1}^m f_{\vec{\alpha}}(\beta_j).$$

Thus P can also be seen as a polynomial in $\vec{x}, \vec{\beta}, y_m$. It is again clear that P now is homogeneous in $\vec{x}$ of degree m. So we see that P and Res have the same homogeneity properties, and P divides Res. Therefore, Res must be a constant multiple of P. To see this, define the function

$$\varphi(\vec{x}, \vec{y}, \vec{\alpha}, \vec{\beta}) := \frac{2\operatorname{Res}(\vec{x}, \vec{y})}{2P(\vec{\alpha}, x_n, \vec{\beta}, y_m)} = \frac{2\operatorname{Res}(\vec{x}, \vec{y})}{x_n^m \prod_{i=1}^n g_{\vec{\beta}}(\alpha_i) + (-1)^{nm} y_m^n \prod_{j=1}^m f_{\vec{\alpha}}(\beta_j)}.$$

Note that this is a polynomial in $F[\vec{\alpha}, \vec{\beta}]$. Since both the numerator and denominator are homogeneous in $\vec{x}$ and $\vec{y}$ of the same respective degrees, we see that $\varphi(z\vec{x}, w\vec{y}, \vec{\alpha}, \vec{\beta}) = \varphi(\vec{x}, \vec{y}, \vec{\alpha}, \vec{\beta})$, where z and w are variables. This means that φ is constant as a polynomial in $(\vec{x}, \vec{y})$. Except for permuting components, changing $\vec{\alpha}$ and $\vec{\beta}$ implies changing $\vec{x}$ and $\vec{y}$, respectively (and vice versa). This means that φ cannot depend polynomially on $\vec{\alpha}$ and $\vec{\beta}$. But φ is a polynomial, so it must be a constant one. Therefore, there is an $s \in F$ such that $\operatorname{Res}(f_{\vec{\alpha}}, g_{\vec{\beta}}) = sP(\vec{\alpha}, x_n, \vec{\beta}, y_m)$. To determine the value of s, note that from (1.3) we see that P has the term $x_n^m y_0^n$ with coefficient 1. From the determinant definition of the resultant, clearly also $\operatorname{Res}(f_{\vec{\alpha}}, g_{\vec{\beta}})$ has $x_n^m y_0^n$ with coefficient 1, which comes from the product of the entries on the diagonal. We conclude that $s = 1$. □

***Corollary* 1.13:** Let f and g be two polynomials over a field F, over which they split. Then $\operatorname{Res}(f, g) = 0$ if and only if f and g have a common root in F.

Proof. Since f and g split, we may write $f(z) = a_n(z-\alpha_1)\cdots(z-\alpha_n)$ and $g(z) = b_m(z-\beta_1)\cdots(z-\beta_m)$, where $a_n, b_m, \alpha_i, \beta_j \in F$, $i = 1, \dots, n$, $j = 1, \dots, m$ and $a_n, b_m \neq 0$, and $n = \deg f$ and $m = \deg g$.
Assume $\operatorname{Res}(f, g) = 0$. By substituting these roots for the indeterminates in Theorem 1.12, we obtain $a_n^m b_m^n \prod_{i=1}^n \prod_{j=1}^m (\alpha_i - \beta_j) = 0$. Since $a_n^m b_m^n \neq 0$, one of the factors $\alpha_i - \beta_j$ equals 0. This means that $\alpha_i = \beta_j$ for some pair (i, j), wherefore f and g have a common root.
The converse is Proposition 1.9. □

1.4 A Formula for the Discriminant

As promised earlier we shall have derived a general formula for the discriminant of a polynomial f in terms of its coefficients by the end of this section. We start with a simple example of the formula for the discriminant of a quadratic polynomial.

Example 1.14: Let F be a field and $ax^2 + bx + c \in F[x]$ be a polynomial, $a \neq 0$. We shall see that the discriminant equals $b^2 - 4ac$. Let α_1 and α_2 be the roots of f in some extension field over F. Then f can be written as $f(x) = a(x - \alpha_1)(x - \alpha_2) = ax^2 - a(\alpha_1 + \alpha_2)x + a\alpha_1\alpha_2$. From this we obtain that $-a(\alpha_1 + \alpha_2) = b$ and $a\alpha_1\alpha_2 = c$. Now, from the definition of the discriminant, we compute

$$\begin{aligned}\Delta(f) &= a^{2\cdot 2-2}(\alpha_1 - \alpha_2)^2 \\ &= a^2((\alpha_1 + \alpha_2)^2 - 4\alpha_1\alpha_2) \\ &= (-a(\alpha_1 + \alpha_2))^2 - 4a(a\alpha_1\alpha_2) \\ &= b^2 - 4ac,\end{aligned} \tag{1.4}$$

which is indeed the well-known formula for the discriminant of a quadratic polynomial.

For this simple example we already needed to use a little 'trick' in the second line of (1.4). It seems likely that this becomes very complicated very quickly for discriminants of higher degree polynomials. It is not even clear whether this is possible! Fortunately, it is and we shall use resultants to show this.
The formal derivative of a polynomial plays a part in the formula. Over $\mathbb{R}$ this is the same as the usual derivative from calculus. More generally, we define it as follows:

Definition 1.15: Let F be a field and let $f(x) := \sum_{i=0}^n a_i x^i \in F[x]$, where $n \in \mathbb{N}_0$. We define the *(formal) derivative of* f, denoted by f' or Df, by

$$f'(x) := Df(x) := \sum_{i=1}^n i a_i x^{i-1}.$$

If $n = 0$, this is the empty sum, which equals 0.

We verify that we have the standard rules for differentiation, namely

$$\begin{aligned} \text{the `sum rule': } (f+g)' &= f' + g' \text{ and} \\ \text{the `product rule': } (fg)' &= f'g + fg', \end{aligned}$$

for any $f, g \in F[x]$, F a field. Indeed, let $f(x) := \sum_{i=0}^{n} a_i x^i$ and $g(x) := \sum_{j=0}^{m} b_j x^j$ be polynomials over F. Assume, without loss of generality, that $n \geq m$ and set $b_j := 0$ for $j > m$. Then we have

$$\begin{aligned} D(f+g) &= D\left(\sum_{i=0}^{n} a_i x^i + \sum_{j=0}^{m} b_j x^j\right) = D\sum_{i=0}^{n}(a_i + b_i)x^i \\ &= \sum_{i=1}^{n} i(a_i + b_i)x^{i-1} = \sum_{i=1}^{n} i a_i x^{i-1} + \sum_{j=1}^{m} j b_j x^{j-1} \\ &= Df + Dg. \end{aligned}$$

By repeated application, this also holds for any finite summation of polynomials. We can now derive the product rule directly:

$$\begin{aligned} D(fg) &= D\left(\left(\sum_{i=0}^{n} a_i x^i\right)\left(\sum_{j=0}^{m} b_j x^j\right)\right) = D\sum_{i=0}^{n}\sum_{j=0}^{m} a_i b_j x^{i+j} \\ &= \sum_{i=0}^{n} D\sum_{j=0}^{m} a_i b_j x^{i+j} = \sum_{i=0}^{n}\sum_{j=0}^{m}(i+j)a_i b_j x^{i+j-1} \\ &= \sum_{i=1}^{n} i a_i x^{i-1}\sum_{j=0}^{m} b_j x^j + \sum_{i=0}^{n} a_i x^i \sum_{j=1}^{m} j b_j x^{j-1} \\ &= (Df)g + fDg. \end{aligned}$$

We also have the following property of derivatives, related to multiple roots:

***Lemma* 1.16:** Let $f \in F[x]$ be a polynomial, F a field. Suppose f has a root α in some extension field E of F. Then α is a multiple root if and only if $f'(\alpha) = 0$.

Proof. For the direct implication, we can factor out $(x-\alpha)^2$. So there exists a polynomial $g \in E[x]$ such that $f(x) = (x-\alpha)^2 g(x)$. By the product rule for derivatives, we have

$$\begin{aligned} f'(x) &= ((x-\alpha)^2)' g(x) + (x-\alpha)^2 g'(x) \\ &= (x^2 - 2\alpha x + \alpha^2)' g(x) + (x-\alpha)^2 g'(x) \\ &= 2(x-\alpha)g(x) + (x-\alpha)^2 g'(x). \end{aligned}$$

From this we see that $f'(\alpha) = 0$.
Conversely, suppose $f'(\alpha) = 0$. There exists a polynomial $g \in E[x]$ such that $f(x) = (x-\alpha)g(x)$. Thus we compute directly:

$$f'(x) = D((x-\alpha)g(x)) = g(x) + (x-\alpha)g'(x).$$

Since $f'(\alpha) = 0$, we see that $g(\alpha) = 0$. Hence there exists an $h \in E[x]$ such that $g(x) = (x-\alpha)h(x)$. Consequently, $f(x) = (x-\alpha)^2 h(x)$. □

At last, we have arrived at the following formula for the discriminant:

***Theorem* 1.17:** Let $f \in F[x]$ be a polynomial of degree $n \geq 2$, F a field with $\operatorname{Char} F \nmid n$. Let a_n be the leading coefficient of f. Then

$$\Delta(f) = (-1)^{\frac{1}{2}n(n-1)} a_n^{-1} \operatorname{Res}(f, f'). \tag{1.5}$$

Proof. First consider the case where f has a multiple root. Then $\Delta(f) = 0$. By Lemma 1.16, f and f' have a root in common, hence the right-hand side of (1.5) is also zero for Proposition 1.9. Whence the formula holds.

Now suppose all roots of f are distinct. Then $\Delta(f) \neq 0$. Lemma 1.16, together with Corollary 1.13, implies that the right-hand side of (1.5) is also non-zero. Let E be a splitting field of f over F. Let $\alpha_1, \ldots, \alpha_n \in E$ be the roots of f. By (1.3) and Theorem 1.12, we see that

$$\operatorname{Res}(f, f') = a_n^{n-1} \prod_{i=1}^{n} f'(\alpha_i). \tag{1.6}$$

Write $f(x) = a_n(x - \alpha_1) \cdots (x - \alpha_n)$. We compute the derivative of f:

$$\begin{aligned}
f'(x) &= a_n(D(x-\alpha_1)) \prod_{j=2}^{n} (x - \alpha_j) + a_n(x - \alpha_1) D \prod_{j=2}^{n} (x - \alpha_j) \\
&= a_n \prod_{j=2}^{n} (x - \alpha_j) + a_n(x-\alpha_1)\left((D(x-\alpha_2)) \prod_{j=3}^{n} (x-\alpha_j) + (x-\alpha_2) D \prod_{j=3}^{n} (x - \alpha_j)\right) \\
&= a_n \prod_{\substack{j=1 \\ j\neq 1}}^{n} (x - \alpha_j) + a_n \prod_{\substack{j=1 \\ j\neq 2}}^{n} (x - \alpha_j) + a_n(x-\alpha_1)(x-\alpha_2) D \prod_{j=3}^{n} (x - \alpha_j) \\
&\vdots \\
&= a_n \prod_{\substack{j=1 \\ j\neq 1}}^{n} (x - \alpha_j) + \cdots + a_n \prod_{\substack{j=1 \\ j\neq n-1}}^{n} (x - \alpha_j) + a_n \left(\prod_{j=1}^{n-1} (x - \alpha_j) \right) D(x - \alpha_n) \\
&= a_n \sum_{i=1}^{n} \prod_{\substack{j=1 \\ j\neq i}}^{n} (x - \alpha_j).
\end{aligned}$$

Note that $\deg f' = n - 1$, since $\operatorname{Char} F \nmid n$. Now we see that for each i, there is exactly one term in the sum that does not contain the factor $x - \alpha_i$. Hence if we substitute α_i for x, then all terms bar one vanish. Consequently, $f'(\alpha_i) = a_n \prod_{\substack{j=1 \\ j\neq i}}^{n} (\alpha_i - \alpha_j)$. Combining this with (1.6), we obtain

$$\begin{aligned}
\operatorname{Res}(f, f') &= a_n^{n-1} \prod_{i=1}^{n} a_n \prod_{\substack{j=1 \\ j\neq i}}^{n} (\alpha_i - \alpha_j) \\
&= a_n^{2n-1} \prod_{i \neq j} (\alpha_i - \alpha_j).
\end{aligned}$$

By (1.1) in the proof of Proposition 1.2, we now see that

$$(-1)^{\frac{1}{2}n(n-1)} a_n^{-1} \operatorname{Res}(f, f') = (-1)^{\frac{1}{2}n(n-1)} a_n^{2n-2} \prod_{i \neq j} (\alpha_i - \alpha_j) = \Delta(f),$$

which is the desired formula. □

As mentioned in section 1.1, we immediately have the following (already proven) consequence:

***Corollary* 1.18:** Let $f \in F[x]$ be a polynomial of degree $n \geq 2$, F a field. Then $\Delta(f) \in F$. □

We end this section with an example, where we again find the discriminant of a quadratic polynomial.

Example 1.19: Let F be a field with $\operatorname{Char} F \neq 2$ and let $f(x) := ax^2 + bx + c \in F[x]$ be a polynomial, $a \neq 0$. Note that $f'(x) = 2ax + b$. We compute the discriminant using Theorem 1.17:

$$\begin{aligned}
\Delta(f) &= (-1)^{\frac{1}{2}\cdot 2 \cdot 1} a^{-1} \begin{vmatrix} a & b & c \\ 2a & b & 0 \\ 0 & 2a & b \end{vmatrix} \\
&= -a^{-1}(ab^2 + 4a^2c - 2ab^2) \\
&= b^2 - 4ac.
\end{aligned}$$

This is indeed the same formula we obtained in Example 1.14.

2 Roots of Third Degree Polynomials

Given a polynomial $f(x) \coloneqq x^3 + px + q$ over a field F, not of characteristic 2 or 3, we have the famous formula of Cardano for the roots of f. This formula dates back to the sixteenth century, hence the methods of deriving it are quite old as well. In this chapter we shall give a more modern approach, using Galois theory, to derive the formula that has been known for so long. This approach is based on the notes from [7].
After this we shall also examine the cases where the characteristic of F equals 2 or 3, using the ideas from [17]. This is where the aforementioned formula fails. The formulas we shall derive are in terms of the usual radicals, which are roots of $x^n - a$ over F, where $\operatorname{Char} F \nmid n$, and radicals that are roots of $x^p - x - a$ over F, where $p = \operatorname{Char} F$. These are the radicals that are normally considered when one speaks of 'solvability by radicals'.

Our general approach is as follows. Firstly, we consider a polynomial $ax^3 + bx^2 + cx + d \in F[x]$ with $a \neq 0$. To obtain the roots, we need to solve the equation $ax^3 + bx^2 + cx + d = 0$ for the indeterminate x. We can always divide by the leading coefficient, hence we shall only consider monic polynomials. Next we use some kind of substitution to obtain a polynomial of the form $f(x) \coloneqq x^3 + px + q$ over F. This is called the *depressed cubic*. To figure out what kinds of radicals will appear in the final formulas, we wish to determine what the splitting field E of f over F looks like. Determining the Galois group of E over F helps with this. Once we have established which radicals we may need, we work out the entire formula itself.

2.1 Deriving Cardano's Formula

Let F be a field, not of characteristic 2 or 3. Consider a cubic polynomial $y^3 + ay^2 + by + c \in F[y]$. To get rid of the quadratic term, we apply the substitution $y = x - \frac{1}{3}a$. Indeed, using the binomial expansion, we get a term $-ax^2$ from $(x - \frac{1}{3}a)^3$, which cancels the term ax^2, obtained from $a(x - \frac{1}{3}a)^2$. We now have the depressed cubic polynomial

$$f(x) \coloneqq x^3 + px + q,$$

where

$$p = b - \frac{1}{3}a^2, \qquad q = \frac{2}{27}a^3 - \frac{1}{3}ab + c.$$

It suffices to find the roots of f: If α satisfies $f(\alpha) = 0$, then $\alpha + \frac{1}{3}a$ is a root of the original polynomial. Let ω denote a primitive third root of unity over F. This exists by Theorem 0.9, as $\operatorname{Char} F \nmid 3$. We shall henceforth assume that $\omega \in F$. This is justified, as we can simply replace F by $F(\omega)$. Assume f is irreducible over F. Denote by E the splitting field of f over F and let $\alpha_1, \alpha_2, \alpha_3 \in E$ be the roots of f. Since f can be written as $f(x) = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3)$ over E, by expanding this product, we obtain:

$$f(x) = x^3 - (\alpha_1 + \alpha_2 + \alpha_3)x^2 + (\alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3)x - \alpha_1\alpha_2\alpha_3.$$

By matching up the coefficients, we get the following equalities:

$$\begin{aligned} \alpha_1 + \alpha_2 + \alpha_3 &= 0, \\ \alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3 &= p, \\ \alpha_1\alpha_2\alpha_3 &= -q. \end{aligned} \tag{2.1}$$

This will be useful later.
Let us investigate what E looks like. By Theorem 0.11, f is separable over F. Indeed, if $\operatorname{Char} F = 0$, we are done. Otherwise, $\operatorname{Char} F > 3$. Since $\deg f = 3$, there is clearly no polynomial $g \in F[x]$ such that $f(x) = g(x^{\operatorname{Char} F})$. Thus E is the splitting field of an irreducible, separable polynomial. Whence E/F is finite Galois by Theorem 0.10. Let $\Delta \coloneqq \Delta(f)$ and

$$\delta \coloneqq \prod_{i<j} (\alpha_i - \alpha_j)$$

so that $\delta^2 = \Delta$. We now have the following theorem about the Galois group of E over F.

***Theorem* 2.1:** Let F be a field, not of characteristic 2 or 3. Let $f \in F[x]$ be irreducible and separable and let E/F be its splitting field. Then the Galois group of E over F is isomorphic to S_3 if $\delta \notin F$ and isomorphic to A_3 if $\delta \in F$. Furthermore, if F contains a primitive third root of unity, then $E = F(\delta)(\xi)$, where ξ is a cube root over $F(\delta)$.

Proof. Since $\deg f = 3$, by Theorem 0.12, the Galois group $\mathrm{Gal}(E/F)$ is a subgroup of the symmetric group S_3. By Theorem 0.13, the degree $[E : F]$ of E over F is a divisor of $6 = |S_3|$. Because f is irreducible, we have $[F(\alpha_i) : F] = 3$, where $i \in \{1, 2, 3\}$. Obviously $F(\alpha_i) \subseteq E$, so $[E : F] \in \{3, 6\}$.
First consider the case where $\delta \notin F$. Then clearly $x^2 - \Delta \in F[x]$ is irreducible and separable, because $\delta \neq -\delta$ as $\mathrm{Char}\, F \neq 2$ and $\delta \neq 0$. Hence the splitting field $F(\delta)$ of this polynomial is a finite Galois extension over F by Theorem 0.10. Thus $[F(\delta) : F] = 2$. Clearly $\delta \in E$, as it consists of multiplication and addition of the roots of f, so $F \subset F(\delta) \subset E$. We saw that $[E : F] \neq 2$, so this last inclusion is indeed strict. Since now $[E : F] = [E : F(\delta)][F(\delta) : F] = 2[E : F(\delta)] \in \{3, 6\}$, we see that $[E : F] = 6$. Therefore, $\mathrm{Gal}(E/F) \cong S_3$. We have also obtained that $[E : F(\delta)] = 3$.
Now suppose $\delta \in F$. We already observed that $E \supseteq F(\alpha_1) \supset F$. Now consider $\mathrm{Gal}(E/F(\alpha_1))$. Since $[F(\alpha_1) : F] = 3$, this group can only have order 1 or 2. Indeed, the $F(\alpha_1)$-automorphisms on E can either be the identity, or the map τ that swaps α_2 and α_3. Suppose $\tau \in \mathrm{Gal}(E/F(\alpha_1))$. Then $\tau(\delta) = \delta$ as $\delta \in F$. But we also have

$$\tau(\delta) = \tau((\alpha_1 - \alpha_2)(\alpha_1 - \alpha_3)(\alpha_2 - \alpha_3)) = (\alpha_1 - \alpha_3)(\alpha_1 - \alpha_2)(\alpha_3 - \alpha_2) = -\delta.$$

Consequently, $2\delta = 0$. Since $\mathrm{Char}\, F \neq 2$, we must have $\delta = 0$. But this contradicts the fact that f is separable. Hence $\tau \notin \mathrm{Gal}(E/F(\alpha_1))$, which means $|\mathrm{Gal}(E/F(\alpha_1))| = 1$ and so $E = F(\alpha_1)$. Hence $[E : F] = 3$.
Since in the latter case $F = F(\delta)$, we have now obtained that $[E : F(\delta)] = 3$. The only subgroup of S_3 of order 3 is the alternating group A_3. Hence $\mathrm{Gal}(E/F(\delta)) \cong A_3$, which is cyclic of order 3. Assume now that F contains a primitive third root of unity. Since $\mathrm{Char}\, F \nmid 3$, Theorem 0.17 implies that $E = F(\delta)(\xi)$ with ξ a cube root over $F(\delta)$. □

The theorem applies immediately to E and f. We saw that $\mathrm{Gal}(E/F(\delta)) \cong A_3$, which is cyclic. So let σ be a generator of $\mathrm{Gal}(E/F(\delta))$, which we can choose to correspond to the permutation $(1\,2\,3)$ on the indices of the roots α_1, α_2 and α_3. Define

$$\begin{aligned} z_1 &:= \alpha_1 + \omega^2\alpha_2 + \omega\alpha_3, \\ z_2 &:= \alpha_1 + \omega\alpha_2 + \omega^2\alpha_3. \end{aligned}$$

