Contents

2 Fields and Field Extensions

Our goal in this chapter is to study the structure of fields, a subclass of rings in which every nonzero element has a multiplicative inverse, and field extensions. Fields arise naturally in studying the solutions to polynomial equations, and we will explore the connections between polynomials and fields in detail. As a particular application of the basic theory of field extensions, we will be able to establish the impossibility of certain classical straightedge-and-compass constructions such as trisecting an arbitrary angle and doubling the cube. We will also discuss at length the structure of fields obtained by "adjoining" roots of polynomials, and in particular the (historically perilous) topic of establishing that every field has an algebraic closure. We finish with a lengthy discussion of some more technical matters regarding separability and transcendence.

2.1 Fields and Vector Spaces

- In the previous chapter, we have already essentially defined fields as a special type of ring. Our first goal is to develop some basic properties of fields, and then to discuss vector spaces over fields.

2.1.1 Definition, Examples, and Basic Properties of Fields

- For simplicity, we will again list the axioms for a field:
- Definition: A field is any set F having two (closed) binary operations $+$ and $\cdot$ that satisfy the nine axioms [F1]-[F9]:

  [F1] The operation $+$ is associative: $a + (b + c) = (a + b) + c$ for any elements a, b, c in F.

  [F2] The operation $+$ is commutative: $a + b = b + a$ for any elements a, b in F.

  [F3] There is an additive identity 0 satisfying $a + 0 = a$ for all a in F.

  [F4] Every element a in F has an additive inverse $-a$ satisfying $a + (-a) = 0$.

  [F5] The operation $\cdot$ is associative: $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ for any elements a, b, c in F.

  [F6] The operation $\cdot$ is commutative: $a \cdot b = b \cdot a$ for any elements a, b in F.

  [F7] There is a multiplicative identity $1 \neq 0$, satisfying $1 \cdot a = a = a \cdot 1$ for all a in F.

  [F8] Every nonzero a in F has a multiplicative inverse a^{-1} satisfying $a \cdot a^{-1} = 1$.

  [F9] The operation $\cdot$ distributes over $+$: $a \cdot (b + c) = a \cdot b + a \cdot c$ for any elements a, b, c in F.

- We have previously mentioned four examples of fields: the rational numbers $\mathbb{Q}$, the real numbers $\mathbb{R}$, the complex numbers $\mathbb{C}$, and the finite fields $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ where p is a prime number. From our discussion of polynomials, we have also established that if F is any field and p is any irreducible polynomial in $F[x]$, then the ring $F[x]/p$ of residue classes modulo p is also a field.
- Here are a few additional examples of fields:
- Example: If F is a field, the collection of rational functions in t with coefficients in F, denoted $F(t)$, forms a field.
  - Remark: We use the letter t to denote the indeterminate rather than x, since we will later want to discuss polynomials in the context of this field of rational functions.
  - Explicitly, the elements of this field are quotients of polynomials $\frac{p}{q}$ where $p, q \in F[t]$ and $q \neq 0$, and where $\frac{p}{q} = \frac{r}{s}$ whenever $ps = rq$.
  - The addition and multiplication operations are defined in the same way as for regular fractions: $\frac{p}{q} + \frac{r}{s} = \frac{ps + qr}{qs}$ and $\frac{p}{q} \cdot \frac{r}{s} = \frac{pr}{qs}$. It is tedious (but straightforward) to verify that these operations are well-defined and satisfy the field axioms.
- Example: The set $S = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$ forms a field, denoted $\mathbb{Q}(\sqrt{2})$ (typically read as "$\mathbb{Q}$ adjoin $\sqrt{2}$").
  - The arithmetic in $\mathbb{Q}(\sqrt{2})$ is as follows: $(a+b\sqrt{2})+(c+d\sqrt{2}) = (a+c)+(b+d)\sqrt{2}$, and $(a+b\sqrt{2})(c+d\sqrt{2}) = (ac + 2bd) + (ad + bc)\sqrt{2}$.
  - Since $\mathbb{Q}(\sqrt{2})$ is clearly closed under subtraction and multiplication, and contains $0 = 0 + 0\sqrt{2}$, it is a subring of $\mathbb{C}$ and hence an integral domain, since it contains 1.
  - To see that $\mathbb{Q}(\sqrt{2})$ is actually a field, we need to show that every element has a multiplicative inverse: this follows by "rationalizing the denominator", since we can write $(a + b\sqrt{2})^{-1} = \frac{a - b\sqrt{2}}{a^2 - 2b^2}$, and as long as one of a, b is nonzero the denominator is also nonzero because $\sqrt{2}$ is irrational.
- Example: The set $S = \{a + bi : a, b \in \mathbb{Q}\}$ forms a field, denoted $\mathbb{Q}(i)$. (As usual, i denotes the imaginary unit with $i^2 = -1$.)
  - The arithmetic in $\mathbb{Q}(i)$ is the same as for regular complex numbers: $(a+bi)+(c+di) = (a+c)+(b+d)i$, and $(a + bi)(c + di) = (ac - bd) + (ad + bc)i$.

- Like with $\mathbb{Q}(\sqrt{2})$ we can see that every nonzero element has a multiplicative inverse, since $(a+bi)^{-1} = \dfrac{a-bi}{a^2+b^2}$, so $\mathbb{Q}(i)$ is a field.

- Both $\mathbb{Q}(\sqrt{2})$ and $\mathbb{Q}(i)$ are special cases of the more general class of quadratic fields, obtained by "adjoining" the square root of something to $\mathbb{Q}$:

- Example: Let D be a squarefree integer not equal to 1. The quadratic field $\mathbb{Q}(\sqrt{D})$ is the set of complex numbers of the form $a+b\sqrt{D}$, where a and b are rational numbers.

  - Remark: An integer is squarefree if it is not divisible by the square of any prime. We lose nothing here by assuming that D is a squarefree integer, since two different integers differing by a square factor would generate the same set of complex numbers $a+b\sqrt{D}$.

  - As in the two cases above, $\mathbb{Q}(\sqrt{D})$ is a field because we can write $(a+b\sqrt{D})^{-1} = \dfrac{a-b\sqrt{D}}{a^2-Db^2}$, and $a^2 - Db^2 \neq 0$ provided that a and b are not both zero because $\sqrt{D}$ is irrational (by the assumption that D is squarefree and not equal to 1).
  - An important quantity related to these quadratic fields is the quadratic field norm $N : \mathbb{Q}(\sqrt{D}) \to \mathbb{Q}$, defined as $N(a+b\sqrt{D}) = a^2 - Db^2 = (a+b\sqrt{D})(a-b\sqrt{D})$.
  - The fundamental property of this field norm is that it is multiplicative: $N(xy) = N(x)N(y)$ for two elements x and y in $\mathbb{Q}(\sqrt{D})$, as can be verified by writing out both sides explicitly and comparing the results.
  - The field norm provides a measure of "size" of an element of $\mathbb{Q}(\sqrt{D})$, in much the same way that the complex absolute value measures the "size" of a complex number. In fact, if $D < 0$, then the field norm of an element $a+b\sqrt{D}$ is the same as the square of its complex absolute value.

- Since fields are commutative integral domains (with 1), we have a number of basic properties of field arithmetic that follow immediately from the axioms:

- Proposition (Basic Arithmetic): The following properties hold in any field F:

  1. The additive identity 0 and the multiplicative identity 1 are unique, as are additive and multiplicative inverses.
  2. Addition has a cancellation law: for any $a, b, c \in F$, if $a+b = a+c$, then $b = c$.
  3. Multiplication has a cancellation law: for any $a, b, c \in F$ with $a \neq 0$, if $ab = ac$ then $b = c$. In particular, $ab = 0$ implies $a = 0$ or $b = 0$.
  4. For any $a \in F$, $0 \cdot a = 0 = a \cdot 0$ and $(-1) \cdot a = -a$.
  5. For any $a, b \in F$, $-(a+b) = (-a) + (-b)$, $(-a) \cdot b = -(a \cdot b) = a \cdot (-b)$, and $(-a) \cdot (-b) = a \cdot b$.
  6. For any positive integers m and n and any $a \in F$, $ma + na = (m+n)a$, $m(na) = (mn)a$, $a^{m+n} = a^m a^n$, and $a^{mn} = (a^m)^n$.

- There is another fundamental quantity attached to a field known as its characteristic:

- Definition: If F is a field, we say F has characteristic p if $p1_F = 0$, and no smaller positive integer multiple of 1 is 0. (Recall that $p1_F = \underbrace{1_F + 1_F + \cdots + 1_F}_{p \text{ times}}$.) If $n1_F \neq 0$ for all $n > 0$, then we say F has characteristic 0.

  - Example: For a prime p, the finite field $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ has characteristic p, while the fields $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ have characteristic 0.
  - Any finite field necessarily has positive characteristic, although infinite fields with positive characteristic also exist.

- Proposition (Positive Characteristic): If the field F has characteristic $p > 0$, then p is a prime.

  - Proof: Suppose F has characteristic $m > 0$ and $m = ab$ for positive integers a, b: then $0 = m1_F = (a1_F) \cdot (b1_F)$.
  - Since F is a field, this implies that one of $a1_F$ and $b1_F$ must be zero, but since m is minimal, the only possibility is that $a = m$ or $b = m$, meaning that m must be prime.

2.1.2 Vector Spaces

- Vector spaces are a central ingredient for studying fields, so we will briefly outline some of the basic properties of vector spaces over an arbitrary field.
- Definition: Let F be a field, and refer to the elements of F as scalars. A vector space over F is a triple $(V, +, \cdot)$ of a collection V of vectors, together with two binary operations, addition of vectors $(+)$ and scalar multiplication of a vector by a scalar $(\cdot)$, satisfying the following axioms:

  [V1] Addition is commutative: $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$ for any vectors $\mathbf{v}$ and $\mathbf{w}$.

  [V2] Addition is associative: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ for any vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$.

  [V3] There exists a zero vector $\mathbf{0}$, with $\mathbf{v} + \mathbf{0} = \mathbf{v} = \mathbf{0} + \mathbf{v}$ for any vector $\mathbf{v}$.

  [V4] Every vector $\mathbf{v}$ has an additive inverse $-\mathbf{v}$, with $\mathbf{v} + (-\mathbf{v}) = \mathbf{0} = (-\mathbf{v}) + \mathbf{v}$.

  [V5] Multiplications are consistent: $\alpha \cdot (\beta \cdot \mathbf{v}) = (\alpha\beta) \cdot \mathbf{v}$ for any scalars α, β and vector $\mathbf{v}$.

  [V6] Addition of scalars distributes: $(\alpha + \beta) \cdot \mathbf{v} = \alpha \cdot \mathbf{v} + \beta \cdot \mathbf{v}$ for any scalars α, β and vector $\mathbf{v}$.

  [V7] Addition of vectors distributes: $\alpha \cdot (\mathbf{v} + \mathbf{w}) = \alpha \cdot \mathbf{v} + \alpha \cdot \mathbf{w}$ for any scalar α and vectors $\mathbf{v}$ and $\mathbf{w}$.

  [V8] The scalar 1 acts like the identity on vectors: $1 \cdot \mathbf{v} = \mathbf{v}$ for any vector $\mathbf{v}$.

  - Remark: We will often omit the $\cdot$ for scalar multiplication, and will frequently assume that the field F is clear from the context.

- Here are a few standard examples of vector spaces:
  - Example: For any positive integer n, the set of all n-tuples of elements from F, denoted F^n, is an F-vector space under componentwise addition and scalar multiplication.
    - Explicitly, the operations in F^n are $\langle a_1, a_2, \ldots, a_n\rangle + \langle b_1, b_2, \ldots, b_n\rangle = \langle a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n\rangle$ and $\alpha \cdot \langle b_1, b_2, \ldots, b_n\rangle = \langle \alpha b_1, \alpha b_2, \ldots, \alpha b_n\rangle$.
    - The additive identity is the zero vector $\langle 0, 0, \ldots, 0\rangle$ and additive inverses are given by negating each component: $-\langle b_1, b_2, \ldots b_n\rangle = \langle -b_1, -b_2, \ldots, -b_n\rangle$.
  - Example: The zero space with a single element $\mathbf{0}$, with $\mathbf{0} + \mathbf{0} = \mathbf{0}$ and $\alpha \cdot \mathbf{0} = \mathbf{0}$ for every $\alpha \in F$, is an F-vector space.
  - Example: The rings $F[x]$ and $F[x]/p$ for any polynomial p are F-vector spaces.
  - Example: The complex numbers $\mathbb{C}$ are a vector space over $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ (in each case via normal addition and multiplication in $\mathbb{C}$).
- Like with rings and fields, vector spaces have some basic arithmetic properties that can be derived immediately from the axioms:
- Proposition (Basic Arithmetic): In any vector space V, the following are true:
  1. The additive identity 0 is unique, as are additive inverses.
  2. Addition has a cancellation law: for any $\mathbf{a}, \mathbf{b}, \mathbf{c} \in V$, if $\mathbf{a} + \mathbf{b} = \mathbf{a} + \mathbf{c}$, then $\mathbf{b} = \mathbf{c}$.
  3. For any $\mathbf{v} \in V$, $0 \cdot \mathbf{v} = \mathbf{0}$, and for any $\alpha \in F$, $\alpha \cdot \mathbf{0} = \mathbf{0}$.
  4. For any $\mathbf{v} \in V$, $(-1) \cdot \mathbf{v} = -\mathbf{v}$, and $-(-\mathbf{v}) = \mathbf{v}$.
- Now we can discuss the structure of vector spaces:
- Definition: A subspace W of a vector space V is a subset of the vector space V which, under the same addition and scalar multiplication operations as V, is itself a vector space.
  - Example: Any vector space V has two obvious subspaces: the zero space and V itself.
  - Example: As a $\mathbb{Q}$-vector space, $\mathbb{R}$ is a subspace of $\mathbb{C}$.
- Definition: Given a set $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ of vectors in a vector space V, we say a vector $\mathbf{w}$ in V is a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ if there exist scalars $a_1, \cdots, a_n$ such that $\mathbf{w} = a_1 \cdot \mathbf{v}_1 + a_2 \cdot \mathbf{v}_2 + \cdots + a_n \cdot \mathbf{v}_n$.

- Example: In $\mathbb{Q}^4$, the vector $\langle 4,0,5,9\rangle$ is a linear combination of $\langle 1,0,0,1\rangle$, $\langle 0,1,0,0\rangle$, and $\langle 1,1,1,2\rangle$, because $\langle 4,0,5,9\rangle = 1\cdot\langle 1,-1,2,3\rangle - 2\cdot\langle 0,1,0,0\rangle + 3\cdot\langle 1,1,1,2\rangle$.
- Example: In $\mathbb{F}_3^2$, the vector $\langle 1,0,2\rangle$ is a linear combination of $\langle 1,1,1\rangle$ and $\langle 2,1,0\rangle$, because $\langle 1,2\rangle = 2\cdot\langle 1,1,1\rangle + 1\cdot\langle 2,1,0\rangle$.
- Non-Example: In $\mathbb{R}^3$, the vector $\langle 0,0,1\rangle$ is not a linear combination of $\langle 1,1,0\rangle$ and $\langle 0,1,1\rangle$ because there exist no scalars a_1 and a_2 for which $a_1\cdot\langle 1,1,0\rangle + a_2\cdot\langle 0,1,1\rangle = \langle 0,0,1\rangle$: this would require a common solution to the three equations $a_1 = 0$, $a_1 + a_2 = 0$, and $a_2 = 1$, and this system has no solution.

- Definition: If V is a vector space and S is a subset, the span of S is defined to be $\text{span}(S) = \{a_1\cdot\mathbf{v}_1+\cdots+a_n\cdot\mathbf{v}_n : a_i \in F, \mathbf{v}_i \in S\}$, the set of all linear combinations of finitely many vectors in S. (Note that $\text{span}(\emptyset) = \{\mathbf{0}\}$.)
  - It is not hard to show that $\text{span}(S)$ is the smallest subspace of V containing S.
  - Example: The span of the set $\{1, x\}$ inside $F[x]$ is the set of linear polynomials (i.e., of the form $a + bx$ for $a, b \in F$).

- Definition: If $\text{span}(S) = V$, we say that S is a spanning set for V: in other words, when every vector in V can be written as a linear combination of the vectors in S.
  - Example: The set $\{1, i\}$ is a spanning set for $\mathbb{C}$ as a vector space over $\mathbb{R}$.
  - Example: The set $\{\langle 1,0,0\rangle, \langle 0,1,0\rangle, \langle 0,0,1\rangle\}$ is a spanning set for F^3 for any field F.
  - Example: The set $\{\langle 1,1\rangle, \langle 2,1\rangle, \langle 3,1\rangle\}$ is a spanning set for $\mathbb{Q}^2$.
  - In the first two examples, it is not hard to see that every vector in V can be written uniquely as a linear combination of the elements of the spanning set. In the third example, however, the linear combinations are not unique (since for example $\langle 4,2\rangle = 2\cdot\langle 2,1\rangle = \langle 1,1\rangle + \langle 3,1\rangle$).

- Definition: If V is a vector space, a subset S of V is linearly independent if, for any distinct vectors $\mathbf{v}_i \in S$ and any scalars $a_i \in F$, $a_1\cdot\mathbf{v}_1 + \cdots + a_n\cdot\mathbf{v}_n = \mathbf{0}$ implies $a_1 = \cdots = a_n = 0$. Otherwise, S is linearly dependent.
  - For a finite set $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, S is linearly independent precisely when the only way to form the zero vector as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is when all the scalar coefficients are zero (the "trivial" linear combination). An infinite set is linearly independent when all its finite subsets are linearly independent.
  - Example: The vectors $\langle 1,1,0\rangle$ and $\langle 0,2,1\rangle$ in $\mathbb{R}^3$ are linearly independent, because $a\cdot\langle 1,1,0\rangle + b\cdot\langle 0,2,1\rangle = \langle 0,0,0\rangle$ implies $a = 0$, $a + 2b = 0$, and $b = 0$, so that $a = b = 0$.
  - Example: The complex numbers $3 - 5i$, $3 - 4i$, and $1 - i$ are $\mathbb{Q}$-linearly dependent because $1(3 - 5i) - 2(3 - 4i) + 3(1 - i) = 0$.
  - Example: The empty set is always linearly independent, in any vector space.
  - The terminology of "linear dependence" arises from the fact that if a set of vectors is linearly dependent, one of the vectors is necessarily a linear combination of the others (i.e., it "depends" on the others).

- If a set of vectors is linearly independent, every vector in their span can be uniquely written as a linear combination:

- Proposition (Characterization of Linear Independence): A set S of vectors is linearly independent if and only if every vector $\mathbf{w}$ in $\text{span}(S)$ may be *uniquely* written as a sum $\mathbf{w} = a_1\cdot\mathbf{v}_1 + \cdots + a_n\cdot\mathbf{v}_n$ for unique scalars $a_1, a_2, \ldots, a_n$ and unique vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ in S (where we view sums as equivalent if additional terms with coefficient 0 are added or removed).
  - Proof: First suppose the decomposition is always unique: then for any $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ in S, $a_1\cdot\mathbf{v}_1 + \cdots + a_n\cdot\mathbf{v}_n = \mathbf{0}$ implies $a_1 = \cdots = a_n = 0$, because $0\cdot\mathbf{v}_1 + \cdots + 0\cdot\mathbf{v}_n = \mathbf{0}$ is by assumption the only decomposition of $\mathbf{0}$.
  - Now suppose that $\mathbf{w} = a_1\cdot\mathbf{v}_1 + \cdots + a_n\cdot\mathbf{v}_n = b_1\cdot\mathbf{v}_1 + \cdots + b_n\cdot\mathbf{v}_n$. Subtracting yields $(a_1 - b_1)\cdot\mathbf{v}_1 + \cdots + (a_n - b_n)\cdot\mathbf{v}_n = \mathbf{w} - \mathbf{w} = \mathbf{0}$, and since $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent, $a_1 - b_1, \cdots, a_n - b_n$ are all zero.

- Definition: A linearly independent set of vectors that spans V is called a basis for V. (The plural of "basis" is "bases".)

- From our characterization of linear independence above, we can see that S is a basis for V if and only if every vector in V can be written uniquely as a linear combination of vectors in S.
- Example: The "standard basis" for F^n consists of the unit coordinate vectors $\langle 1, 0, \ldots, 0, 0\rangle$, $\langle 0, 1, \ldots, 0, 0\rangle$, ... , $\langle 0, 0, \ldots, 0, 1\rangle$.
- Example: The set $\{1, i\}$ is a basis for $\mathbb{C}$ over $\mathbb{R}$, as is the set $\{1 + i, 2 - 3i\}$.
- Example: If p has degree n, then the set $\{1, x, x^2, \ldots, x^{n-1}\}$ is a basis for $F[x]/p$.
- Non-Example: The vectors $\langle 1, 1, 0\rangle$ and $\langle 1, 1, 1\rangle$ are not a basis for $\mathbb{Q}^3$ since they do not span $\mathbb{Q}^3$.
- Non-Example: The vectors $\langle 1, 0, 0\rangle$, $\langle 0, 1, 0\rangle$, $\langle 0, 0, 1\rangle$, $\langle 1, 1, 1\rangle$ are not a basis for $\mathbb{Q}^3$ since they are not linearly independent.

- A basis for a vector space can always be obtained by removing vectors from a spanning set, or by adding vectors to a linearly independent set:
- Theorem (Spanning Sets): If V is a vector space, then any spanning set for V contains a basis of V.
  - If the spanning set is finite, then the idea is to throw away linearly dependent vectors one at a time until the resulting set is linearly independent. The collection of elements which we have not thrown away will always be a spanning set (since removing a dependent element will not change the span). By an easy induction argument, this process will eventually terminate, and the end result will be a linearly independent spanning set.
  - In the event that the spanning set is infinite, the argument relies on a result known as Zorn's lemma (equivalent to the axiom of choice).
  - Zorn's lemma says that if S is a nonempty partially-ordered set such that every chain has an upper bound in S, then S contains a maximal element[1].
  - Here is the Zorn's lemma argument: Let $\mathcal{F}$ be the collection of all linearly-independent subsets of V, partially ordered by inclusion, and note that $\mathcal{F}$ is not empty since it contains the empty set. If $\mathcal{C}$ is any chain in $\mathcal{F}$, then the union of all the elements of $\mathcal{C}$ is an upper bound for $\mathcal{C}$ and is linearly independent: any linear dependence in the union would imply a linear dependence in one of the elements in the chain (linear dependences involve only finitely many vectors, so we may take the maximum of the subsets in which all vectors appear). Thus, by Zorn's lemma, $\mathcal{F}$ contains a maximal element. Finally, we observe that a maximal linearly-independent subset is in fact a basis (otherwise, we could adjoin an element not in the span, contradicting maximality).
- Theorem (Building-Up Theorem): Given any linearly independent set of vectors in V, there exists a basis of V containing those vectors. In short, any linearly independent set of vectors can be extended to a basis.
  - The idea (roughly speaking) is to start with the given linearly independent set, and then append linearly independent vectors to S one at a time until a basis for V is obtained.
  - If V is "finite-dimensional" (see below), then this procedure will always terminate in a finite number of steps. In the case where V is "infinite-dimensional", the argument again relies on Zorn's lemma.
- Theorem (Bases of Vector Spaces): Every vector space has a basis, and any two bases have the same number of elements.
  - The existence of bases follows from either of the theorems given above[2]. To show that any two bases have the same number of elements is more difficult, and can be done by first proving the following "replacement theorem":
  - Theorem (Replacement Theorem): Suppose that $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for V and $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ is a linearly independent subset of V. Then there is a reordering of the basis S, say $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n\}$ such that for each $1 \leq k \leq m$, the set $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k, \mathbf{a}_{k+1}, \mathbf{a}_{k+2}, \ldots, \mathbf{a}_n\}$ is a basis for V. Equivalently, the elements $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ can be used to successively replace the elements of the basis, with each replacement remaining a basis of V.

[1] A chain is a totally-ordered subset of S, an upper bound for a subset is an element greater than or equal to all elements of the subset, and a maximal element is an element such that no element is strictly greater than it.

[2] It has been proven that the statement "every vector space has a basis" is actually equivalent to the axiom of choice (under the Zermelo-Frankel axioms of set theory), so in fact appealing to the axiom of choice, or equivalently Zorn's lemma, is necessary to establish this theorem.

- By applying the replacement theorem appropriately, one can deduce the following useful result:
- Corollary: Suppose V has a basis with n elements. If $m > n$, then any set of m vectors of V is linearly dependent. In particular, any two bases must have the same number of elements.

- Definition: If V is an F-vector space, the number of elements in any basis of V is called the dimension of V and is denoted $\dim_F(V)$. If $\dim_F(V)$ is finite, V is finite-dimensional; otherwise, V is infinite-dimensional[3].
  - Example: $\dim_F(F^n) = n$, since the standard unit vectors form a basis.
  - Example: $\dim_F(F[x]) = \infty$ since the set $\{1, x, x^2, \dots\}$ is a basis.
  - Example: $\dim_F(F[x]/p) = \deg(p)$ since the set $\{1, x, x^2, \dots, x^{\deg(p)-1}\}$ is a basis.
  - Example: The dimension of the zero space is 0, because the empty set (containing 0 elements) is a basis.
  - Example: $\dim_{\mathbb{R}}(\mathbb{C}) = 2$ since the set $\{1, i\}$ is a basis.
  - Example: $\dim_{\mathbb{C}}(\mathbb{C}) = 2$ since the set $\{1\}$ is a basis.
  - Example: $\dim_{\mathbb{Q}}(\mathbb{C}) = \infty$ since any finite-dimensional vector space over $\mathbb{Q}$ necessarily has only countably many elements, and $\mathbb{C}$ is uncountable. Alternatively, $\mathbb{C}$ contains a transcendental number π, so the set $\{1, \pi, \pi^2, \pi^3, \dots\}$ is $\mathbb{Q}$-linearly independent since otherwise π would be a root of a polynomial with rational coefficients.
  - As the last three examples indicate, the dimension of a vector space depends intrinsically on its associated field of scalars.

- We can also study the structure-preserving maps on vector spaces, which are the vector-space equivalent of homomorphisms:

- Definition: If V and W are vector spaces having the same scalar field F, we say a function $T : V \to W$ is a linear transformation if it respects addition of vectors and scalar multiplication: in other words, if $T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2)$ and $T(\alpha\mathbf{v}) = \alpha T(\mathbf{v})$ for any vectors $\mathbf{v}, \mathbf{v}_1, \mathbf{v}_2 \in V$ and any scalar $\alpha \in F$. If T is a linear transformation that is also a bijection, then T is a (vector space) isomorphism.
  - Example: If A is any $m \times n$ matrix, the map $T : F^m \to F^n$ given by $T(\mathbf{v}) = A\mathbf{v}$ is a linear transformation; indeed, these are all the linear transformations from F^m to F^n.
  - Example: If V is the vector space of differentiable functions and W is the vector space of real-valued functions, the derivative map D sending a function to its derivative is a linear transformation from V to W.
  - Example: If V is the vector space of all continuous functions on $[a, b]$, then the integral map $I(f) = \int_a^b f(x)\,dx$ is a linear transformation from V to $\mathbb{R}$.
  - Example: The transpose map is a linear transformation from $M_{m\times n}(F)$ to $M_{n\times m}(F)$ for any field F and any positive integers m, n: in fact, it is an isomorphism.
  - Example: For any $a \in F$, the evaluation at a map on $F[x]$, defined by $T(p) = p(a)$, is a linear transformation from $F[x]$ to F.
  - Example: If V and W are any vector spaces, the zero map sending all elements of V to the zero vector in W is a linear transformation from V to W.
  - Example: If V is any vector space, the identity map sending all elements of V to themselves is a linear transformation from V to V. The identity map is an isomorphism of V with itself.

- Also like with rings, we have the natural notion of kernel and image for linear transformations:

- Definition: If $T : V \to W$ is a linear transformation, then the kernel of T, denoted $\ker(T)$, is the set of elements $\mathbf{v} \in V$ with $T(\mathbf{v}) = \mathbf{0}$, and the image of T, denoted $\mathrm{im}(T)$, is the set of elements $\mathbf{w} \in W$ such that there exists $\mathbf{v} \in V$ with $T(\mathbf{v}) = \mathbf{w}$.
  - It is easy to verify from the definitions that the kernel and image are subspaces of V and W, respectively.

[3] In general, we will not need to concern ourselves with the cardinality of the basis for an infinite-dimensional vector space, and merely refer to all of these infinite cardinalities as ∞.

- Like with ring homomorphisms, it is not hard to show that $\ker(T) = \{\mathbf{0}\}$ if and only if T is one-to-one. Thus, we see that T is an isomorphism if and only if $\ker(T) = \{\mathbf{0}\}$ and $\text{im}(T) = W$.

- Like with ring homomorphisms, linear transformations have various basic properties:

- Proposition (Properties of Linear Transformations): If $T : V \to W$ is linear, then the following hold:

  1. $T(\mathbf{0}_V) = \mathbf{0}_W$.
  2. For any $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ and $a_1, \ldots, a_n \in F$, $T(a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) = a_1T(\mathbf{v}_1) + \cdots + a_nT(\mathbf{v}_n)$.
  3. $T : V \to W$ is linear if and only if for any $\mathbf{v}_1$ and $\mathbf{v}_2$ in V and any scalar α, $T(\mathbf{v}_1 + \alpha\mathbf{v}_2) = T(\mathbf{v}_1) + \alpha T(\mathbf{v}_2)$.
     - Proofs: Straightforward from the definition.
  4. T is characterized by its values on a basis of V: for any basis $B = \{\mathbf{v}_i\}$ of V and any vectors $\{\mathbf{w}_i\} \in W$, there exists a unique linear transformation $T : V \to W$ such that $T(\mathbf{v}_i) = \mathbf{w}_i$ for each i.
     - Proof: The fact that the values of T are determined by its values on the basis follows from property (2) above, since any any vector $\mathbf{v}$ in V can be written as $\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_n\mathbf{v}_n$ for (unique) vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in B and scalars $a_1, \ldots, a_n$.
     - Conversely, suppose that we are given the values $T(\mathbf{v}_i) = \mathbf{w}_i$ for each $\mathbf{v}_i \in B$. Then it is not hard to see that the map $T : V \to W$ defined by setting $T(a_1\mathbf{v}_{i_1} + a_2\mathbf{v}_{i_2} + \cdots + a_n\mathbf{v}_{i_n}) = a_1\mathbf{w}_{i_1} + \cdots + a_n\mathbf{w}_{i_n}$ is a well-defined linear transformation from V to W, and by the remark above, it must be unique.
  5. If T is an isomorphism, then T preserves linear independence and span (i.e., if S is a linearly independent set then so is $T(S)$, and likewise for a spanning set).
     - Proof: Straightforward from the definition.
  6. Two vector spaces V and W are isomorphic if and only if they have the same dimension. In particular, any finite-dimensional vector space V with scalar field F is isomorphic to F^n, where $n = \dim_F V$.
     - Proof: By (5), isomorphisms preserve linear independence, so two vector spaces can only be isomorphic if they have the same dimension.
     - For the other direction, choose a basis $\{\mathbf{v}_i\}_{i \in I}$ for V and a basis $\{\mathbf{w}_i\}_{i \in I}$ for W. Then by (4), there exists a unique linear transformation $T : V \to W$ with $T(\mathbf{v}_i) = \mathbf{w}_i$ for each $i \in I$. It is then a straightforward check that T is an isomorphism.

- There is a well-defined notion of a quotient vector space, but we will not bother to develop this notion. However, we can still give the analogue of the first isomorphism theorem, which is extremely important:

- Theorem (Nullity-Rank): For any linear transformation $T : V \to W$, $\dim(\ker(T)) + \dim(\text{im}(T)) = \dim(V)$.

  - The dimension of the kernel is called the nullity, while the dimension of the image is called the rank (whence the name "nullity-rank theorem").
  - Proof: Let $\beta = \{\mathbf{w}_i\}_{i \in I}$ be a basis for $\text{im}(T)$ in W. By definition, there exist $\{\mathbf{v}_i\}_{i \in I}$ in V such that $T(\mathbf{v}_i) = \mathbf{w}_i$ for each $i \in I$
  - Also, let $\alpha = \{\mathbf{a}_j\}_{j \in J}$ be a basis for $\ker(T)$. We claim that the set of vectors $S = \{\mathbf{v}_i\}_{i \in I} \cup \{\mathbf{a}_j\}_{j \in J}$ is a basis for V.
  - To see that S spans V, let $\mathbf{v}$ be an element of V. Since $T(\mathbf{v}) \in \text{im}(T)$, there exist scalars $b_1, \ldots, b_k$ and $\mathbf{v}_1, \ldots, \mathbf{v}_k$ such that $T(\mathbf{v}) = \sum_{j=1}^{k} b_j\mathbf{w}_j$.
  - Then $T\left[\mathbf{v} - \sum_{j=1}^{k} b_j\mathbf{v}_j\right] = T(\mathbf{v}) - \sum_{j=1}^{k} b_jT(\mathbf{v}_j) = \sum_{j=1}^{k} b_j\mathbf{w}_j - \sum_{j=1}^{k} b_j\mathbf{w}_j = \mathbf{0}$.
  - This means $\mathbf{v} - \sum_{j=1}^{k} b_j\mathbf{v}_j$ is in $\ker(T)$, so it can be written as a sum $\sum_{i=1}^{l} c_i\mathbf{a}_i$ for some scalars c_i and some $\mathbf{a}_1, \ldots, \mathbf{a}_l \in \alpha$: then $\mathbf{v} = \sum_{j=1}^{k} b_j\mathbf{v}_j + \sum_{i=1}^{l} c_i\mathbf{a}_i \in \text{span}(S)$, so S spans V.

- To see that S is linearly independent, if we had a dependence $\mathbf{0} = \sum_{j=1}^{k} b_j \mathbf{v}_j + \sum_{i=1}^{l} c_i \mathbf{a}_i$, applying T to both sides would yield $\mathbf{0} = T(\mathbf{0}) = \sum_{j=1}^{k} b_j T(\mathbf{v}_j) + \sum_{i=1}^{l} c_i T(\mathbf{a}_i) = \sum_{j=1}^{k} b_j \mathbf{w}_j$.
- Since the $\mathbf{w}_j$ are linearly independent, all the coefficients b_j must be zero. Then $\mathbf{0} = \sum_{i=1}^{l} c_i \mathbf{a}_i$, but now since the $\mathbf{a}_i$ are linearly independent, all the coefficients c_i must also be zero.

2.2 Subfields and Field Extensions

- As with other algebraic structures like vector spaces and rings, a natural first step in studying the structure of fields is to study subfields.
- Definition: If F is a field, we say a subset S of F is a subfield if S is itself a field under the same operations as F. If F is a subfield of the field K, we say that K is an extension field of F.
  - Notation: We often write K/F ("K over F") to symbolize that K is an extension field of F. (It is not the quotient of K by F!)
  - Example: $\mathbb{Q}$ is a subfield of $\mathbb{R}$, which is a subfield of $\mathbb{C}$.
  - Example: For any squarefree integer $D \neq 1$, $\mathbb{Q}(\sqrt{D})$ is a subfield of $\mathbb{C}$.
  - Example: $\mathbb{F}_2$ is a subfield of $\mathbb{F}_2[x]/(x^2 + x + 1)$ since the latter is also a field.
- We can also exploit the structure of vector spaces to study the structure of fields. A fundamental observation is that if K is an extension field of F, then K is an F-vector space (under the addition and multiplication of K).
- Definition: If K is an extension field of F, the degree $[K : F]$ (also called the relative degree or occasionally the index) is the dimension $\dim_F(K)$ of K as an F-vector space. The extension K/F is finite if it has finite degree; otherwise, the extension is infinite.
  - Example: We have $[\mathbb{C} : \mathbb{R}] = 2$, $[\mathbb{Q}(\sqrt{D}) : \mathbb{Q}] = 2$, and $[\mathbb{R} : \mathbb{Q}] = \infty$. The first two are finite extensions, while the third is infinite.

2.2.1 Examples of Subfields

- Like with subrings, it is not necessary to verify most of the field axioms to show that a subset is actually a subfield:
- Proposition (Subfield Criterion): A subset S of a field F is a subfield if and only if S contains 0 and 1, and is closed under subtraction and division. In other words, for any $a, b, c \in S$ with $c \neq 0$, we have $a - b \in S$ and $a \cdot c^{-1} \in S$.
  - Equivalently, S is a subfield if and only if it is a subring that contains 1 and is closed under multiplicative inverses.
  - Proof: First suppose S is a subfield: then it contains an additive identity 0_S. We have $0_S + 0_S = 0_S = 0_S + 0_F$ by [F3] in S and in F, and so by additive cancellation in F we see $0_S = 0_F$. In a similar way we see that $1_S = 1_F$, and that the additive and multiplicative inverses in S agree with those in F; the statements $a - b \in S$ and $a \cdot c^{-1} \in S$ are then immediate.
  - Conversely, suppose S is nonempty and closed under subtraction and division. The axioms [F1], [F2], [F5], [F6], and [F9], follow immediately from the corresponding properties of F.
  - Setting $b = a$ shows $a - a = 0 \in S$ yielding [F3], and then setting $a = 0$ yields $0 - b = -b \in S$ yielding [F4].
  - Setting $c = a$ shows $a \cdot a^{-1} = 1 \in S$ yielding [F7], and then setting $a = 1$ yields $c^{-1} \in S$ yielding [F8].

- Remark: Note that closure under subtraction and division is equivalent to closure under addition, multiplication, additive inverses, and multiplicative inverses. (It is sometimes easier to check these four properties independently, rather than combining them.)

- Here are a few more examples and non-examples of subfields:

- Non-Example: The set $S = \{a + b\sqrt[3]{2} : a, b \in \mathbb{Q}\}$ is not a subfield of $\mathbb{R}$, where $\sqrt[3]{2}$ denotes the real cube root of 2.

  - This set is not closed under multiplication (so it is not even a subring): the element $\sqrt[3]{2} \cdot \sqrt[3]{2} = \sqrt[3]{4}$ is not in the set, because we cannot write $\sqrt[3]{4} = a + b\sqrt[3]{2}$ for any rational numbers a and b. This fact may seem obvious, but it is not so easy to prove directly!
  - Here is one argument: if $\sqrt[3]{4} = a + b\sqrt[3]{2}$ then multiplying by $\sqrt[3]{2}$ yields $2 = a\sqrt[3]{2} + b\sqrt[3]{4}$ and plugging in for $\sqrt[3]{4}$ then yields $2 = a\sqrt[3]{2} + b(a + b\sqrt[3]{2}) = ab + (a + b^2)\sqrt[3]{2}$. Since $\sqrt[3]{2}$ is irrational and a, b are rational, the coefficient of $\sqrt[3]{2}$ must be 0 so that $a = -b^2$. But this does not work since it yields $-a^3 = 2$, which is impossible if a is rational.

- Example: The set $S = \{a + b\sqrt[3]{2} + c\sqrt[3]{4} : a, b, c \in \mathbb{Q}\}$ is a subfield of $\mathbb{R}$, denoted $\mathbb{Q}[\sqrt[3]{2}]$. (We use square brackets, like with the polynomial ring $F[x]$, because $\mathbb{Q}[\sqrt[3]{2}]$ is the collection of polynomials in $\sqrt[3]{2}$.)

  - It is a straightforward calculation to see that S is closed under addition, additive inverses, and multiplication (so it is a subring). It is less clear why every nonzero element in S possesses a multiplicative inverse.
  - In fact, one may verify that $\dfrac{1}{a + b\sqrt[3]{2} + c\sqrt[3]{4}} = \dfrac{(a^2 - 2bc) + (2c^2 - ab)\sqrt[3]{2} + (b^2 - ac)\sqrt[3]{4}}{a^3 + 2b^3 + 4c^3 - 6abc}$, and that the denominator is never zero for $a, b, c \in \mathbb{Q}$ except when $a = b = c = 0$.
  - Explicitly: since every term in the denominator has degree 3, by multiplying through by a common denominator we may assume that a, b, c are relatively prime integers. Then a must be even since the other terms all have even coefficients; cancelling the common factor of 2 then shows b must be even, and then cancelling again shows c must be even: contradiction.
  - Using a similar calculation, we can show that the set $\{1, \sqrt[3]{2}, \sqrt[3]{4}\}$ is $\mathbb{Q}$-linearly independent and is therefore a basis for $\mathbb{Q}[\sqrt[3]{2}]$. Thus, we see that $[\mathbb{Q}[\sqrt[3]{2}] : \mathbb{Q}] = 3$.

- Non-Example: The set $S = \{a + b\sqrt{2} + c\sqrt{3} : a, b, c \in \mathbb{Q}\}$ does not form a field.

  - This set is not closed under multiplication, since $\sqrt{6} = \sqrt{2} \cdot \sqrt{3}$ is not in S. Like in the example above, this is not so easy to prove directly.
  - Here is one argument: if $\sqrt{6} = a + b\sqrt{2} + c\sqrt{3}$ then rearranging yields $\sqrt{6} - c\sqrt{3} = a + b\sqrt{2}$. Squaring both sides yields $(6 + 3c^2) - 6c\sqrt{2} = (a^2 + 2b^2) + 2ab\sqrt{2}$. Since $\sqrt{2}$ is irrational this requires $2ab = -6c$ and $6 + 3c^2 = a^2 + 2b^2$. Solving the first equation for c yields $c = -ab/3$, and then plugging into the second equation yields $18 + a^2b^2 = 3a^2 + 6b^2$. But this can be rearranged and factored as $(a^2 - 6)(b^2 - 3) = 0$, which has no rational solutions for a, b.

- Example: The set $S = \{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} : a, b, c, d \in \mathbb{Q}\}$ forms a field, denoted $\mathbb{Q}[\sqrt{2}, \sqrt{3}]$.

  - As with $\mathbb{Q}[\sqrt[3]{2}]$, it is easy to see that S is a subring: the hard part is the existence of multiplicative inverses.
  - One can "rationalize denominators" repeatedly to compute multiplicative inverses in S: explicitly, the multiplicative inverse of $a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}$ can be computed to be

    $$\frac{(a^3-2ab^2-3ac^2+12bcd-6ad^2)+(-2a^2b+2b^3-3bc^2+6acd-6bd^2)\sqrt{2}+(-a^2c-2b^2c+3c^3+4abd-6cd^2)\sqrt{3}+(2abc-a^2d+2b^2d-3c^2d+6d^3)\sqrt{6}}{a^4-4a^2b^2-6a^2c^2-12a^2d^2+48abcd+4b^4-12b^2c^2-24b^2d^2+9c^4-36c^2d^2+36d^4}$$

    and one can similarly show that the denominator is never zero unless $a = b = c = d = 0$.
  - Using a similar calculation, we can show that the set $\{1, \sqrt{2}, \sqrt{3}, \sqrt{6}\}$ is $\mathbb{Q}$-linearly independent and is therefore a basis for $\mathbb{Q}[\sqrt{2}, \sqrt{3}]$. Thus, we see that $[\mathbb{Q}[\sqrt{2}, \sqrt{3}] : \mathbb{Q}] = 4$.