Then we see that $\sigma(z_1) = \omega z_1$ and $\sigma(z_2) = \omega^2 z_2$. It now follows that σ fixes z_1^3 and z_2^3. Since σ generates $\mathrm{Gal}(E/F(\delta))$, all elements of this Galois group fix z_1^3 and z_2^3. This means that z_1^3 and z_2^3 are in the fixed field of $\mathrm{Gal}(E/F(\delta))$ in E. By the fundamental theorem of Galois theory, this fixed field is $F(\delta)$.
Since ω is a root of $x^3 - 1 = (x - 1)(x^2 + x + 1)$, we see that $\omega^2 + \omega + 1 = 0$. Now observe

$$\begin{aligned} z_1 + z_2 &= z_1 + z_2 + \alpha_1 + \alpha_2 + \alpha_3 \\ &= 3\alpha_1 + (\omega^2 + \omega + 1)(\alpha_2 + \alpha_3) \\ &= 3\alpha_1. \end{aligned}$$

Therefore, $\alpha_1 = \frac{1}{3}(z_1 + z_2)$. Similarly,

$$\begin{aligned} \omega z_1 + \omega^2 z_2 &= 3\alpha_2 + (\omega^2 + \omega + 1)(\alpha_1 + \alpha_3) = 3\alpha_2, \\ \omega^2 z_1 + \omega z_2 &= 3\alpha_3 + (\omega^2 + \omega + 1)(\alpha_1 + \alpha_2) = 3\alpha_3, \end{aligned}$$

and so $\alpha_2 = \frac{1}{3}(\omega z_1 + \omega^2 z_2)$ and $\alpha_3 = \frac{1}{3}(\omega^2 z_1 + \omega z_2)$. Since we know that $z_1^3, z_2^3 \in F(\delta)$, we shall now try to find elements $a_j, b_j \in F$ so that $z_j^3 = a_j + b_j\delta$ for $j = 1, 2$. We do the computation only for z_1, as for z_2 it will be very similar. We expand

$$\begin{aligned} z_1^3 &= (\alpha_1 + \omega^2\alpha_2 + \omega\alpha_3)^3 \\ &= \alpha_1^3 + \omega^6\alpha_2^3 + \omega^3\alpha_3^3 + 3\alpha_1^2(\omega^2\alpha_2 + \omega\alpha_3) + 3\omega^4\alpha_2^2(\alpha_1 + \omega\alpha_3) + 3\omega^2\alpha_3^2(\alpha_1 + \omega^2\alpha_2) + 6\omega^3\alpha_1\alpha_2\alpha_3 \\ &= (\alpha_1^3 + \alpha_2^3 + \alpha_3^3) + 6\alpha_1\alpha_2\alpha_3 + 3\omega(\alpha_1^2\alpha_3 + \alpha_1\alpha_2^2 + \alpha_2\alpha_3^2) + 3\omega^2(\alpha_1^2\alpha_2 + \alpha_2^2\alpha_3 + \alpha_1\alpha_3^2). \qquad (2.2) \end{aligned}$$

We tackle each term individually. Write

$$\begin{aligned} t_1 &\coloneqq \alpha_1^2\alpha_3 + \alpha_1\alpha_2^2 + \alpha_2\alpha_3^2, \\ t_2 &\coloneqq \alpha_1^2\alpha_2 + \alpha_2^2\alpha_3 + \alpha_1\alpha_3^2 \end{aligned}$$

so that

$$z_1^3 = (\alpha_1^3 + \alpha_2^3 + \alpha_3^3) + 6\alpha_1\alpha_2\alpha_3 + 3\omega t_1 + 3\omega^2 t_2.$$

We compute the first term. The expansion of $(\alpha_1 + \alpha_2 + \alpha_3)^3$ follows from the same computation as z_1^3 with ω and ω^2 both replaced by 1. Using this expansion, we obtain:

$$\begin{aligned} \alpha_1^3 + \alpha_2^3 + \alpha_3^3 &= (\alpha_1 + \alpha_2 + \alpha_3)^3 - 3t_1 - 3t_2 - 6\alpha_1\alpha_2\alpha_3 \\ &= -3(\alpha_1^2\alpha_2 + \alpha_1^2\alpha_3 + \alpha_1\alpha_2^2 + \alpha_2^2\alpha_3 + \alpha_1\alpha_3^2 + \alpha_2\alpha_3^2) + 6q \\ &= -3[(\alpha_1 + \alpha_2 + \alpha_3)(\alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3) - 3\alpha_1\alpha_2\alpha_3] + 6q \\ &= 9\alpha_1\alpha_2\alpha_3 + 6q = -3q. \end{aligned} \tag{2.3}$$

The term $6\alpha_1\alpha_2\alpha_3$ of (2.2) is simply $-6q$. For the last two terms, we first expand δ:

$$\begin{aligned} \delta &= (\alpha_1 - \alpha_2)(\alpha_1 - \alpha_3)(\alpha_2 - \alpha_3) \\ &= \alpha_1^2\alpha_2 - \alpha_1^2\alpha_3 - \alpha_1\alpha_2\alpha_3 + \alpha_1\alpha_3^2 - \alpha_1\alpha_2^2 + \alpha_1\alpha_2\alpha_3 + \alpha_2^2\alpha_3 - \alpha_2\alpha_3^2 \\ &= \alpha_1^2\alpha_2 + \alpha_1\alpha_3^2 + \alpha_2^2\alpha_3 - (\alpha_1^2\alpha_3 + \alpha_1\alpha_2^2 + \alpha_2\alpha_3^2). \end{aligned} \tag{2.4}$$

Here we recognize the last line to be $t_2 - t_1$. From (2.3) we infer $-3(t_1 + t_2) = 6\alpha_1\alpha_2\alpha_3 - 3q$ and hence $t_1 + t_2 = 3q$. We now have two simultaneous equations:

$$\begin{cases} t_2 - t_1 = \delta, \\ t_2 + t_1 = 3q. \end{cases} \tag{2.5}$$

From the first one, we obtain $t_1 = t_2 - \delta$. Substituting this into the second one yields $2t_2 - \delta = 3q$ and so $t_2 = \frac{1}{2}(3q + \delta)$. Now we see that $t_1 = \frac{1}{2}(3q - \delta)$.
All terms have been treated and we can now rewrite (2.2) as

$$\begin{aligned} z_1^3 &= -3q - 6q + \frac{3}{2}\omega(3q - \delta) + \frac{3}{2}\omega^2(3q + \delta) \\ &= -9q + \frac{9}{2}q(\omega + \omega^2) + \frac{3}{2}\delta(\omega^2 - \omega) \\ &= -\frac{27}{2}q + \frac{3}{2}\delta(\omega^2 - \omega). \end{aligned}$$

Here we used that $1 + \omega + \omega^2 = 0$, which yields $\omega + \omega^2 = -1$. An analogous computation shows that $z_2^3 = -\frac{27}{2}q - \frac{3}{2}\delta(\omega^2 - \omega)$. Now choose a cube root of the polynomial $x^3 - z_1^3 \in F(\delta)[x]$ such that

$$z_1 = \sqrt[3]{-\frac{27}{2}q + \frac{3}{2}\delta(\omega^2 - \omega)}.$$

Since the roots are z_1, ωz_1 and $\omega^2 z_1$, choosing a different root amounts to permuting the roots of f. Indeed, choosing ωz_1 instead of z_1 is just applying σ to z_1, which is permuting the roots of f according to $(1\,2\,3)$. Similarly, if $\delta \notin F$, choosing to adjoin $-\delta$ to F instead of δ results in swapping z_1 and z_2. This ultimately boils down to swapping α_2 and α_3, which corresponds to the permutation $(2\,3)$. Note that $\{(1\,2\,3), (2\,3)\}$ generates S_3, which we expected the Galois group of E/F to be when $\delta \notin F$.

Remark 2.2: When using n-th root symbols such as $\sqrt{}$ and $\sqrt[3]{}$, it is generally unclear *which* root is meant. We customarily choose one particular root to be represented by such a symbol and henceforth related roots are chosen such that they commute with multiplication. For example, let $x^3 - a$ and $x^3 - b$ be polynomials over F. Let $\sqrt[3]{a}$ and $\sqrt[3]{b}$ denote any one of their roots, respectively. We note that $\sqrt[3]{a}\sqrt[3]{b}$ is a root of $x^3 - ab$. We shall then, implicitly, let $\sqrt[3]{ab}$ denote the root $\sqrt[3]{a}\sqrt[3]{b}$. Since $\omega \in F$, all these roots are in $F(\sqrt[3]{a}, \sqrt[3]{b})$, hence no issues arise.

Since there are three possibilities for choosing a cube root of z_1^3 and another three for z_2^3, there are a total of nine combinations. However, since f only has three roots, merely three of these combinations are correct. Fortunately, we have the following, so that we can choose the correct combinations:

$$\begin{aligned} z_1z_2 &= (\alpha_1+\omega^2\alpha_2+\omega\alpha_3)(\alpha_1+\omega\alpha_2+\omega^2\alpha_3) \\ &= \alpha_1^2+\alpha_2^2+\alpha_3^2+(\omega+\omega^2)(\alpha_1\alpha_2+\alpha_1\alpha_3+\alpha_2\alpha_3) \\ &= \alpha_1^2+\alpha_2^2+\alpha_3^2-p \\ &= (\alpha_1+\alpha_2+\alpha_3)^2-2\alpha_1\alpha_2-2\alpha_1\alpha_3-2\alpha_2\alpha_3-p \\ &= -3p. \end{aligned} \tag{2.6}$$

This restricts our choice for a cube root of z_2^3 to one, once we have chosen one for z_1^3.

Remark 2.3: It is not immediately clear that we can still choose the cube roots correctly if $p=0$. The following happens: If $p=0$, then either $z_1=0$ or $z_2=0$. Indeed, assuming $p=0$, f equals x^3+q. Now observe that α_1, $\omega\alpha_1$ and $\omega^2\alpha_1$ are three distinct roots of f (recall that f is irreducible, so $q\neq 0$). Thus, say, $\alpha_2=\omega\alpha_1$ and $\alpha_3=\omega^2\alpha_1$. Then $z_1=\alpha_1+\omega^2(\omega\alpha_1)+\omega(\omega^2\alpha_1)=3\alpha_1$ and $z_2=\alpha_1+\omega(\omega\alpha_1)+\omega^2(\omega^2\alpha_1)=(1+\omega+\omega^2)\alpha_1=0$. Thus the only choice for the cube root of z_2^3 is 0. Therefore, any choice for a cube root of z_1^3 is correct. Thus, if $z_1z_2=0$, we can simply choose any cube root for one of the z_j^3's and let the other be 0. So the relation $z_1z_2=-3p$ still works in this case.

Finally, we will clean up the formula so that it be expressed only in terms of p and q. First we rid the formulas for z_j of the primitive third roots of unity. Observe that

$$(\omega^2-\omega)^2=\omega^4-2\omega^3+\omega^2=\omega+\omega^2-2=-3.$$

We now see that $\delta(\omega^2-\omega)$ is a root of $x^2+3\Delta$. Write $\sqrt{-3\Delta}$ for this root. Next we wish to express Δ in terms of p and q, which can be done by Theorem 1.17. First note that $f'(x)=3x^2+p$. The theorem provides the formula for the discriminant:

$$\begin{aligned} \Delta &= (-1)^{\frac{1}{2}\cdot 3\cdot 2}\cdot 1^{-1}\cdot \operatorname{Res}(f,f') \\ &= -\begin{vmatrix} 1 & 0 & p & q & 0 \\ 0 & 1 & 0 & p & q \\ 3 & 0 & p & 0 & 0 \\ 0 & 3 & 0 & p & 0 \\ 0 & 0 & 3 & 0 & p \end{vmatrix} \\ &= -\begin{vmatrix} 1 & 0 & p & q \\ 0 & p & 0 & 0 \\ 3 & 0 & p & 0 \\ 0 & 3 & 0 & p \end{vmatrix} - 3\begin{vmatrix} 0 & p & q & 0 \\ 1 & 0 & p & q \\ 3 & 0 & p & 0 \\ 0 & 3 & 0 & p \end{vmatrix} \\ &= -p\begin{vmatrix} 1 & p & q \\ 3 & p & 0 \\ 0 & 0 & p \end{vmatrix} - 3\left(-\begin{vmatrix} p & q & 0 \\ 0 & p & 0 \\ 3 & 0 & p \end{vmatrix} + 3\begin{vmatrix} p & q & 0 \\ 0 & p & q \\ 3 & 0 & p \end{vmatrix}\right) \\ &= -p(p^2-3p^2)+3p^3-9(p^3+3q^2) \\ &= -4p^3-27q^2. \end{aligned}$$

If we now pick a cube root of z_1^3 and let the cube root of z_2^3 be determined by $z_1z_2=-3p$, we can combine this all to find Cardano's formula for all the roots of f, expressed in terms of its coefficients:

$$\begin{aligned} \alpha_i &= \frac{1}{3}(\omega^{i-1}z_1+\omega^{1-i}z_2) \\ &= \frac{\omega^{i-1}}{3}\sqrt[3]{-\frac{27}{2}q+\frac{3}{2}\sqrt{-3\Delta}}+\frac{\omega^{1-i}}{3}\sqrt[3]{-\frac{27}{2}q-\frac{3}{2}\sqrt{-3\Delta}}, \qquad i=1,2,3. \end{aligned}$$

Remark 2.4: The formula holds for *any* depressed cubic f, not just irreducible, separable ones. If f is separable, the elements α_i above are clearly distinct roots. If $\Delta(f)=0$, then $\alpha_2=\alpha_3$ yields a root with multiplicity 2. Now $p=0$ if and only if $q=0$. If they are both zero, clearly 0 is a triple root, which the formula supplies. If they are non-zero, then $\alpha_2=\frac{1}{2}(\omega+\omega^2)\alpha_1\neq\alpha_1$ and so the formula still provides all roots of f.

Remark 2.5: We only needed the primitive third root of unity to be in the field for the derivation of the formula; not for using the formula.

Lastly, we will illustrate Cardano's formula with an example.

Example 2.6: Let $f(x) \coloneqq x^3 + 9x - 6 \in \mathbb{Q}[x]$. It is already of the depressed cubic form, so we can apply Cardano's formula with $p = 9$ and $q = -6$. Firstly, we compute the discriminant: $\Delta(f) = -4 \cdot 9^3 - 27 \cdot (-6)^2 = -3888$. From Corollary 1.5 we see that we will get one real root and two complex ones. Now $-3\Delta(f) = 11\ 664 = 108^2$. We take $\sqrt{-3\Delta(f)} = 108$. The radicands of the cube roots now become $\frac{27 \cdot 6}{2} \pm \frac{3}{2} \cdot 108 = 81 \pm 162$, hence $243 = 3^5$ and $-81 = -3^4$, respectively. Let $\sqrt[3]{243}$ and $\sqrt[3]{-81}$ denote the real cube roots. We then have

$$\sqrt[3]{243}\,\sqrt[3]{-81} = 3\sqrt[3]{9} \cdot -3\sqrt[3]{3} = -3^2\,\sqrt[3]{9 \cdot 3} = -9\sqrt[3]{27} = -3 \cdot 9.$$

Hence these roots are a correctly chosen combination. Note that $\sqrt[3]{27} = 3$, because this is the only real cube root of 27. The formula now yields the roots $\sqrt[3]{9} - \sqrt[3]{3}$, $\omega\sqrt[3]{9} - \omega^2\sqrt[3]{3}$ and $\omega^2\sqrt[3]{9} - \omega\sqrt[3]{3}$. They lie in the field $\mathbb{Q}(\omega, \sqrt[3]{3})$.

2.2 A Formula in Characteristic 2

Let F be a field of characteristic 2. Because $2 \nmid 3$, by Theorem 0.9, there exists a primitive third root of unity ω in some extension field of F. As in the previous section, we assume that $\omega \in F$. Let $f(x) \coloneqq x^3 + px + q \in F[x]$ be irreducible over F. To derive the roots, it suffices to consider only polynomials of the depressed form, which we saw in the previous section. Indeed, if we start with a polynomial $y^3 + ay^2 + by + c \in F[y]$, the same substitution works in characteristic 2: Substituting $y = x - \frac{1}{3}a$ now boils down to $y = x + a$. We then obtain f with $p = a^2 + b$ and $q = ab + c$.
Let E be the splitting field of f over F and let $\alpha_1, \alpha_2, \alpha_3 \in E$ be the roots of f. By expanding $(x - \alpha_1)(x - \alpha_2)(x - \alpha_3)$, we obtain the equalities (2.1).
We again define the following quantities:

$$\begin{aligned}
z_1 &\coloneqq \alpha_1 + \omega^2\alpha_2 + \omega\alpha_3, \\
z_2 &\coloneqq \alpha_1 + \omega\alpha_2 + \omega^2\alpha_3, \\
t_1 &\coloneqq \alpha_1^2\alpha_3 + \alpha_1\alpha_2^2 + \alpha_2\alpha_3^2, \\
t_2 &\coloneqq \alpha_1^2\alpha_2 + \alpha_2^2\alpha_3 + \alpha_1\alpha_3^2, \\
\delta &\coloneqq \prod_{i<j}(\alpha_i - \alpha_j).
\end{aligned}$$

By the same calculation as in the previous section, we obtain the equalities (2.5). So we have $t_2 - t_1 = q$ and $t_1 + t_2 = \delta$. In fact, because $-t_1 = +t_1$, these are equal: $\delta = q$. Therefore, $\Delta(f) = q^2$ and hence the discriminant is always a square. When $\operatorname{Char} F \neq 2, 3$, we saw that this means that $\operatorname{Gal}(E/F) \cong A_3$. However, unfortunately, this is generally not the case in characteristic 2.
From (2.2) we obtain $z_1^3 = q + \omega t_1 + \omega^2 t_2$. Similarly, we find $z_2^3 = q + \omega t_2 + \omega^2 t_1$.
We try to determine the Galois group of E/F. First note that, by Theorem 0.11, f is separable: There is clearly no polynomial $g \in F[x]$ such that $f(x) = g(x^2)$. Now E is the splitting field of an irreducible, separable polynomial, hence E/F is a finite Galois extension. We know as well that $\operatorname{Gal}(E/F)$ is a subgroup of S_3 by the Theorem 0.12. Note that $[F(\alpha_1) : F] = 3$, as f is irreducible. If $f(x)/(x - \alpha_1)$ does not split over $F(\alpha_1)$, then $[F(\alpha_1, \alpha_2) : F(\alpha_1)] = 2$. By the relation $\alpha_1 + \alpha_2 + \alpha_3 = 0$, we see that $E = F(\alpha_1, \alpha_2)$. Hence $[E : F]$ is either 3 or 6, in which cases it is isomorphic to A_3 and S_3, respectively. Using the idea from [4], we shall obtain two theorems about $\operatorname{Gal}(E/F)$ and E. First define the polynomial $\rho(x) \coloneqq (x - z_1^3)(x - z_2^3) = x^2 + (z_1^3 + z_2^3)x + z_1^3 z_2^3$. We claim that $\rho \in F[x]$. In fact, $\rho(x) = x^2 + qx + p^3$. For the linear coefficient we have:

$$\begin{aligned}
z_1^3 + z_2^3 &= q + \omega t_1 + \omega^2 t_2 + q + \omega t_2 + \omega^2 t_1 \\
&= 2q + (\omega + \omega^2)(t_1 + t_2) \\
&= t_1 + t_2 = q.
\end{aligned}$$

For the constant coefficient we obtain $z_1^3 z_2^3 = (z_1 z_2)^3 = (-3p)^3 = p^3$ by (2.6). The polynomial ρ is called the *quadratic resolvent* of f. Example 1.14 shows that $\Delta(\rho) = q^2 - 4p^3 = q^2$. This is precisely the discriminant of f, wherefore we now have the following:

***Theorem* 2.7:** Let F be a field of characteristic 2, containing a primitive third root of unity. Let $f \in F[x]$ be irreducible and separable and let $\rho \in F[x]$ be the quadratic resolvent of f. Let E/F be the splitting field of f. Then the Galois group of E over F is isomorphic to A_3 if and only if ρ is reducible over F.

Proof. Suppose $\mathrm{Gal}(E/F) \cong A_3$. We know that $(1\,2\,3)$ generates A_3, so we let $\sigma \in \mathrm{Gal}(E/F)$ correspond to this permutation on the indices of the roots of f. Consider $\alpha_1^2\alpha_3$. Applying σ once and twice yields $\alpha_2^2\alpha_1$ and $\alpha_3^2\alpha_2$, respectively. Adding these three elements together yields t_1. Consequently, σ merely permutes the terms of t_1, which means that $\sigma(t_1) = t_1$. Thus, t_1 is fixed by A_3, wherefore $t_1 \in F$. An analogous argument shows that $t_2 \in F$. It now follows that $z_1^3, z_2^3 \in F$. Thus ρ splits over F.
Conversely, suppose ρ is reducible over F. The roots of f are distinct, as f is separable and irreducible. Therefore, $\Delta(f) \neq 0$. As $\Delta(\rho) = \Delta(f)$, also the roots of ρ are distinct. Let $\tau \in S_3$ be a transposition. A direct computation shows that $\tau(t_1) = t_2$, which means that $\tau(z_1^3) = z_2^3 \neq z_1^3$. Since $z_1^3 \in F$, it follows that $\tau \notin \mathrm{Gal}(E/F)$ and so $\mathrm{Gal}(E/F) \ncong S_3$. As we noted earlier, the only remaining possibility for the Galois group of E/F is A_3. □

The following theorem provides a more useful representation of E.

***Theorem* 2.8:** Let F be a field of characteristic 2, containing a primitive third root of unity. Let $f \in F[x]$ be irreducible and separable and let E/F be the splitting field of f. Then E equals $F(z_j^3)(\xi)$, with $j \in \{1, 2\}$ and z_j as above, and ξ a cube root over $F(z_j^3)$. That is, E is obtained by adjoining either root of the quadratic resolvent of f and then some cube root.