- The computations in the examples above have the virtue of being explicit, but seem far more complicated than necessary to justify the (seemingly) simple statements that the sets $\mathbb{Q}[\sqrt[3]{2}]$ and $\mathbb{Q}[\sqrt{2}, \sqrt{3}]$ are fields, and they also seem unlikely to generalize well to further examples.
    - *Ad hoc* calculations like these are hard to extend to a general theory, so we will need to develop other techniques for studying subfields and field extensions.

2.2.2 Properties of Subfields

- As an immediate corollary of the subfield criterion, we see that the intersection of subfields is also a subfield:
- Proposition (Intersection of Subfields): If F is a field, then the intersection of any nonempty collection of subfields of F is also a subfield of F.
    - Proof: Let $S = \bigcap_{i \in I} F_i$ where the F_i are subfields of F. Then by the subfield criterion, $0, 1 \in F_i$ for all $i \in I$, so S contains 0 and 1.
    - Furthermore, for any $a, b, c \in S$ with $c \neq 0$, we have $a, b, c \in F_i$ for all i. Thus, $a - b \in F_i$ and $a \cdot c^{-1} \in F_i$ for all i by the subfield criterion, and therefore $a - b \in S$ and $a \cdot c^{-1} \in S$, so S is a subfield.
- Like with vector spaces and span, if we have a subset S of a field, we would like to understand the structure of the subfield of F "generated by" the elements of S.
    - If F is a field and S is a subset of F, a natural choice is to define "the subfield generated by S" to be the smallest subfield of F containing S.
    - *A priori*, it is not obvious that there is such a smallest subfield. However, since the intersection of any nonempty collection of subfields is also a subfield, per the above proposition, and since S is always contained in at least one subfield (namely F itself), we can equivalently define the subfield $E \subseteq F$ generated by S to be the intersection of all subfields containing S.
    - Although this definition is clearly well-posed, we have not really described what the elements in this subfield E actually are.
    - If $x_1, x_2, \ldots, x_n \in S$, then since E is closed under addition and multiplication and contains 1, we see that any polynomial with integer coefficients in $x_1, x_2, \ldots, x_n$ must be in S as well. And since E is closed under division, it must in fact contain any "rational function" (i.e., quotient of one polynomial by another) of $x_1, x_2, \ldots, x_n$.
    - On the other hand, one can verify that the collection of all such rational functions in elements of S actually does form a field (since the sum, product, additive inverse, and multiplicative inverse of rational functions are also rational functions), so this collection is the desired field E.
- Definition: If F is a field and S is a subset of F, we define the subfield of F generated by S to be the intersection of all subfields of F containing S.
    - We will frequently be interested in extensions of subfields: if S is a subset of F and E is a subfield of F, then we define $E(S)$ to be the smallest subfield containing E and S.
- A particular special case is the subfield generated by 1:
- Definition: If F is a field, the prime subfield of F is the subfield generated by 1. (It is sometimes written F'.)
    - Any subfield of F contains 1, so the subfield generated by 1 will be the "smallest" subfield of F, and will be contained in every other subfield of F.
    - The structure of the prime subfield will depend on the characteristic of F:
- Proposition (Prime Subfield): If F has characteristic $p > 0$, then the prime subfield of F is (isomorphic to) $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$, and if F has characteristic 0, then the prime subfield of F is (isomorphic to) $\mathbb{Q}$.
    - Proof: Let E be the prime subfield of F. If F has characteristic $p > 0$, consider the map $\varphi : (\mathbb{Z}/p\mathbb{Z}) \to E$ defined by $\varphi(\overline{a}) = a1_F$. This map is well-defined by the assumption on the characteristic (since $p1_F = 0$ in a field of characteristic p).

- Furthermore, it is easy to see that $\varphi(\overline{a}+\overline{b}) = \varphi(\overline{a}) + \varphi(\overline{b})$, $\varphi(\overline{a}\overline{b}) = \varphi(\overline{a})\varphi(\overline{b})$, and that φ has an inverse map defined by $\varphi^{-1}(a1_F) = \overline{a}$. Thus, φ is a ring isomorphism and E is isomorphic to $\mathbb{Z}/p\mathbb{Z}$.
- If F has characteristic 0, then instead consider the map $\varphi : \mathbb{Q} \to E$ defined by $\varphi(a/b) = (a1_F) \cdot (b1_F)^{-1}$. This map is well-defined by the assumption that $b1_F \neq 0_F$ whenever $b \neq 0$.
- As above it is straightforward to see that $\varphi(a+b) = \varphi(a) + \varphi(b)$, $\varphi(ab) = \varphi(a)\varphi(b)$, and that φ has an inverse map defined by $\varphi^{-1}[(a1_F) \cdot (b1_F)^{-1}] = a/b$. Thus, φ is a ring isomorphism and E is isomorphic to $\mathbb{Q}$.

- As we noted above, every subfield of F contains the prime subfield of F, which is to say, every subfield of F is an extension field of the prime subfield. We can therefore always denote the subfield of F generated by S as $E(S)$, where E is the prime subfield of F.
  - The round parentheses are intended to indicate that we are closing under field operations, in contrast to square brackets where we only close under ring operations.
  - Thus, for example, the subfield of $\mathbb{R}$ generated by $\sqrt[3]{2}$ (the "rational functions in $\sqrt[3]{2}$ with rational coefficients") is denoted $\mathbb{Q}(\sqrt[3]{2})$, in contrast to the set $\mathbb{Q}[\sqrt[3]{2}]$ of polynomials in $\sqrt[3]{2}$.
  - As it happens, these two sets turn out to be the same, because $\mathbb{Q}[\sqrt[3]{2}]$ is actually a field, but as we discussed, this is not a trivial statement to establish. Furthermore, as we will see, there exist real numbers α with the property that $\mathbb{Q}(\alpha) \neq \mathbb{Q}[\alpha]$ (an example being the case where α is the transcendental number π).

2.2.3 Simple Extensions, Minimal Polynomials

- Let us return to the example of the field $F = \mathbb{Q}[\sqrt[3]{2}] = \{a+b\sqrt[3]{2}+c\sqrt[3]{4} : a,b,c \in \mathbb{Q}\}$ with a different approach to show it is (in fact) a field:
  - For shorthand, write $\overline{x} = \sqrt[3]{2}$: then every element of F has the form $a + b\overline{x} + c\overline{x}^2$.
  - Addition is performed in the obvious way: $(a+b\overline{x}+c\overline{x}^2)+(d+e\overline{x}+f\overline{x}^2) = (a+d)+(b+e)\overline{x}+(c+f)\overline{x}^2$.
  - For multiplication, we can use the distributive law to compute $(a + b\overline{x} + c\overline{x}^2) \cdot (d + e\overline{x} + f\overline{x}^2) = ad + (ae + bd)\overline{x} + (af + be + cd)\overline{x}^2 + (bf + ce)\overline{x}^3 + cf\overline{x}^4$.
  - Since $\overline{x}^3 = (\sqrt[3]{2})^3 = 2$, we see $(a + b\overline{x} + c\overline{x}^2) \cdot (d + e\overline{x} + f\overline{x}^2) = (ad + 2bf + 2ce) + (ae + bd + 2cf)\overline{x} + (af + be + cd)\overline{x}^2$.
  - Now notice this is precisely the same description as the arithmetic in the polynomial quotient ring $\mathbb{Q}[x]/(x^3-2)$: thus, since the map sending $a+b\overline{x}+c\overline{x}^2$ to $a+b\sqrt[3]{2}+c\sqrt[3]{4}$ is clearly a bijection, the rings $\mathbb{Q}[\sqrt[3]{2}]$ and $\mathbb{Q}[x]/(x^3-2)$ are isomorphic.
  - Furthermore, because the polynomial $x^3 - 2$ is irreducible in $\mathbb{Q}[x]$ (either because it has degree 3 and no rational roots, or by Eisenstein's criterion with $p = 2$), we know that $\mathbb{Q}[x]/(x^3-2)$ is a field.
  - Therefore, $\mathbb{Q}[\sqrt[3]{2}]$ is a field as well, since it is ring-isomorphic to a field.

- We can generalize the analysis in this example to the class of field extensions generated by a single element:
- Definition: If K/F is a field extension, we say that K is a simple extension if $K = F(\alpha)$ for some $\alpha \in K$: in other words, if K is generated over F by the single element α.
  - Example: $\mathbb{C}$ is a simple extension of $\mathbb{R}$, generated by the element i. (In fact, $\mathbb{C}$ is generated over $\mathbb{R}$ by any non-real complex number.)
  - Example: The rational function field $F(x)$ is a simple extension of F, generated by the element x.

- It is not always obvious whether a given extension has a single generator (although we will later be able to characterize simple extensions). In many cases, if we have a list of generators for the extension, there is often some combination of generators that generates the field by itself:
- Example: Show that the field $\mathbb{Q}[\sqrt{2}, \sqrt{3}]$ is a simple extension of $\mathbb{Q}$ generated by the element $\alpha = \sqrt{2} + \sqrt{3}$.

- We can see that $\alpha^2 = 5 + 2\sqrt{6}$ and $\alpha^3 = 11\sqrt{2} + 9\sqrt{3}$, so $\sqrt{6} = \frac{1}{2}(\alpha^2 - 5)$, $\sqrt{2} = \frac{1}{2}(\alpha^3 - 9\alpha)$, and $\sqrt{3} = \frac{1}{2}(11\alpha - \alpha^3)$.
- Therefore, every element in the field is a rational function (in fact, a polynomial) with rational coefficients in α, meaning that α is a generator for the field.

- The structure of the simple extension $K = F(\alpha)$ will depend on the nature of the element α: specifically, on whether α is the root of some polynomial with coefficients in F.

- Definition: If K/F is a field extension, we say that the element $\alpha \in K$ is algebraic over F if α is the root of some nonzero polynomial $p \in F[x]$. Otherwise, if α is not a root of any nonzero polynomial in $F[x]$, we say α is transcendental over F.
  - Example: The elements $\sqrt{2}$, $\sqrt[3]{2}$, i, and $2 - 3i$ of $\mathbb{C}$ are algebraic over $\mathbb{Q}$, since they are roots of the polynomials $x^2 - 2$, $x^6 - 4$, $x^4 - 1$, and $(x-2)^2 + 9$ respectively.
  - Example: The elements e and π of $\mathbb{R}$ are transcendental over $\mathbb{Q}$ (neither of these statements is easy to prove, and we will not prove them!).
  - Example: The element t in the field of rational functions $F(t)$ is transcendental over F, since it does not satisfy any polynomial with coefficients in F. (This fact is implicit in the definition of the polynomial ring $F[t]$.)
  - Example: If p is any irreducible polynomial over F, then the element $\overline{x}$ in the polynomial quotient ring $K = F[x]/p$ is algebraic over F, because $p(\overline{x}) = 0$ in K.

- An algebraic element is by definition a root of some nonzero polynomial, and may be a root of many different polynomials (for example, $\sqrt{2}$ is a root of each of $x^2 - 2$, $x^3 + x^2 - 2x - 2$, and $x^4 - 4$). However, all of these polynomials are multiples of an essentially unique monic polynomial:

- Proposition (Minimal Polynomials): If K/F is a field extension and $\alpha \in K$ is algebraic over F and nonzero, then there exists a unique monic irreducible polynomial $m \in F[x]$ such that $m(\alpha) = 0$. This polynomial is called the minimal polynomial of α over F, and is the monic polynomial of smallest positive degree having α as a root; furthermore, any other polynomial having α as a root is divisible by m.
  - Proof: If α is algebraic, consider the set of all nonzero polynomials in $F[x]$ having α as a root. By hypothesis, S is nonempty, so by the well-ordering axiom, S contains a polynomial of minimal positive degree.
  - It is easy to see that if $m(\alpha) = 0$, then any F-multiple of m also has α as a root, so we may divide m by its leading coefficient to see that m is monic.
  - We claim that m is irreducible: if m had a factorization $m = pq$ with $0 < \deg p, \deg q < \deg m$, then by evaluating both sides at α we would see $0 = m(\alpha) = p(\alpha)q(\alpha)$.
  - Since K is a field, this implies $p(\alpha) = 0$ or $q(\alpha) = 0$ so that one of p, q has α as a root and is therefore in S. But this is a contradiction, since m was assumed to be an element of minimal degree in S: thus, m is irreducible.
  - Furthermore, if $b(\alpha) = 0$ then applying the division algorithm to b and m shows that $b = qm + r$ for some q, r with $\deg r < \deg m$. Evaluating both sides at α and rearranging then yields $r(\alpha) = b(\alpha) - q(\alpha)m(\alpha) = 0$, so since $\deg r < \deg m$ we must have $r = 0$.
  - For the uniqueness of m, if there were another such polynomial m', then by the above we would have $m'|m$ and $m|m'$ so that m and m' are associates. But since both m and m' are monic, they are equal.

- Example: The minimal polynomial of $\sqrt{2}$ over $\mathbb{Q}$ is $x^2 - 2$, and the minimal polynomial of $\sqrt[3]{2}$ over $\mathbb{Q}$ is $x^3 - 2$.
  - Both of these observations follow because these polynomials do have the required root and are irreducible.

- The nature of the simple extension $F(\alpha)$ will depend on whether α is algebraic or transcendental over F. Explicitly:

- Theorem (Simple Extensions): Suppose K/F is a simple extension with $K = F(\alpha)$. If α is algebraic over F with minimal polynomial $m(x)$ then K is isomorphic to the field $F[x]/m(x)$, while if α is transcendental over F then K is isomorphic to the field $F(t)$ of rational functions in t.
  - The idea of the proof is simply to show that the map associating $\overline{x}$ (or t, as appropriate) with α is a well-defined ring isomorphism.
  - Proof: First suppose α is algebraic over F with minimal polynomial $m(x)$. Consider the map $\varphi : F[x]/m(x) \to K$ given by mapping $\overline{p(x)}$ to $p(\alpha)$. This map is well defined because if $\overline{p} = \overline{q}$ in $F[x]/m(x)$, then m divides $q - p$, so since α is a root of m, we see that $p(\alpha) = q(\alpha)$.
  - Furthermore, it is easy to see that
$$\begin{aligned} \varphi(\overline{p} + \overline{q}) &= \varphi(\overline{p+q}) = (p+q)(\alpha) = p(\alpha) + q(\alpha) = \varphi(\overline{p}) + \varphi(\overline{q}) \\ \varphi(\overline{p} \cdot \overline{q}) &= \varphi(\overline{pq}) = (pq)(\alpha) = p(\alpha)q(\alpha) = \varphi(\overline{p})\varphi(\overline{q}) \end{aligned}$$
    so φ respects addition and multiplication.
  - Furthermore, φ is injective, because $\varphi(\overline{p}) = \varphi(\overline{q})$ implies $\varphi(\overline{p-q}) = 0$ implies $(p-q)(\alpha) = 0$, and by our discussion above this implies m divides $p - q$ so that $\overline{p} = \overline{q}$ in $F[x]/m(x)$.
  - All of this shows that the image of the map φ (i.e., the set of elements of K of the form $\varphi(\overline{p})$ for some $\overline{p}$) is ring-isomorphic to $F[x]/m(x)$, since φ yields a bijection between $F[x]/m(x)$ and this subring of K.
  - But $F[x]/m(x)$ is a field since m is irreducible, so the image of φ is a subfield of K containing α and F. But by definition of $F(\alpha)$, this means it must actually be $F(\alpha) = K$, as claimed.
  - If α is transcendental over F, the argument is similar, except we instead use the map $\varphi : F(t) \to K$ sending $\dfrac{p(t)}{q(t)}$ to $\dfrac{p(\alpha)}{q(\alpha)}$. This map is well defined because $q(\alpha) \neq 0$ whenever q is not the zero polynomial by the assumption that α is transcendental.
  - It is easy to see that φ respects addition and multiplication, and is injective (the latter because α is transcendental).
  - Surjectivity follows by a similar argument as above: $F(t)$ is isomorphic as a ring to the image of φ, and since $F(t)$ is a field, we conclude that the image of φ is a subfield of K containing α and F, hence is equal to $K = F(\alpha)$.

- Using the description of simple extensions we can easily compute the extension degree, and characterize when $F(\alpha) = F[\alpha]$:

- Corollary (Simple Extension Degrees): Suppose K/F is a simple extension with $K = F(\alpha)$. If α is algebraic over F with minimal polynomial $m(x)$ then $[F(\alpha) : F] = \deg m$, and $F(\alpha)$ is spanned (as an F-vector space) by $\{1, \alpha, \alpha^2, \ldots, \alpha^{\deg m - 1}\}$, while if α is transcendental over F then $[F(\alpha) : F] = \infty$. Furthermore, $F(\alpha) = F[\alpha]$ if and only if α is algebraic over F.
  - Proof: If α is algebraic over F with minimal polynomial $m(x)$, then K is isomorphic to $F[x]/m(x)$. Suppose $\deg m = n$.
  - From our discussion of residue classes in $F[x]/m(x)$, we know (via an application of the division algorithm) that every residue class can be written uniquely in the form $b_0 + b_1\overline{x} + \cdots + b_{n-1}\overline{x}^{n-1}$ for unique elements $b_i \in F$.
  - Equivalently, this says that the set $\{1, \overline{x}, \overline{x}^2, \ldots, \overline{x}^{n-1}\}$ is an F-basis for $F[x]/m(x)$. Applying the isomorphism between K and $F[x]/m(x)$ shows that the set $\{1, \alpha, \alpha^2, \ldots, \alpha^{n-1}\}$ is an F-basis for $F(\alpha)$, so $[F(\alpha) : F] = n$. Furthermore, we see immediately that $F(\alpha) = F[\alpha]$ in this case.
  - If α is transcendental over F, then the set $\{1, \alpha, \alpha^2, \ldots\}$ is linearly independent over F, as any nontrivial linear dependence $c_0 + c_1\alpha + \cdots + c_n\alpha^n = 0$ would imply that α is the root of some nonzero polynomial in $F[x]$, but this cannot occur because α is transcendental.
  - Since K contains an infinite F-linearly independent set we see $[K : F] = \infty$. Furthermore, $F(\alpha)$ contains elements that are not polynomials in α (namely, any rational function that is not a polynomial). For example, we cannot have $\alpha^{-1} = p(\alpha)$ since this would imply $1 - \alpha p(\alpha) = 0$ so that α would be a root of a nonzero polynomial. Therefore, $\alpha^{-1} \in F(\alpha)$ is not in $F[\alpha]$, so $F(\alpha) \neq F[\alpha]$ in this case.

- Example: Show that the field $\mathbb{Q}(\sqrt[8]{2})$ has degree 8 over $\mathbb{Q}$ and find a basis.
  - Observe that $\sqrt[8]{2}$ is a root of the polynomial $x^8 - 2$ over $\mathbb{Q}$, and this polynomial is irreducible by Eisenstein's criterion with $p = 2$.
  - Therefore, $x^8 - 2$ is the minimal polynomial of $\sqrt[8]{2}$, so by our results on simple extensions we see that $[\mathbb{Q}(\sqrt[8]{2}) : \mathbb{Q}] = 8$ and that the set $\{1, 2^{1/8}, 2^{2/8}, \ldots, 2^{7/8}\}$ is a basis.
- We can also see that if α and β have the same minimal polynomial over F, then the corresponding field extensions $F(\alpha)$ and $F(\beta)$ are isomorphic:
- Corollary: If α and β are two elements in K/F with equal minimal polynomials, then the fields $F(\alpha)$ and $F(\beta)$ are isomorphic. Explicitly, there is an isomorphism $\varphi : F(\alpha) \to F(\beta)$ that fixes F (i.e., sends every element in F to itself) and sends α to β.
  - Proof: Both fields are isomorphic to $F[x]/m(x)$ where m is the common minimal polynomial, and so $F(\alpha)$ is isomorphic to $F(\beta)$ since the composition of isomorphisms is an isomorphism. Writing this map down explicitly shows that it sends α to β and fixes every element of F.
- Example: If $\alpha = \sqrt{2}$ and $\beta = -\sqrt{2}$, then α and β both have the minimal polynomial $x^2 - 2$ over $\mathbb{Q}$, so $F(\alpha)$ is isomorphic to $F(\beta)$.
  - Explicitly, the isomorphism maps $a + b\sqrt{2} \in \mathbb{Q}(\sqrt{2})$ to the element $a + b(-\sqrt{2}) \in \mathbb{Q}(-\sqrt{2})$.
  - In fact, these fields are equal (as subfields of $\mathbb{R}$ or of $\mathbb{C}$) because $\sqrt{2} \in \mathbb{Q}(-\sqrt{2})$ and $-\sqrt{2} \in \mathbb{Q}(\sqrt{2})$. Thus, we may alternatively view this isomorphism as a map from $\mathbb{Q}(\sqrt{2})$ to itself, acting via $\sigma(a + b\sqrt{2}) = a - b\sqrt{2}$.
- In the example above, we saw that $F(\alpha)$ was actually equal to $F(\beta)$ as a set. This is not necessarily the case in general:
- Example: If $\alpha = \sqrt[3]{2}$ is the real cube root of 2, and $\beta = e^{2\pi i/3}\sqrt[3]{2}$ is a nonreal cube root of 2, then α and β both have the minimal polynomial $x^3 - 2$ over $\mathbb{Q}$, and so $F(\alpha)$ is isomorphic to $F(\beta)$.
  - As an explicit complex number, notice that $e^{2\pi i/3} = (-1 + i\sqrt{3})/2$, and the cube of this number is indeed 1 (one may either note that $e^{2\pi i} = 1$, or simply cube it explicitly). Thus, $\beta^3 = 2$ as claimed.
  - However, as subfields of $\mathbb{C}$, $F(\alpha)$ is not equal to $F(\beta)$, because $F(\alpha)$ is a subfield of $\mathbb{R}$ while $F(\beta)$ is not. Nonetheless, these two fields have precisely the same algebraic structure.
- The point is that the different roots of the minimal polynomial are "algebraically indistinguishable", in the sense that the resulting extension fields have the same algebraic structure.
  - This does not mean that the fields are "the same", since we may sometimes be able to distinguish these fields in some other ("non-algebraic") way.
  - In the example above with $\mathbb{Q}(\sqrt[3]{2})$ and $\mathbb{Q}(e^{2\pi i/3}\sqrt[3]{2})$, we used information about the field $\mathbb{R}$ (which involves using additional continuous operations, rather than intrinsic algebraic properties of the field $\mathbb{Q}$) to distinguish these two fields.

2.2.4 Algebraic Extensions

- Now that we have described simple extensions, we can expand our focus to other field extensions. A natural class of extensions are those in which every element is algebraic:
- Definition: The field extension K/F is algebraic if every $\alpha \in K$ is algebraic over F: in other words, if every α is a root of a nonzero polynomial in $F[x]$.
- From our description of simple extensions it is easy to characterize when a simple extension is algebraic:
- Proposition (Simple Algebraic Extensions): A simple extension $F(\alpha)/F$ is algebraic if and only if it has finite degree. Furthermore, if $[F(\alpha) : F] = n$, then every element in $F(\alpha)$ satisfies a nonzero polynomial of degree at most n in $F[x]$.

- Proof: If $F(\alpha)$ is algebraic then $\alpha \in F(\alpha)$ is algebraic over F. If the minimal polynomial for α has degree n, then as we showed earlier, $[F(\alpha) : F] = n$, so the extension has finite degree.
- Conversely, suppose $F(\alpha)/F$ has degree n and let $\beta \in F(\alpha)$. Observe the set $\{1, \beta, \beta^2, \ldots, \beta^n\}$ has $n+1$ elements, so since $F(\alpha)$ only has dimension n over F, it must be linearly dependent.
- In other words, there exist $c_i \in F$, not all zero, with $c_0 + c_1\beta + \cdots + c_n\beta^n = 0$: thus, β is the root of a nonzero polynomial in $F[x]$, so β is algebraic over F. The second statement is also immediate.

- Corollary: Finite-degree extensions are algebraic.
  - We will remark that there exist algebraic extensions of infinite degree, so the converse of this result is not true.
  - Proof: By the previous proposition, any element of a degree-n extension satisfies a polynomial of degree at most n, and is therefore algebraic.

- Next we analyze extensions with a finite number of generators: $K = F(\alpha_1, \alpha_2, \ldots, \alpha_n)$ for some $\alpha_i \in K$.
  - If any of the α_i are transcendental over F, then clearly K is non-algebraic since it contains a transcendental element.
  - On the other hand, if all of the α_i are algebraic over F, it seems reasonable to hypothesize that K itself will also be algebraic over F, since every element of K is a combination of algebraic elements.
  - The key idea is to observe that we can obtain K as a "chain" of simple extensions by adjoining the α_i one at a time.
  - To illustrate, suppose $K = F(\alpha, \beta, \gamma)$. Then K contains $F(\alpha)$ along with β, and so it contains the extension field $F(\alpha)(\beta)$.
  - But $F(\alpha, \beta)$ is by definition the smallest subfield of K containing α and β, so $F(\alpha, \beta)$ is contained in $F(\alpha)(\beta)$.
  - On the other hand, since $F(\alpha, \beta)$ contains $F(\alpha)$ and β, since $F(\alpha)(\beta)$ is the smallest subfield of K containing $F(\alpha)$ and β, we see that $F(\alpha)(\beta)$ is contained in $F(\alpha, \beta)$.
  - Thus, $F(\alpha, \beta) = F(\alpha)(\beta)$, so $F(\alpha, \beta)/F(\alpha)$ is a simple extension.
  - In the same way, we can see that $K = F(\alpha, \beta, \gamma)/F(\alpha, \beta)$ is also a simple extension, so we can obtain K from F using a chain of 3 simple extensions $F(\alpha, \beta, \gamma)/F(\alpha, \beta)/F(\alpha)/F$.
  - In order to show that the resulting field K will be algebraic if each α_i is algebraic, we need to know how extension degrees behave in "towers" of field extensions:

- Theorem (Degrees in Towers): If L/K and K/F are both field extensions, then so is L/F, and $[L : F] = [L : K] \cdot [K : F]$ (where if one side is infinite, then so is the other). In particular, $[K : F]$ divides $[L : F]$.
  - Although we will not need it, in fact a more general statement is true: if V is a K-vector space and K/F is a field extension, then (under the same operations) V is also an F-vector space, and $\dim_F V = [K : F] \cdot \dim_K V$. The theorem is the special case where V is the K-vector space L.
  - Proof: First suppose that $[K : F] = n$ with basis $\{a_1, a_2, \ldots, a_n\}$ and $[L : K] = m$ with basis $\{v_1, v_2, \ldots, v_m\}$ are both finite. We claim that the set β of the mn pairwise products $a_i v_j$ for $1 \leq i \leq n$ and $1 \leq j \leq m$ is a basis for L/F.
  - First observe that no two of these pairwise products are equal; to see this suppose $a_i v_j = a_k v_l$ so that $a_i v_j - a_k v_l = 0$. If $j \neq l$ then v_j, v_l would be K-linearly dependent (contrary to our assumption), and if $j = l$ then cancelling v_j (which is nonzero since it is a basis element) would yield $a_i = a_k$.
  - To see that β is a spanning set, for any $w \in L$ by the hypothesis that $\{v_1, v_2, \ldots, v_m\}$ spans L/K, we may write $w = b_1 v_1 + \cdots + b_m v_m$ for some $b_i \in K$. Furthermore, since the $b_i \in K$, by the hypothesis that $\{a_1, a_2, \ldots, a_n\}$ spans K/F, we may write $b_i = c_{i,1} a_1 + \cdots + c_{i,n} a_n$ for some $c_{i,j} \in F$.
  - Now substituting in the expressions for the b_i in terms of the $c_{i,j}$ and the a_i to the expression for w yields

$$\begin{aligned} w &= b_1 v_1 + \cdots + b_m v_m \\ &= (c_{1,1} a_1 + \cdots + c_{1,n} a_n) v_1 + \cdots + (c_{m,1} a_1 + \cdots + c_{m,n} a_n) v_m \\ &= c_{1,1} a_1 v_1 + \cdots + c_{m,n} a_n v_m \end{aligned}$$

  and therefore w is an F-linear combination of the elements of β, meaning that β is a spanning set.

- To see that β is linearly independent, suppose we had a linear combination $c_{1,1}a_1v_1 + \cdots + c_{m,n}a_nv_m = 0$ for some $c_{i,j} \in F$.
- By the above calculation (in reverse) if we set $b_i = c_{i,1}a_1 + \cdots + c_{i,n}a_n$ then $b_i \in K$ for each i and $b_1v_1 + \cdots + b_mv_m = 0$. Since the v_i are linearly independent over K, this means $b_i = 0$ for each i.
- Then since $b_i = c_{i,1}a_1 + \cdots + c_{i,n}a_n$ and the a_i are linearly independent over F, we conclude that $c_{i,j} = 0$ for each i, j, and so β is also linearly independent, hence a basis.
- For the infinite-degree cases, if $[K : F] = \infty$ then any basis of K/F is an infinite linearly-independent subset of L/F, meaning that $[L : F] = \infty$ as well.
- Likewise, if $[L : K] = \infty$, then any basis of L/K is an infinite K-linearly independent subset, which is also clearly F-linearly independent (since any linear dependence over F would also hold over K), and so $[L : F] = \infty$ again.
- Finally, if $[L : F] = \infty$, then at least one of $[L : K]$ and $[K : F]$ must be infinite, since if both are finite then our proof above shows that $[L : F]$ is also finite.

- Corollary (Finite Algebraic Extensions): If K/F is a field extension with $K = F(\alpha_1, \alpha_2, \ldots, \alpha_n)$, then K/F is algebraic if and only if each of the α_i are algebraic over F. In this case, $[K : F] \leq \prod_{i=1}^{n}[F(\alpha_i) : F]$, and every element of K is a polynomial (with coefficients from F) in the α_i.
  - Proof: If any of the α_i are transcendental over F then K is not algebraic over F, so now suppose each of the α_i are algebraic.
  - As noted earlier, we may obtain K as a chain of simple extensions $K/F(\alpha_1, \ldots, \alpha_{n-1})/\cdots/F(\alpha_1, \alpha_2)/F(\alpha_1)/F$.
  - By hypothesis, for each $1 \leq i \leq n$, α_i is algebraic over F, so α_i is also algebraic over $F(\alpha_1, \ldots, \alpha_{i-1})$, since the minimal polynomial for α_i over F may also be thought of as a polynomial over $F(\alpha_1, \ldots, \alpha_{i-1})$.
  - Therefore, since a simple extension is algebraic if and only if it has finite degree, we see that $[F(\alpha_1, \ldots, \alpha_i) : F(\alpha_1, \ldots, \alpha_{i-1})]$ is finite for each i.
  - Then by the multiplicativity of extension degrees (and a trivial induction), we conclude that $[K : F] = \prod_{i=1}^{n}[F(\alpha_1, \ldots, \alpha_i) : F(\alpha_1, \ldots, \alpha_{i-1})]$ is finite. Since finite-degree extensions are algebraic, this means K/F is algebraic as claimed.
  - For the second statement, consider the minimal polynomial $m(x)$ of α_i over F and the minimal polynomial $m'(x)$ of α_i over $F(\alpha_1, \ldots, \alpha_{i-1})$. Since $m(x)$ is also a polynomial in $F(\alpha_1, \ldots, \alpha_{i-1})$ having α_i as a root, by properties of minimal polynomials we see that $m'(x)$ divides $m(x)$, so $\deg m' \leq \deg m$. Converting to a statement about extension degrees yields $[F(\alpha_1, \ldots, \alpha_i) : F(\alpha_1, \ldots, \alpha_{i-1})] \leq [F(\alpha_i) : F]$, and then taking the product from $i = 1$ to n yields $[K : F] \leq \prod_{i=1}^{n}[F(\alpha_i) : F]$.
  - For the last statement, since $E[\beta] = E(\beta)$ when β is algebraic over E, by an easy induction we see that every element of K is a polynomial in the α_i.
  - Remark: More explicitly, every element of K is an F-linear combination of elements of the form $\alpha_1^{c_1}\alpha_2^{c_2}\cdots\alpha_n^{c_n}$, where each c_i is an integer with $0 \leq c_i \leq [F(\alpha_i) : F]$. This also follows by a straightforward induction, using the fact that every element of $E(\beta)$ is of the form $b_0 + b_1\beta + \cdots + b_{d-1}\beta^{d-1}$ where $[E(\beta) : E] = d$, as both the base case and inductive step.

- We can also show that every finite-degree extension is generated by a finite set of algebraic elements, and that an algebraic extension of an algebraic extension is also algebraic:

- Corollary (Characterization of Finite Extensions): If K/F is a field extension, then K/F has finite degree if and only if $K = F(\alpha_1, \ldots, \alpha_n)$ for some elements $\alpha_1, \ldots, \alpha_n \in K$ that are algebraic over F.
  - Proof: We already showed that $F(\alpha_1, \ldots, \alpha_n)/F$ is finite if $\alpha_1, \ldots, \alpha_n$ are algebraic over F. For the forward direction, suppose K/F has finite degree: then by definition, K has a finite basis $\{\alpha_1, \ldots, \alpha_n\}$ as an F-vector space, and so $K = F(\alpha_1, \ldots, \alpha_n)$.
  - Furthermore, since $F(\alpha_i)$ is a subfield of the finite-degree extension K/F, we see that $[F(\alpha_i) : F]$ is also finite (by the multiplicativity of extension degrees) and thus α_i is algebraic over F for each i, as required.

- Corollary (Towers of Algebraic Extensions): If L/K is an algebraic extension, and K/F is an algebraic extension, then L/F is an algebraic extension.

- Proof: Let $\alpha \in L$: then since α is algebraic over K it is the root of some polynomial $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ with the $a_i \in K$.
- Since K/F is also algebraic, each of the a_i are algebraic over F, and so the extension $E = F(a_0, a_1, \ldots, a_n)$ has finite degree over F. Furthermore, $E(\alpha)/E$ also has finite degree, because α is the root of a nonzero polynomial in $E[x]$.
- Thus, since $E(\alpha)/E$ and E/F both have finite degree, so does $E(\alpha)/F$: this means α satisfies a polynomial of finite degree over F, so α is algebraic over F. This holds for all $\alpha \in L$, so L is algebraic over F.

- We can also extend these results on degree to general "composite fields":
- Definition: If K_1 and K_2 are subfields of K, the composite field K_1K_2 is the intersection of all subfields of K containing both K_1 and K_2.
  - We can also consider composites of an arbitrary collection of subfields (namely, the intersection of all subfields containing every field in the collection).
  - Like with subfields generated by a set, it is easy to see that the composite field is the smallest subfield of K that contains both K_1 and K_2.
- Proposition (Degrees of Composites): If K_1/F and K_2/F are both finite-degree subextensions of K/F, then $[K_1K_2 : F] \leq [K_1 : F] \cdot [K_2 : F]$. If the degrees $[K_1 : F]$ and $[K_2 : F]$ are relatively prime, then equality always holds.
  - Proof: Suppose K_1/F has basis $\alpha_1, \ldots, \alpha_n$ and K_2/F has basis $\beta_1, \ldots, \beta_m$.
  - Then K_1K_2 contains F and each of $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m$ hence it contains $F(\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m)$. On the other hand, $F(\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m)$ contains both K_1 and K_2, hence K_1K_2.
  - Therefore, $K_1K_2 = F(\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m) = K_1(\beta_1, \ldots, \beta_m)$. Thus, $\beta_1, \ldots, \beta_m$ span K_1K_2/K_1, so $[K_1K_2 : K_1] \leq [K_2 : F]$. Then $[K_1K_2 : F] = [K_1K_2 : K_1] \cdot [K_1 : F] \leq [K_1 : F] \cdot [K_2 : F]$ as claimed.
  - For the second statement, note that $[K_1K_2 : F]$ is divisible by both $[K_1 : F] = n$ and $[K_2 : F] = m$. If m and n are relatively prime, this implies $[K_1K_2 : F] \geq mn$, so since $[K_1K_2 : F] \leq mn$ from above, we must have equality.

2.2.5 Examples of Small-Degree Field Extensions

- By using our results on simple and composite extensions, along with the multiplicativity of field degrees in towers, we can often say a great deal about extensions of small degree:
- Proposition (Quadratic Extensions): Suppose F is a field of characteristic not equal to 2 and K/F is a quadratic extension (i.e., degree 2). Then $K = F(\alpha)$ for any $\alpha \in K$ not in F, and in fact we can take $K = F(\beta)$ for some element with $\beta^2 \in F$ and $\beta \notin F$.
  - The last statement says (essentially) that $K = F(\sqrt{D})$ for some $D \in F$ that is not a square in F. (It is hard to be more precise than this, because it is difficult to give a clear definition for what "$\sqrt{D}$" means that does not end up being circular.)
  - In particular, the quadratic extensions of $\mathbb{Q}$ (inside $\mathbb{C}$) are precisely the extensions $\mathbb{Q}(\sqrt{D})$ that we have previously described.
  - Proof: Suppose K/F is a quadratic extension. If $\alpha \in K$ is not in F, then the set $\{1, \alpha\}$ is F-linearly independent, and since $[K : F] = 2$ it must therefore be a basis for K. Thus, $K = F(\alpha)$.
  - For the second statement, consider the minimal polynomial for any $\alpha \in K$ not in F: since $K = F(\alpha)$ and $[K : F] = 2$ we see that the minimal polynomial for α has degree 2: say, $x^2 + bx + c$.
  - Then $\alpha^2 + b\alpha + c = 0$, so completing the square (here is where we require the characteristic not to be equal to 2, since we must divide by 2 to do this) and setting $\beta = \alpha + b/2$ yields $(\alpha + b/2)^2 + (c - b^2/4) = 0$. Setting $\beta = \alpha + b/2$ shows that $\beta^2 = (b^2 - 4c)/4 \in F$. Furthermore, β is not in F since otherwise this would imply that $\alpha = \beta - b/2$ was in F.
  - Thus, $K = F(\beta)$ for an element β with $\beta^2 \in F$ and $\beta \notin F$, as claimed.

- Example: Determine the degree of $L = \mathbb{Q}(\sqrt{2}, \sqrt{3})$ over $\mathbb{Q}$.
  - The field $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ is a finite algebraic extension of $\mathbb{Q}$, and since $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = [\mathbb{Q}(\sqrt{3}) : \mathbb{Q}] = 2$ we see that the extension degree is at least 2 and at most 4.
  - In particular, by the remark following the corollary on finite algebraic extensions, we see that $\mathbb{Q}(\sqrt{2}, \sqrt{3}) = \mathbb{Q}[\sqrt{2}, \sqrt{3}] = \{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} : a, b, c, d \in \mathbb{Q}\}$, which establishes that the latter ring is in fact a field. (Notice how much simpler this argument is than the explicit calculations we performed earlier!)
  - In fact, since $\mathbb{Q}(\sqrt{2})$ is a subfield of $\mathbb{Q}(\sqrt{2}, \sqrt{3})$, the extension degree $[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}]$ must in fact be divisible by 2, so it is either 2 or 4. To determine which of these cases holds we need to compute $[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}(\sqrt{2})]$, which is either 1 or 2 since $[\mathbb{Q}(\sqrt{3}) : \mathbb{Q}] = 2$.
  - If $[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}(\sqrt{2})] = 1$, then the degree of the minimal polynomial of $\sqrt{3}$ over $\mathbb{Q}(\sqrt{2})$ is 1, which is to say, $\sqrt{3} \in \mathbb{Q}(\sqrt{2})$.
  - But this is not true: if $\sqrt{3} = a + b\sqrt{2}$ for $a, b \in \mathbb{Q}$ then squaring yields $3 = (a^2 + 2b^2) + 2ab\sqrt{2}$, so since $\sqrt{2}$ is irrational, one of a, b would be zero (otherwise we could write $\sqrt{2} = (3 - a^2 - 2b^2)/(2ab)$). However, we cannot have $a = \sqrt{3}$ or $b\sqrt{2} = \sqrt{3}$ because $\sqrt{3}$ and $\sqrt{6}$ are also irrational.
  - Therefore, we must have $[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}] = 4$. This tells us in particular that $[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}(\sqrt{2})] = 2$, and so for example this means that the minimal polynomial for $\sqrt{3}$ over $\mathbb{Q}(\sqrt{2})$ must have degree 2. Since $\sqrt{3}$ is a root of $x^2 - 3$, that means the polynomial $x^2 - 3$ is irreducible in $\mathbb{Q}(\sqrt{2})$.
  - Furthermore, since $\{1, \sqrt{2}, \sqrt{3}, \sqrt{6}\}$ is a spanning set for $\mathbb{Q}(\sqrt{2}, \sqrt{3})$, the fact that $[\mathbb{Q}(\sqrt{2}, \sqrt{3}) : \mathbb{Q}] = 4$ tells us that this set is a basis (and thus linearly independent), which is also not so easy to prove directly.

- Example: Determine the degree of $L = \mathbb{Q}(\sqrt{3}, \sqrt[3]{3})$ over $\mathbb{Q}$.
  - If we let $K_1 = \mathbb{Q}(\sqrt[3]{3})$ and $K_2 = \mathbb{Q}(\sqrt{3})$, then $L = K_1K_2$. Furthermore, $[K_1 : \mathbb{Q}] = 3$ and $[K_2 : \mathbb{Q}] = 2$ since K_1 is generated by a root of the irreducible polynomial $x^3 - 3$ and K_2 is generated by a root of the irreducible polynomial $x^2 - 3$.
  - Then from our result on the degree of a composite extension, we know that $[L : \mathbb{Q}] \leq [K_1 : \mathbb{Q}] \cdot [K_2 : \mathbb{Q}] = 6$.
  - Furthermore, since K_1 and K_2 are both subfields of L, we see that $[L : \mathbb{Q}]$ is divisible by both $[K_1 : \mathbb{Q}] = 2$ and $[K_2 : \mathbb{Q}] = 3$, and hence by 6. Therefore, since $[L : \mathbb{Q}] \leq 6$, the only possibility is to have $[L : \mathbb{Q}] = 6$.
  - Another approach is to observe that L contains the element $\alpha = \sqrt{3}/\sqrt[3]{3} = 3^{1/6}$. But since $\sqrt{3} = \alpha^3$ and $\sqrt[3]{3} = \alpha^2$ we conclude that $L = \mathbb{Q}(\alpha)$.
  - Then since α is a root of the (Eisenstein) irreducible polynomial $x^6 - 3$, we see that $L = \mathbb{Q}(\alpha)$ has degree 6 over $\mathbb{Q}$.