Proof. First we show that $\mathrm{Gal}(E/F(z_j^3)) \cong A_3$. We have two cases.
If ρ is reducible, then $F(z_j^3) = F$ and so $\mathrm{Gal}(E/F(z_j^3)) = \mathrm{Gal}(E/F)$. The latter is isomorphic to A_3 by Theorem 2.7.
If ρ is irreducible, then $\mathrm{Gal}(E/F) \cong S_3$ by Theorem 2.7. We also have $[F(z_j^3) : F] = 2$. Clearly $F \subset F(z_j^3) \subset E$. We now show that f is still irreducible over $F(z_j^3)$. Suppose to the contrary that f is reducible over $F(z_j^3)$. Then f splits into a linear and quadratic factor, which means one of the roots of f lies in $F(z_j^3)$. Now, $F(z_j^3)$ is a normal extension of F, as its degree is 2. Since f is irreducible over F and has a root in $F(z_j^3)$, it follows that f splits over $F(z_j^3)$. However, this conflicts with the fact that $\deg f = 3$. So this cannot occur and f must be irreducible over $F(z_j^3)$. Consequently, $[F(z_j^3)(\alpha_1) : F(z_j^3)] = 3$. Since clearly $F(z_j^3)(\alpha_1) \subseteq E$, and $[F(z_j^3)(\alpha_1) : F] = [E : F] = 6$, we must have $E = F(z_j^3)(\alpha_1)$. Whence, $[E : F(z_j^3)] = 3$ and so $\mathrm{Gal}(E/F(z_j^3)) \cong A_3$.
Since A_3 is cyclic and $\mathrm{Char}\, F \nmid 3$, Theorem 0.17 implies that $E = F(z_j^3)(\xi)$, where ξ is a cube root over $F(z_j^3)$. □

Now that we have established what E looks like, we shall continue with expressing the roots of f in terms of radicals. First we express the roots in terms of z_1 and z_2.

$$\begin{aligned} z_1 + z_2 &= 2\alpha_1 + \omega(\alpha_2 + \alpha_3) + \omega^2(\alpha_3 + \alpha_2) \\ &= \alpha_1(\omega + \omega^2) = \alpha_1. \end{aligned}$$

Note that we used the relation $\alpha_1 + \alpha_2 + \alpha_3 = 0$. Similarly,

$$\begin{aligned} \omega z_1 + \omega^2 z_2 &= 2\alpha_2 + \omega(\alpha_1 + \alpha_3) + \omega^2(\alpha_3 + \alpha_1) = \alpha_2, \\ \omega^2 z_1 + \omega z_2 &= 2\alpha_3 + \omega(\alpha_2 + \alpha_1) + \omega^2(\alpha_1 + \alpha_2) = \alpha_3. \end{aligned}$$

Next we need to express z_j^3 in terms of radicals. Since f is separable, $\Delta(f) = q^2 \neq 0$ and so q^2 is invertible. Consider $q^{-2}\rho(qx)$ and observe that

$$q^{-2}\rho(qx) = q^{-2}[(qx)^2 + q(qx) + p^3] = x^2 + x + q^{-2}p^3 = \wp_2(x) + q^{-2}p^3.$$

Take one of the roots $\wp_2^{-1}(q^{-2}p^3)$. Using Remark 0.15, we find that $q\wp_2^{-1}(q^{-2}p^3)$ and $q\wp_2^{-1}(q^{-2}p^3)+q$ are the roots of $q^{-2}\rho$ and hence of ρ. Without loss of generality, $z_1^3 = q\wp_2^{-1}(q^{-2}p^3)$. (Swapping z_1 and z_2 ultimately boils down to swapping α_2 and α_3.) Fix a cube root $\sqrt[3]{q\wp_2^{-1}(q^{-2}p^3)}$ and use the relation $z_1z_2 = p$ to choose the corresponding cube root $\sqrt[3]{q\wp_2^{-1}(q^{-2}p^3)+q}$, which still works when $p = 0$ by Remark 2.3. We have now obtained the general formula for the roots of f:

$$\alpha_i = \omega^{i-1}\sqrt[3]{q\wp_2^{-1}(q^{-2}p^3)} + \omega^{1-i}\sqrt[3]{q\wp_2^{-1}(q^{-2}p^3)+q}, \qquad i = 1,2,3.$$

Remark 2.9: The formula holds for any separable, depressed cubic f. Separability is equivalent to $q \neq 0$, because $\Delta(f) = q^2$. If $q = 0$, then f becomes $x^3 + px$. The formula suggests 0 is a triple root, which is only true if $p = 0$ (cf. Remark 2.4). Indeed, 0 is a root and $\sqrt{p}$ is the other root, with multiplicity 2. In this case we would have to allow radicals of the form $x^\ell - a$, where $\ell = \operatorname{Char} F$ (cf. Definition 0.14 (i)).

Remark 2.5 applies here, i.e., we do not need the primitive third root of unity in the base field to apply the formula.

We end this section with an example.

Example 2.10: Consider the polynomial $x^3 + x + 1 \in \mathbb{F}_2$. In this case $p = q = 1$, thus we first need to find $\wp_2^{-1}(1)$, i.e., a root of x^2+x+1. We have seen that ω satisfies the relation $\omega^2+\omega+1 = 0$, hence we choose $\wp_2^{-1}(1) = \omega$. Next we need to choose cube roots $\sqrt[3]{1\cdot\omega}$ and $\sqrt[3]{1\cdot\omega+1}$ such that their product equals 1. Since $(\sqrt[3]{\omega})^9 = 1$ and $\omega \neq 1$, it follows that $\sqrt[3]{\omega}$ is a primitive ninth root of unity. We denote this by ζ_9. Observe also that $(\sqrt[3]{\omega+1})^9 = (\omega+1)^3 = \omega^3+\omega^2+\omega+1 = 1$, hence this cube root is also a primitive ninth root of unity. Since $\sqrt[3]{\omega}\sqrt[3]{\omega+1}$ must equal 1, we must have that $\sqrt[3]{\omega+1} = \zeta_9^8$, because then the product $\zeta_9\zeta_9^8 = 1$. The roots of x^3+x+1, which lie in $\mathbb{F}_2(\zeta_9) \cong \mathbb{F}_{2^9}$, are now given by

$$\omega^{i-1}\zeta_9 + \omega^{1-i}\zeta_9^8, \qquad i = 1,2,3.$$

2.3 A Formula in Characteristic 3

Let F be a field of characteristic 3. We start with a polynomial $y^3+ay^2+by+c \in F[y]$. Assume that $a \neq 0$. We apply the substitution $y = x + \frac{b}{a}$ to get rid of the linear term:

$$\begin{aligned}&\left(x+\frac{b}{a}\right)^3 + a\left(x+\frac{b}{a}\right)^2 + b\left(x+\frac{b}{a}\right) + c\\ &= x^3 + \frac{b^3}{a^3} + ax^2 + 2bx + \frac{b^2}{a} + bx + \frac{b^2}{a} + c\\ &= x^3 + ax^2 + \frac{b^3}{a^3} + \frac{2b^2}{a} + c.\end{aligned}$$

We now have a polynomial $\tilde{f}(x) := x^3 + \tilde{p}x^2 + \tilde{q}$, where $\tilde{p} := a$ and $\tilde{q} := \frac{b^3}{a^3} + \frac{2b^2}{a} + c$, which looks rather similar to the depressed cubic. If we consider $F(x)$ instead of $F[x]$, then x has an inverse. Assume further that $\tilde{q} \neq 0$. Define the polynomial f by

$$f(x) := \frac{x^3}{\tilde{q}}\tilde{f}(x^{-1}) = x^3 + \frac{\tilde{p}}{\tilde{q}}x + \frac{1}{\tilde{q}}.$$

Now setting $p := \frac{\tilde{p}}{\tilde{q}}$ and $q := \frac{1}{\tilde{q}}$ yields the form of the depressed cubic: $f(x) = x^3 + px + q$. By the assumption $\tilde{q} \neq 0$, we see that 0 is not a root of $\tilde{f}$. Observe that α is a root of f if and only if α^{-1} is a root of $\tilde{f}$. Indeed,

$$\begin{aligned}\tilde{f}(\alpha^{-1}) &= \alpha^{-3} + \tilde{p}\alpha^{-2} + \tilde{q}\\ &= \tilde{q}\alpha^{-3}\left(\frac{1}{\tilde{q}} + \frac{\tilde{p}}{\tilde{q}}\alpha + \alpha^3\right)\\ &= \tilde{q}\alpha^{-3}(\alpha^3 + p\alpha + q).\end{aligned}$$

The last factor we recognize as $f(\alpha)$. Since $\tilde{q}\alpha^{-3} \neq 0$, we see that $\tilde{f}(\alpha^{-1}) = 0$ precisely when $f(\alpha) = 0$. In the case that $\tilde{q} = 0$, the polynomial $\tilde{f}$ becomes $x^3 + \tilde{p}x^2$. So its roots are $-\tilde{p}$ and 0, with multiplicity 2.

Assume now that f is both irreducible and separable. Let E be the splitting field of f over F and let $\alpha_1, \alpha_2, \alpha_3 \in E$ be the roots of f. Let

$$\delta := \prod_{i<j} (\alpha_i - \alpha_j).$$

We now have the following theorem about E and the Galois group of E over F.

***Theorem* 2.11:** Let F be a field of characteristic 3. Let $f \in F[x]$ be irreducible and separable and let E/F be the splitting field of f. Then the Galois group of E over F is isomorphic to S_3 if $\delta \notin F$ and isomorphic to A_3 if $\delta \in F$. Moreover, $E = F(\delta)(\xi)$, where ξ is a third Artin-Schreier root over $F(\delta)$.

Proof. The first part of the proof of Theorem 2.1 works unalteredly to prove the first statement. It also yields that $[E : F(\delta)] = 3$, which implies that $\mathrm{Gal}(E/F(\delta)) \cong A_3$. Since $|A_3| = \mathrm{Char}\, F = 3$, Theorem 0.18 shows that $E = F(\delta)(\xi)$ with ξ a third Artin-Schreier root over $F(\delta)$. □

We can now continue to find a formula for the roots of f using radicals. By expanding the product $(x - \alpha_1)(x - \alpha_2)(x - \alpha_3)$, we once more have the equalities (2.1). Let us compute the discriminant of f. Since $\mathrm{Char}\, F = 3$, it divides the degree of f. So we, unfortunately, cannot use Theorem 1.17 to determine the discriminant. We will have to do a direct computation in terms of the roots. We again define the quantities t_1 and t_2:

$$t_1 := \alpha_1^2\alpha_3 + \alpha_1\alpha_2^2 + \alpha_2\alpha_3^2,$$
$$t_2 := \alpha_1^2\alpha_2 + \alpha_2^2\alpha_3 + \alpha_1\alpha_3^2.$$

By (2.5) we have $t_2 - t_1 = \delta$ and $t_1 + t_2 = 3q = 0$. It now follows that

$$\Delta(f) = \delta^2 = (t_2 - t_1)^2 = (t_1 + t_2)^2 - 4t_1t_2 = -t_1t_2.$$

Hence it suffices to compute t_1t_2, which we do directly:

$$\begin{aligned}
t_1t_2 &= (\alpha_1^2\alpha_3 + \alpha_1\alpha_2^2 + \alpha_2\alpha_3^2)(\alpha_1^2\alpha_2 + \alpha_2^2\alpha_3 + \alpha_1\alpha_3^2)\\
&= \alpha_1^4\alpha_2\alpha_3 + \alpha_1^2\alpha_2^2\alpha_3^2 + \alpha_1^3\alpha_3^3 + \alpha_1^3\alpha_2^3 + \alpha_1\alpha_2^4\alpha_3 + \alpha_1^2\alpha_2^2\alpha_3^2 + \alpha_1^2\alpha_2^2\alpha_3^2 + \alpha_2^3\alpha_3^3 + \alpha_1\alpha_2\alpha_3^4\\
&= \alpha_1\alpha_2\alpha_3(\alpha_1^3 + \alpha_2^3 + \alpha_3^3) + 3\alpha_1^2\alpha_2^2\alpha_3^2 + \alpha_1^3\alpha_2^3 + \alpha_1^3\alpha_3^3 + \alpha_2^3\alpha_3^3\\
&= \alpha_1\alpha_2\alpha_3(\alpha_1 + \alpha_2 + \alpha_3)^3 + (\alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3)^3\\
&= p^3
\end{aligned}$$

We have obtained that $\Delta(f) = -p^3$ and hence $\delta = \sqrt{-p^3} = p\sqrt{-p}$ for a suitable chosen square root of $-p$ (see Remark 2.2). Consequently, $\delta \in F$ if and only if $-p$ is a square in F. With $\sqrt{-p}$ fixed, we can now determine the formula for the roots of f. We have the following:

$$\frac{f(x\sqrt{-p}) - q}{-p\sqrt{-p}} = \frac{-p\sqrt{-p}x^3 + p\sqrt{-p}x}{-p\sqrt{-p}} = x^3 - x = \wp_3(x).$$

Now if we plug in $\frac{\alpha_1}{\sqrt{-p}}$, we obtain

$$\wp_3\left(\frac{\alpha_1}{\sqrt{-p}}\right) = \frac{q}{p\sqrt{-p}},$$

wherefore, with a suitably chosen radical,

$$\alpha_1 = \sqrt{-p}\wp_3^{-1}\left(\frac{q}{p\sqrt{-p}}\right).$$

Using Remark 0.15, we get the other third Artin-Schreier roots. The formula for all the roots is now

$$\alpha_i = \sqrt{-p}\wp_3^{-1}\left(\frac{q}{p\sqrt{-p}}\right) + (i-1)\sqrt{-p}, \qquad i = 1, 2, 3. \tag{2.7}$$

The roots of the polynomial $\tilde{f}$ are obtained by inverting the roots: α_i^{-1}, $i = 1, 2, 3$.
Regarding the original polynomial $y^3 + ay^2 + by + c$, we assumed that $a \neq 0$. The case where $a = 0$ is now immediate: The polynomial becomes $y^3 + by + c$, which is already a depressed cubic. Hence (2.7), with p and q replaced by b and c, respectively, yields the roots directly.

Remark 2.12: Also this formula holds for any separable, depressed cubic. Indeed, since $\Delta(f) = p^3$, separability is equivalent with $p \neq 0$, which is precisely when we have a problem in the formula. If $p = 0$, then the depressed cubic becomes $x^3 + q$, in which case $\sqrt[3]{-q}$ is the triple root (cf. Remark 2.9).

3 Formally Real Fields

In section 1.2 we found some useful properties of the discriminant for polynomials over $\mathbb{R}$. These properties rely on the facts that $\mathbb{R}$ admits an order relation $>$ and that $\mathbb{C} = \mathbb{R}(\boldsymbol{i})$ is algebraically closed. Since $\mathbb{R}$ is a rather specific field, we wish to take the properties of $\mathbb{R}$ that we need and try to define a type of field based on these properties alone. That way we make fewer unnecessary assumptions and get a stronger result. Thus the main objective of this chapter is to generalize the fields $\mathbb{R}$ and $\mathbb{C}$. We typically look at a specific property of $\mathbb{R}$ and then consider the fields that share this property. Subsequently we show that some of those fields are one and the same, resulting in fields with several of the properties of $\mathbb{R}$. This will yield some interesting results. In particular, we shall find generalized statements of the fundamental theorem of algebra and Theorem 1.3.

3.1 Ordered Fields

We find it convenient to define an order of a field F in terms of a subset P of F that represents the positive elements. This subset will yield an order relation $>$, like the one on $\mathbb{Q}$ and $\mathbb{R}$. The converse is obtained by taking the set of positive elements with respect to $>$, e.g., $\mathbb{Q}_{>0}$ and $\mathbb{R}_{>0}$. Although we primarily look at ordered fields, we shall also see an example of an ordered ring. Hence we define an order for any ring:

Definition 3.1: Let R be a ring with $1 \neq 0$. An *order* on R is a subset P satisfying the following properties:

(i) For each $a \in R$ exactly one of the following three possibilities is true: $a \in P$, $a = 0$ or $-a \in P$;

(ii) If $a, b \in P$, then also $a + b \in P$ and $ab \in P$, i.e., P is closed under addition and multiplication.

We say that P *orders* R. We also call P a set of *positive elements* of R and the elements of P are called *positive*. The elements of $-P := \{-a \mid a \in P\}$ are *negative*.
If such a subset of R exists, then we call R an *ordered ring*. If R is a field, then it is called an *ordered field*. We shall usually say that (R, P) is an ordered ring (or field) with the meaning that R is an ordered ring (or field) and P a set of positive elements of R.

The set P induces an order relation $>$ that satisfies similar properties as $>$ or $\mathbb{R}$. We show this in the following lemma:

***Lemma* 3.2:** Let (R, P) be an ordered ring. For $a, b \in R$, define $b > a$ (or $a < b$) to mean $b - a \in P$. We show that $>$ satisfies the following properties for all $a, b, c \in R$:

(a) Exactly one of $a > b$, $a = b$ and $a < b$ is true (trichotomy);

(b) If $a > b$ and $b > c$, then $a > c$ (transitivity);

(c) If $a > b$, then $a + c > b + c$;

(d) If $a > b$ and $c > 0$, then $ac > bc$.

Proof. By (i), $a - b$ satisfies one and only one of the possibilities $a - b \in P$, $a - b = 0$ and $b - a \in P$. These conditions are, respectively, equivalent to $a > b$, $a = b$ and $a < b$. Hence trichotomy is satisfied.
If $a > b$ and $b > c$, then $a - b$ and $b - c$ are elements of P. Since P is closed under addition, $(a - b) + (b - c) = a - c \in P$. Thus $a > c$.
If $a > b$, then $a - b \in P$. Note that $(a + c) - (b + c) = a - b \in P$. Thus $a + c > b + c$.
Finally, if $a > b$ and $c > 0$, then $a - b$ and c are elements of P. Since P is closed under multiplication, also $(a - b)c = ac - bc \in P$. Thus $ac > bc$. □

We have now seen that a set of positive elements yields an order relation $>$ with the four properties (a-d). As mentioned above, the converse is also true:

***Lemma* 3.3:** Let R be a ring and let $>$ be a binary relation on R satisfying the four properties (a-d). Also define $P := \{a \in F \mid a > 0\}$. Then P is an order on R.

Proof. For any element $a \in R$, apply the trichotomy property to a and 0 to obtain property (i) from Definition 3.1: Only one of $a > 0$, $a = 0$ or $a < 0$ holds. The first two immediately yield $a \in P$ and $a = 0$, respectively. The last one yields $a - a < -a$ by (c) and so $0 < -a$, which means $-a \in P$.

Let $a, b \in P$, so $a > 0$ and $b > 0$. We show that $a + b \in P$. By property (i), $-b \notin P$. Clearly $-b \neq 0$, hence trichotomy of $>$ implies that $0 > -b$. Transitivity of $>$ now yields that $a > -b$. Using (c), we have $a + b > -b + b = 0$. Whence $a + b \in P$.
Let again $a, b \in P$. Since $b > 0$, property (d) implies that $ab > 0b = 0$. It follows that $ab \in P$. □

From the two lemmas above we conclude that these two notions of an order are indeed equivalent.

Remark 3.4: When an ordered ring (F, P) is given, we usually assume the order relation $>$ (and $<$) to be defined as in Lemma 3.2. We also use the symbol $\geq$ (or $\leq$) to mean: $a \geq b$ if and only if $a > b$ or $a = b$. (The relation $\leq$ is then in particular a total order; see Definition A.2.)

Examples of ordered fields are $\mathbb{Q}$ and $\mathbb{R}$. In [14, p. 450], the example $\mathbb{R}[x]$ of an ordered ring is mentioned. We expand on this idea and show that in fact $R(x)$ is an ordered field for any ordered field R.

***Lemma* 3.5:** Let (R, P) be an ordered, commutative ring. Then R is an integral domain and its field of quotients is an ordered field, whose order extends P.