- Example: Determine the degree of $L = \mathbb{Q}(\sqrt[3]{2}, e^{2\pi i/3}\sqrt[3]{2})$ over $\mathbb{Q}$.
  - If we let $K_1 = \mathbb{Q}(\sqrt[3]{2})$ and $K_2 = \mathbb{Q}(e^{2\pi i/3}\sqrt[3]{2})$, then $L = K_1K_2$. Furthermore, from our earlier discussion of these fields, we know that $[K_1 : \mathbb{Q}] = [K_2 : \mathbb{Q}] = 3$, since both fields are generated by an element whose minimal polynomial over $\mathbb{Q}$ is $x^3 - 2$.
  - Then $[L : \mathbb{Q}] = [L : K] \cdot [K : \mathbb{Q}]$ so $[L : \mathbb{Q}]$ is divisible by 3, and we also know that $[L : \mathbb{Q}] \leq [K_1 : \mathbb{Q}] \cdot [K_2 : \mathbb{Q}] = 9$.
  - We might expect $[L : \mathbb{Q}]$ to be 9, but in fact, it is not!
  - To see this, observe that L also contains the element $\zeta = e^{2\pi i/3} = (-1 + i\sqrt{3})/2$, and one can verify that $\zeta^2 + \zeta + 1 = 0$. Thus, ζ is a root of the polynomial $x^2 + x + 1$, which is irreducible over $\mathbb{Q}$, and so for $K_3 = \mathbb{Q}(\zeta)$ we have $[K_3 : \mathbb{Q}] = 2$.
  - Since $\mathbb{Q}(\zeta)$ is also a subfield of L, we see that $[L : \mathbb{Q}]$ is divisible by 2. Since it is also divisible by 3 and ≤ 9, the only possibility is for $[L : \mathbb{Q}] = 6$.
  - In fact, it is not hard to see that $L = \mathbb{Q}(\sqrt[3]{2}, \zeta) = K_1K_3$. (With this description, it is much easier to see that $[L : \mathbb{Q}] = 6$.)
  - Remark: One way to explain why the degree of the composite is strictly less than the product of the field degrees is that the minimal polynomial of $e^{2\pi i/3}\sqrt[3]{2}$, namely $x^3 - 2$, is not irreducible over K_1. Alternatively, the basis $\{1, e^{2\pi i/3}\sqrt[3]{2}, e^{4\pi i/3}\sqrt[3]{4}\}$ for $K_2/\mathbb{Q}$ is not linearly independent over K_1, because $2 + (\sqrt[3]{4})[e^{2\pi i/3}\sqrt[3]{2}] + (\sqrt[3]{2})[e^{4\pi i/3}\sqrt[3]{4}] = 0$.

- If K/F is any field extension, we can in particular consider the collection of all elements of K that are algebraic over F.
- Proposition (Algebraic Elements): If K/F is any field extension and $\alpha, \beta \in K$ are algebraic over F, then so are $\alpha \pm \beta$, $\alpha\beta$, and α^{-1} (the latter presuming $\alpha \neq 0$). In particular, the collection of all elements of K that are algebraic over F is a subfield of K.
    - Proof: The field $F(\alpha, \beta)$ has finite degree over F when α and β are both algebraic over F, hence $F(\alpha, \beta)$ is algebraic over F. Then every element in $F(\alpha, \beta)$ is algebraic over F, including (in particular) $\alpha \pm \beta$, $\alpha\beta$, and α^{-1}.
    - The second statement follows immediately from the first one, upon applying the subfield criterion.
- We can use this observation to construct infinite algebraic extensions:
- Example: Consider the collection $\overline{\mathbb{Q}}$ of all elements of $\mathbb{C}$ that are algebraic over $\mathbb{Q}$. Show that every element of $\overline{\mathbb{Q}}$ has finite degree over $\mathbb{Q}$, but that $\overline{\mathbb{Q}}/\mathbb{Q}$ is an infinite (algebraic) extension.
    - The first statement follows immediately from our discussion of algebraic elements.
    - To show that $[\overline{\mathbb{Q}} : \mathbb{Q}] = \infty$, notice that (for any positive integer n) the element $\sqrt[n]{2}$ is contained in $\overline{\mathbb{Q}}$, hence the entire field $\mathbb{Q}(\sqrt[n]{2})$ is a subfield of $\overline{\mathbb{Q}}$.
    - Because $\sqrt[n]{2}$ has minimal polynomial $x^n - 2$ (irreducible by Eisenstein), we see that $[\mathbb{Q}(\sqrt[n]{2}) : \mathbb{Q}] = n$.
    - Therefore, $[\overline{\mathbb{Q}} : \mathbb{Q}] \geq [\mathbb{Q}(\sqrt[n]{2}) : \mathbb{Q}] = n$ for every positive integer n, so we must have $[\overline{\mathbb{Q}} : \mathbb{Q}] = \infty$.
    - We will also remark that although $\overline{\mathbb{Q}}$ is much larger than $\mathbb{Q}$, it still only has a countably infinite number of elements: every element of $\overline{\mathbb{Q}}$ is a root of a monic polynomial with rational coefficients, and there are only countably infinitely many such polynomials (and each one has a finite number of roots).
    - In particular, because $\mathbb{R}$ and $\mathbb{C}$ are both uncountable, there exist (very many!) transcendental real and complex numbers. Nonetheless, it is typically quite difficult to prove that any particular transcendental number (like e or π) is actually transcendental.

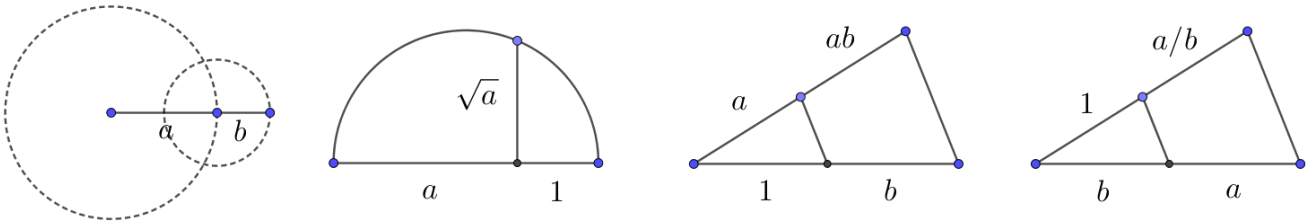
2.2.6 Classical Geometric Constructions

- One major aspect of classical Euclidean geometry, per Euclid, is concerned with describing geometric constructions using straightedge and compass.
    - Among various problems that can be solved with straightedge and compass are: bisecting (or trisecting) a segment, bisecting an angle, projecting a point onto a line, or drawing a line parallel to a given line passing through a given point.
- As an application of some of our results on degrees in field extensions, we can establish the impossibility of several classical geometric problems, originally posed by the ancient Greeks:
    - **Doubling the Cube**: Is it possible to construct, with straightedge and compass, a cube whose volume is twice that of a given cube?
    - **Trisecting an Angle**: Given an arbitrary angle, is it possible to trisect it with straightedge and compass? (In other words, to construct an angle with 1/3 the measure of the given angle.)
    - **Squaring the Circle**: Given a circle, is it possible using straightedge and compass to construct a square with the same area as the circle?
- In order to discuss these problems, we must first translate the allowed operations of straightedge-and-compass constructions into algebraic language.
    - A straightedge is an (unmarked) straight segment of arbitrary length, and may be used to draw the line between two given points.
    - A compass may be used to draw a circle with center at one given point passing through another given point.

- If two lines, a line and a circle, or two circles intersect, we may draw a new point where they intersect.

- Each of these problems begins with two given points: by translating and rescaling, we may assume the distance is 1 and the points are $(0,0)$ and $(1,0)$.
  - Any distance is determined by its length in terms of this unit distance, so we may view distances as elements of $\mathbb{R}$, and view points as elements of the Cartesian plane $\mathbb{R}^2$.
  - We say a distance d is constructible if, starting with the points $(0,0)$ and $(1,0)$, we can use some sequence of straightedge-and-compass constructions to create two points whose distance is d.
  - It is a standard Euclidean construction to project a point onto a line. Thus, if we can construct a point $(a,b) \in \mathbb{R}^2$, then we can construct both of the numbers a and b. Conversely, if we can construct both the lengths a and b, then we may construct the point (a,b) by drawing $x=a$ and $y=b$ and finding their intersection.
  - Thus, we say a point $(x,y) \in \mathbb{R}^2$ is constructible if both its coordinates are constructible lengths.
  - Any problem of constructibility then reduces to determining whether the appropriate lengths are constructible.

- Proposition (Constructible Lengths): If a and b are constructible lengths, then so are $a \pm b$, ab, a/b, and $\sqrt{a}$.
  - Proof: Each of these is a standard construction from Euclidean geometry, which we briefly illustrate:



- From the proposition we can immediately see that the set of constructible lengths (and their negatives) is a subfield of $\mathbb{R}$, and so we can construct all of $\mathbb{Q}$, and we may also take arbitrary square roots (possibly iteratively). Explicitly, suppose that all of the lengths in our constructions so far lie in the field F: we want to know what kind of field extension we may obtain by performing another construction step.
  - First, we may draw a line through two constructible points $P=(a,b)$ and $Q=(c,d)$. It is straightforward to verify that an equation for this line is $(c-a)(y-b)=(d-b)(x-a)$, which has the form $Ax+By=C$ for A,B,C rational functions in terms of a,b,c,d. Thus, if $a,b,c,d \in F$, then $A,B,C \in F$ as well.
  - Second, we may find the intersection of two lines. If the coefficients of the lines are elements of a field F, then so are the coefficients of the intersection points, since the solution to two simultaneous equations $Ax+By=C$ and $A'x+B'y=C'$ with $A,B,C,A',B',C' \in F$ will also have $x,y \in F$ by basic linear algebra.
  - We may also draw a circle with a given center and radius: the equation of such a circle has the form $(x-h)^2+(y-k)^2=r^2$ where $h,k,r \in F$. We can see, again, that all of the coefficients of the equation of the circle lie in F.
  - Third, we may find the intersection of a line and a circle. If the line has equation $Ax+By=C$ and the circle has equation $(x-h)^2+(y-k)^2=r^2$, then by solving for x or y in the equation of the line and plugging into the equation of the circle, we end up with a quadratic equation for the other variable. Thus, both x and y lie in a quadratic extension of F.
  - Fourth, we may find the intersection of two circles $(x-h)^2+(y-k)^2=r^2$ and $(x-h')^2+(y-k')^2=(r')^2$. By subtracting the two equations, we may equivalently intersect the circle $(x-h)^2+(y-k)^2=r^2$ with the line $2(h'-h)x+2(k'-k)y=r^2-(r')^2-h^2+(h')^2-k^2+(k')^2$. (This is simply the line passing through the two intersection points of the circles, presuming they do intersect.) Thus, by the previous analysis, x and y again both lie in a quadratic extension of F.
  - Since these are the only possible operations, we see that every operation either yields another element in F or an element in a quadratic extension of F.

- Proposition (Constructibility): The element $\alpha \in \mathbb{R}$ is constructible if and only if the field $\mathbb{Q}(\alpha)$ can be obtained by a sequence of quadratic extensions of $\mathbb{Q}$. In particular, if α is constructible, then $[\mathbb{Q}(\alpha) : \mathbb{Q}]$ is a power of 2.
    - Proof: From our proposition on constructible lengths, and our characterization of quadratic extensions (as being obtained via taking square roots), we can see that any α in an extension field of $\mathbb{Q}$ obtained by a sequence of quadratic is in fact constructible.
    - The converse follows from our discussion of the possible extension fields obtained by each step of the construction, since each individual field extension is either trivial or quadratic.
- Corollary: None of the three classical Greek problems (doubling the cube, trisecting an angle, and squaring the circle) can be solved using straightedge-and-compass constructions.
    - Proof: Doubling the cube is possible if and only if $\sqrt[3]{2}$ is constructible. However, as we have discussed, $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = 3$, which is not a power of 2. Thus, $\sqrt[3]{2}$ is not constructible.
    - If the angle θ can be constructed with straightedge and compass, then by orienting the angle from the positive x-axis and intersecting the corresponding ray with the unit circle, then $\cos\theta$ is constructible. Conversely, if $\cos\theta$ is constructible, then the angle θ can be obtained in the same way by intersecting the line $y = \cos\theta$ with the unit circle.
    - We will show that $\cos 20^\circ$ is not constructible. The triple angle formula for cosine states $\cos 3\theta = 4\cos^3\theta - 3\cos\theta$, so taking $\theta = 20^\circ$, and writing $\alpha = 2\cos 20^\circ$, yields $\frac{1}{2} = \alpha^3/2 - 3\alpha/2$, so that $\alpha^3 - 3\alpha - 1 = 0$.
    - By the rational root test, $x^3 - 3x - 1$ has no rational roots and is therefore irreducible (since it has degree 3). Therefore, $[\mathbb{Q}(\alpha) : \mathbb{Q}] = 3$, and so α, and thus $\alpha/2 = \cos 20^\circ$, is not constructible. From our discussion, we conclude that an angle of 20° is not constructible.
    - Finally, squaring the circle requires constructing $\sqrt{\pi}$. Since π is transcendental, π itself is not even constructible (let alone $\sqrt{\pi}$).
- Another classical constructibility question is: which n-gons are constructible?
    - Since the interior angle of a regular n-gon is $\pi - 2\pi/n$, whose cosine is $-\cos(2\pi/n)$, this question is equivalent to asking: for which n is the number $\cos(2\pi/n)$ constructible?
    - We will return to this problem later once we discuss cyclotomic extensions, but we will mention that there are standard constructions for an equilateral triangle ($n = 3$) and a regular pentagon ($n = 5$).
    - From the addition and subtraction formulas for cosine, and the half-angle formulas, we can then see that $\cos 3^\circ$ is constructible (corresponding to a 120-gon). Since $\cos 20^\circ$ is not constructible, this means that $\cos 1^\circ$ and $\cos 2^\circ$ are not constructible, so the smallest constructible integer-valued angle is 3°.
- As a final remark, all of the constructions we have described rely on an unmarked straightedge. By using different tools, it is possible to give solutions to some of these classical problems.
    - For example, if one wishes to use a ruler (a device that allows one to mark off specific lengths, while positioning the ruler arbitrarily), then there do in fact exist ruler-and-compass constructions for doubling the cube and for trisecting an arbitrary angle.
    - Alternatively, by using a formalization of the operations allowed in origami (paper folding), it can also be shown that there exist origami constructions for doubling the cube and trisecting an arbitrary angle.
    - However, a marked ruler and origami constructions can only create algebraic distances, and therefore squaring the circle is still impossible, even with these additional tools.

2.3 Splitting Fields

- We now continue investigating the connections between fields and roots of polynomials.

- We have already shown that if p is an irreducible polynomial in $F[x]$, then F has a field extension that contains a root of p: explicitly, in the extension $K = F[t]/p(t)$, the element $\bar{t} \in K$ has the property that $p(\bar{t}) = 0$.
- We may extend this observation to any polynomial p as follows: first find the factorization of p over $F[x]$, and choose any irreducible factor $q(x)$. Then construct the field extension $K = F[t]/q(t)$, and like before, observe that the element $\bar{t} \in K$ has the property that $q(\bar{t}) = 0$. Finally, since q divides p in $F[x]$, we also have $p(\bar{t}) = 0$.
- Thus, we see that if p is any polynomial in $F[x]$, then there exists a field extension of F that contains a root of p.

- We will now extend this argument to show that there is a field extension that contains "all the roots" of p, and in fact there is a well-defined notion of a "smallest" such field.

2.3.1 Splitting Fields

- Definition: If K is an extension field of F, the polynomial $p(x) \in F[x]$ splits completely (or factors completely) in $K[x]$ if there exist $c, r_1, r_2, \ldots, r_n \in K$ such that $p(x) = c(x - r_1)(x - r_2) \cdots (x - r_n)$ in $K[x]$.
  - Example: The polynomial $x^4 - 1 \in \mathbb{R}[x]$ splits completely over $\mathbb{C}$ as $(x - 1)(x + 1)(x - i)(x + i)$.
  - Example: The polynomial $x^2 - 5 \in \mathbb{Q}[x]$ splits completely over $\mathbb{C}$ as $(x - \sqrt{5})(x + \sqrt{5})$. In fact, it also splits completely with the same factorization over $\mathbb{R}$, or over $\mathbb{Q}(\sqrt{5})$.
- We would like to show that there is always some extension field of F in which $p(x)$ splits completely.
  - From the above discussion, we can construct a field extension K_1/F in which $p(x)$ has at least one root r_1. By the factor theorem, if p has degree n, then we can then write $p(x) = (x - r_1) \cdot p_1(x)$ for a polynomial $p_1(x) \in K_1[x]$ of degree $n - 1$ and some element $r_1 \in K_1$.
  - Now applying the argument to $p_1(x)$ over $K_1[x]$ shows that there exists a field extension K_2/K_1 in which $p_1(x)$ has at least one root r_2, so like before we can write $p_1(x) = (x - r_2) \cdot p_2(x)$ for a polynomial $p_2(x) \in K_2[x]$ of degree $n - 2$ and some element $r_2 \in K_2$.
  - By iterating this argument we eventually obtain a tower of field extensions $K_n/K_{n-1}/\cdots/K_1/K$, where $p_{i-1}(x) = (x - r_i)p_i(x)$ where $p_i(x) \in K_i[x]$ has degree $n - i$. Then p_n has degree 0 so it is some constant c, and so we obtain $p(x) = c(x - r_1)(x - r_2) \cdots (x - r_n)$ for some $c, r_1, r_2, \ldots, r_n \in K$.
- If $p \in F[x]$ splits completely over K, then p also splits completely over any extension field of K (as we saw above with the example of $x^2 - 5$, which splits completely over $\mathbb{Q}(\sqrt{5})$ and also in its field extensions $\mathbb{R}$ and $\mathbb{C}$).
  - It is therefore natural to ask: what is the "smallest possible" field extension of F in which p splits completely?
  - When we discussed simple extensions inside the extension K/F, we defined the field $F(\alpha)$ to be the intersection of all subfields of K containing F and α.
  - It might seem valid to define this "smallest possible" field extension of F in which p splits completely to be the intersection of all extension fields K/F in which p splits completely. But in fact, this definition only makes sense when all of these extension fields are themselves subsets of some larger field.
  - We can illustrate the difficulties with an example: consider the polynomial $p(x) = x^2 + 4 \in \mathbb{R}[x]$.
  - We can see that $p(x) = x^2 + 4$ splits completely over $\mathbb{C}$ as $p(x) = (x - 2i)(x + 2i)$, and $p(x)$ also splits completely over the field extension $\mathbb{R}[t]/(t^2 + 4)$ as $p(x) = (x - \bar{t})(x + \bar{t})$. Since both of these fields are degree-2 extensions of $\mathbb{R}$, they both seem valid candidates for the "smallest possible" field extension of $\mathbb{R}$ in which p splits completely.
  - It does not really make sense to ask what "the intersection" of $\mathbb{C}$ and $\mathbb{R}[t]/(t^2 + 4)$ is, without specifying the manner in which these two fields are to be considered as subsets of some larger collection.
  - We can avoid this particular issue (and although it seems minor, it is actually very important!) by instead posing the definition entirely within the field K itself.

- Definition: If K/F is a field extension, we say that K is a splitting field for the polynomial $p(x) \in F[x]$ if p splits completely over K, and p does not split completely over any proper subfield of K.
    - If p splits completely over K as $p(x) = c(x-r_1)(x-r_2)\cdots(x-r_n)$, then by the remainder theorem, any subfield of K in which p splits completely must contain $r_1, \ldots, r_n$, and hence contain $F(r_1, \ldots, r_n)$.
    - On the other hand, clearly $p(x)$ does split completely over $F(r_1, \ldots, r_n)$, so saying that p splits completely in K but not over any proper subfield is equivalent to saying that $K = F(r_1, r_2, \ldots, r_n)$. (In particular, the definition is well-posed.)
    - From our construction above, we can see that splitting fields always exist: simply choose an extension L/F in which $p(x)$ splits completely as $p(x) = c(x-r_1)(x-r_2)\cdots(x-r_n)$, and then take $K = F(r_1, r_2, \ldots, r_n)$.
- Example: $\mathbb{Q}(\sqrt{D})$ is a splitting field for the polynomial $p(x) = x^2 - D$ over $\mathbb{Q}$.
    - Like with the example above, this follows immediately because $p(x) = (x+\sqrt{D})(x-\sqrt{D}) \in \mathbb{Q}(\sqrt{D})[x]$ and $\mathbb{Q}(\sqrt{D}) = \mathbb{Q}(\sqrt{D}, -\sqrt{D})$.
- Example: $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ is a splitting field for the polynomial $p(x) = (x^2-2)(x^2-3)$ over $\mathbb{Q}$.
    - We can see that $p(x)$ splits completely over $K = \mathbb{Q}(\sqrt{2}, \sqrt{3})$ because $p(x) = (x-\sqrt{2})(x+\sqrt{2})(x-\sqrt{3})(x+\sqrt{3})$ in $K[x]$.
    - Furthermore, $K = \mathbb{Q}(\sqrt{2}, \sqrt{3}) = \mathbb{Q}(\sqrt{2}, -\sqrt{2}, \sqrt{3}, -\sqrt{3})$, so K is indeed a splitting field.
- We can give an upper bound on the degree of a splitting field, by formalizing the arguments we gave above:
- Proposition (Degree of Splitting Fields): For any field F and any (nonzero) polynomial $p \in F[x]$ of degree n, there exists a splitting field K/F for p, and $[K : F] \leq n!$.
    - Proof: This is simply a formalization of the discussion above. Explicitly, we use induction on n: for the base case $n = 1$ with $p(x) = ax + b = a(x + b/a)$, we have a single root $r_1 = -b/a \in F$. Thus, $K = F$ is a splitting field.
    - For the inductive step, assume that any polynomial of degree $n-1$ over any field has a splitting field extension of degree at most $(n-1)!$, and let $p \in F[x]$ have degree n.
    - Choose any irreducible factor (in $F[x]$) q of p of degree $k \leq n$ and set $K_1 = F[t]/q(t)$. Then $[K_1 : F] = k$ by our results on simple extensions.
    - Furthermore, $q(\overline{t}) = 0$ in K_1, so since q divides p we have $p(\overline{t}) = 0$ in K_1. Hence by the factor theorem we may write $p(x) = (x - \overline{t}) \cdot p_1(x)$ for a polynomial $p_1(x) \in K_1[x]$ of degree $n-1$ and some element $r_1 \in K_1$.
    - By the induction hypothesis, there exists a splitting field L for $p_1(x)$ over K_1 of degree at most $(n-1)!$. Then $[L : F] = [L : K_1] \cdot [K_1 : F] \leq (n-1)! \cdot k \leq n!$, and $p(x)$ splits completely in L, say as $p(x) = c(x-r_1)(x-r_2)\cdots(x-r_n)$.
    - The subfield $F(r_1, r_2, \ldots, r_n)$ of L is then a splitting field for p, and its degree is at most $[L : F] \leq n!$, as required.
- We have also seen that a given polynomial may (in a sense) have several "different" splitting fields.
    - To summarize, we saw that $\mathbb{C}$ is a splitting field for x^2+4 over $\mathbb{R}$, since $x^2+4 = (x+2i)(x-2i)$ in $\mathbb{C}[x]$, and $\mathbb{C} = \mathbb{R}(2i, -2i)$.
    - But the field $K = \mathbb{R}[t]/(t^2+4)$ is also a splitting field for x^2+4 over $\mathbb{R}$, since $x^2+4 = (x-\overline{t})(x+\overline{t})$ in $K[x]$.
    - The key is that these two fields are isomorphic (as fields), with an explicit isomorphism being the one that associates $\overline{t}$ with $2i$ (extended in the natural way).
    - Thus, both of these splitting fields have the same structure. This turns out to be true for arbitrary splitting fields, although it is actually easier to prove a slightly stronger result:

- Theorem (Uniqueness of Splitting Fields): Let $\varphi : E \to F$ be an isomorphism of fields with $p(x) = a_0 + a_1 x + \cdots + a_n x^n \in E[x]$, and set $q(x) = \varphi(a_0) + \varphi(a_1)x + \cdots + \varphi(a_n)x^n \in F[x]$ to be the polynomial obtained by applying φ to the coefficients of p. If K/E is a splitting field for p, and L/F is a splitting field for q, then the isomorphism φ extends to an isomorphism $\sigma : K \to L$ (i.e., $\sigma|_E = \varphi$, or explicitly, for any $\alpha \in E$ we have $\sigma(\alpha) = \varphi(\alpha)$). In particular, any two splitting fields for p are isomorphic.
    - Proof: The second statement follows from the first by taking φ to be the identity map, since in that case the first statement says that if K/E and L/E are both splitting fields of p, then K and L are isomorphic.
    - To prove the first statement, we induct on the degree n of p. For the base case $n = 1$, as we have already observed, the splitting field of any degree-1 polynomial over a field is simply the field itself. Thus, $K = E$ and $L = F$, so the desired map σ is simply φ.
    - Now suppose the result holds for polynomials of degree $n-1$, and let p have degree n. Choose any monic irreducible factor $a(x) = c_0 + c_1 x + \cdots + c_m x^m$ of p, and set $b(x) = \varphi(c_0) + \varphi(c_1)x + \cdots + \varphi(c_m)x^m$. It is then straightforward to verify that $b(x)$ divides $q(x)$ and that $b(x)$ is also irreducible in $F[x]$ (one can simply do these calculations explicitly, or show first that φ extends to an isomorphism of $E[x]$ with $F[x]$, so that it preserves factorizations and hence irreducibility).
    - Since every root of $a(x)$ is a root of $p(x)$ we see that K contains every root of a, and similarly L contains every root of b.
    - Choose any root r of $a(x)$ and any root s of $b(x)$, and consider the map $\tilde{\varphi} : F(r) \to E(s)$ defined by mapping the polynomial $d_0 + d_1 r + \cdots + d_{m-1} r^{m-1} \mapsto \varphi(d_0) + \varphi(d_1)s + \cdots + \varphi(d_{m-1})s^{m-1}$ for $d_i \in F$. This map is actually well-defined on all of $F(r)$ because r is algebraic over F with minimal polynomial $a(x)$ of degree m, so $\{1, r, \ldots, r^{m-1}\}$ is an F-basis of $F(r)$.
    - It is a straightforward (though tedious) calculation to check that $\tilde{\varphi}$ is a ring (hence field) isomorphism.
    - By the factor theorem, since r is a root of p and s is a root of q, we may write $p(x) = (x - r)p'(x)$ and $q(x) = (x - s)q'(x)$ for some polynomials $p' = c(x - r_2)\cdots(x - r_n)$ and $q' = d(x - s_2)\cdots(x - s_n)$ of degree $n - 1$. In particular, p' splits completely over K and q' splits completely over L.
    - Since K is the splitting field of p, we see that $K = F(r, r_2, \ldots, r_n) = F(r)(r_2, \ldots, r_n)$, and so in fact K is the splitting field of p' over $F(r)$. Likewise, L is the splitting field of q' over $E(s)$.
    - Finally, by the induction hypothesis, since we have an isomorphism $\tilde{\varphi} : F(r) \to E(s)$, we may lift it to obtain an isomorphism $\sigma : K \to L$, as required.
- Because splitting fields are unique up to isomorphism, we will refer to "the" splitting field of $p(x)$ over F.

2.3.2 Examples of Splitting Fields

- In general, it can be quite difficult to compute an explicit description of a splitting field, because it requires knowing information about the factorization and the precise nature of the roots of $p(x)$, along with any sort of algebraic relations among the roots.
    - As such, for the moment we will primarily focus on finding splitting fields over $\mathbb{Q}$, since we have irreducibility criteria that can apply to polynomials of arbitrarily large degree in $\mathbb{Q}[x]$.
- Example: Find the splitting field for $p(x) = x^2 + 1$ over $\mathbb{Q}$, over $\mathbb{F}_2$, and over $\mathbb{F}_3$.
    - Over $\mathbb{Q}$, we can see that $\mathbb{Q}(i)$ is a splitting field because $p(x) = (x+i)(x-i) \in \mathbb{Q}[x]$ and $\mathbb{Q}(i) = \mathbb{Q}(i, -i)$.
    - Over $\mathbb{F}_2$, the field $\mathbb{F}_2$ itself is actually already a splitting field because $p(x) = (x+1)^2 \in \mathbb{F}_2[x]$.
    - Over $\mathbb{F}_3$, the polynomial is irreducible (since it has degree 2 and no roots in $\mathbb{F}_3$), so any splitting field must be of degree at least 2 over $\mathbb{F}_3$. On the other hand, in the degree-2 field extension $K = \mathbb{F}_3[t]/p(t)$, we can factor $p(x)$ as $p(x) = (x - \bar{t})(x + \bar{t})$, and $K = \mathbb{F}_3(\bar{t}, -\bar{t})$, so we see that K is a splitting field for p.
    - Remark: More generally, for any quadratic polynomial $p(x) \in F[x]$, one can show that if p has a root in F, then both its roots are in F (because their sum is an element of F), so F itself is a splitting field. Otherwise, if p is irreducible, then p does not split completely over F, but does split completely over the quadratic extension $F[t]/p(t)$: thus, $F[t]/p(t)$ will be a splitting field.

- Example: Show that $\mathbb{Q}(\sqrt[3]{2}, \zeta_3)$ is the splitting field for the polynomial $p(x) = x^3 - 2$ over $\mathbb{Q}$, where $\zeta_3 = e^{2\pi i/3} = (-1 + i\sqrt{3})/2$ denotes a nonreal cube root of unity.
  - As we have mentioned previously, this field $K = \mathbb{Q}(\sqrt[3]{2}, \zeta_3)$ is also equal to $\mathbb{Q}(\sqrt[3]{2}, \zeta_3\sqrt[3]{2})$, and has degree 6 over $\mathbb{Q}$.
  - We can see that $p(x)$ splits completely over K because $p(x) = (x - \sqrt[3]{2})(x - \zeta_3\sqrt[3]{2})(x - \zeta_3^2\sqrt[3]{2})$ in $K[x]$. (One may compute the third root of $p(x)$ using polynomial division once the roots $\sqrt[3]{2}$ and $\zeta_3\sqrt[3]{2}$ are identified, or by directly observing that $\zeta_3^2\sqrt[3]{2}$ is also a root.)
  - Thus, we see that $L = \mathbb{Q}(\sqrt[3]{2}, \zeta_3\sqrt[3]{2}, \zeta_3^2\sqrt[3]{2})$ is a splitting field for $p(x)$ over $\mathbb{Q}$.
  - Notice that L contains both generators $\sqrt[3]{2}$ and $\zeta_3 = (\zeta_3\sqrt[3]{2})/(\sqrt[3]{2})$ of $K/\mathbb{Q}$, so L contains K. On the other hand, K contains all three generators $\sqrt[3]{2}$, $\zeta_3\sqrt[3]{2}$, and $\zeta_3^2\sqrt[3]{2}$ of $L/\mathbb{Q}$, so K contains L. Thus, $K = L$ is a splitting field for $p(x)$ as claimed.

- Example: Find the splitting field for $p(x) = x^4 + 64$ over $\mathbb{Q}$.
  - As it happens, this polynomial factors over $\mathbb{Q}$ as $p(x) = (x^2 - 4x + 8)(x^2 + 4x + 8)$, and using the quadratic formula we can see that the roots of these two quadratics are $\pm 2 \pm 2i$.
  - Therefore, we can see that $\mathbb{Q}(i)$ is a splitting field for p, since it is the subfield of $\mathbb{C}$ generated by the roots of p. Notice in particular that, although $p(x)$ has degree 4, the degree of the splitting field is only 2.

- Example: If n is a positive integer, show that the splitting field of the polynomial $x^n - 1$ over $\mathbb{Q}$ is of the form $\mathbb{Q}(\zeta_n)$ where ζ_n is the complex number $\zeta_n = e^{2\pi i/n} = \cos(2\pi/n) + i\sin(2\pi/n)$.
  - To see this, first observe that $\zeta_n = e^{2\pi i/n}$ has the property that $\zeta_n^n = e^{2\pi i} = 1$, and so ζ_n is a root of $q(x)$ over $\mathbb{C}$.
  - Furthermore, for each integer k with $0 \le k \le n-1$, we see that $\zeta_n^k = e^{2\pi ik/n} = \cos(2\pi k/n) + i\sin(2\pi k/n)$ also has the property that $(\zeta_n^k)^n = 1^k = 1$ and so ζ_n^k is also a root of $q(x)$ over $\mathbb{C}$.
  - The n complex numbers ζ_n^k for $0 \le k \le n-1$ are distinct as elements of $\mathbb{C}$ (geometrically, they represent n equally spaced points around the unit circle $|z| = 1$ in the complex plane).
  - Thus, by the factor theorem we obtain the factorization $q(x) = (x-1)(x-\zeta_n)(x-\zeta_n^2)\cdots(x-\zeta_n^{n-1})$, and so the splitting field for $q(x)$ over $\mathbb{Q}$ is $\mathbb{Q}(1, \zeta_n, \zeta_n^2, \ldots, \zeta_n^{n-1})$. This field clearly contains $\mathbb{Q}(\zeta_n)$, but since $\mathbb{Q}(\zeta_n)$ contains each of the generators $1, \zeta_n, \ldots, \zeta_n^{n-1}$, it is equal to $\mathbb{Q}(\zeta_n)$ as claimed.

- Definition: The splitting field $\mathbb{Q}(\zeta_n)$ arising in the example above is called the cyclotomic field of nth roots of unity.
  - It is a nontrivial problem (and one to which we will return later) to compute the degree $[\mathbb{Q}(\zeta_n) : \mathbb{Q}]$, which is equivalent to determining the degree of the minimal polynomial of ζ_n over $\mathbb{Q}$.
  - In the case where $n = p$ is a prime, however, we can compute it now:

- Proposition (Prime Cyclotomic Fields): If p is a prime, the degree $[\mathbb{Q}(\zeta_p) : \mathbb{Q}]$ is equal to $p - 1$.
  - Proof: As noted above, the degree $[\mathbb{Q}(\zeta_p) : \mathbb{Q}]$ is equal to the degree of the minimal polynomial of ζ_p over $\mathbb{Q}$.
  - Since $\zeta_p \neq 1$, and since $x - 1$ divides $x^p - 1$, by the factor theorem we see that ζ_p is a root of the polynomial $\Phi_p(x) = \dfrac{x^p - 1}{x - 1} = x^{p-1} + x^{p-2} + \cdots + x + 1$.
  - We claim that $\Phi_p(x)$ is irreducible over $\mathbb{Q}$, and is therefore the minimal polynomial of ζ_p.
  - To show this, observe that $\Phi_p(x)$ is irreducible if and only if $\Phi_p(x+1)$ is irreducible (since any factorization $\Phi_p(x+1) = a(x)b(x)$ would yield a factorization $\Phi_p(x) = a(x-1)b(x-1)$ and vice versa).
  - We can compute $\Phi_p(x+1) = \dfrac{(x+1)^p - 1}{(x+1) - 1} = \dfrac{1}{x} \cdot \displaystyle\sum_{k=1}^{p} \binom{p}{k} x^k = \sum_{k=1}^{p} \binom{p}{k} x^{k-1} = x^{p-1} + px^{p-2} + \cdots + p$.

- Each of the binomial coefficients $\binom{p}{k} = \frac{p!}{k!(p-k)!}$ with $0 < k < p$ is divisible by p (since there is a p in the numerator but not the denominator) and the constant term of $\Phi_p(x+1)$ is not divisible by p^2.
- Thus, $\Phi_p(x+1)$ is irreducible over $\mathbb{Q}$ by Eisenstein's criterion (with prime p), and so $\Phi_p(x)$ is also irreducible over $\mathbb{Q}$.
- Therefore, $\Phi_p(x)$ is the minimal polynomial of ζ_p, and $[\mathbb{Q}(\zeta_p) : \mathbb{Q}] = \deg \Phi_p = p - 1$.

- By using some of our results about cyclotomic fields, we can compute certain other splitting fields along with their degrees:

- Example: If p is a prime, find the splitting field K for $q(x) = x^p - 3$ over $\mathbb{Q}$ and compute the degree $[K : \mathbb{Q}]$.
  - We begin by observing that the roots of $q(x)$ in $\mathbb{C}$ are $\zeta_p^k \cdot \sqrt[p]{3}$ for $0 \leq k \leq p-1$, since each of these is a root of q and they are all distinct. (Here, as elsewhere, $\sqrt[p]{3}$ represents the real pth root of 3.)
  - Therefore, the splitting field for p over $\mathbb{Q}$ is $K = \mathbb{Q}(\sqrt[p]{3}, \zeta_p \sqrt[p]{3}, \ldots, \zeta_p^{p-1} \sqrt[p]{3}) = \mathbb{Q}(\sqrt[p]{3}, \zeta_p)$, since both fields contains the generators for the other.
  - Notice that K is the composite of the fields $E = \mathbb{Q}(\sqrt[p]{3})$ and $F = \mathbb{Q}(\zeta_p)$, and so $[K : \mathbb{Q}] \leq [E : \mathbb{Q}] \cdot [F : \mathbb{Q}]$.
  - We showed that $[F : \mathbb{Q}] = p-1$ above, and for $[E : \mathbb{Q}]$, because $x^p - 3$ is irreducible over $\mathbb{Q}$ by Eisenstein's criterion (with prime 3), $x^p - 3$ is necessarily the minimal polynomial of $\sqrt[p]{3}$, so $[\mathbb{Q}(\sqrt[p]{3}) : \mathbb{Q}] = p$.
  - Therefore, $[K : \mathbb{Q}] \leq p(p-1)$.
  - However, since E and F are both subfields of K, $[K : \mathbb{Q}]$ is divisible by both $[E : \mathbb{Q}] = p$ and $[F : \mathbb{Q}] = p - 1$, and thus (since they are relatively prime) by their product. We must therefore have equality, meaning that $[K : \mathbb{Q}] = p(p-1)$.
  - Remark: Because K/E has degree $p-1$ and is generated by ζ_p (which is a root of the degree-$(p-1)$ polynomial $\Phi_p(x) \in \mathbb{Q}[x]$), we obtain the nontrivial fact that $\Phi_p(x)$ is irreducible over $\mathbb{Q}(\sqrt[p]{3})$. By the same reasoning, we can also deduce that $x^p - 3$ is irreducible over $\mathbb{Q}(\zeta_p)$.

- Example: Find the splitting field K for $p(x) = x^8 - 2$ over $\mathbb{Q}$ and compute the degree $[K : \mathbb{Q}]$.
  - As in the previous example, we can see that the roots of $p(x)$ in $\mathbb{C}$ are $\zeta_8^k \cdot \sqrt[8]{2}$ for $0 \leq k \leq 7$.
  - Therefore, the splitting field for p over $\mathbb{Q}$ is $K = \mathbb{Q}(\sqrt[8]{2}, \zeta_8 \sqrt[8]{2}, \ldots, \zeta_8^7 \sqrt[8]{2}) = \mathbb{Q}(\sqrt[8]{2}, \zeta_8)$, since both fields contains the generators for the other.
  - We can compute $\zeta_8 = \cos(2\pi/8) + i\sin(2\pi/8) = \sqrt{2}/2 + i\sqrt{2}/2$, and so since $\sqrt{2} = (\sqrt[8]{2})^4$, we see that K contains $\sqrt{2} \cdot \zeta_8 - 1 = i$.
  - Then, since K contains i and $\sqrt[8]{2}$, and because $\mathbb{Q}(\sqrt[8]{2}, i)$ contains the generators of K, we in fact have $K = \mathbb{Q}(\sqrt[8]{2}, i)$.
  - By the multiplicativity of field degrees, $[K : \mathbb{Q}] = [K : \mathbb{Q}(\sqrt[8]{2})] \cdot [\mathbb{Q}(\sqrt[8]{2}) : \mathbb{Q}]$.
  - Because $x^8 - 2$ is irreducible over $\mathbb{Q}$, it is necessarily the minimal polynomial of $\sqrt[8]{2}$, and so $[\mathbb{Q}(\sqrt[8]{2}) : \mathbb{Q}] = 8$.
  - To compute $[K : \mathbb{Q}(\sqrt[8]{2})]$, notice that $\mathbb{Q}(\sqrt[8]{2})$ is a subfield of $\mathbb{R}$ but K is not, since it contains the nonreal number i. On the other hand, since $K/\mathbb{Q}(\sqrt[8]{2})$ is generated by i, the extension degree is at most the degree of the minimal polynomial of i over $\mathbb{Q}$, which is 2.
  - Thus, the only possibility is to have $[K : \mathbb{Q}(\sqrt[8]{2})] = 2$, and so $[K : \mathbb{Q}] = 16$.

2.3.3 Algebraic Closures

- As we have shown, for any polynomial $p \in F[x]$, there exists a field extension K/F with the property that K contains all of the roots of p.
  - A natural extension of this question is: does there exist a field extension K/F with the property that K contains all of the roots of *every* polynomial $p \in F[x]$?
  - One example of such an extension is $\mathbb{C}/\mathbb{R}$, since every polynomial in $\mathbb{R}[x]$ splits completely over $\mathbb{C}$. (This statement is equivalent to the fundamental theorem of algebra.)

- Given an arbitrary field F, we would like to construct an analogous extension that contains all of the roots of all polynomials in $F[x]$: this extension represents the closure of F under algebraic operations (i.e., solving polynomials) and is called the algebraic closure of F.

- Definition: If F is a field, the field $\overline{F}$ is an algebraic closure of F if $\overline{F}$ is algebraic over F and every polynomial in $F[x]$ splits completely over $\overline{F}$.
  - Initially, it seems natural that such an extension would exist (since we may construct towers of extensions having roots of more and more polynomials of larger and larger degrees), but in fact this question is substantially more delicate than it might seem.
  - Intuitively, we would like to think of the algebraic closure of F as the composite of all of the splitting fields of the polynomials in $F[x]$.
  - However, the composite of two arbitrary fields is not defined: we have only defined the composite of two subfields of a larger field.
  - Thus, saying that the algebraic closure is "the composite of all of the splitting fields" presupposes the existence of some larger field that contains all of these splitting fields, and this is entirely circular since this larger field is precisely what the algebraic closure would be!