Proof. If $a, b \in R$ are non-zero, then either ab or $-ab$ is an element of P. Indeed, if $a, b \in P$ or $-a, -b \in P$, then $ab = (-a)(-b) \in P$. If $a, -b \in P$ or $-a, b \in P$, then $a(-b) = (-a)b = -ab \in P$. Since $0 \notin P$, there are no zero divisors in R.
Let F be the field of quotients of R. Then define the set Q as follows:

$$Q := \left\{ \frac{a}{b} \in F \;\middle|\; \text{there exist } \tilde{a}, \tilde{b} \in P \text{ such that } \frac{a}{b} = \frac{\tilde{a}}{\tilde{b}} \right\}.$$

We show that Q orders F. First we show that Q is closed under addition and multiplication. Let $\frac{a_1}{b_1}, \frac{a_2}{b_2} \in Q$. Then there exist elements $\tilde{a}_1, \tilde{b}_1, \tilde{a}_2, \tilde{b}_2 \in P$ such that $\frac{a_1}{b_1} = \frac{\tilde{a}_1}{\tilde{b}_1}$ and $\frac{a_2}{b_2} = \frac{\tilde{a}_2}{\tilde{b}_2}$. Firstly, for addition, we have

$$\frac{a_1}{b_1} + \frac{a_2}{b_2} = \frac{\tilde{a}_1}{\tilde{b}_1} + \frac{\tilde{a}_2}{\tilde{b}_2} = \frac{\tilde{a}_1\tilde{b}_2 + \tilde{a}_2\tilde{b}_1}{\tilde{b}_1\tilde{b}_2}.$$

Since P satisfies (ii) of Definition 3.1, we have $\tilde{a}_1\tilde{b}_2 + \tilde{a}_2\tilde{b}_1 \in P$ and $\tilde{b}_1\tilde{b}_2 \in P$. Consequently, $\frac{a_1}{b_1} + \frac{a_2}{b_2} \in Q$. Similarly, multiplication yields

$$\frac{a_1}{b_1}\frac{a_2}{b_2} = \frac{\tilde{a}_1}{\tilde{b}_1}\frac{\tilde{a}_2}{\tilde{b}_2} = \frac{\tilde{a}_1\tilde{a}_2}{\tilde{b}_1\tilde{b}_2}.$$

Since $\tilde{a}_1\tilde{a}_2, \tilde{b}_1\tilde{b}_2 \in P$, also $\frac{a_1}{b_1}\frac{a_2}{b_2} \in Q$.
Now let $\frac{a}{b} \in F$. We show that property (i) of Definition 3.1 holds. If $\frac{a}{b} = 0$, then $a = 0$ and so neither $\frac{a}{b}$ nor $-\frac{a}{b}$ is an element of Q, because $0 \notin P$. Next suppose $a \neq 0$. If $a, b \in P$, then clearly $\frac{a}{b} \in Q$. If $a \in P$ and $b \notin P$, then $-b \in P$. Thus $\frac{a}{-b} = -\frac{a}{b} \in Q$. Similarly, if $a \notin P$ and $b \in P$, then $-\frac{a}{b} \in Q$. Finally, if both $a, b \notin P$, then $-a, -b \in P$ and so $\frac{-a}{-b} = \frac{a}{b} \in Q$. We are only left to show that $Q \cap -Q$ is empty. Since $0 \notin Q$, this statement follows from the fact that Q is closed under addition. Thus we conclude Q orders F.
Clearly $P \hookrightarrow Q$ via the usual identification $R \ni a \mapsto \frac{a}{1} \in F$, thus Q extends P. □

***Proposition* 3.6:** Let (R, P_R) be an ordered field. Then $R(x)$ is an ordered field.

Proof. We first show that $R[x]$ is an ordered ring. Define $P_{R[x]}$ on $R[x]$ as follows:

$$P_{R[x]} := \{ f \in R[x] \mid \text{the leading coefficient } a_n \text{ of } f \text{ satisfies } a_n \in P_R \}.$$

Here the leading coefficient of 0 is 0. Since P_R is an order, we obtain property (i) from Definition 3.1 for $P_{R[x]}$ immediately. Let $f, g \in P_{R[x]}$ and let a and b be their respective leading coefficients. Then $a, b \in P_R$. The leading coefficient of $f + g$ is a, $a + b$ or b, depending on whether $\deg f > \deg g$, $\deg f = \deg g$ and $\deg f < \deg g$, respectively. In any case, the elements $a, b, a + b \in P_R$ by (ii) and so $f + g \in P_{R[x]}$. The product fg has, by Lemma 3.5, non-zero leading coefficient ab, which is in P_R by (ii), thus $fg \in P_{R[x]}$. Whence $P_{R[x]}$ is indeed an order on $R[x]$.
Since $R(x)$ is the field of quotients of $R[x]$, Lemma 3.5 implies that $R(x)$ is an ordered field, whose order P extends the one on $R[x]$. □

Let P_R, $P_{R[x]}$ and P be as in the proof of Proposition 3.6. As usual we identify elements in R with constant polynomials in $R[x]$ and hence in $R(x)$. Then $P_{R[x]}$, and hence P, extends P_R. The polynomial $x \in R(x)$ is then 'infinite' with respect to the elements of R: For each $a \in R$, we have $x > a$. To see this, note that for each $a \in R$, the polynomial $x - a$ has leading coefficient $1 \in P_R$, wherefore $x - a \in P_{R[x]}$ and so $x - a = \frac{x-a}{1} \in P$. If $>$ denotes the relation induced by P, we see that $x > a$ for every $a \in R$.
In this case we also see that the order P_R of the subfield R of $R(x)$ satisfies $P_R = P \cap R$. In fact, this always works: A subfield K of an ordered field (F, P) is ordered by $P \cap K$. We call the order obtained by $P \cap K$ the *induced order.* We prove this in the following theorem:

***Theorem* 3.7:** Let (F, P) be an ordered field and let $K \subseteq F$ be a subfield. Then $P' := P \cap K$ orders K.

Proof. Let $a \in K$ and suppose that $a \neq 0$. We show that either $a \in P'$ or $-a \in P'$. Since K is a subfield, $a \in F$. By (i) from Definition (3.1), we have either $a \in P$ or $-a \in P$. This clearly implies that either $a \in P'$ or $-a \in P'$. The fact that P and $-P$ are disjoint immediately implies that P' and $-P'$ are disjoint.
Next let $a, b \in P'$. Then $a, b \in P$ and hence $a + b \in P$. Since K is a field, $a + b \in K$ and so $a + b \in P'$. An analogous argument shows that $ab \in P'$. We conclude that P' is an order on K. □

Before we continue, here are some useful properties of ordered fields.

***Proposition* 3.8:** Let (F, P) be an ordered field. Then the following hold:

(I) If $a, b \in -P$, then $ab \in P$.

(II) For every $a \in F^*$ the square $a^2 \in P$;

(III) $1 \in P$;

(IV) F has characteristic 0;

(V) If $a \in P$, then $a^{-1} \in P$.

Proof. Let $a, b \in -P$. Then $-a, -b \in P$ and so $(-a)(-b) = ab \in P$.
Let $a \in F^*$. Then either a or $-a$ is an element of P. Either way we have $a^2 = aa = (-a)(-a)$ and so $a^2 \in P$. Thus (II) holds.
Since $1^2 = 1$, (II) implies that $1 \in P$.
Suppose $\operatorname{Char} F = p$ with $p > 0$. Since $1 \in P$, also $\overbrace{1 + \cdots + 1}^{p \text{ terms}} = 0 \in P$, which contradicts property (i) of Definition 3.1. Thus $\operatorname{Char} F = 0$ and (IV) is satisfied.
Finally, let $a \in P$ and note that $a \neq 0$. By (II), the square $(a^{-1})^2 \in P$. Whence $a(a^{-1})^2 = a^{-1} \in P$. □

3.1.1 Order-Isomorphisms

When we wish to compare ordered fields with each other, regular homomorphisms do not suffice, for they might not translate the order of one ordered field to the other. Thus, to relate ordered fields to one another, we need kinds of morphisms between them that respect their orders. We define these as follows:

Definition 3.9: Let (K, P) and (L, Q) be ordered fields. A monomorphism $\varphi : K \to L$ is called an *order-monomorphism* if $\varphi(P) \subseteq Q$. This means that if $a > b$ in K, then $\varphi(a) > \varphi(b)$ in L. Indeed, $a > b$ means that $a - b \in P$, which implies that $\varphi(a - b) \in Q$. Since φ is also a homomorphism, $\varphi(a - b) = \varphi(a) - \varphi(b)$. Therefore, $\varphi(a) > \varphi(b)$.
If φ is also surjective, then φ is called an *order-isomorphism.*
If F is a subfield of K and L, and φ the identity on F, i.e., φ is also an F-isomorphism, then φ is called an *order-F-isomorphism.*
Similar terminology applies to other morphisms.

***Proposition* 3.10:** Let (K, P) and (L, Q) be ordered fields and let $\varphi : K \to L$ be an order-isomorphism. Then $\varphi(P) = Q$ and $\varphi^{-1} : L \to K$ is also an order-isomorphism.

Proof. Suppose there is an element $b \in Q \setminus \varphi(P)$. Since φ is surjective and $b \neq 0$, there is an element $a \in -P$ such that $\varphi(a) = b$. Then $-a \in P$ and since φ is order preserving, $-b = \varphi(-a) \in Q$, which is a contradiction. Thus $\varphi(P) = Q$.
From basic field theory we know that φ^{-1} is an isomorphism. We also see that $\varphi^{-1}(Q) = P$, thus φ^{-1} is also an order-isomorphism. □

***Proposition* 3.11:** Every order-homomorphism is injective.

Proof. Let (K, P) and (L, Q) be ordered fields and let $\varphi : K \to L$ be an order-homomorphism. Since $1 \in P$, we have $\varphi(1) \in Q$ and hence $\varphi(1) \neq 0$. Therefore, $\ker \varphi \neq K$. One of the basic properties of a homomorphism is that its kernel is an ideal. The only ideals of a field are $\{0\}$ and the field itself. Therefore, the only remaining possibility is that $\ker \varphi = \{0\}$. Thus φ is injective. □

3.2 Formally Real Fields

Instead of looking at the order property $\mathbb{R}$ admits, we focus on another property. In $\mathbb{R}$ one cannot write -1 as a sum of squares. We consider fields in which this is also not possible and prove some useful attributes of them. We mainly lay a basis here for the next section, where these fields will play a role.

Notation: Let F be a field. We denote the set of sums of squares by Σ_F. We also write Σ_F^* for $\Sigma_F \setminus \{0\}$.

Definition 3.12: Let F be a field. We call F *formally real* if $-1 \notin \Sigma_F$, i.e., -1 cannot be written as a sum of squares.

From the definition we see that every formally real field must have characteristic 0, as otherwise -1 is a sum of the squares $1 = 1^2$. The definition also immediately implies that every subfield K of a formally real field F is again formally real. Since if any sum of squares in K equals -1, then clearly this same sum also equals -1 in F, contradicting the fact that F is formally real.
The following proposition establishes some properties of the sums of squares.

***Proposition* 3.13:** Let F be a field. Then Σ_F is closed under addition and multiplication. Furthermore, if $s \in \Sigma_F^*$, then also $s^{-1} \in \Sigma_F^*$.

Proof. Let $\sum_{i=0}^n a_i^2$ and $\sum_{i=0}^m b_i^2$ be elements of Σ_F. We may assume that $m = n$, because $0^2 = 0 \in \Sigma_F$. Clearly $\sum_{i=0}^n a_i^2 + \sum_{j=0}^n b_j^2 = \sum_{i=0}^n (a_i^2 + b_i^2) \in \Sigma_F$, thus Σ_F is closed under addition. For multiplication we have

$$\left(\sum_{i=0}^n a_i^2\right)\left(\sum_{j=0}^n b_j^2\right) = \sum_{i=0}^n \sum_{j=0}^n (a_i b_j)^2.$$

It now follows that this is an element of Σ_F, because Σ_F is closed under addition.
Lastly, let $s \in \Sigma_F^*$. Since every non-zero square is in Σ_F^*, $(s^{-1})^2 \in \Sigma_F^*$. That fact that Σ_F is closed under multiplication now implies that $s(s^{-1})^2 = s^{-1}$ is again a sum of squares. □

There is also an immediate relation to ordered fields. Namely, we have the following:

***Proposition* 3.14:** Every ordered field (F, P) is formally real.

Proof. Suppose F is not formally real. Then $-1 \in \Sigma_F^*$. Since P is closed under addition, property (II) of Proposition 3.8 implies that $\Sigma_F^* \subseteq P$. Thus $-1 \in P$. But by property (III) of Proposition 3.8, $1 \in P$. So this leads to a contradiction. We conclude that F must be formally real. □

Interestingly enough, this proposition has a converse. This is proved in the next section: see Corollary 3.35.
Another well-known and useful property of $\mathbb{Q}$ and $\mathbb{R}$ is that if a sum of squares equals zero, then all terms must be zero. This is a property that also formally real fields have. In fact, we have the following:

***Proposition* 3.15:** Let F be a field. The following are equivalent:

(i) F is formally real;

(ii) For $a_0, \ldots, a_n \in F$, where $n \in \mathbb{N}_0$, if $\sum_{i=0}^n a_i^2 = 0$, then $a_i = 0$ for all $0 \leq i \leq n$.

Proof. Assume F is formally real. We prove (ii) by contradiction. Suppose there exist elements $a_0, \ldots, a_n \in F$, not all zero, such that $\sum_{i=0}^n a_i^2 = 0$. Without loss of generality, $a_0 \neq 0$. Multiply both sides by a_0^{-2} and subtract $a_0^2 a_0^{-2} = 1$ from both sides to obtain

$$\sum_{i=1}^n \left(\frac{a_i}{a_0}\right)^2 = -1.$$

Therefore, $-1 \in \Sigma_F$, which is a contradiction.
For the converse implication, we again argue by contradiction. Suppose there exist elements $a_0, \dots, a_n \in F$ such that $\sum_{i=0}^n a_i^2 = -1$. Taking 1, which is a non-zero square, to the left-hand side yields a sum of non-zero squares that equals zero. This is a contradiction, so F must be formally real. □

3.2.1 Algebraic Extensions

We already saw that every subfield of a formally real field is again formally real. The same is clearly not true for algebraic extensions: Adjoining a root of $x^2 + 1$ compromises the quality of being formally real. There is still a large group of algebraic extensions of a formally real field that are formally real. The following two theorems show this.

***Theorem* 3.16:** Let (F, P) be an ordered field. Let α be a root of a polynomial $x^2 - \eta \in F[x]$, where $\eta \in P$. Then $F(\alpha)$ is formally real.

Proof. Note that F is formally real by Proposition 3.14. If $\alpha \in F$, then $F(\alpha) = F$ and we are done. So suppose $\alpha \notin F$. Then $[F(\alpha) : F] = 2$. Let $n \in \mathbb{N}_0$ and let $a_i + \alpha b_i \in F(\alpha)$, where $a_i, b_i \in F$ and $i = 0, \dots, n$, such that

$$\sum_{i=0}^{n} (a_i + \alpha b_i)^2 = 0.$$

We show that each term is zero. Expanding each term $(a_i + \alpha b_i)^2 = a_i^2 + \eta b_i^2 + 2\alpha a_i b_i$ on the left-hand side and then splitting the sum yields

$$\sum_{i=0}^{n} (a_i^2 + \eta b_i^2) + 2\alpha \sum_{i=0}^{n} a_i b_i = 0.$$

If $\sum_{i=0}^n a_i b_i \neq 0$, then

$$\alpha = \frac{-\sum_{i=0}^n (a_i^2 + \eta b_i^2)}{2\sum_{i=0}^n a_i b_i} \in F,$$

which contradicts the fact that $\alpha \notin F$. Thus $\sum_{i=0}^n a_i b_i = 0$. Hence we have $\sum_{i=0}^n (a_i^2 + \eta b_i^2) = 0$. We rewrite the summation as $\sum_{i=0}^n a_i^2 + \eta \sum_{i=0}^n b_i^2$. Suppose $\sum_{i=0}^n b_i^2 \neq 0$. Then we have

$$-\eta = \frac{\sum_{i=0}^n a_i^2}{\sum_{i=0}^n b_i^2}.$$

The right-hand side is an element of P by Proposition 3.8. Thus $-\eta \in P$. But this is a contradiction, as $\eta \in P$. Therefore, we must have that $\sum_{i=0}^n b_i^2 = 0$. Consequently, $b_i = 0$ for each i by Proposition 3.15, because F is formally real. In that case, also all a_i are zero for the same reason. Therefore, all terms $(a_i + \alpha b_i)^2$ of the original summation are zero. Thus $F(\alpha)$ is formally real by Proposition 3.15. □

The proof of the following theorem is based on the proof of Lemma 2 from [12, p. 653].

***Theorem* 3.17:** Let F be a formally real field. Let $f \in F[x]$ be an irreducible polynomial of odd degree and let α be a root of f. Then $F(\alpha)$ is formally real.

Proof. Let n be the degree of f. We prove the statement by induction on n.
If $n = 1$, then $\alpha \in F$ and $F(\alpha) = F$ is formally real. Thus the base case has been taken care of.
Let $n > 1$. Assume that every extension $F(\beta)$ of F, where β is the root of an irreducible polynomial of odd degree less than n, is formally real. This is our inductive hypothesis. We argue by contradiction that $F(\alpha)$ is formally real. Suppose it is not. The elements of $F(\alpha)$ can be represented by polynomials over F of degree less than n with α substituted for the indeterminate. Since, by assumption, $F(\alpha)$ is not formally real, there exist polynomials $g_0, \dots, g_m \in F[x]$, each of degree less than n, with $m \in \mathbb{N}_0$, such that

$$\sum_{i=0}^{m} g_i(\alpha)^2 = -1.$$

By the isomorphism $F(\alpha) \cong F[x]/\langle f \rangle$, there exists a polynomial $p \in F[x]$ such that

$$\sum_{i=0}^{m} g_i(x)^2 = -1 + p(x)f(x). \tag{3.1}$$

Let $d \coloneqq \max_i\{\deg g_i\}$ and let $I \coloneqq \{i \mid \deg g_i = d\}$. For each $j \in I$, let a_j be the leading coefficient of g_j. Then the leading coefficient of the left-hand side of (3.1) is $\sum_{j \in I} a_j^2$. Since each $a_j \neq 0$, the sum is non-zero, for F is formally real. Hence the degree of the left-hand side is $2d$, which is even. The right-hand side has degree $2d = \deg(pf) = \deg p + \deg f = \deg p + n$. Since n is odd, it follows that $\deg p$ is odd as well. Because $2d < 2n$, it also follows that $\deg p + n < 2n$ and so $\deg p < n$. Therefore, p has an irreducible factor $q \in F[x]$ with odd degree strictly less than n. Now let γ be a root of q. Then

$$\sum_{i=0}^{m} g_i(\gamma)^2 = -1 + p(\gamma)f(\gamma) = -1.$$

But the left-hand side is the sum of squares of elements of $F(\gamma)$. Since $F(\gamma)$ is formally real by the inductive hypothesis, this leads to a contradiction. Thus $F(\alpha)$ must be formally real. □

3.3 Real Closed Fields

To establish properties of formally real fields, it is efficient to consider the 'largest' formally real fields in order that it suffice to prove these properties for them. For then we can extend those properties to their subfields, which we saw are formally real. However, to be able to say something about every formally real field, we must show that each of them is contained in such a largest one. We do this in section 3.3.2. We shall also obtain a generalization of the fundamental theorem of algebra along the way in section 3.3.1. Finally, we generalize the statements we obtained in section 1.2.
We start with the formal definition:

Definition 3.18: Let F be a formally real field. We call F *real closed* if every proper algebraic extension of F is not formally real.

We can immediately relate these fields to the ordered fields from section 3.1:

***Theorem* 3.19:** Let R be a real closed field. Then Σ_R^* orders R and this order is unique.

Proof. Since R is formally real, every sum of non-zero squares is again non-zero. Hence Proposition 3.13 implies that Σ_R^* is closed under addition and multiplication.
Let $a \in R$ be non-zero. Suppose $a \notin \Sigma_R^*$. We show that $-a \in \Sigma_R^*$. The polynomial $x^2 - a \in R[x]$ is now irreducible over R. If α is a root of this polynomial, then $R(\alpha)$ is a proper algebraic extension of R of degree 2. By definition of R, $R(\alpha)$ is not formally real. Hence there exist elements $a_i + \alpha b_i \in R(\alpha)$, where $a_i, b_i \in R$, $i = 0, \ldots, n$ and not all b_i are zero, such that

$$\sum_{i=0}^{n} (a_i + \alpha b_i)^2 = -1.$$

By splitting the summation, we obtain

$$\sum_{i=0}^{n} (a_i^2 + a b_i^2) + 2\alpha \sum_{i=0}^{n} a_i b_i = -1.$$

If $\sum_{i=0}^{n} a_i b_i \neq 0$, then

$$\alpha = \frac{-1 - \sum_{i=0}^{n}(a_i^2 + a b_i^2)}{2\sum_{i=0}^{n} a_i b_i}.$$

This means that $\alpha \in R$, which is a contradiction. Therefore, we must have that $\sum_{i=0}^{n} a_i b_i = 0$. From this now follows that $\sum_{i=0}^{n}(a_i^2 + a b_i^2) = -1$. Solving for $-a$ yields

$$-a = \frac{1 + \sum_{i=0}^{n} a_i^2}{\sum_{i=0}^{n} b_i^2}.$$

Note that indeed $\sum_{i=0}^{n} b_i^2 \neq 0$ as not all b_i are zero and R is formally real. By Proposition 3.13, $-a \in \Sigma_R^*$. Finally, since $0 \notin \Sigma_R^*$, it follows that Σ_R^* and $-\Sigma_R^*$ are disjoint. We conclude that Σ_R^* orders R.
To show that this order is unique, suppose P also orders R. Property (II) of Proposition 3.8 implies that $\Sigma_R^* \subseteq P$. Now let $s \in P$ and consider a root $\sqrt{s}$ of $x^2 - s$. By Theorem 3.16, $R(\sqrt{s})$ is a formally real field. Because R is real closed, this cannot be a proper extension of R, which means that $R(\sqrt{s}) = R$. So s is a square in R, wherefore $s \in \Sigma_R^*$. We conclude that $P = \Sigma_R^*$. □

***Corollary* 3.20:** Let R be a real closed field. Then every sum of squares is again a square.

Proof. Clearly $0 \in \Sigma_R$ is a square, hence we are left to prove that every element of Σ_R^* is a square. Let $s \in \Sigma_R^*$ and let $\sqrt{s}$ be a root of $x^2 - s$. Since (R, Σ_R^*) is an ordered field, Theorem 3.16 implies that $R(\sqrt{s})$ is formally real. By definition, $R(\sqrt{s})$ cannot be a proper extension of R, hence $R(\sqrt{s}) = R$. Thus s is a square in R. □

Note in particular the analogy with $\mathbb{R}$, in that also in $\mathbb{R}$ every positive element is a square. Indeed, as we shall see later, $\mathbb{R}$ is a real closed field.

3.3.1 A Generalization of the Fundamental Theorem of Algebra

In order to continue proving that every formally real field is contained in a real closed one, we need another type of field. The idea of these fields arose from Theorems 3.16 and 3.17: We want an ordered field that is maximal with respect to those theorems in the sense that all squares of positive elements and all roots of odd polynomials have been adjoined. Of course, for each square root that we adjoin, we need to adjoin another square root, so this is an unending process. We therefore directly define a field that captures this maximality:

Definition 3.21: Let R be a field. We call R *virtually real* if it admits an order $P \subseteq R$ and satisfies the following two conditions:

(A) Every element $p \in P$ is a square;

(B) Every polynomial over R of odd degree has a root in R.

Note that every virtually real field is formally real by Proposition 3.14, as it is an ordered field. Not coincidentally is $\mathbb{R}$ again an example (see also the first part of the proof of Theorem 3.30).
The term 'virtually real' is not standard terminology, nor does there seem to be standard terminology for this kind of field. Fortunately, we shall show that they are equivalent to real closed fields, hence the non-standard terminology will not be necessary for long. We start with one implication:

***Proposition* 3.22:** Every real closed field R is virtually real.

Proof. By Theorem 3.19 and Corollary 3.20, R admits an order and satisfies (A).
Let $f \in R[x]$ be a polynomial of odd degree. Then f has an irreducible factor g of odd degree. Let α be a root of g. Then, by Theorem 3.17, $R(\alpha)$ is formally real. Since R is real closed, $R = R(\alpha)$ and so $\alpha \in R$, yielding (B). □

The converse implication is proved in Corollary 3.27. It is a consequence of a generalization of the fundamental theorem of algebra. We first need a couple of lemmas before we can prove that theorem.