- Let us instead examine another feature of $\mathbb{C}$: not only is it the algebraic closure of $\mathbb{R}$, it is the algebraic closure of itself.
  - This follows by the observation that every polynomial in $\mathbb{C}[x]$ splits completely over $\mathbb{C}$ (which is, again, simply the fundamental theorem of algebra).
  - This tells us that $\mathbb{C}$ has no nontrivial algebraic extensions: if $L/\mathbb{C}$ were an algebraic extension, any element $\alpha \in L$ would be a root of its minimal polynomial in $\mathbb{C}[x]$, but the only irreducible polynomials in $\mathbb{C}[x]$ are linear polynomials.
  - In other words, $\mathbb{C}$ is "algebraically closed". To be precise:

- Definition: The field F is algebraically closed if every polynomial in $F[x]$ has a root in F.
  - By the factor theorem and a trivial induction, if every polynomial in $F[x]$ has a root in F, then in fact it must split completely over F.
  - Equivalently, by the same logic as given above for $\mathbb{C}$, a field is algebraically closed whenever it has no nontrivial algebraic extensions.
  - Based on the similarity of the names, and the fact that $\mathbb{C}$ is both an algebraic closure (namely, of $\mathbb{R}$) and is itself algebraically closed, it is reasonable to guess that algebraic closures are algebraically closed. This is in fact true:

- Proposition (Algebraic Closures are Algebraically Closed): If F is any field, then any algebraic closure $\overline{F}$ is algebraically closed. Symbolically, $\overline{\overline{F}} = \overline{F}$.
  - Proof: Suppose that $p(x) \in \overline{F}[x]$ is a polynomial and α is any root of $p(x)$ in $\overline{\overline{F}}$. Then $\overline{F}(\alpha)$ is an algebraic extension of $\overline{F}$, and $\overline{F}$ is an algebraic extension of F.
  - We have previously shown that an algebraic extension of an algebraic extension is algebraic, so applying it to $\overline{F}(\alpha)/\overline{F}$ and $\overline{F}/F$ shows that $\overline{F}(\alpha)/F$ is algebraic, which is to say, α is algebraic over F.
  - But since $\overline{F}$ contains all elements algebraic over F, we see $\alpha \in \overline{F}$, so $\overline{\overline{F}} = \overline{F}$.

- One approach to showing that every field has an algebraic closure is to show that every field F is a subfield of an algebraically closed field L: if we can do this, then the subfield of L consisting of all elements algebraic over F is an algebraic closure of F. (Recall that we showed previously that the collection of all algebraic elements is a subfield.)

- Theorem (Algebraic Closures): If F is a field, then F is a subfield of an algebraically closed field.
  - The proof of this theorem requires invoking Zorn's lemma (equivalent to the axiom of choice) and technically also uses colimits, so we will instead just sketch the argument.

- First observe that in any commutative ring R with 1, if I is any ideal of R, then there exists a proper ideal M (i.e., with $M \neq R$) containing I that is maximal under containment (i.e., so that there is no other ideal J with $M \subsetneq J \subsetneq R$). This fact is a straightforward consequence of Zorn's lemma[4].
- Furthermore, if M is a maximal ideal of R, then R/M is a field, because if $\overline{r}$ were any nonzero nonunit in R/M, then the ideal of R generated by r and M would be strictly larger than M but still a proper ideal of R, contradicting maximality of M.
- Next, take R to be a polynomial ring in infinitely many variables X_f, indexed by the polynomials $f(x) \in F[x]$ of positive degree. (The elements of R are the polynomials involving finitely many of the X_f, with coefficients from F.)
- Let I be the smallest ideal of R containing all of the elements $f_i(X_{f_i})$, for each polynomial $f_i \in F[x]$ of positive degree.
- Then I is a proper ideal of R, because if not, 1 would be an element of I and so there would exist a relation of the form $r_1 f_1(X_{f_1}) + r_2 f_2(X_{f_2}) + \cdots + r_n f_n(X_{f_n}) = 1$ for some irreducible $f_i \in F[x]$ of positive degree and some elements $r_i \in R$. If we take K to be the splitting field of $f_1 f_2 \cdots f_n$ and choose a root $\alpha_i \in K$ of each f_i, then evaluating both sides of this relation at $X_{f_1} = \alpha_1, \ldots, X_{f_n} = \alpha_n$ yields $0 = 1$, which is impossible.
- Thus, I is a proper ideal of R so it is contained in some maximal ideal M. The quotient ring $L = R/M$ is then a field that is an extension of F (since F embeds in L as the images of the constant polynomials). Every polynomial $f(x) \in F[x]$ of positive degree then has a root in L since $f(\overline{X_f}) = \overline{0}$ in the quotient ring (this is because $f(X_f) \in M$ since it is in I).
- Unfortunately this is not quite enough to say that L is algebraically closed, because although every polynomial in $F[x]$ now has a root, there may exist polynomials in $L[x]$ that have no roots.
- To deal with this issue, we iterate the construction to obtain an infinite sequence of fields $F \subseteq L_1 \subseteq L_2 \subseteq L_3 \subseteq \cdots$, where every polynomial in $L_i[x]$ has at least one root in L_{i+1}.
- We may then take the union of this infinite sequence of fields (technically, we actually take a colimit) to obtain a field $\overline{F}$. Each element of this field is contained in some L_i: thus, any polynomial with coefficients from $\overline{F}$ has all its coefficients from some L_i, and this polynomial has a root in L_{i+1} (hence in $\overline{F}$). Thus, $\overline{F}$ is algebraically closed.

- Corollary: If F is a field, then there exists an algebraic closure $\overline{F}$ of F. Furthermore, the algebraic closure $\overline{F}$ is unique up to isomorphism.
  - Proof: By the previous theorem, F is a subfield of an algebraically closed field L. Then the collection of all elements of L that are algebraic over F is a subfield of L, and is an algebraic closure of F.
  - For the uniqueness, one may use an argument similar to the one we used to establish that splitting fields are unique up to isomorphism.
  - More explicitly, by a similar argument as used for splitting fields (along with an invocation of Zorn's lemma), one may show that if K/F is algebraic and L/K is also algebraic, then there exists an embedding of K into $\overline{F}$, and an embedding extending this one that embeds L into $\overline{F}$. (By "an embedding of E into F" we mean a map that is an isomorphism of E with a subfield of F.)
  - Now suppose that E_1 and E_2 are both algebraic closures of F: by applying the above result, we obtain an embedding of E_1 into E_2, and so E_1 is isomorphic to a subfield of E_2. But then E_2 is an algebraic extension of (a field isomorphic to) E_1, but E_1 has no nontrivial algebraic extensions: thus, the embedding of E_1 into E_2 is actually an isomorphism.

- Since $\mathbb{C}$ is algebraically closed by the fundamental theorem of algebra, by the argument above it contains an algebraic closure of any of its subfields.
  - In particular, this means that we can always view any question about algebraic extensions of $\mathbb{Q}$ as taking place inside of $\mathbb{C}$ (as, in fact, we have already implicitly been doing).
  - Furthermore, we also see that the set $\overline{\mathbb{Q}}$ of elements of $\mathbb{C}$ that are algebraic over $\mathbb{Q}$ is an algebraically closed field.

[4]Specifically, if we let $\mathcal{F}$ be the set of all proper ideals of R containing I, then $\mathcal{F}$ is nonempty since $I \in \mathcal{F}$. Also, for any chain $\mathcal{C}$ with indexing set J, the union $\cup_j C_j$ is an upper bound for C: this follows by noting that the union of a chain of ideals is also an ideal (since it contains 0 and is closed under subtraction and R-multiplication) and that it is proper since it cannot contain 1, as otherwise some C_j would contain 1 hence equal R, contradicting the assumption that each C_j is proper. Hence every chain has an upper bound, so by Zorn's lemma, there exists a maximal element of $\mathcal{F}$, as claimed.

2.4 Separability and Transcendence

- In this section we discuss two additional topics related to the structure of general field extensions: separability and transcendence.
    - These topics arise frequently in the context of number theory and algebraic geometry.

2.4.1 Separable and Inseparable Polynomials

- As we have shown, for any field F and any polynomial $p \in F[x]$, there exists an extension field K/F that contains all the roots of p.
    - In many cases, the roots of a polynomial will be distinct. However, there certainly exist cases in which polynomials have "repeated roots", such as $p(x) = x^3$ or $p(x) = x^2(x-1)^2$.
    - None of these polynomials is irreducible, and it is difficult (and as we will explain, with good reason!) to find examples of irreducible polynomials with repeated roots.
- Definition: If F is a field with $q \in F[x]$, and the factorization of $q(x) = c(x-r_1)^{d_1}(x-r_2)^{d_2}\cdots(x-r_k)^{d_k}$ with the $d_i \geq 1$, we say that d_i is the multiplicity of r_i. Furthermore, r_i is a simple root if $d_i = 1$, and is a repeated root (or multiple root) if $d_i \geq 2$. If all of the roots of q are simple, then we say q is separable, and otherwise q is inseparable.
    - Example: The polynomial $x^2(x-1)^2(x^2+1)$ has two repeated roots (0 and 1) and two simple roots (i and $-i$) over $\mathbb{Q}$, and is inseparable.
    - Example: The polynomial $x^3 + 4x$ has three simple roots (0, $2i$, and $-2i$) over $\mathbb{Q}$, and is separable.
    - Example: Over $F = \mathbb{F}_2(t)$, the field of rational functions in t with coefficients in $\mathbb{F}_2$, the polynomial $q(x) = x^2 - t$ is irreducible (it has no roots in F since there is no rational function whose square is t). Nonetheless, q has a repeated root $t^{1/2}$, because in $\overline{F}$ the polynomial $q(x)$ factors as $q(x) = (x - t^{1/2})^2$, and so q is inseparable.
- As a first goal, we can give a necessary condition for when a polynomial has repeated roots.
    - Recall from calculus that we can test whether a polynomial has a "double root" at r by testing whether $q(r) = q'(r) = 0$. By the factor theorem, this is equivalent to saying that q and q' are both divisible by $x - r$.
    - We can formulate a similar test over any field, since we may give a purely algebraic definition of the derivative:
- Definition: If $q(x) = \sum_{k=0}^{n} a_k x^k$ is a polynomial in $F[x]$, its derivative is the polynomial $q'(x) = \sum_{k=0}^{n} k a_k x^{k-1}$.
    - It is a straightforward calculation to verify that the standard differentiation rules apply: $(f+g)'(x) = f'(x) + g'(x)$ and $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$. (For the product rule, the easiest method is to check it for products of monomials and then apply the distributive law, since both sides are additive.)
    - Example: In $\mathbb{C}[x]$, the derivative of $x^6 - 4x^2 + x$ is $6x^5 - 8x + 1$.
    - Example: In $\mathbb{F}_p[x]$, the derivative of $x^{p^2} - x$ is $p^2 x^{p^2-1} - 1 = -1$. Notice here that although the degree of the original polynomial is p^2, the degree of its derivative is 0.
- Proposition (Derivatives and Separability): Let F be a field and $q \in F[x]$. Then r is a repeated root of q (in a splitting field) if and only if $q(r) = q'(r) = 0$. Furthermore, the polynomial $q(x)$ is separable if and only if $q(x)$ and $q'(x)$ are relatively prime in $F[x]$.
    - Proof: First suppose that $q(x)$ has a repeated root r in some extension K/F: then $q(x) = (x-r)^2 s(x)$ for some $s(x) \in K[x]$.
    - Taking the derivative yields $q'(x) = 2(x-r)s(x) + (x-r)^2 s'(x) = (x-r)\cdot[2s(x) + (x-r)s'(x)]$. Thus, q' is also divisible by $x - r$ in $K[x]$. By the factor theorem, we conclude that $q(r) = q'(r) = 0$.

- Conversely, if $q(r) = q'(r) = 0$, then by the factor theorem $x - r$ divides $q(x)$, so we may write $q(x) = (x-r)a(x)$. Then $q'(x) = a(x) + (x-r)a'(x)$, so $q'(r) = a(r)$. Thus $a(r) = 0$ and so $x - r$ divides $a(x)$: then $q(x)$ is divisible by $(x-r)^2$ so r is a repeated root.
- For the second statement, any root of a common factor of q and q' is a multiple root (by the above) and conversely any repeated root of q yields a nontrivial common factor of q and q' in $F[x]$ (namely, the minimal polynomial of the repeated root).

- In characteristic 0, this result implies that every irreducible polynomial is separable:
- Corollary (Separability in Characteristic 0): If F is a field of characteristic 0 and $q(x) \in F[x]$ is irreducible, then $q(x)$ is separable.
  - Proof: From the result above, we know that q is separable if and only if q and q' have a common factor in $F[x]$. Since q is irreducible in $F[x]$, up to associates the only possible common factors are q and 1.
  - In characteristic 0, if q has degree n then q' has degree $n-1$, so q cannot divide q'. Thus, the only possibility is for q and q' to be relatively prime, meaning that q is separable.
- In positive characteristic, as we have already noted, there can exist inseparable irreducible polynomials.
  - As we noted earlier, over $F = \mathbb{F}_2(t)$, the polynomial $q(x) = x^2 - t$ is irreducible and also inseparable, because it has a repeated root $t^{1/2}$ that is not in F. Note in this case that $q'(x) = 2x = 0$ is identically zero, so indeed q and q' have a common divisor of positive degree (namely, q itself).
  - Indeed, by degree considerations, the case where q' is the zero polynomial is the only case in which we can have an inseparable irreducible polynomial, since if $q' \neq 0$ then since $\deg q > \deg q'$, it is not possible for q to divide q'.
  - From the definition of derivative, we can see that if $q(x) = \sum_{k=0}^{n} a_n x^n$ then $q'(x) = \sum_{k=0}^{n} n a_n x^{n-1}$ is zero if and only if $na_n = 0$ for each n, and this is true precisely when the only nonzero coefficients of q are in degrees that are divisible by p.
  - Equivalently, this means that $q(x) = s(x^p)$ for some polynomial $s \in F[x]$.
  - Thus, there is an inseparable irreducible polynomial over F precisely when there is a polynomial $s \in F[x]$ with the property that $s(x^p)$ is irreducible.

- To examine this property in more detail requires a (very useful!) result on field arithmetic in characteristic p:
- Proposition ("Freshman" Binomial Theorem): If the field F has characteristic $p > 0$, then $(a+b)^p = a^p + b^p$ for any $a, b \in F$.
  - Proof: By an induction argument, the binomial theorem holds in any field: thus, $(a+b)^p = \sum_{n=0}^{p} \binom{p}{n} a^n b^{p-n}$.
  - For each $0 < n < p$, the binomial coefficient $\binom{p}{n} = \frac{p!}{n!(p-n)!}$ is an integer divisible by p (since there is a p in the numerator but not in the denominator, and p is prime), so $\binom{p}{n} = 0$ in the field F.
  - Therefore, all terms in the sum except those for $n = 0$ and $n = p$ are zero, whence $(a+b)^p = a^p + b^p$ for any $a, b \in F$ as claimed.

- Definition: If F is a field of characteristic p, the Frobenius endomorphism is the map $\varphi : F \to F$ defined by $\varphi(a) = a^p$.
  - It is easy to see that φ respects multiplication (i.e., $\varphi(ab) = \varphi(a)\varphi(b)$) and by the proposition above it also respects addition.
  - Furthermore, φ is injective, because $\varphi(a) = \varphi(b)$ implies $a^p = b^p$ so that $(a-b)^p = 0$ (since φ distributes over addition, and $(-1)^p = -1$ in $\mathbb{F}_p$ for any prime p). Since F is a field, this is equivalent to saying $a - b = 0$.
  - By iterating the additivity of φ, we can see that $(a_0 + a_1 x + \cdots + a_n x^n)^p = a_0^p + a_1^p x^p + \cdots + a_n^p x^{np}$.

- Applying this in reverse, we can see that if all of the coefficients of $s(x)$ are pth powers in F, then $s(x^p)$ is a pth power, and therefore cannot be irreducible. We give fields with this property a name:

- Definition: If F is a field of characteristic p, and every element of F is a pth power (i.e., $F^p = F$) then we say F is a perfect field. (Fields of characteristic 0 are also considered perfect fields.)
  - Equivalently, F is a perfect field if the Frobenius endomorphism is surjective. But since it respects addition and multiplication and is injective, this is the same as saying that the Frobenius map is an isomorphism of F with itself.
  - Example: If F is a finite field, then F is perfect. This follows by the fact that the Frobenius map is an injective map from a finite set to itself, hence is also surjective.

- Proposition (Separability and Perfect Fields): If F is a perfect field, then every irreducible polynomial in $F[x]$ is separable. Inversely, if F is not perfect, then there exists an irreducible inseparable polynomial in $F[x]$.
  - Proof: As we showed above, every field of characteristic 0 is separable, so now assume F has characteristic p.
  - If $q(x)$ is an irreducible inseparable polynomial in $F[x]$, then as we have already discussed, q' must be the zero polynomial, so $q(x) = s(x^p)$ for some polynomial $s \in F[x]$.
  - If F is perfect, then every coefficient of s is a pth power, so we may write $s(x) = a_0^p + a_1^p x + \cdots + a_n^p x^n$. But then $q(x) = s(x^p) = a_0^p + a_1^p x^p + \cdots + a_n^p x^{np} = (a_0 + a_1 x + \cdots + a_n x^n)^p$ is not irreducible, which is a contradiction.
  - For the inverse statement, suppose F is not perfect: then there exists some element $\alpha \in F$ that is not a pth power in F.
  - Now consider the polynomial $q(x) = x^p - \alpha$: if we set $\beta = \alpha^{1/p}$ (inside a splitting field for q) then in $F(\beta)$ we may write $q(x) = x^p - \beta^p = (x - \beta)^p$ so q is inseparable.
  - We claim that q is also irreducible in $F[x]$. To see this, suppose it had a factorization $q(x) = c(x)d(x)$ in $F[x]$: then from the factorization above, up to constant factors in F we must have $c(x) = (x - \beta)^d$ for some $0 < d < p$.
  - Expanding out $c(x)$ shows that $c(x) = x^d - d\beta x^{d-1} + \cdots + (-1)^d \beta^d$. In particular, we see that the coefficient $d\beta$ is in F, and because $d \neq 0$ in F (since $0 < d < p$) we would have $\beta \in F$. But this contradicts the assumption that α is not a pth power in F. Thus, q is an irreducible inseparable polynomial over the non-perfect field F.

- As an application of our results, we can show that there exists a finite field with p^n elements, and that it is unique up to isomorphism:

- Theorem (Existence and Uniqueness of Finite Fields): For any prime p and any positive integer n, there exists a finite field of degree n over $\mathbb{F}_p$, and this field has p^n elements. Furthermore, any two finite fields with p^n elements are isomorphic.
  - Proof: Consider the polynomial $q(x) = x^{p^n} - x$ over $\mathbb{F}_p$, and let K be its splitting field.
  - We see that $q'(x) = p^n x^{p^n - 1} - 1 = -1$: thus, q is separable and so it has precisely p^n roots in K.
  - If r and s are any two roots of q in K, then $r^{p^n} = r$ and $s^{p^n} = s$. We can then see that $(rs)^{p^n} = r^{p^n} s^{p^n} = rs$, and $(r - s)^{p^n} = r^{p^n} - s^{p^n} = r - s$, and if $r \neq 0$ then $(r^{-1})^{p^n} = (r^{p^n})^{-1} = r^{-1}$.
  - These three calculations show that if r and s are roots of q, then so are rs, $r - s$, and r^{-1}. Together with the trivial observations that 0 and 1 are roots of q, this says that the set of roots of q is a subfield of K.
  - But since K is generated (as a field) by the set of roots of q, this means that the set of roots is all of K.
  - It is easy to see that if K/F has dimension n, then the number of elements of K/F is $(\#F)^n = p^n$ (since each coefficient for each element in a basis has $\#F = p$ possible choices), so this tells us that $[K : \mathbb{F}_p] = n$.
  - For the uniqueness, now suppose that $[K : \mathbb{F}_p] = n$ and let $S = \{u_1, \ldots, u_{p^n - 1}\}$ be the set of nonzero elements of K. For any nonzero $r \in K$, multiplication by r is an injective function on S (since r is a unit) and hence is a bijection. Thus, the elements $\{ru_1, \ldots, ru_{p^n - 1}\}$ are the same as the elements $\{u_1, \ldots, u_{p^n - 1}\}$, though possibly in a different order.

- In particular, the products of these collections of elements are equal: multiplying out and collecting terms yields $r^{p^n-1}(u_1 \cdots u_{p^n-1}) = u_1 \cdots u_{p^n-1}$ so cancelling the nonzero elements yields $r^{p^n-1} = 1$.
- We conclude that every element in K (including 0) is a root of the polynomial $x^{p^n} - x$, and so K is contained in the splitting field for this polynomial. But as we have just shown above, the splitting field of $x^{p^n} - x$ over $\mathbb{F}_p$ already has p^n elements, so it must be equal to K. Since splitting fields are unique up to isomorphism, we are done.