***Lemma* 3.23:** Let F and K be fields and let $\varphi : F \to K$ be a monomorphism. Let $f \in F[x]$ be a polynomial. Then α is a root of f if and only if $\varphi(\alpha)$ is a root of $\varphi(f)$, where $\varphi(f)$ is the polynomial f with φ applied to each coefficient.

Proof. Since φ is in particular a homomorphism, we obtain $\varphi(f)(\varphi(\alpha)) = \varphi(f(\alpha))$. Using the fact that $\varphi(0) = 0$ and φ is injective, we see that $\varphi(f)(\varphi(\alpha)) = 0$ if and only if $f(\alpha) = 0$. □

***Lemma* 3.24:** Let R be formally real and let $\boldsymbol{i}$ be a root of $x^2 + 1$ in an algebraic closure of R. Then the map that sends $\boldsymbol{i}$ to $-\boldsymbol{i}$ on $R(\boldsymbol{i})$ is an R-automorphism.

Proof. Clearly $\boldsymbol{i} \notin R$, for otherwise R were not formally real. Thus $x^2 + 1$ is irreducible over R and thus $[R(\boldsymbol{i}) : R] = 2$. Observe that $R(\boldsymbol{i})$ is the splitting field of the separable polynomial $x^2 + 1$, so $R(\boldsymbol{i})/R$ is a Galois extension by Theorem 0.10. Since $\boldsymbol{i}$ and $-\boldsymbol{i}$ are the roots of $x^2 + 1$, the Galois group of $R(\boldsymbol{i})$ over R contains the identity and the map that sends $\boldsymbol{i}$ to $-\boldsymbol{i}$. Thus this is an R-automorphism. □

Notation: If R is a formally real field, we will henceforth denote a root of x^2+1 in an algebraic closure of R by $\boldsymbol{i}$. By analogy with $\mathbb{C} = \mathbb{R}(\boldsymbol{i})$, we also denote the R-automorphism that sends $\boldsymbol{i}$ to $-\boldsymbol{i}$ on $R(\boldsymbol{i})$ by $\overline{\cdot}$, e.g., $\overline{\boldsymbol{i}} = -\boldsymbol{i}$.

***Lemma* 3.25:** Let (R, P) be a virtually real field. Then every element of $R(\boldsymbol{i})$ is a square. Moreover, $R(\boldsymbol{i})$ does not have any algebraic extension of degree 2.

Proof. Let $>$ denote the order induced on R by P. If $a \in R$ and $a \geq 0$, then a is a square by assumption. If $a < 0$, then $-a > 0$ and hence there exists a $b \in R$ such that $b^2 = -a$. Then $(b\boldsymbol{i})^2 = -a(-1) = a$. Now let $a + b\boldsymbol{i} \in R(\boldsymbol{i})$, $a, b \in R$ with $b \neq 0$. We want elements $c, d \in R$ such that $(c + d\boldsymbol{i})^2 = a + b\boldsymbol{i}$. Expanding the left-hand side yields $c^2 - d^2 + 2cd\boldsymbol{i}$ and hence we have the simultaneous equations

$$\begin{cases} c^2 - d^2 = a, \\ \qquad 2cd = b. \end{cases}$$

Since $b \neq 0$, it follows that $d \neq 0$. Hence the second equation yields $c = \frac{b}{2d}$. We substitute this into the first equation to obtain $\frac{b^2}{4d^2} - d^2 = a$, which in turn yields

$$d^4 + ad^2 - \frac{b^2}{4} = 0.$$

We have a quadratic equation in d^2. The discriminant is $a^2 - 4 \cdot (-\frac{b^2}{4}) = a^2 + b^2$. This is clearly positive, hence this is a square itself. Let $\sqrt{a^2+b^2}$ denote the positive square root of $a^2 + b^2$. From the quadratic formula we obtain that

$$d^2 = \frac{-a + \sqrt{a^2+b^2}}{2}$$

is a solution. Suppose $-a + \sqrt{a^2+b^2} < 0$, then certainly $-a < 0$ as $\sqrt{a^2+b^2} > 0$. Thus $a + \sqrt{a^2+b^2} > 0$, hence we may multiply the first inequality with this element due to property (d) on page 20. We obtain $-a^2 + (\sqrt{a^2+b^2})^2 < 0$ and so $b^2 < 0$, which is a contradiction. Therefore, $-a + \sqrt{a^2+b^2} > 0$. Since also $\frac{1}{2} > 0$, d^2 has a square root in R. Now $c^2 = a + d^2 = \frac{a + \sqrt{a^2+b^2}}{2}$. We show that $a + \sqrt{a^2+b^2}$ is positive. Suppose $-a - \sqrt{a^2+b^2} > 0$. Then, again by property (d), $(-a + \sqrt{a^2+b^2})(-a - \sqrt{a^2+b^2}) > 0$ and so $a^2 - a^2 - b^2 = -b^2 > 0$, which is a contradiction. Thus $c^2 > 0$, wherefore also c^2 has a square root in R. Thus we have found solutions c and d in R. We conclude that every element of $R(\boldsymbol{i})$ has a square root.
For the second statement, let $f(x) := x^2 + px + q \in R(\boldsymbol{i})$ be a polynomial. By the foregoing, there is an element $r \in R(\boldsymbol{i})$ such that $r^2 = p^2 - 4q$. Then

$$f\left(\frac{-p+r}{2}\right) = \frac{p^2 + r^2 - 2pr}{4} + p\left(\frac{-p+r}{2}\right) + q = 0.$$

Thus any quadratic polynomial is reducible over $R(\boldsymbol{i})$ and hence there cannot be an algebraic extension of $R(\boldsymbol{i})$ of degree 2. □

We are now ready to state and prove a generalization of the fundamental theorem of algebra. The proof is based on the proof of Theorem 5.2 from [11].

***Theorem* 3.26** (Generalized Fundamental Theorem of Algebra)**:** Let R be a virtually real field. Then $R(\boldsymbol{i})$ is algebraically closed.

Proof. Let $C := R(\boldsymbol{i})$. We need to show that every polynomial in $C[x]$ splits over C. If $f \in C[x]$, then $f\overline{f}$ is a polynomial over R. Indeed, $\overline{f\overline{f}} = \overline{f}f = f\overline{f}$, hence $f\overline{f}$ is fixed by $\overline{\cdot}$. Since this is an R-automorphism by Lemma 3.24, it follows that $f\overline{f} \in R[x]$. Thus if $\alpha \in C$ is a root of $f\overline{f}$, then α is a root of f or $\overline{f}$. If $f(\alpha) \neq 0$, then α is a root of $\overline{f}$. Therefore, by Lemma 3.23, $\overline{\alpha}$ is a root of $\overline{\overline{f}} = f$. Thus it suffices to consider polynomials over R.
Let $f \in R[x]$ be a non-constant polynomial. If $\deg f$ is odd, then f has a root in R by the assumption that R is virtually real. Thus assume that f has even degree. Let E be the splitting field of $f(x)(x^2+1)$ over R. Because $\boldsymbol{i}$ is a root of $f(x)(x^2+1)$, the splitting field contains C. Since $\operatorname{Char} R = 0$, $f(x)(x^2+1)$ is separable. Hence, by Theorem 0.10, E/R is a Galois extension. By assumption, the degree is even.

Thus write $[E : R] = 2^n k$ with $n \in \mathbb{N}$ and k odd. Then the Galois group of E/R has order $2^n k$. By the first Sylow theorem, $\mathrm{Gal}(E/R)$ has a subgroup G of order 2^n. Let F be the fixed field of G in E. Since $R \subseteq F \subseteq E$, we now have $[E : F] = 2^n$ and $[F : R] = k$. If $k > 1$, then there is an irreducible polynomial g over R with odd degree. However, this contradicts the fact that odd polynomials have a root in R. Hence $k = 1$. Consequently, $F = R$ and $[E : R] = 2^n$.
We shall now show that $n = 1$. Suppose $n > 1$. Then $\mathrm{Gal}(E/C)$ has order 2^{n-1}, which is divisible by 2. Whence, by Cauchy's theorem, $\mathrm{Gal}(E/C)$ has a subgroup H of order 2. The corresponding fixed field C^H then has degree 2 over C. But this contradicts Lemma 3.25. Thus $n = 1$ and hence $[E : C] = 1$. Since $C \subseteq E$, we conclude that $E = C$. □

As a consequence of this theorem, we obtain the converse of Proposition 3.22. The proof is based on the proof of Theorem 11.2 from [12].

***Corollary* 3.27:** Let (R, P) be a virtually real field. Then R is real closed.

Proof. Theorem 3.26 yields that $C := R(\boldsymbol{i})$ is algebraically closed. Thus C is an algebraic closure of R, which means that every algebraic extension of R can be embedded in C, i.e., we can deem algebraic extensions of R as subfields of C. Thus if E is a proper algebraic extension of R, then $E \subseteq C$ and so $[E : R] = 2$. Consequently, $E = C$ and hence not formally real. Whence R does not have a proper algebraic extension that is formally real, wherefore R is real closed. □

Although we have proven the generalized fundamental theorem of algebra, we have not actually proven the fundamental theorem of algebra itself. By Corollary 3.27, it suffices to show that $\mathbb{R}$ is virtually real. In Proposition 1.6 we already proved that (B) from Definition 3.21 holds for $\mathbb{R}$. However, here we already used that all roots are contained in $\mathbb{C}$, i.e., that $\mathbb{C}$ is algebraically closed. Thus, lest our proof be circular, we cannot use this theorem. The approach we use is by analogy with the *intermediate value theorem*. We consider fields in which a similar property is true—on $\mathbb{R}$ they coincide—and show that they are equivalent to virtually real fields. We then, for completeness sake, give an analytical proof of the intermediate value theorem for $\mathbb{R}$, which yields the desired result.

Definition 3.28: Let F be an ordered field. We say that F has the *intermediate value property* if the following holds: If $f \in F[x]$ and $a, b \in F$, with $a < b$, such that $f(a)f(b) < 0$, then there exists a $c \in F$ such that $f(c) = 0$ and $a < c < b$.

The following theorem shows the equivalence between fields with this property and virtually real fields. The proof of the first implication is based on the proof of Proposition 3.9 from [3, p. 315].

***Theorem* 3.29:** Let (R, P) be an ordered field. Then R is virtually real if and only if it has the intermediate value property.

Proof. Suppose that R is virtually real. By Theorem 3.26, $R(\boldsymbol{i})$ is algebraically closed. Let $f \in R[x]$ be a polynomial and let $a, b \in R$, with $a < b$, such that $f(a)f(b) < 0$. Since f splits entirely over $R(\boldsymbol{i})$ and $[R(\boldsymbol{i}) : R] = 2$, f splits into irreducible quadratic and linear factors over R. Any such quadratic factor is of the form $g(x) := x^2 + px + q$. Completing the square yields $g(x) = (x + \frac{p}{2})^2 + q - \frac{p^2}{4}$. If $q - \frac{p^2}{4} < 0$, then $\frac{p^2}{4} - q$ has a square root $\sqrt{\frac{p^2}{4} - q}$ in R and so $-\frac{p}{2} + \sqrt{\frac{p^2}{4} - q}$ is a root of g in R, which contradicts the irreducibility of g. Hence $q - \frac{p^2}{4} > 0$. Now for every $\alpha \in R$ we have $g(a) = (a + \frac{p}{2})^2 + q - \frac{p^2}{4} > 0$. Thus if f consisted of only quadratic factors, then $f(a)f(b) > 0$, which is false. Thus f must have a linear factor $x - c$ with $a < c < b$, and therefore a root $c \in R$. Whence R has the intermediate value property.
For the converse implication assume R has the intermediate value property. Let $a \in P$ and consider $g(x) := x^2 - a \in R[x]$. We have $g(0) = -a < 0$ and $g(a + 1) = a^2 + a + 1 > 0$, thus $g(0)g(a + 1) < 0$ and so g has a root by the intermediate value property. This means that each positive element is a square in R.
Now let $f(x) := x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in R[x]$ be a polynomial of odd degree bigger than 1. Let $|\cdot| : R \to P \cup \{0\}$ be defined as follows: For $a \in R$, set $|a| := a$ if $a \in P \cup \{0\}$ and $|a| := -a$ otherwise. Observe that $|ab| = |a|\,|b|$ and $\left|a^{-1}\right| = |a|^{-1}$ for all $a, b \in R$. Consider $h : R^* \to P \cup \{0\}$ defined by:

$$h(x) := \left|\frac{f(x) - x^n}{x^n}\right| = \left|\frac{a_{n-1}}{x} + \cdots + \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n}\right|.$$

Choose $p \in P$ such that $p > n \max\{1, |a_0|, \ldots, |a_{n-1}|\}$. Then $h(p) < 1$ and so $|f(p) - p^n| < |p^n| = p^n$. If $f(p) < 0$, then $|f(p) - p^n| = -(f(p) - p^n) = p^n - f(p)$. The inequality $p^n - f(p) < p^n$ now yields that $-f(p) < 0$, which is a contradiction. Thus $f(p) > 0$.
Next notice that also $h(-p) < 1$. Thus $|f(-p) - (-p)^n| = |f(-p) + p^n| < |(-p)^n| = p^n$. If $f(-p) > 0$, then $|f(-p) + p^n| = f(-p) + p^n < p^n$, which implies that $f(-p) < 0$. Hence we have a contradiction. Thus $f(-p) < 0$. Consequently, $f(p)f(-p) < 0$ and hence, by the intermediate value property, f has a root c in R with $-p < c < p$.
We conclude that R is virtually real. □

***Theorem* 3.30** (Fundamental Theorem of Algebra)**:** The complex numbers $\mathbb{C}$ are algebraically closed.

Proof. Since $\mathbb{C} = \mathbb{R}(\boldsymbol{i})$, it suffices to show that $\mathbb{R}$ has the intermediate value property. For then Theorem 3.29 implies that $\mathbb{R}$ is virtually real. Subsequently, Theorem 3.26 shows that $\mathbb{R}(\boldsymbol{i})$ is algebraically closed.

Remark 3.31: The intermediate value property for $\mathbb{R}$ is exactly the intermediate value theorem for polynomials. The proof given here involves some analysis. It uses the *completeness* of $\mathbb{R}$: Every subset of $\mathbb{R}$ that is bounded above has a least upper bound. We also use that polynomials are continuous. That is to say, a function $f : \mathbb{R} \to \mathbb{R}$ is continuous at $\xi \in \mathbb{R}$ if for each $\varepsilon > 0$ there exists a $\delta > 0$ such that, if $r \in (\xi - \delta, \xi + \delta)$, then $f(r) \in (f(\xi) - \varepsilon, f(\xi) + \varepsilon)$. Now f is called continuous if it is continuous at each point.

Let $f \in \mathbb{R}[x]$ and let $a, b \in \mathbb{R}$, with $a < b$, such that $f(a)f(b) < 0$. Without loss of generality, we assume that $f(a) < 0 < f(b)$. We show that f has a root between a and b. Let $N := \{r \in (a, b) \mid f(r) < 0\}$. Clearly b is an upper bound of N, hence there exists a least upper bound u of N. We show that $f(u) = 0$. We argue by contradiction and have two cases.
First suppose $f(u) > 0$. Settings $\varepsilon := f(u)$, by continuity, we obtain a positive δ such that, for each $r \in (u - \delta, u + \delta)$, we have $f(r) \in (f(u) - \varepsilon, f(u) + \varepsilon)$. This means in particular that $f(r)$ is positive for $u - \delta < r < u + \delta$. This implies that if $n \in N$, then $n \leq u - \delta$, because $n < u$ and $n \notin (u - \delta, u)$ by the foregoing. This, however, contradicts the definition of u, because now $u - \delta$ is a smaller upper bound of N. Thus $f(u)$ cannot be positive.
For the second case we suppose that $f(u) < 0$. Now set $\varepsilon := -f(u) > 0$. By continuity, there exists a $\delta > 0$ such that, for each $r \in (u - \delta, u + \delta)$, we have $f(r) \in (f(u) - \varepsilon, f(u) + \varepsilon)$. In particular, $f(r) < 0$ for $u - \delta < r < u + \delta$. Clearly now $u \neq b$, so $u < b$. Thus $(u, b) \cap (u, u + \delta)$ is non-empty. Let η be an element of this intersection, then $\eta \in N$ while $\eta > u$. This again contradicts the definition of u.
By trichotomy, the only remaining possibility is $f(u) = 0$. Since we also see that $a < u < b$, we have the desired result. □

Because we have introduced various, eventually equivalent, fields, we sum up the equivalences in the following theorem:

***Theorem* 3.32:** Let R be an ordered fields. Then the following are equivalent:

(i) R is real closed;

(ii) R is virtually real;

(iii) $R(\boldsymbol{i})$ is algebraically closed;

(iv) R has the intermediate value property.

Proof. We have already seen that (i) $\iff$ (ii) $\iff$ (iv) and (ii) $\implies$ (iii). We have proven the implication (iii) $\implies$ (i) in the proof of Corollary 3.27. This yields the implication (ii) $\impliedby$ (iii). □

3.3.2 Real Closures

We are now ready to show that every formally real field is contained in a real closed one, which is algebraic over it. As mentioned above, we can then deduce that every formally real field admits an order. After this we obtain a more general statement, which says that every ordered field has a real closed, algebraic extension that extends its order and is unique up to order-isomorphism. The latter will allow us to generalize the statements about discriminants in section 1.2. For the first statement we need a lemma.

***Lemma* 3.33:** Let F be a field and let $\mathcal{C}$ be a set of algebraic extensions of F that is totally ordered by inclusion, i.e., $\mathcal{C}$ is a chain. Then $\bigcup \mathcal{C}$ is an algebraic extension of F.

Proof. We first show that $\bigcup \mathcal{C}$ is a field. Let $a, b \in \bigcup \mathcal{C}$. Then there exist $C, C' \in \mathcal{C}$ such that $a \in C$ and $b \in C'$. Since $\mathcal{C}$ is a chain, without loss of generality, $C' \subseteq C$, hence $a, b \in C$. Now let addition and multiplication of a and b in $\bigcup \mathcal{C}$ be defined via the addition and multiplication in C. This makes $\bigcup \mathcal{C}$ into a field.
By definition of $\mathcal{C}$, $F \subseteq \bigcup \mathcal{C}$. Now let $a \in \bigcup \mathcal{C}$. Then $a \in C$ for some $C \in \mathcal{C}$. Since C is an algebraic extension of F, a is algebraic over F. So each element of $\bigcup \mathcal{C}$ is algebraic over F, which means that $\bigcup \mathcal{C}$ is an algebraic extension of F. □

***Theorem* 3.34:** Let F be a formally real field. Then there exists a real closed, algebraic extension of F.

Proof. For this proof we use Zorn's lemma. Let $\overline{F}$ be an algebraic closure of F and let $\mathcal{F}$ be the set of all formally real, algebraic extensions of F that are contained in $\overline{F}$. The reason why we need to limit ourselves to formally real, algebraic extension in an algebraic closure is that otherwise $\mathcal{F}$ might not be a set (see [1]). Let $\mathcal{C} \subseteq \mathcal{F}$ be a non-empty chain. By Corollary A.7, it suffices to show that $\bigcup \mathcal{C}$ is a formally real, algebraic extension of F. Lemma 3.33 shows that $\bigcup \mathcal{C}$ is an algebraic extension of F.
We are left to show that $\bigcup \mathcal{C}$ is formally real. Let $a_0, \ldots, a_n \in \bigcup \mathcal{C}$ such that $\sum_{i=0}^{n} a_i^2 = 0$. For each i there is a $C_i \in \mathcal{C}$ such that $a_i \in C_i$. Because $\{C_0, \ldots, C_n\} \subseteq \mathcal{C}$ is a finite chain, the sets can, possibly after relabeling, be ordered by inclusion: $C_0 \subseteq C_1 \subseteq \cdots \subseteq C_n$. Hence $a_i \in C_n$ for all i. Now $\sum_{i=0}^{n} a_i^2$ is a sum of squares in C_n that equals zero. Since C_n is formally real, we must have that $a_i = 0$ for each i. By Proposition 3.15, $\bigcup \mathcal{C}$ is formally real.
We conclude that $\bigcup \mathcal{C} \in \mathcal{F}$. Therefore, by Zorn's lemma, $\mathcal{F}$ contains a maximal element R. By construction, R is real closed. □

We now obtain the converse of Proposition 3.14:

***Corollary* 3.35:** Every formally real field F admits an order.

Proof. By Theorem 3.34, there exists a real closed, algebraic extension R of F. By Theorem 3.19, (R, Σ_R^*) is an ordered field. Finally, Theorem 3.7 yields that the induced order $\Sigma_R^* \cap F$ orders F. □

Next we work toward extending the result of the previous theorem to any ordered field such that the real closed field extends its order.

Definition 3.36: Let (F, P) be an ordered field. We call a field R a *real closure* of F if R is a real closed, algebraic extension of F, whose order extends the order of F.

Example 3.37: The set of real numbers $\mathbb{R}$ is *not* a real closure of $\mathbb{Q}$: Despite the fact that $\mathbb{R}$ is a real closed extension of $\mathbb{Q}$, whose order extends the order of $\mathbb{Q}$, it is not algebraic over $\mathbb{Q}$. In Example 3.42 we shall see that any real closure of $\mathbb{Q}$ is isomorphic to the field of elements of $\mathbb{R}$ that are algebraic over $\mathbb{Q}$, i.e., $\overline{\mathbb{Q}} \cap \mathbb{R}$, where $\overline{\mathbb{Q}}$ is an algebraic closure of $\mathbb{Q}$.

We shall first show that every ordered field has a real closure and then that these closures are unique up to order-isomorphism. We first need the following lemma:

***Lemma* 3.38:** Let (F, P) be an ordered field. Take for each $p \in P$ a square root $\sqrt{p}$ and let $E := F(\{\sqrt{p} \mid p \in P\})$. Then E is formally real.

Proof. We prove by contradiction. Suppose E is not formally real. Then there exist elements $a_0, \ldots, a_n \in E$ such that $\sum_{i=0}^{n} a_i^2 = -1$. Then this sum shows that $F(a_0, \ldots, a_n)$ is not formally real. However, since $F(a_, \ldots, a_n)$ is a finite extension of F, repeatedly applying Theorem 3.16 shows that $F(a_0, \ldots, a_n)$ is formally real. Thus we have reached a contradiction. Whence E must be formally real. □

The idea of the proof of the following theorem came from the proof of Theorem 179 from [2].

***Theorem* 3.39:** Every ordered field (F, P) has a real closure.

Proof. Let (F, P) be an ordered field and let $E := F(\{\sqrt{p} \mid p \in P\})$, where $\sqrt{p}$ is either square root of p. By Lemma 3.38, E is formally real. Thus, by Theorem 3.34, there exists a real closed field R containing

E. Then R is a real closed, algebraic extension of F, so we are left to show that the order of F extends to R. Let $Q := \Sigma_R^* \cap E$ be the induced order on E. Let $p \in P$. By definition of E, p is a square in E and so $p \in Q$. Since $Q \subseteq \Sigma_R^*$, also $p \in \Sigma_R^*$. Therefore, $P \subseteq \Sigma_R^*$. Thus the order of R extends P. We conclude that R is a real closure of F. □

***Theorem* 3.40:** Let (F, P) be an ordered field and let R_1 and R_2 be two real closures of F. Then there exists an order-F-isomorphism $\varphi : R_1 \to R_2$.

Proof. Using Zorn's lemma, we shall construct a real closed field contained in R_1 that is mapped via an order-F-monomorphism into R_2. We then argue that this field must be R_1 itself and that the order-F-monomorphism is surjective, yielding the desired statement.

Let $\iota : F \to R_2$ be the inclusion. Let $\mathcal{F}$ be the set of pairs $((E, Q), f)$ with $F \subseteq E \subseteq R_1$, $P \subseteq Q \subseteq \Sigma_{R_1}^*$ and $f : E \to R_2$ an order-F-monomorphism. Define a partial order $\preceq$ on $\mathcal{F}$ by setting $((E_1, Q_1), f_1) \preceq ((E_2, Q_2), f_2)$ if and only if

$$E_1 \subseteq E_2, \quad Q_1 \subseteq Q_2 \quad \text{and} \quad f_2|_{E_1} = f_1$$

From the reflexive and antisymmetric properties of the inclusion we find that $\preceq$ is reflexive and antisymmetric as well. For transitivity, from the inclusions $E_1 \subseteq E_2 \subseteq E_3$ and $Q_1 \subseteq Q_2 \subseteq Q_3$ clearly follow $E_1 \subseteq E_3$ and $Q_1 \subseteq Q_3$. Furthermore, we see that if f_3 extends f_2 and f_2 extends f_1, then f_3 extends f_1. Thus $\preceq$ partially orders $\mathcal{F}$.

Note that $((F, P), \iota) \in \mathcal{F}$, so $\mathcal{F} \neq \emptyset$. Let $\mathcal{C} := \{((E_i, Q_i), f_i) \mid i \in I\}$ be a non-empty chain in $\mathcal{F}$, where I is some index set. Define $\mathcal{E} := \bigcup_{i \in I} E_i$ and $\mathcal{Q} := \bigcup_{i \in I} Q_i$. By Lemma 3.33, $\mathcal{E}$ is an algebraic extension of F, which is clearly contained in R_1. Next we show that $\mathcal{Q}$ orders $\mathcal{E}$.

We first show that $\mathcal{Q}$ is closed under addition and multiplication. Let $a, b \in \mathcal{Q}$. Then there exist $i, j \in I$ such that $a \in Q_i$ and $b \in Q_j$. Since $\mathcal{C}$ is a chain, without loss of generality, $Q_j \subseteq Q_i$. Hence $a, b \in Q_i$. Since Q_i is an order for E_i, it follows that $a + b$ and ab are elements of Q_i. Consequently, $a + b, ab \in \mathcal{Q}$.

We now prove that for any $c \in \mathcal{E}$, either $c = 0$, $c \in \mathcal{Q}$ or $-c \in \mathcal{Q}$, and that these are mutually exclusive. Let $c \in \mathcal{E}$. Then there exists a $k \in I$ such that $c \in E_k$. If $c \neq 0$, then either $c \in Q_k$ or $-c \in Q_k$. Therefore, $c \in \mathcal{Q}$ or $-c \in \mathcal{Q}$. Observe that $0 \notin Q_i$ for every $i \in I$, hence $0 \notin \mathcal{Q}$. Since $\mathcal{Q}$ is closed under addition, it follows that $\mathcal{Q}$ and $-\mathcal{Q}$ are disjoint. We conclude that $\mathcal{Q}$ orders $\mathcal{E}$.

We now define $f : \mathcal{E} \to R_2$. Let $a \in \mathcal{E}$. Then there exists an $i \in I$ such that $a \in E_i$. Now set $f(a) := f_i(a)$. To see that f is well-defined, suppose we used some other index $j \in I$ with $a \in E_j$. Because $\mathcal{C}$ is a chain, either f_i extends f_j or vice versa. Either way, $f_i(a) = f_j(a)$. Hence the choice of the index is inconsequential.

We just need to show that f is an order-F-monomorphism. First we show that f extends ι. Let $e \in F$. There is some $k \in I$ such that $e \in E_k$. Then, by the definitions of f and f_k, $f(e) = f_k(e) = e$.

Now let $a, b \in \mathcal{E}$. Then there are again $i, j \in I$ such that $a \in E_i$ and $b \in E_j$. Then, without loss of generality, $E_j \subseteq E_i$ and $f_i|_{E_j} = f_j$. Because f_i is a homomorphism, $f(a+b) = f_i(a+b) = f_i(a)+f_i(b) = f(a) + f(b)$ and similarly $f(ab) = f(a)f(b)$. We show that f preserves order. Let $p \in \mathcal{Q}$. Then $p \in Q_i$ for some $i \in I$. Since f_i is order-preserving, we obtain $f(p) = f_i(p) \in \Sigma_{R_2}^*$. Therefore, $f(\mathcal{Q})$ is contained in $\Sigma_{R_2}^*$. Thus we see that f is an order-F-homomorphism into R_2. By Proposition 3.11, f is an order-F-monomorphism.

We have now obtained that $((\mathcal{E}, \mathcal{Q}), f) \in \mathcal{F}$. By construction, it is an upper bound of $\mathcal{C}$. Therefore, by Zorn's lemma, $\mathcal{F}$ has a maximal element $((K, O), \varphi)$.

We show that K is real closed via contradiction. So suppose K is not real closed. By Corollary 3.27, either some element in O is not a square, or there is a polynomial of odd degree over K that has no root in K. Either way, we obtain a proper algebraic extension $L := K(\alpha)$ of K, where α is (1) a root of a polynomial $x^2 - o \in K[x]$, where o is not a square in K, or (2) a root of an irreducible polynomial of odd degree over K. Let $p \in K[x]$ be the minimal polynomial of α over K. Consider $\varphi(p) \in R_2[x]$. Since R_2 is real closed, Proposition 3.22 implies that $\varphi(p)$ has a root β in R_2. Lemma 3.23 implies that $\beta \notin \operatorname{im} \varphi$, for otherwise p were reducible over K. Since p is irreducible and $\varphi : K \to \operatorname{im} \varphi$ is an isomorphism, by Theorem 4.6.1 from [6], φ extends to an isomorphism $\tilde{\varphi} : L \to \operatorname{im}(\varphi)(\beta)$, sending α to β. Hence $\tilde{\varphi} : L \to R_2$ is an F-monomorphism. Next we define an order on L. Let $\mathcal{O}$ be the set $\tilde{\varphi}^{-1}(\Sigma_{R_2}^* \cap \operatorname{im} \tilde{\varphi})$. That is, the induced order when identifying L with its image under $\tilde{\varphi}$. Thus $\mathcal{O}$ orders L by Theorem 3.7. We still need to show that $\mathcal{O}$ extends O. Observe that $\operatorname{im} \varphi \subseteq \operatorname{im} \tilde{\varphi}$. In particular, $\varphi(O) \subseteq \operatorname{im} \tilde{\varphi}$. Moreover, since φ is an order-monomorphism, $\operatorname{im} \varphi \subseteq \Sigma_{R_2}^*$. Hence $\varphi(O) \subseteq \Sigma_{R_2}^* \cap \operatorname{im} \tilde{\varphi}$. Therefore, $O \subseteq \mathcal{O}$. We can now conclude that $((L, \mathcal{O}), \tilde{\varphi}) \in \mathcal{F}$, while $((K, O), \varphi) \prec ((L, \mathcal{O}), \tilde{\varphi})$. But this

contradicts the fact that $((K, O), \varphi)$ is a maximal element of $\mathcal{F}$. Therefore, K must be real closed.
By the construction of K, R_1 is a formally real, algebraic extension of K. Therefore, $K = R_1$ and hence $O = \Sigma_{R_1}^*$ by Theorem 3.19. Also, $\varphi : R_1 \to R_2$ is an order-F-monomorphism. All that remains is to show that φ is surjective. Consider the ordered subfield $(\operatorname{im}\varphi, \varphi(\Sigma_{R_1}^*))$ of R_2. Note that every element in $\varphi(\Sigma_{R_1}^*)$ is a square. Let $q \in (\operatorname{im}\varphi)[x]$ be a polynomial of odd degree. Then there exists a polynomial $\hat{q} \in R_1[x]$ such that $q = \varphi(\hat{q})$. Since R_1 is real closed, $\hat{q}$ has a root $\gamma \in R_1$. By Lemma 3.23, $\varphi(\gamma)$ is a root of q. Thus q has a root in $\operatorname{im}\varphi$ and so, by Corollary 3.27, $\operatorname{im}\varphi$ is real closed. It now follows that $R_2 = \operatorname{im}\varphi$. Therefore, φ is an order-F-isomorphism between R_1 and R_2. □

Notation: We can now unambiguously write $\widetilde{F}$ for the real closure of an ordered field (F, P).

Example 3.41: We exemplify here that two real closures of an ordered field need not be equal, but are merely order-isomorphic. Pick a square root $\sqrt{2}$ of 2 and consider $\mathbb{Q}(\sqrt{2})$. Since $\mathbb{Q}$ is formally real, so is $\mathbb{Q}(\sqrt{2})$ and hence admits an order. There are two choices: Either $\sqrt{2}$ is a positive element or $\sqrt{2}$ is a negative element. Let P_+ be the order on $\mathbb{Q}(\sqrt{2})$ where $\sqrt{2}$ is positive and P_- where it is negative. Let R_+ be the real closure of $(\mathbb{Q}(\sqrt{2}), P_+)$ and R_- the real closure of $(\mathbb{Q}(\sqrt{2}), P_-)$. Since R_+ and R_- both extend the order of $\mathbb{Q}$, they must both be real closures of $\mathbb{Q}$. These real closures are distinct: $\sqrt{-\sqrt{2}}$ is an element of R_-, but not of R_+, for $-\sqrt{2}$ is not positive there. They are still order-$\mathbb{Q}$-isomorphic by Theorem 3.40. In fact, the restriction of this order-$\mathbb{Q}$-isomorphism to $\mathbb{Q}(\sqrt{2})$ must send $\sqrt{2}$ to itself or $-\sqrt{2}$. The former is not order-preserving, hence it must be the $\mathbb{Q}$-automorphism that sends $\sqrt{2}$ to $-\sqrt{2}$.

Example 3.42: The real closure $\widetilde{\mathbb{Q}}$ of $\mathbb{Q}$ is isomorphic to the field of elements of $\mathbb{R}$ that are algebraic over $\mathbb{Q}$. That is to say, $\widetilde{\mathbb{Q}} \cong \overline{\mathbb{Q}} \cap \mathbb{R}$, where $\overline{\mathbb{Q}}$ is an algebraic closure of $\mathbb{Q}$. Let $\overline{\mathbb{Q}}_{\mathbb{R}} := \overline{\mathbb{Q}} \cap \mathbb{R}$. The induced order from $\mathbb{R}$ orders $\overline{\mathbb{Q}}_{\mathbb{R}}$, which clearly extends the order of $\mathbb{Q}$. We show that $\overline{\mathbb{Q}}_{\mathbb{R}}$ is virtually real. Let $\alpha \in \overline{\mathbb{Q}}_{\mathbb{R}}$ be positive. Let $f \in \overline{\mathbb{Q}}_{\mathbb{R}}[x]$ be such that $f(\alpha) = 0$. Then consider $g(x) := f(x^2) \in \overline{\mathbb{Q}}_{\mathbb{R}}[x]$. Any square root $\sqrt{\alpha}$ is a root of g and hence lies in $\overline{\mathbb{Q}}_{\mathbb{R}}$. Now let $h \in \overline{\mathbb{Q}}_{\mathbb{R}}[x]$ be of odd degree. Since $\mathbb{R}$ is real closed, h has a root $\beta \in \mathbb{R}$. Now $\overline{\mathbb{Q}}_{\mathbb{R}}(\beta)$ is algebraic over $\overline{\mathbb{Q}}_{\mathbb{R}}$, which in turn is algebraic over $\mathbb{Q}$. Thus, by Theorem 4.4.4 from [6], $\overline{\mathbb{Q}}_{\mathbb{R}}(\beta)$ is algebraic over $\mathbb{Q}$, which implies that $\overline{\mathbb{Q}}_{\mathbb{R}}(\beta) = \overline{\mathbb{Q}}_{\mathbb{R}}$. Consequently, h has a root in $\overline{\mathbb{Q}}_{\mathbb{R}}$. We see that $\overline{\mathbb{Q}}_{\mathbb{R}}$ is real closed and hence a real closure of $\mathbb{Q}$. Thus $\overline{\mathbb{Q}}_{\mathbb{R}} \cong \widetilde{\mathbb{Q}}$ by Theorem 3.40.

3.3.3 Discriminants

In section 1.2 we obtained some properties of the discriminant specifically for polynomials over $\mathbb{R}$. There we used the facts that $\mathbb{R}$ admits an order and $\mathbb{C}$ is algebraically closed. Since we now have generalized this to real closed fields, we also obtain generalizations of these statements. First we introduce some terminology by analogy with $\mathbb{R}$ and $\mathbb{C}$.

Definition 3.43: Let R be a formally real field and let $C := \widetilde{R}(\boldsymbol{i})$. We call an element in $\widetilde{R}$ *formally real* and an element in C *formally complex*. An element in $C \setminus \widetilde{R}$ is then called *non-real formally complex*.
For an element $a + b\boldsymbol{i} \in C$, where $a, b \in \widetilde{R}$, we define $\operatorname{Re}(a + b\boldsymbol{i}) := a$ and $\operatorname{Im}(a + b\boldsymbol{i}) := b$. These are called the *formally real* and *formally imaginary* parts of $a + b\boldsymbol{i}$, respectively.
The map $\overline{\cdot}$ is called *formally complex conjugation*, the element $\overline{a}$ is called the *formally complex conjugate* of a, with $a \in C$, and the pair $a, \overline{a}$ is called a *formally complex pair.*

Remark 3.44: We may occasionally drop the word ‘formally’ if the meaning remains clear.

Definition 3.45: Let (R, P) be an ordered field. We define the *signum* function $\operatorname{sgn} : R \to \{-1, 0, 1\}$ by

$$\operatorname{sgn}(r) := \begin{cases} 1 & \text{if } r \in P, \\ 0 & \text{if } r = 0, \\ -1 & \text{if } -r \in P. \end{cases}$$

***Lemma* 3.46:** Let R be a formally real field and let $f \in R[x]$. The formally complex roots of f come in complex pairs. That is to say, $\alpha \in \widetilde{R}(\boldsymbol{i})$ is a root of f if and only if $\overline{\alpha}$ is a root of f.

Proof. By Lemma 3.24, formally complex conjugation is an R-automorphism on $R(\boldsymbol{i})$. Then, since $\overline{f} = f$, Lemma 3.23 implies the desired result. □

***Theorem* 3.47** (Generalized Theorem 1.3)**:** Let R be a formally real field. Let $f \in R[x]$ be a polynomial of degree $n \geq 2$ with distinct roots $\alpha_1, \ldots, \alpha_n \in \widetilde{R}(\boldsymbol{i})$. Let r be half of the number of formally complex roots of f (note that $r \in \mathbb{N}_0$ by Lemma 3.46). Then $\operatorname{sgn}(\Delta(f)) = (-1)^r$. □

The proof is entirely analogous to the proof of Theorem 1.3.
We also get generalizations of Corollaries 1.4 and 1.5. The proofs of these are also analogous to the proofs of the original corollaries.

***Corollary* 3.48** (Generalized Corollary 1.4)**:** Let R be formally real and let $f \in R[x]$ be a polynomial of degree 2. Let α_1 and α_2 be the roots of f in $\widetilde{R}(\boldsymbol{i})$. Then the following hold:

- $\Delta(f) = 0$ if and only if $\alpha_1 = \alpha_2$;
- $\Delta(f) > 0$ if and only if $\alpha_1, \alpha_2 \in \widetilde{R}$ and $\alpha_1 \neq \alpha_2$;
- $\Delta(f) < 0$ if and only if $\alpha_1, \alpha_2 \in \widetilde{R}(\boldsymbol{i}) \setminus \widetilde{R}$ and $\alpha_1 = \overline{\alpha_2}$. □

***Corollary* 3.49** (Generalized Corollary 1.5)**:** Let R be formally real and let $f \in R[x]$ be a polynomial of degree 3. Let α_1, α_2 and α_3 be the roots of f in $\widetilde{R}(\boldsymbol{i})$. Then the following hold:

- $\Delta(f) = 0$ if and only if $\alpha_i = \alpha_j$ for some $i \neq j$;
- $\Delta(f) > 0$ if and only if $\alpha_1, \alpha_2, \alpha_3 \in \widetilde{R}$ and are all distinct;
- $\Delta(f) < 0$ if and only if one root is formally real and the other two form a formally complex pair. □

4 Polynomials with Formally Real Roots

In this chapter we take an in-depth look at polynomials over a formally real field R, whose roots lie in $\widetilde{R}$. In particular, we look at cubic polynomials with positive discriminants. Corollary 3.49 states that there indeed are three distinct formally real roots. Yet we shall see that the cube roots in Cardano's formula are non-real formally complex. When such a polynomial is irreducible over R, we shall see that it is impossible to express these roots using only radicals of formally real elements and so the complex radicals are necessary. This is the *casus irreducibilis.*

After this we generalize this concept to polynomials of higher degrees over R, whose roots are all formally real. We shall see that there are only very specific cases where the roots could be expressed using radicals of formally real elements only.

4.1 The Casus Irreducibilis

Let R be a formally real field with an order relation $>$ and let $C := \widetilde{R}(\boldsymbol{i})$ be its algebraic closure, where $\boldsymbol{i}$ is a root of x^2+1. Consider a third degree polynomial $f(x) := x^3 + px + q \in R[x]$ in the depressed cubic form. We may apply Cardano's formula to find the roots:

$$\frac{\omega^{i-1}}{3}\sqrt[3]{-\frac{27}{2}q+\frac{3}{2}\sqrt{-3\Delta}}+\frac{\omega^{1-i}}{3}\sqrt[3]{-\frac{27}{2}q-\frac{3}{2}\sqrt{-3\Delta}}, \qquad i=1,2,3,$$

where $\Delta := \Delta(f) = -4p^3 - 27q^2$, ω is a primitive third root of unity and the product of the cube roots equals $-3p$. Looking at the radicand -3Δ of the square root, we observe that if the discriminant of f is positive, then the radicand is negative. Thus $\sqrt{-3\Delta}$ does not lie in $\widetilde{R}$, because every square in $\widetilde{R}$ is non-negative. This means that the radicand of each cube root is non-real formally complex and, consequently, the cube roots themselves must also be non-real. The latter is true, because $\widetilde{R}$ is closed under multiplication, which means that the cube of a formally real element is again formally real. The peculiar part is that Corollary 3.49 asserts that all roots of f are formally real, despite their being expressed using formally complex radicals (i.e., radicals of formally complex elements). If specifically $f \in \mathbb{Q}[x]$ is reducible over $\mathbb{Q} \subseteq R$, then these complex radicals are overly complicated: One can use the rational root test to find a root of f in $\mathbb{Q}$. The other two roots can then be found with the quadratic formula, which uses only formally real radicals. However, when f is irreducible over $\mathbb{Q}$, the casus irreducibilis states that the roots of f cannot be expressed using formally real radicals alone.

We start with an old example of a reducible polynomial, which we got from [5, p. 18].

Example 4.1: Consider the polynomial $g(x) := x^3 - 15x - 4 \in \mathbb{Q}[x]$. The discriminant is $-4(-15)^3 - 27(-4)^2 = 13\,068 > 0$, so there are three distinct real roots. We compute $-3\Delta(g) = -39\,204 = (198\boldsymbol{i})^2$. We choose $198\boldsymbol{i}$ for a square root of $-3\Delta(g)$. The radicands of the cube roots in Cardano's formula are now $-\frac{27}{2}(-4) \pm \frac{3}{2} \cdot 198\boldsymbol{i} = 54 \pm 297\boldsymbol{i} = 3^3(2 \pm 11\boldsymbol{i})$. With suitably chosen cube roots, a root of g is

$$\tfrac{1}{3}\sqrt[3]{3^3(2+11\boldsymbol{i})}+\tfrac{1}{3}\sqrt[3]{3^3(2-11\boldsymbol{i})}=\sqrt[3]{2+11\boldsymbol{i}}+\sqrt[3]{2-11\boldsymbol{i}},$$

where the latter cube roots are chosen such that $\sqrt[3]{2 \pm 11\boldsymbol{i}} = \frac{1}{3}\sqrt[3]{3^3(2 \pm 11\boldsymbol{i})}$. In 1550 Bombelli observed that $(2 \pm \boldsymbol{i})^3 = 2 \pm 11\boldsymbol{i}$. Note that $3(2+\boldsymbol{i}) \cdot 3(2-\boldsymbol{i}) = 45 = -3(-15)$, hence these are correctly chosen cube roots. Therefore, we obtain the root $\frac{1}{3} \cdot 3(2+\boldsymbol{i}) + \frac{1}{3} \cdot 3(2-\boldsymbol{i}) = 4$ of g. Of course this root could have more easily been obtained via the rational root test.

In this example we notice that the cube roots are each other's complex conjugates. This is always the case when there are three real roots of $f \in R[x]$, where f is again as before. Since each cube root in Cardano's formula is then non-real, their formally imaginary parts must be equal in size and opposite in sign in order that they cancel when adding them together. That is to say, if $a + b\boldsymbol{i}$ and $c + d\boldsymbol{i}$ are correctly chosen cube roots in Cardano's formula, then $a + b\boldsymbol{i} + c + d\boldsymbol{i} \in \widetilde{R}$ and hence $d = -b$. Because the cube roots are correctly chosen, $(a + b\boldsymbol{i})(c - b\boldsymbol{i}) = -3p$ and so $a(-b) + bc = 0$ as $p \in \widetilde{R}$. Then $b \neq 0$ implies that $a = c$. Thus $c + d\boldsymbol{i} = a - b\boldsymbol{i} = \overline{a + b\boldsymbol{i}}$.

One may wonder, why we do not just compute a and b above to find the cube root via cubing $a + b\boldsymbol{i}$ and equaling this to the radicand of the cube root. This works well for square roots after all. In fact, we did just that in the proof of Lemma 3.25, where we showed that every element of C is a square. The following proposition shows why this attempt is futile for cube roots.

***Proposition* 4.2:** Let R be a formally real field and let $f(x) := x^3 + px + q \in R[x]$ be an irreducible polynomial with positive discriminant $\Delta = -4p^3 - 27q^2$. Then solving the equation

$$\frac{1}{3}\sqrt[3]{-\frac{27}{2}q + \frac{3}{2}\sqrt{-3\Delta}} = a + b\boldsymbol{i} \tag{4.1}$$

for $a, b \in \widetilde{R}$ amounts to finding a root of f.

Proof. First we cube both sides of (4.1) to obtain

$$\frac{1}{27}\left(-\frac{27}{2}q + \frac{3}{2}\sqrt{-3\Delta}\right) = a^3 - 3ab^2 + \boldsymbol{i}(3a^2b - b^3).$$

Let $\sqrt{3\Delta}$ be the square root such that $\boldsymbol{i}\sqrt{3\Delta} = \sqrt{-3\Delta}$ and rewrite the left-hand side as follows:

$$\frac{1}{27}\left(-\frac{27}{2}q + \frac{3}{2}\boldsymbol{i}\sqrt{3\Delta}\right) = -\frac{q}{2} + \frac{\boldsymbol{i}\sqrt{3\Delta}}{2\cdot 9} = -\frac{q}{2} + \boldsymbol{i}\sqrt{-\frac{p^3}{27} - \frac{q^2}{4}}.$$

Note that the radicand of the last square root is positive and hence the square root lies in $\widetilde{R}$. For convenience, write $\varsigma := \sqrt{-\frac{p^3}{27} - \frac{q^2}{4}}$. Thus we obtain the simultaneous equations

$$\begin{cases} a^3 - 3ab^2 = -\frac{q}{2}, \\ 3a^2b - b^3 = \varsigma. \end{cases}$$

We shall need to divide by a, hence we need to show that $a \neq 0$. If $a = 0$, then $q = 0$ and so $f(0) = 0$, which means that f is not irreducible. This is a contradiction, so $a \neq 0$. As suggested in [16], we solve for b^2 and b. This yields

$$b^2 = \frac{a^3 + \frac{q}{2}}{3a}, \qquad b = \frac{b^3 + \varsigma}{3a^2}. \tag{4.2}$$

To find b^3, we multiply these two together:

$$b^3 = b^2 b = \frac{a^3 + \frac{q}{2}}{3a}\frac{b^3 + \varsigma}{3a^2} = \frac{b^3(a^3 + \frac{q}{2}) + \varsigma(a^3 + \frac{q}{2})}{9a^3}.$$

Solving for b^3 now yields

$$b^3 = \frac{\varsigma(a^3 + \frac{q}{2})}{8a^3 - \frac{q}{2}}.$$

We can now express b in terms of a by plugging b^3 into (4.2):

$$b = \frac{\frac{\varsigma(a^3 + \frac{q}{2})}{8a^3 - \frac{q}{2}} + \varsigma}{3a^2} = \frac{\varsigma(a^3 + \frac{q}{2}) + \varsigma(8a^3 - \frac{q}{2})}{3a^2(8a^3 - \frac{q}{2})} = \frac{9a^3\varsigma}{3a^2(8a^3 - \frac{q}{2})} = \frac{6a\varsigma}{16a^3 - q}.$$

By squaring this equation, we obtain another expression for b^2. Combining this with (4.2), we obtain:

$$b^2 = \left(\frac{6a\varsigma}{16a^3 - q}\right)^2 = \frac{36a^2\varsigma^2}{256a^6 - 32qa^3 + q^2} = \frac{a^3 + \frac{q}{2}}{3a}.$$

We reduce the last equality as follows:

$$\begin{aligned} (256a^6 - 32qa^3 + q^2)(a^3 + \tfrac{q}{2}) &= 3a \cdot 36a^2\varsigma^2 \\ 256a^9 + 96qa^6 - 15q^2a^3 + \tfrac{q^3}{2} &= 108a^3(-\tfrac{p^3}{27} - \tfrac{q^2}{4}) \\ 256a^9 + 96qa^6 + (4p^3 + 12q^2)a^3 + \tfrac{q^3}{2} &= 0 \\ (2a)^9 + 3q(2a)^6 + (p^3 + 3q^2)(2a)^3 + q^3 &= 0. \end{aligned}$$

Next substitute $(2a)^3 = y - q$ to obtain the depressed cubic form in y:

$$\begin{aligned}(y-q)^3 + 3q(y-q)^2 + (p^3 + 3q^2)(y-q) + q^3 &= 0 \\ y^3 + p^3 y - p^3 q &= 0.\end{aligned}$$

Observe that $p \neq 0$, for otherwise $\Delta = -27q^2 \leq 0$. Thus we can divide by $-p^3$:

$$\frac{y^3}{-p^3} - y + q = 0.$$

Finally substitute $y = -pz$ to obtain

$$\frac{(-pz)^3}{-p^3} - (-pz) + q = z^3 + pz + q = 0.$$

Thus we see that we need to find a root of the original polynomial to solve (4.1). □

We now move on to formally stating and proving the casus irreducibilis over formally real fields.
Recall that an algebraic extension E of a field F is a *radical extension* of F if there exists a field tower

$$F = E_0 \subset E_1 \subset \cdots \subset E_k = E, \tag{4.3}$$

such that $E_i = E_{i-1}(r_i)$ with r_i a radical over E_{i-1} for each i. An element that is contained in a radical extension is *expressible by radicals.* A polynomial $f \in F[x]$ is *solvable by radicals* if its splitting field over F is contained in a radical extension of F. The corresponding definitions for formally real radicals are as follows:

Definition 4.3: Let R be a formally real field. We call a radical extension E of R a *formally real radical extension* if $E \subseteq \widetilde{R}$. An element contained in such an extension is *expressible by formally real radicals.* A polynomial $f \in R[x]$ is *solvable by formally real radicals* if its splitting field over R is contained in a formally real radical extension of R.

In this chapter we shall generally only work with formally real fields, which have characteristic 0. Thus all radicals are roots of polynomials of the form $x^n - a \in F[x]$ for some $n \in \mathbb{N}$. In fact, we only need radicals of prime degree, i.e., when n is a prime number. The following proposition shows this.

***Proposition* 4.4:** Let F be a field of characteristic 0. Let $d := p_1 \cdots p_k$, where $k \in \mathbb{N}$ and each $p_i \in \mathbb{N}$ is prime. Suppose $F(r_k)$ is a proper radical extension of F, where r_k is a root of $x^d - r_0 \in F[x]$. Then there exists a tower of radical extensions $F = E_0 \subset E_1 \subset \cdots \subset E_k = F(r_k)$ such that $E_i = E_{i-1}(r_i)$, where r_i is a root of $x^{p_i} - r_{i-1}$ for $1 \leq i \leq k$.

Proof. If $k = 1$, then we are done. So suppose $k > 1$. Define $r_i := r_k^{p_{i+1}\cdots p_k}$ for each $1 \leq i < k$. Then r_i is a root of $x^{p_i} - r_{i-1}$ for $1 \leq i \leq k$, because $(r^{p_{i+1}\cdots p_k})^{p_i} - r_{i-1} = r^{p_i \cdots p_k} - r^{p_i \cdots p_k} = 0$. Then setting $E_0 := F$ and $E_i := E_{i-1}(r_i)$ for $1 \leq i \leq k$ yields the tower of radical extensions

$$F = E_0 \subset E_1 \subset \cdots \subset E_k.$$

Note also that $E_k = F(r_k)$, hence we have the desired field tower. □

By applying Proposition 4.4 to each extension of (4.3), we see that every radical extension can be realized as a tower of radical extensions, where each extension is obtained by adjoining a radical of prime degree to the preceding extension.
Before we can state the main theorem of this section, we need some properties of the polynomials that realize the radicals: the polynomials of the form $x^p - a$. We have the following proposition about their reducibility, whose proof is based on that of Proposition 4.2.6 from [5].

***Proposition* 4.5:** Let F be a field of characteristic 0 and let p be a prime number. If $f(x) := x^p - a \in F[x]$ is reducible over F, then f has a root in F.

Proof. If $a = 0$, then clearly $0 \in F$ is a root. So suppose $a \neq 0$. Let α be a root of f in some extension field of F. We know that the remaining roots are $\zeta\alpha, \dots, \zeta^{p-1}\alpha$, where ζ is a primitive p-th root of unity. Thus in the splitting field E of f over F we have $f(x) = (x-\alpha)(x-\zeta\alpha)\cdots(x-\zeta^{p-1}\alpha)$. We assume that f is reducible over F and prove it has a root in F. There exists an irreducible polynomial $g \in F[x]$ such that $f = gh$, where $h \in F[x]$. If the leading coefficient of g is b, then the leading coefficient of h is b^{-1} as f is monic. Thus we can multiply g by b^{-1} and h by b to obtain a product of monic polynomials without changing the resulting product, i.e., we may assume, without loss of generality, that g is monic. Let m be the degree of g and note that $m < p$. Then there are numbers $n_1, \dots, n_m \in \mathbb{Z}/p$ such that $\zeta^{n_1}\alpha, \dots, \zeta^{n_m}\alpha$ are the roots of g. Since factorization is unique in E, g must be equal to $(x-\zeta^{n_1}\alpha)\cdots(x-\zeta^{n_m}\alpha)$. Because $g \in F[x]$, the constant term $\zeta^{n_1}\cdots\zeta^{n_m}\alpha^m$ must lie in F as well. Let $\xi := \zeta^{n_1+\dots+n_m}$ and notice that $\xi^p = 1$. The $\gcd(m,p) = 1$, for p is prime. Thus there exist $s, t \in \mathbb{Z}$ such that $sm + tp = 1$. We now find

$$\xi^s\alpha = \xi^s\alpha^{sm+tp} = (\xi\alpha^m)^s(\alpha^p)^t.$$

Because $\xi\alpha^m \in F$ and $\alpha^p = a \in F$, we see that $\xi^s\alpha \in F$. Now observe that $(\xi^s\alpha)^p = 1^s a = a$, thus $\xi^s\alpha$ is a root of f lying in F. □

We also need the following two lemmas.

***Lemma* 4.6:** Let R be a formally real field. There is no primitive p-th root of unity contained in R, where p is an odd prime.

Proof. We prove by contradiction. Suppose $\zeta \in R$ is a primitive p-th root of unity. Since $\zeta = (\zeta^{\frac{p+1}{2}})^2$, we see that $\zeta > 0$. If $\zeta > 1$, then $\zeta^2 > \zeta > 1$. This implies that $\zeta^2 > 1$. From there we obtain that $\zeta^3 > \zeta$ and so $\zeta^3 > 1$. Repeating this argument eventually yields that $\zeta^p > \zeta^{p-1} > 1$ and so $\zeta^p > 1$. But $\zeta^p = 1$, therefore $1 > 1$, which contradicts the trichotomy. An analogous argument yields the same contradiction if $\zeta < 1$. □

***Lemma* 4.7:** Let R be a formally real field. Let p be a prime number and let $\alpha \in \widetilde{R}$ such that $\alpha^p \in R$. If $\alpha \notin R$, then $x^p - \alpha^p$ is irreducible over R. Furthermore, $[R(\alpha) : R] = p$.

Proof. By Proposition 4.5, it suffices to show that there is no root of $x^p - \alpha^p$ contained in R. Suppose that there is a root $\beta \in R$. Since $\alpha \notin R$, we see that $\alpha \neq 0$. Thus $\frac{\beta}{\alpha} \in \widetilde{R}$. Now observe that $(\frac{\beta}{\alpha})^p = \frac{\beta^p}{\alpha^p} = 1$ while $\frac{\beta}{\alpha} \neq 1$. Thus $\frac{\beta}{\alpha}$ is a p-th root of unity. If ζ is a primitive p-th root of unity, then $\langle\zeta\rangle$, the group generated by ζ, is a group of p elements: the p-th roots of unity. Hence each element has either order 1 or order p, for p is prime. Since $\frac{\beta}{\alpha} \neq 1$, it must have order p in this group. Therefore, it is a primitive p-th root of unity. If p is odd, then this contradicts Lemma 4.6. Thus $p = 2$ and so $\frac{\beta}{\alpha} = -1$, which implies that $\beta = -\alpha$, contradicting the assumption that $\alpha \notin R$. Therefore, we see that R does not contain a root of $x^p - \alpha^p$ and hence this polynomial is irreducible over R. We immediately obtain that $[R(\alpha) : R] = p$. □

We can now state and prove the casus irreducibilis over any formally real field. The idea of the proof came from [15, p. 178].

***Theorem* 4.8** (Casus Irreducibilis)**:** Let R be a formally real field and let $f \in R[x]$ be an irreducible polynomial of degree 3 with positive discriminant. Then f is not solvable by formally real radicals.

Proof. Let $\Delta := \Delta(f)$. Since Δ is positive, the radical extension $R(\sqrt{\Delta})$ is formally real. Moreover, $[R(\sqrt{\Delta}) : R] \in \{1, 2\}$, depending on whether Δ is a square in R. Since $\deg f = 3$, f is still irreducible over $R(\sqrt{\Delta})$. Let E be the splitting field of f over R. From Theorem 2.1 we find that $[E : R(\sqrt{\Delta})] = 3$—note that this part of the proof of this theorem does not require that the field contain a primitive third root of unity. Moreover, if α is a root of f, then $[R(\sqrt{\Delta}, \alpha) : R(\sqrt{\Delta})] = 3$. Thus we see that $E = R(\sqrt{\Delta}, \alpha)$, because E contains $R(\sqrt{\Delta}, \alpha)$ and both extensions are of degree 3 over $R(\sqrt{\Delta})$. We assume that f is solvable by formally real radicals to reach a contradiction. We obtain a tower of algebraic extensions

$$R \subseteq R(\sqrt{\Delta}) = E_0 \subset E_1 \subset \cdots \subset E_{k-1} \subset E_k,$$

where $E_i = E_{i-1}(r_i)$ with r_i a formally real radical over E_{i-1} of prime degree p_i for each i. Note that, by Lemma 4.7, $[E_i : E_{i-1}] = p_i$ for each i. We claim that there is such a tower, such that f has no root in E_i for each $i < k$, but f splits over E_k. Moreover, we claim that E_k is obtained from E_{k-1} by adjoining a cube root of a non-cube element of E_{k-1}.

Firstly, we simply take sufficiently few radical extensions so that f splits only over the last one, which will be E_k. Secondly, suppose E_i contains a root α of f, but E_{i-1} does not. We show that $i = k$. We have $E_{i-1} \subset E_{i-1}(\alpha) \subseteq E_i$. Since f is irreducible over E_{i-1}, we see that $[E_{i-1}(\alpha) : E_{i-1}] = 3$ and therefore we find $p_i = [E_i : E_{i-1}] = 3[E_i : E_{i-1}(\alpha)]$. Since p_i is prime, we must have that $p_i = 3$ and hence $E_i = E_{i-1}(\alpha)$. Since $E_i \supseteq R(\sqrt{\Delta}, \alpha)$, we see that f splits over E_i. This means that $i = k$, by the assumption that f only splits over the extension E_k.

Finally, as the foregoing showed that $p_k = 3$, we find that E_k is obtained from E_{k-1} by adjoining a cube root of some $a \in E_{k-1}$. Furthermore, E_k is the splitting field of f over E_{k-1}, whence it is a normal extension. Therefore, E_k contains all roots of $x^3 - a \in E_{k-1}[x]$. These roots are $\sqrt[3]{a}$, $\omega\sqrt[3]{a}$ and $\omega^2\sqrt[3]{a}$, where ω is a primitive third root of unity. Because $a \neq 0$, this means that $\omega = \frac{\omega\sqrt[3]{a}}{\sqrt[3]{a}} \in E_k$, which contradicts Lemma 4.6. Therefore, we can conclude that no such tower exists and hence f is not solvable by formally real radicals. □

The classical casus irreducibilis is the case where $R = \mathbb{Q}$ considered as a subfield of $\mathbb{R}$.

4.2 A Generalization to Higher Degree Polynomials

In the previous section we have only considered the case of polynomials of degree 3 with formally real roots. We know from the quadratic formula that any polynomial of degree 2 with distinct formally real roots is solvable by formally real radicals: One extends the base field with the square root of the positive discriminant. As it turns out, if a polynomial is solvable by formally real radicals, then those radicals are all square roots. This is one of the main statements that we shall prove in this section. We primarily follow Cox's approach (see section 8.6 A-B of [5]).

We shall first generalize Theorem 4.8 to extensions of degree p for any odd prime, not just 3. To do this, we need to define a notion of combining two subfields:

Definition 4.9: Let M be a field with subfields K and L. We define the *compositum* of K and L to be the smallest subfield of M that contains both K and L, denoted by $\langle K, L\rangle$. Formally, this is defined by the intersection of all such subfields:

$$\langle K, L\rangle \coloneqq \bigcap_{\substack{F \text{ a field} \\ K,L \subseteq F \subseteq M}} F.$$

We immediately have the following property of the compositum.

***Proposition* 4.10:** Let K and L be subfields of a field M. Then $\langle K, L\rangle = K$ if and only if $L \subseteq K$.

Proof. The direct implication is trivial, as $L \subseteq \langle K, L\rangle = K$. Now we prove the converse implication. Note that $K \subseteq K$ and $L \subseteq K$. Hence, since $\langle K, L\rangle$ is the smallest subfield of M that contains both K and L, it follows that $\langle K, L\rangle \subseteq K$. Because $K \subseteq \langle K, L\rangle$ by definition, we see that $\langle K, L\rangle = K$. □

We also have the following relation to radical extensions.

***Proposition* 4.11:** Let F, K, L and M be fields such that $F \subseteq K, L$ and $K, L \subseteq M$ and K is a radical extension of F. Then $\langle K, L\rangle$ is a radical extension of L. Moreover, if

$$F = E_0 \subset E_1 \subset \cdots \subset E_k = K,$$

where $E_i = E_{i-1}(r_i)$ and r_i a radical over E_{i-1} for $1 \leq i \leq k$, is the tower corresponding to the radical extension K/F, then $\langle K, L\rangle = L(r_1, \ldots, r_k)$.

Proof. We first show the last equality. Since $F(r_1, \ldots, r_k) = K \subseteq \langle K, L\rangle$, we see that $L(r_1, \ldots, r_k) \subseteq \langle K, L\rangle$. Because $F \subseteq L$, we see that $K \subseteq L(r_1, \ldots, r_k)$. Because also $L \subseteq L(r_1, \ldots, r_k)$, it follows that $\langle K, L\rangle \subseteq L(r_1, \ldots, r_k)$ by the definition of the compositum. Now define $L_0 \coloneqq L$ and $L_i \coloneqq L_{i-1}(r_i)$ for $1 \leq i \leq k$. We see that $L_k = L(r_1, \ldots, r_k) = \langle K, L\rangle$ and hence we have the tower

$$L = L_0 \subset L_1 \subset \cdots \subset L_k = \langle K, L\rangle,$$

wherefore $\langle K, L\rangle$ is a radical extension of L. □

With the notation as in the proof above, it follows immediately that if M is formally real and K is a formally real radical extension of F, then also $\langle K, L\rangle$ is a formally real radical extension of L.
Using the foregoing, we now have the following generalization of Theorem 4.8 thanks to [5, p. 221]:

***Theorem* 4.12:** Let R be a formally real field and let $L \subseteq \widetilde{R}$ be a Galois extension of R of degree p, where p is an odd prime. Then L is not contained in a formally real radical extension of R.

Proof. Let $r \in \widetilde{R}$ be a formally real radical of prime degree q over R, i.e., $r \notin R$ and $r^q \in R$. By Lemma 4.7, the degree $[R(r) : R]$ is equal to q. We first show via contradiction that $r \notin L$. Suppose $r \in L$. Then we have $p = [L : R] = [L : R(r)][R(r) : R] = q[L : R(r)]$. Since p is prime, we must have that $p = q$, which means that q is odd. Note that $x^q - r^q$ is irreducible over R by Lemma 4.7 and that L/R is a normal extension. Because L contains a root of $x^q - r^q$, it follows that $x^q - r^q$ splits over L. Therefore, L must contain a primitive q-th root of unity. But q is an odd prime and L is formally real, so this contradicts Lemma 4.6. Whence $r \notin L$.
Lemma 4.7 implies that $[L(r) : L] = q$. Since both L and $R(r)$ are intermediate fields of $L(r)/R$, we have the equality $[L(r) : L][L : R] = [L(r) : R(r)][R(r) : R]$. From this we find

$$[L(r) : R(r)] = \frac{[L(r) : L][L : R]}{[R(r) : R]} = \frac{q[L : R]}{q} = [L : R] = p. \tag{4.4}$$

This means that adding any formally real radical of prime degree to both L and R does not affect the degree. Now let K be any formally real radical extension of R. By Proposition 4.4, we may assume that K can be obtained by adding radicals of prime degree. That is to say, we have the tower

$$R = E_0 \subset E_1 \subset \cdots \subset E_k = K,$$

where $E_i = E_{i-1}(r_i)$ with r_i a formally real radical over E_{i-1} of degree p_i, where p_i is prime. Since K and L are contained in $\widetilde{R}$, we can use the compositum of K and L. Proposition 4.11 implies that $\langle K, L\rangle = L(r_1, \ldots, r_k)$. Since $K = R(r_1, \ldots, r_k)$, we can apply (4.4) repeatedly to obtain

$$\begin{aligned}[\langle K, L\rangle : K] &= [L(r_1, \ldots, r_k) : R(r_1, \ldots, r_k)] \\ &= [L(r_1, \ldots, r_{k-1}) : R(r_1, \ldots, r_{k-1})] \\ &\,\vdots \\ &= [L : R] = p.\end{aligned}$$

This shows that $\langle K, L\rangle$ is a proper extension of K. Thus $K \neq \langle K, L\rangle$, which implies that $L \nsubseteq K$ for Proposition 4.10. We conclude that L is not contained in any formally real radical extension of R, because K was an arbitrary one. □

To see that this theorem indeed generalizes Theorem 4.8, let f be an irreducible cubic polynomial over a formally real field F with positive discriminant. Take $R := F(\sqrt{\Delta(f)})$ and L the splitting field of f over R. We see that L is a Galois extension of R of degree 3, hence L is not contained in a formally real radical extension of R by the theorem above. Thus f is not solvable by formally real radicals. So Theorem 4.12 indeed implies Theorem 4.8.
As we mentioned at the start of this section, there is an even stronger version of the casus irreducibilis for polynomials of higher degrees. For this we shall need a lemma, which is based on the following well-known generalized statement of the first Sylow theorem (see also Theorem 2.15.3 from [6]):

***Theorem* 4.13** (Sylow)**:** Let G be a group of order $n \in \mathbb{N}$ and let p be a prime dividing n. Then G contains a subgroup of order p^k for each $k \in \mathbb{N}_0$ such that $p^k \mid |G|$. □

***Lemma* 4.14:** Let p be a prime number, $n \in \mathbb{N}$ and let G be a group of order p^n. Then there exists a chain of subgroups

$$\{e\} = G_n < G_{n-1} < \cdots < G_1 < G_0 = G$$

such that $|G_i| = p^{n-i}$ for each i.

Proof. Since p^{n-1} divides $|G|$, there is a subgroup G_1 of G of order p^{n-1}. Now p^{n-2} divides the order of G_1, so it contains a subgroup G_2 of order p^{n-2}. Repeating this argument, and setting $G_0 := G$, yields the desired chain of subgroups. □

We also need the following lemma:

***Lemma* 4.15:** Let R be a formally real field. If E/R is a formally real extension of degree 2, then $E = R(\sqrt{a})$ for some positive $a \in R$.

Proof. Since E/R is finite, it is algebraic. Then $E = R(\alpha)$ for some $\alpha \in E$. The minimal polynomial of α over R is of the form $f(x) := x^2 + px + q$. Since $\operatorname{Char} R \neq 2$, the quadratic formula shows that we need only adjoin the square root of the discriminant to R. Since E is formally real, the roots of f are formally real, wherefore $\Delta(f) > 0$. Thus $E = R(\sqrt{\Delta(f)})$, as desired. □

Finally, we state and prove a generalization of the casus irreducibilis as it is done in [5, p. 222]. The original results were proved by Hölder ([9]) in 1891 and Isaacs ([10]) in 1985, independently.

***Theorem* 4.16** (Generalized Casus Irreducibilis)**:** Let R be a formally real field and let $f \in R[x]$ be an irreducible polynomial, whose splitting field L is contained in $\widetilde{R}$. Thus f has only formally real roots. The following are equivalent:

(i) Some root of f can be expressed by formally real radicals over R;

(ii) All roots of f can be expressed by formally real square roots over R;

(iii) L is a formally real radical extension of R;

(iv) $[L : R]$ is a power of 2.

Proof. We prove the following implications:

$$\text{(iv)} \implies \text{(iii)} \implies \text{(i)} \implies \text{(iv)} \implies \text{(ii)} \implies \text{(i)}.$$

The implications (iii) $\implies$ (i) and (ii) $\implies$ (i) are immediate.
Assume (iv). We show that both (iii) and (ii) hold. Note that f is separable, as R has characteristic 0. Since L is the splitting field of f, it is a Galois extension of R. Therefore, $|\operatorname{Gal}(L/R)| = [L : R] = 2^n$ for some $n \in \mathbb{N}$. By Lemma 4.14, we get a chain of subgroups of $\operatorname{Gal}(L/R)$

$$\{e\} = G_n < G_{n-1} < \cdots < G_1 < G_0 = \operatorname{Gal}(L/R)$$

such that $|G_i| = 2^{n-i}$ for each i. By the fundamental theorem of Galois theory, we obtain a chain of subfields

$$R = L_0 \subset L_1 \subset \cdots \subset L_n = L,$$

where $L_i := L^{G_i}$, the fixed field of G_i in L, for each i. By the Galois correspondence, $G_i = \operatorname{Gal}(L/L_i)$ for each i. Thus $[L : L_i] = |G_i| = 2^{n-i}$. Since $[L : R] = [L : L_i][L_i : R]$, we have $[L_i : R] = \frac{[L:R]}{[L:L_i]} = \frac{2^n}{2^{n-i}} = 2^i$. Consequently, $[L_i : L_{i-1}] = 2$ for every i. Therefore, by Lemma 4.15, each extension L_i/L_{i-1} is obtained by adjoining a formally real square root of an element of L_{i-1}. We conclude that L is a formally real radical extension of R, in which only square roots appear. Thus (iii) and (ii) hold.
The only implication we are left to prove is (i) $\implies$ (iv). Let α be a root of f, which is contained in a formally real radical extension K of R. We prove by contradiction. So suppose that $[L : R]$ is not a power of 2 and let p be an odd prime dividing $[L : R]$. Thus p divides $|\operatorname{Gal}(L/R)|$. To arrive at a contradiction, we shall use an element of $\operatorname{Gal}(L/R)$ of order p that does not fix α. We now show that such an element exists.
By Cauchy's theorem, there exists an element $\sigma \in \operatorname{Gal}(L/R)$ of order p. Let $n := \deg f$ and let $\alpha =: \alpha_1, \alpha_2, \ldots, \alpha_n \in L$ be the roots of f. Since $L = R(\alpha_1, \ldots, \alpha_n)$ and σ is not the identity, we see that $\sigma(\alpha_i) \neq \alpha_i$ for some i. Because f is irreducible over R, by Theorem 0.5, there exists an R-automorphism $\varphi : L \to L$ sending α to α_i. Then $\varphi^{-1}\sigma\varphi$ has order p in $\operatorname{Gal}(L/R)$ and

$$\varphi^{-1}\sigma\varphi(\alpha) = \varphi^{-1}\sigma(\alpha_i) \neq \varphi^{-1}(\alpha_i) = \alpha.$$

Thus we have $\psi := \varphi^{-1}\sigma\varphi \in \operatorname{Gal}(L/R)$ of order p that does not fix α. We can now derive a contradiction. Let $M := L^{\langle\psi\rangle}$, the fixed field of the group generated by ψ in L. By the fundamental theorem of Galois theory, $M \subseteq L$ and

$$[L : M] = |\operatorname{Gal}(L/M)| = |\langle\psi\rangle| = p.$$

Thus, by Theorem 4.12, L does not lie in a formally real radical extension of M.
Because $\psi(\alpha) \neq \alpha$, it follows that $\alpha \notin M$. Since L is the splitting field, $\alpha \in L$. We have

$$p = [L : M] = [L : M(\alpha)][M(\alpha) : M].$$

Since p is prime and $[M(\alpha) : M] > 1$, it follows that $[M(\alpha) : M] = p$ and $[L : M(\alpha)] = 1$. Consequently, $L = M(\alpha)$. Because both K and M are contained in $\widetilde{R}$, we have the compositum $\langle K, M\rangle$ of K and M. As $L = M(\alpha)$ and $\alpha \in K$, we see that $L \subseteq \langle K, M\rangle$. Since R is contained in K and M, and K is a formally real radical extension of R, by Proposition 4.11, $\langle K, M\rangle$ is a formally real radical extension of M. Thus L is contained in a formally real radical extension of M, which contradicts our previous argument. We conclude that $[L : R]$ must be a power of 2. □

We immediately obtain that the polynomials that *could* be solvable by formally real radicals are the ones whose degrees are a power of 2:

***Corollary* 4.17:** Let R be a formally real field and let $f \in R[x]$ be an irreducible polynomial that splits over $\widetilde{R}$. If $\deg f$ is not a power of 2, then f is not solvable by formally real radicals.

Proof. Let $L \subseteq \widetilde{R}$ be the splitting field of f over R and let $\alpha \in L$ be a root of f. Because f is irreducible over R, $[R(\alpha) : R] = \deg f$ and hence $\deg f$ divides $[L : R]$. If $\deg f$ is not a power of 2, neither is $[L : R]$. The equivalence (i) $\iff$ (iv) shows that none of the roots of f are expressible by formally real radicals. □

We also have a relation to the constructibility with compass and straightedge of the roots of polynomials over $\mathbb{Q}$. Recall that an element $\rho \in \mathbb{R}$ is constructible if and only if there exists a tower of extensions $\mathbb{Q} = E_0 \subset E_1 \subset \cdots \subset E_k$ such that $\rho \in E_k$ and $[E_i : E_{i-1}] = 2$ for each i.

***Corollary* 4.18:** Let $f \in \mathbb{Q}[x]$ be an irreducible polynomial, whose splitting field E lies in $\mathbb{R}$. If f has at least one root that is expressible by real radicals, then all roots of f are constructible with compass and straightedge.

Proof. From the implications (i) $\implies$ (iv) $\implies$ (ii) of Theorem 4.16 we obtain, as in the proof, a chain of extensions

$$\mathbb{Q} = E_0 \subset E_1 \subset \cdots \subset E_k = E,$$

where $[E_i : E_{i-1}] = 2$ for each i. Thus each root of f is constructible with compass and straightedge. □

4.3 Non-Algebraic Real Solutions to Cubic Polynomials

Let f be an irreducible polynomial over a subfield of $\mathbb{R}$, whose discriminant is positive. In section 4.1 we saw, in particular, that we need complex numbers to express the roots f by radicals. Although in this case complex radicals are necessary, there is a way to express the roots in purely real quantities: using trigonometric functions. This is of course not an algebraic solution. The solution is based on the triple-angle formula for the cosine:

$$4\cos^3(\theta) - 3\cos(\theta) = \cos(3\theta) \tag{4.5}$$

for all $\theta \in \mathbb{R}$. This means that $\cos(\theta)$ is a root of the polynomial $4x^3 - 3x - \cos(3\theta) \in \mathbb{Q}(\cos(3\theta))[x]$. Suppose the discriminant of this polynomial is positive, so that there are three real roots. If we substitute $\theta + \frac{2k\pi}{3}$ for θ with $k \in \mathbb{Z}$, then we see that $\cos(\theta + \frac{2k\pi}{3})$ is a root of

$$4x^3 - 3x - \cos(3(\theta + \tfrac{2k\pi}{3})) = 4x^3 - 3x - \cos(3\theta + 2k\pi) = 4x^3 - 3x - \cos(3\theta).$$

Therefore, all roots are given by $\cos(\theta + \frac{2k\pi}{3})$ with $k = 0, 1, 2$. We shall use this to derive the roots of any depressed cubic polynomial with distinct real roots.
Let F be a subfield of $\mathbb{R}$ and let $f(x) := x^3 + px + q \in F[x]$ be a polynomial with positive discriminant. Thus $-4p^3 - 27q^2 > 0$. First we note that $p < 0$, for otherwise $\Delta(f) \leq 0$ as $q^2 \geq 0$. The idea, which originally comes from François Viète, is to substitute $u\cos(\theta)$ for x with suitably chosen u in order that the equation $f(x) = 0$ reduce to the triple-angle identity (4.5). The constant term shall then coincide

with $-\cos(3\theta)$, yielding a solution. Executing this substitution yields $u^3\cos^3(\theta)+pu\cos(\theta)+q$. Looking at (4.5), we get the simultaneous equations

$$\begin{cases} u^3 = 4, \\ pu = -3. \end{cases}$$

From the second equation we find $4 = -\frac{4p}{3}u$. Combining this with the first equation yields $u^3 = -\frac{4p}{3}u$. Lest the polynomial become constant, we require that u be non-zero. Thus we obtain $u^2 = -\frac{4p}{3}$. Since $p < 0$, the right-hand side is positive and so we can take a real square root $u := \sqrt{-\frac{4p}{3}} = 2\sqrt{-\frac{p}{3}} \in \mathbb{R}$. By applying the substitution, we obtain the equation

$$-\frac{8p}{3}\sqrt{-\frac{p}{3}}\cos^3(\theta) + 2p\sqrt{-\frac{p}{3}}\cos(\theta) + q = 0.$$

Since $p \neq 0$, we can divide both sides by $-\frac{2p}{3}\sqrt{-\frac{p}{3}}$ to obtain

$$4\cos^3(\theta) - 3\cos(\theta) - \frac{3q}{2p\sqrt{-\frac{p}{3}}} = 0.$$

From (4.5) follows that $\cos(3\theta) = \frac{3q}{2p\sqrt{-\frac{p}{3}}}$. To see that such a θ in fact exists, we need to show that the right-hand side lies in $[-1, 1]$. We show, in fact, that it lies in $(-1, 1)$ by showing that $|3q| < \left|2p\sqrt{-\frac{p}{3}}\right|$. We start by squaring both sides. We obtain:

$$\begin{aligned} 9q^2 &< 4p^2 \cdot -\frac{p}{3} \\ 27q^2 &< -4p^3 \\ 0 &< -4p^3 - 27q^2. \end{aligned}$$

The latter is true, because the right-hand side is precisely the discriminant of f. Thus we see that $\frac{3q}{2p\sqrt{-\frac{p}{3}}}$ lies in the image of the cosine. We take $\arccos : [-1, 1] \to [0, \pi]$ as usual, which yields

$$\theta = \frac{1}{3}\arccos\left(\frac{3q}{2p\sqrt{-\frac{p}{3}}}\right).$$

Finally, we invert the substitution $x = u\cos(\theta)$ with $u = 2\sqrt{-\frac{p}{3}}$ to find all three roots of f, which are displayed in (4.6).

***Theorem* 4.19:** Let $f(x) := x^3 + px + q \in \mathbb{R}[x]$ be a polynomial with positive discriminant. Then the roots of f are given by the following expression:

$$2\sqrt{-\frac{p}{3}}\cos\left(\frac{1}{3}\arccos\left(\frac{3q}{2p\sqrt{-\frac{p}{3}}}\right) + \frac{2k\pi}{3}\right), \qquad k = 0, 1, 2, \tag{4.6}$$

which is entirely real throughout. □

To exemplify this formula, we use the example from the Introduction:

Example 4.20: Let $f(x) := x^3 - 3x + 1 \in \mathbb{Q}[x]$. As we saw in the Introduction, the discriminant of f is 81 and f is irreducible over $\mathbb{Q}$. Thus we are in the situation of the casus irreducibilis and so we apply the formula from Theorem 4.19 with $p = -3$ and $q = 1$. We find for the arccosine

$$\arccos\left(\frac{3 \cdot 1}{2 \cdot -3\sqrt{-\frac{-3}{3}}}\right) = \arccos\left(-\frac{1}{2}\right) = \frac{2\pi}{3}.$$

From here we obtain the exact solutions $2\cos(\frac{2\pi}{9})$, $2\cos(\frac{8\pi}{9})$ and $2\cos(\frac{14\pi}{9})$.

Appendix A Zorn's Lemma

Definition A.1: Let S be a set. We call a binary relation $\leq$ on S a *partial order* if it satisfies the following properties for all $a, b, c \in S$:

(i) $a \leq a$ (reflexivity);

(ii) If $a \leq b$ and $b \leq a$, then $a = b$ (antisymmetry);

(iii) If $a \leq b$ and $b \leq c$, then $a \leq c$ (transitivity).

A set that admits a partial order is called *partially ordered.* We usually say that $(S, \leq)$ is a partially ordered set.

Definition A.2: Let S be a set. We call a partial order $\leq$ a *total order* if for every $a, b \in S$ we have $a \leq b$ or $b \leq a$. A set that admits a total order is called *totally ordered.* We then say that $(S, \leq)$ is a totally ordered set.

Remark A.3: We may write $b \geq a$ for $a \leq b$.

Definition A.4: Let $(S, \leq)$ be a partially ordered set. A subset $C \subseteq S$ is called a *chain* if $(C, \leq)$ is totally ordered.

Definition A.5: Let $(S, \leq)$ be a partially ordered set and let $C \subseteq S$ be a chain in S. We call $u \in S$ an *upper bound* of C if $c \leq u$ for every $c \in C$.

Definition A.6: Let $(S, \leq)$ be a partially ordered set. A *maximal element* of S is an element $m \in S$ such that, if $s \in S$ satisfies $s \geq m$, then $s = m$.

We now have the follow result from set theory. We shall not give a proof here, as it is equivalent to one of the axioms of set theory (namely, the Axiom of Choice). The following formulation of Zorn's Lemma is based on the formulation in [14, p. 880].

***Zorn's Lemma*:** Let $(S, \leq)$ be a non-empty, partially ordered set. If every non-empty chain has an upper bound in S, then S contains a maximal element. □

***Corollary* A.7:** Let S be a set and let $\mathcal{S}$ be a non-empty set of subsets of S. Then $(\mathcal{S}, \subseteq)$ is a partially ordered set. If for every non-empty chain $\mathcal{C}$ in $\mathcal{S}$ the union $\bigcup \mathcal{C} = \bigcup_{C \in \mathcal{C}} C$ is an element of $\mathcal{S}$, then $\mathcal{S}$ contains a maximal element.

Proof. The properties of a partial order are immediate for $\subseteq$. Let $\mathcal{C}$ be a non-empty chain in $\mathcal{S}$. Let $C \in \mathcal{C}$. For each $c \in C$, obviously $c \in \bigcup \mathcal{C}$. Hence $C \subseteq \bigcup \mathcal{C}$ and so $\bigcup \mathcal{C}$ is an upper bound of $\mathcal{C}$. Since, by assumption, $\bigcup \mathcal{C} \in \mathcal{S}$, we see that every non-empty chain has an upper bound in $\mathcal{S}$. Whence, by Zorn's lemma, $\mathcal{S}$ contains a maximal element. □

References

[1] A. Blass, *Why is the collection of all algebraic extensions of F not a set?* Mathematics Stack Exchange `http://math.stackexchange.com/q/1259874`, 2015.

[2] P. L. Clark, *Field Theory.* `http://math.uga.edu/~pete/FieldTheory.pdf`.

[3] P. M. Cohn, *Algebra*, vol. 3, John Wiley & Sons, 2 ed., 1991.

[4] K. Conrad, *Galois Groups of Cubics and Quartics in All Characteristics.* `http://www.math.uconn.edu/~kconrad/blurbs/galoistheory/cubicquarticchar2.pdf`.

[5] D. A. Cox, *Galois Theory*, Pure and Applied mathematics: a Wiley-Interscience series of texts, monographs, and tracts, Wiley-Interscience, 2004.

[6] G. Ehrlich, *Fundamental Concepts of Abstract Algebra*, Dover Publications, 2011.

[7] R. Friedman, *More Notes on Galois Theory.* `http://www.math.columbia.edu/~rf/moregaloisnotes.pdf`, 2013.

[8] H. Gross and P. Hafner, *Über die Eindeutigkeit des reellen Abschlusses eines angeordneten Körpers*, Commentarii Mathematici Helvetici, 44 (1969), pp. 491–494.

[9] O. Hölder, *Über den Casus Irreducibilis bei der Gleichung dritten Grades*, Mathematische Annalen, 38 (1891), pp. 307–312.

[10] I. M. Isaacs, *Solution of Polynomials by Real Radicals*, The American Mathematical Monthly, 92 (1985), pp. 571–575.

[11] N. Jacobson, *Basic Algebra I*, W. H. Freeman and Company, 2 ed., 1985.

[12] ———, *Basic Algebra II*, W. H. Freeman and Company, 2 ed., 1989.

[13] M. Knebusch, *On the Uniqueness of Real Closures and the Existence of Real Places*, Commentarii Mathematici Helvetici, 47 (1972), pp. 260–269.

[14] S. Lang, *Algebra*, vol. 211 of Graduate Texts in Mathematics, Springer-Verslag, 3 ed., 2002.

[15] B. L. van der Waerden, *Moderne Algebra I*, Springer-Verslag, 1930.

[16] J. Villanueva, *On the Casus Irreducibilis of Solving the Cubic Equation.* `http://archives.math.utk.edu/ICTCM/VOL24/C036/paper.pdf`.

[17] F. Voloch and P. Monsky, *Solving the cubic by "radicals" in characteristics 2 and 3.* MathOverflow `http://mathoverflow.net/q/81997`, 2011.