2.4.2 Separable and Inseparable Extensions

- We can also extend these notions of separability and inseparability to algebraic elements by considering their minimal polynomials:
- Definition: If K/F is a field extension, then $\alpha \in K$ is separable over F if α is algebraic over K and its minimal polynomial $m(x)$ over F is a separable polynomial. We say K/F itself is separable if every $\alpha \in K$ is separable over F, and K/F is inseparable if it is not separable.
  - Example: Any algebraic element in an extension of characteristic 0 is separable, so algebraic extensions of characteristic-0 fields are separable. More generally, any algebraic element in an extension K/F where F is a perfect field is separable, so algebraic extensions of perfect fields are separable.
  - Example: The element $t \in \mathbb{F}_2(t^{1/2})$ is not separable over $\mathbb{F}_2(t)$, since its minimal polynomial is the inseparable polynomial $p(x) = x^2 - t$.
  - Example: The element $t \in \mathbb{F}_2(t^{1/3})$ is separable over $\mathbb{F}_2(t)$, since its minimal polynomial is the separable polynomial $p(x) = x^3 - t$.
- The inverse notion to a separable element is of an inseparable element that is "as inseparable as possible", where all of the roots of its minimal polynomial are the same:
- Definition: If K/F is a field extension, then $\alpha \in K$ is purely inseparable over F if α is algebraic over K and its minimal polynomial $m(x)$ over F has only α as a root. We say K/F itself is purely separable if every $\alpha \in K$ is purely inseparable over F.
  - Example: The element $t \in \mathbb{F}_2(t^{1/2})$ is purely inseparable over $\mathbb{F}_2(t)$: its minimal polynomial is the inseparable polynomial $m(x) = x^2 - t$, which factors as $m(x) = (x - t^{1/2})^2$ over $\mathbb{F}_2(t^{1/2})$, and this polynomial has only $t^{1/2}$ as a root.
  - Example: The element $t \in \mathbb{F}_5(t^{1/25})$ is purely inseparable over $\mathbb{F}_5(t)$: its minimal polynomial is the inseparable polynomial $m(x) = x^{25} - t$, which can be seen to factor as $m(x) = (x - t^{1/25})^{25}$ over $\mathbb{F}_5(t^{1/25})$, and this polynomial has only $t^{1/25}$ as a root. This factorization also shows why m is irreducible, since no lower power $(x - t^{1/25})^k$ for $1 \leq k \leq 24$ actually yields a polynomial with coefficients in $\mathbb{F}_5(t)$. (Alternatively, this polynomial is Eisenstein-irreducible with prime t.)
- We can have various properties of inseparability and purely inseparable extensions:
- Proposition (Properties of Inseparability): Let L/K and K/F be field extensions of characteristic p.

1. If $q(x) \in F[x]$ is an irreducible inseparable polynomial, then $q(x) = q_{\text{sep}}(x^{p^k})$ for a unique positive integer k and a unique irreducible separable polynomial $q_{\text{sep}}(x) \in F[x]$.
   - Proof: As we showed earlier, an irreducible polynomial q is inseparable if and only if its derivative q' is the zero polynomial. Equivalently, this means every monomial term in q must have the corresponding power of x divisible by p: this means $q(x) = q_1(x^p)$ for some polynomial $q_1 \in F[x]$.
   - If q_1 is separable, then it must necessarily be irreducible since otherwise any factorization of $q_1(x) = f(x)g(x)$ would give a factorization of $q(x) = q_1(x^p) = f(x^p)g(x^p)$.
   - Otherwise, if q_1 is inseparable, then by the argument above, we must have $q_1(x) = q_2(x^p)$ for some $q_2(x) \in F[x]$. By iterating this argument (or equivalently, by a trivial induction), eventually we must obtain a polynomial $q_k(x)$ that is separable and irreducible. Then $q(x) = q_{\text{sep}}(x^{p^k})$ as claimed.
2. The element $\alpha \in K/F$ is purely inseparable if and only if there exists some positive integer k such that $\alpha^{p^k} \in F$.

- Proof: First suppose α is purely inseparable and let $m(x) \in F[x]$ be the minimal polynomial of α over F.
- Then $m(x)$ is an irreducible purely inseparable polynomial, so $m(x) = q_{\text{sep}}(x^{p^k})$ for some separable polynomial q_{sep} by (1).
- If q_{sep} had two distinct roots r_1 and r_2, then (in an appropriate splitting field) m would have roots s_1 and s_2 satisfying $s_1^{p^k} = r_1$ and $s_2^{p^k} = r_2$. But since the pth-power map is injective and $r_1 \neq r_2$, this would mean that $s_1 \neq s_2$ and thus that m has two distinct roots, contradicting the assumption that m was purely inseparable.
- Conversely, if $\alpha^{p^k} \in F$, then α is a root of the polynomial $q(x) = x^{p^k} - \alpha^{p^k} = (x - \alpha)^{p^k}$ in $K[x]$.
- Then the minimal polynomial of α over F must therefore divide q, but since q has only one root α, that means m also has only one root α. Thus, α is purely inseparable.
- Remark: This result shows that the examples we gave of inseparable elements above are essentially the only possible ones.

3. The extension K/F is purely inseparable if and only if the minimal polynomial of each $\alpha \in K$ over F is of the form $m_\alpha(x) = x^{p^k} - d$ for some nonnegative integer k and some $d \in F$.
   - Proof: The forward direction follows immediately from (2) above. The reverse direction follows from the observation above that $m_\alpha(x) = (x - \alpha)^{p^k}$ inside K, so m_α has only the single root α.
4. The extension K/F is purely inseparable if and only if K/F is algebraic and the only elements of K separable over F are the elements of F.
   - Proof: If K/F is purely inseparable, then by (3) above any $\alpha \in K$ has minimal polynomial of the form $m_\alpha(x) = x^{p^k} - d = (x - \alpha)^{p^k}$ in K. Such a polynomial cannot be separable unless $k = 0$, in which case it has the form $m_\alpha(x) = x - d$, implying $\alpha \in F$.
   - Conversely, suppose K/F is algebraic the only elements of K separable over F are the elements of F.
   - For any $\alpha \in K$ consider its minimal polynomial, which by hypothesis must be inseparable: then by (1) it has the form $q(x) = q_{\text{sep}}(x^{p^k})$ for some positive integer k, where q_{sep} is separable.
   - But then the minimal polynomial of α^{p^k} is $q_{\text{sep}}(x)$, which is separable. Therefore, α^{p^k} must be an element of F, and then α is purely inseparable by (2).
   - Remark: This result is the reason for the terminology of "purely inseparable": all elements of the extension, other than the elements of the ground field F themselves, are inseparable over F.
5. The extension L/F is purely inseparable if and only if L/K and K/F are purely inseparable.
   - Proof: If L/F is purely inseparable, then by (2), for any $\alpha \in L \backslash F$ we have $\alpha^{p^k} \in F$ for some positive integer k.
   - In particular this holds for any $\alpha \in K \backslash F$, so K/F is purely inseparable.
   - Furthermore, if $\alpha \in L \backslash F$ then since $\alpha^{p^k} \in F$ we have $\alpha^{p^k} \in K$, so L/K is purely inseparable by (2).
   - Conversely, suppose L/K and K/F are purely inseparable.
   - Then by (2), for any $\alpha \in L$ we have $\alpha^{p^{k_1}} \in K$ for some k_1, and also if $\beta = \alpha^{p^{k_1}}$ we have $\beta^{p^{k_2}} \in F$ for some k_2.
   - But then $\alpha^{p^{k_1+k_2}} = \beta^{p^{k_1}} \in F$, so by (2) again, this means α is purely inseparable, so L/F is purely inseparable.
6. The composite of purely inseparable extensions over F is also purely inseparable over F.
   - Proof: Suppose K is a composite of purely inseparable extensions of F.
   - Then any $\gamma \in K$ is of the form $\gamma = \dfrac{p(\alpha_1, \ldots, \alpha_i)}{q(\alpha_{i+1}, \ldots, \alpha_{i+j})} \in K$ where $\alpha_1, \ldots, \alpha_i, \alpha_{i+1}, \ldots, \alpha_{i+j}$ are purely inseparable elements over F and p, q are polynomials with coefficients in F.
   - Then by (2), there exist integers $k_1, \ldots, k_{i+j}$ such that $\alpha_l^{p^{k_l}} \in F$ for each $1 \leq l \leq i + j$. If $M = \max(k_l)$, then $\alpha_l^{p^M} \in F$ for each l.
   - Then γ^{p^M} is a rational function with coefficients from F in the elements $\alpha_l^{p^M} \in F$, so $\gamma^{p^M} \in F$.
   - Hence by (2), γ is purely inseparable over F, and so K/F is purely inseparable.

7. If K/F has finite degree and $K = F(\alpha_1, \ldots, \alpha_k)$, then K/F is purely inseparable if and only if each α_i is purely inseparable over F.
   - Proof: If K/F is purely inseparable, then by (5) each of the extensions $F(\alpha_i)$ is purely inseparable, so each α_i is inseparable.
   - Conversely, if each of the $F(\alpha_i)$ is purely inseparable, then by (6) so is their composite field $K = F(\alpha_1, \ldots, \alpha_k)$.
8. Every finite-degree purely inseparable extension has degree equal to a power of p.
   - Proof: This follows from applying the degree tower formula to (7) and by noting that any simple purely inseparable extension has degree equal to a power of p by (3).

- Most of these properties also hold analogously for separable extensions, although in order to prove them we must use some facts about embeddings of fields into their algebraic closures:

- Proposition (Properties of Separability): Let L/K and K/F be field extensions of characteristic p.

1. If α is algebraic over F, then α is separable over F if and only if there are $[F(\alpha) : F]$ different embeddings of $F(\alpha)/F$ into $\overline{F}/F$.
   - Note that an embedding of $F(\alpha)/F$ into $\overline{F}/F$ is an injective ring homomorphism $\sigma : F(\alpha) \to \overline{F}$ such that σ fixes F (i.e., $\sigma(x) = x$ for all $x \in F$).
   - The idea of this result is that the embeddings of $F(\alpha)/F$ into $\overline{F}/F$ count the number of different roots that the minimal polynomial of α has inside $\overline{F}$.
   - Proof: Suppose α is separable over F, let its minimal polynomial be $m(x)$ of degree n, and let L be the splitting field of m over F.
   - First observe that $\sigma(\alpha)$ must also be a root of $m(x)$ inside $\overline{F}$: this follows simply by noting that $m(\sigma(\alpha)) = \sigma(m(\alpha)) = \sigma(0) = 0$.
   - Now let β be any root of $m(x)$. By the theorem on the uniqueness of splitting fields, the identity map on F extends to an isomorphism of L with itself that maps α to β, since the identity map sends the minimal polynomial of α to the minimal polynomial of β (since they have the same minimal polynomial $m(x)$).
   - In particular, restricting this isomorphism to $F(\alpha)$ yields an embedding of $F(\alpha)$ into $\overline{F}$ whose image is $F(\beta)$.
   - Since any map $\sigma : F(\alpha) \to \overline{F}$ fixing F is completely determined by the value of $\sigma(\alpha)$, we see that this correspondence yields a bijection between embeddings of $F(\alpha)/F$ into $\overline{F}/F$ with the distinct roots β of $m(x)$.
   - The result then follows immediately, since α is separable if and only if $m(x)$ has $\deg(m) = [F(\alpha) : F]$ distinct roots.
2. If K/F has finite degree, then there are at most $[K : F]$ different embeddings of K/F into $\overline{F}/F$.
   - Our goal is to prove, in a moment, that equality holds if and only if K/F is separable.
   - Proof: Induct on the number $n = 1$ of generators of K/F. The case $n = 1$, where $K = F(\alpha_1)$ was shown above, since in this situation the embeddings are in bijection with the distinct roots of the minimal polynomial of the generator α_1.
   - For the inductive step, suppose the result holds for extensions having k generators and suppose $K = F(\alpha_1, \ldots, \alpha_{k+1})$, and set $E = F(\alpha_1, \ldots, \alpha_k)$, so that $K = E(\alpha_{k+1})$.
   - Then any embedding of K/F into $\overline{E} = \overline{F}$ is determined by the image of E, which has at most $[E : F]$ possible choices by the induction hypothesis, and the image of α_{k+1}, which has at most $[K : E]$, the degree of the minimal polynomial of α_{k+1} over E, possible choices once the image of E is determined.
   - Therefore, the number of embeddings is at most $[K : E] \cdot [E : F] = [K : F]$, as claimed.
3. If α is algebraic over F, then α is separable over F if and only if $F(\alpha)$ is separable over F.
   - Proof: Trivially, if $F(\alpha)/F$ is separable then α is separable over F.
   - Now suppose α is separable over F and suppose there were an inseparable element $\beta \in F(\alpha)$.
   - Then by (1), the number $n_{F(\beta)/F}$ of embeddings of $F(\beta)/F$ is strictly less than $[F(\beta) : F]$.

- Also, by (2), the number of embeddings $n_{F(\alpha)/F(\beta)}$ of $F(\alpha)/F(\beta)$ into $\overline{F(\beta)} = \overline{F}$ is at most $[F(\alpha) : F(\beta)]$.
- Therefore, since any embedding of $F(\alpha)/F$ is determined uniquely by the embeddings of $F(\beta)/F$ and $F(\alpha)/F(\beta)$, the number of embeddings $n_{F(\alpha)/F}$ of $F(\alpha)/F$ is at most $n_{F(\alpha)/F(\beta)} \cdot n_{F(\beta)/F} < [F(\alpha) : F(\beta)] \cdot [F(\beta) : F] = [F(\alpha) : F]$.
- This is a contradiction, since $F(\alpha)/F$ is separable and so by (1), the total number of embeddings equals $[F(\alpha) : F]$.

4. If L/F is separable, then L/K and K/F are separable.
   - Proof: First suppose that L/F is separable, so that every element of L is separable over F. Then because K is a subset of L, this means every element of K is separable over F, so K/F is separable.
   - Furthermore, for any $\alpha \in L$, if $m_F(x)$ is the minimal polynomial of α over F, then the minimal polynomial $m_K(x)$ of α over K divides it, since $m_F(\alpha) = 0$ in K. All roots of $m_F(x)$ are distinct since α is separable, so all roots of $m_K(x)$ are also distinct. Thus, L/K is separable.
5. If K/F has finite degree, then K/F is separable if and only if there are exactly $[K : F]$ different embeddings of K/F into $\overline{F}/F$.
   - Proof: Suppose $K = F(\alpha_1, \ldots, \alpha_k)$, and let $E_i = F(\alpha_1, \ldots, \alpha_i)$ for each $0 \leq i \leq k$.
   - By (2), the total number of different embeddings of K/F into $\overline{F}/F$ is at most $[K : F]$ by (2).
   - If K is separable, then by (4), each subextension $E(\alpha_{i+1})/E$ is separable, and then by (3), the argument in (1), and a trivial induction, this means the number of embeddings of $E(\alpha_{i+1})/F$ into $\overline{F}/F$ is equal to $[E(\alpha_{i+1}) : E] \cdot [E : F] = [E(\alpha_{i+1}) : F]$, since each embedding of $E(\alpha_{i+1})/F$ is realized by an embedding of E/F along with an embedding of $E(\alpha_{i+1})/E$.
   - Thus, taking $i = k$ yields that the number of embeddings of K/F into $\overline{F}/F$ is equal to $[K : F]$, as required.
   - Inversely, if K is not separable, then it contains some inseparable element β.
   - Then $F(\beta)/F$ has fewer than $[F(\beta) : F]$ embeddings into $\overline{F}/F$ by (1). Since the number of embeddings of $K/F(\beta)$ is at most $[K : F(\beta)]$ by (2), by the same argument as above, the total number of embeddings of K/F into $\overline{F}$ is strictly fewer than $[K : F(\beta)] \cdot [F(\beta) : F] = [K : F]$.
6. If K/F is separable, then α is separable over K if and only if α is separable over F.
   - Proof: If α is separable over F then by the argument in (4), it is separable over K.
   - First suppose K/F has finite degree. Suppose α is separable over K and consider the tower $K(\alpha)/K/F$.
   - By (3), $K(\alpha)/K$ is separable, and then by (5), the number of embeddings of $K(\alpha)/K$ into $\overline{K} = \overline{F}$ is equal to $[K(\alpha) : K]$. Also by (5), the number of embeddings of K/F into $\overline{F}$ is $[K : F]$.
   - Furthermore, it is easy to see by composing the appropriate maps that if we have an embedding of K/F into $\overline{F}$ and an embedding of $K(\alpha)/K$ into $\overline{K} = \overline{F}$, then it yields a unique embedding of $K(\alpha)/F$ into $\overline{F}$.
   - Therefore, the total number of embeddings of $K(\alpha)/F$ into $\overline{F}$ equals $[K(\alpha) : K] \cdot [K : F] = [K(\alpha) : F]$. Hence by (5) again, $K(\alpha)/F$ is separable, so α is separable over F.
   - The general case follows from the finite-degree case by noting that if the minimal polynomial for α over K is $m(x) = b_d x^d + \cdots + b_0$ for $b_i \in K$, then α is separable over $F(b_0, \ldots, b_d)$ since it is separable over K and has the same minimal polynomial over both fields, and $F(b_0, \ldots, b_d)/F$ has finite degree.
7. If K/F has finite degree and $K = F(\alpha_1, \ldots, \alpha_k)$, then K/F is separable if and only if each α_i is separable over F.
   - Proof: This follows by repeatedly applying (6) to the tower $K/F(\alpha_1, \ldots, \alpha_{k-1})/\cdots/F(\alpha_1)/F$.
8. If L/K and K/F are separable, then L/F is separable.
   - Proof: If L/K has finite degree then this follows by writing $L = K(\alpha_1, \ldots, \alpha_k)$ and applying (7). The general case follows from the finite-degree case because for each $\alpha \in L$, α is separable over F if and only if $K(\alpha)/F$ is separable over L by (3).
9. The composite of separable extensions is separable.

- Proof: If the extensions have finite degree then this follows from (7) by writing $K_1 = F(\alpha_1, \ldots, \alpha_k)$ and $K_2 = F(\beta_1, \ldots, \beta_l)$ and noting that $K_1K_2 = F(\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_l)$.
- The general case follows from the finite-degree case since any element of a composite extension is a rational function in finitely many elements from the given fields.

- Using the properties of separable extensions we can define the "separable closure" of F inside K/F:
- Definition: If K/F is a field extension, we define the maximal separable extension F^{sep} of F inside K to be the composite of all separable extensions of F inside K.
  - The elements of F^{sep} consist of all $\alpha \in K$ that are separable over F: all such elements are in this composite since $F(\alpha)/F$ is separable by property (3) of separable extensions.
  - From this observation, we can see that F^{sep} is indeed the largest subfield of K that is separable over F, whence the name.
  - Also, any element of K not in F^{sep} is inseparable over F hence also over F^{sep}: by property (4) of purely inseparable extensions, this means K/F^{sep} is purely inseparable.
  - Indeed, F^{sep} is the only subfield E of K that is separable over F such that K/E is purely inseparable: any proper subfield of F^{sep} will not have the property that K/E is purely inseparable, since there exist elements of K not in E that are not purely inseparable over E (namely, any element of F^{sep} not in E).

- Using the separable closure, we can define a notion of separable and inseparable degree for extensions:
- Definition: If K/F is algebraic, the separable degree $[K : F]_{\text{sep}}$ is defined to be the degree $[F^{\text{sep}} : F]$, while the inseparable degree $[K : F]_{\text{insep}}$ is defined to be the degree $[K : F^{\text{sep}}]$.
  - The product of the separable degree and the inseparable degree is the regular degree $[K : F]$.
  - Also, since composites and separable extensions of separable extensions are separable, the separable degree (and hence also the inseparable degree) is multiplicative in towers.
  - From our properties of purely inseparable extensions, the inseparable degree $[K : F]_{\text{insep}}$ is either ∞ or a power of the characteristic.

- For simple extensions, we can calculate the separable and inseparable degree using the minimal polynomial of a generator:
- Proposition (Separable Degree of Simple Extension): Suppose α is algebraic over F with minimal polynomial $m(x) = m_{\text{sep}}(x^{p^k})$ where k is a nonnegative integer and $m_{\text{sep}}(x)$ is a separable polynomial. Then $F^{\text{sep}} = F(\alpha^{p^k})$, so that $[F(\alpha) : F]_{\text{sep}} = \deg(m_{\text{sep}})$ and $[F(\alpha) : F]_{\text{insep}} = p^k$.
  - Proof: Observe that α^{p^k} is a root of m_{sep} since $m_{\text{sep}}(\alpha^{p^k}) = m(\alpha) = 0$, so α^{p^k} is separable over F. Thus, $F(\alpha^{p^k})$ is separable over F by property (3) of separable extensions.
  - Furthermore, since $K/F(\alpha^{p^k})$ is generated by α, and $\alpha^{p^k} \in F(\alpha^{p^k})$, by properties (3) and (7) of purely inseparable extensions we see that $K/F(\alpha^{p^k})$ is purely inseparable.
  - But this means $F(\alpha^{p^k})$ must be F^{sep} by the uniqueness property we noted above.
  - For the degree calculations we have $[F(\alpha^{p^k}) : F] = \deg(m_{\text{sep}})$ since m_{sep} is the minimal polynomial of α^{p^k} over F, and also $[F(\alpha) : F(\alpha^{p^k})] = p^k$ since $x^{p^k} - \alpha^{p^k}$ is the minimal polynomial of α over $F(\alpha^{p^k})$.

- Example: For $F = \mathbb{F}_p(t)$ and $K = F(\alpha)$ where α is a root of the irreducible polynomial $q(x) = x^{2p} - tx^p + t$, we have $[K : F]_{\text{sep}} = 2$ and $[K : F]_{\text{insep}} = p$.
  - Note that q is irreducible in $F[x]$ since it is Eisenstein at t. Explicitly, F^{sep} is generated by a root of $m_{\text{sep}}(x) = x^2 - tx + t$.

2.4.3 Transcendental Extensions and Transcendence Degree

- We now discuss in more detail the structure of transcendental extensions.
  - If K/F is any field extension, let E be the field of elements algebraic over F inside K. Then, since algebraic extensions of algebraic extensions are algebraic, any element of K/E not in E must be transcendental over F.
  - Our goal is to describe how to analyze this "transcendental part" of the extension.
  - To describe the elements of K, the idea is to identify a minimal set of independent generators for K/E, in analogy with the situation in vector spaces.
  - Here, we do not merely need the generators to be linearly independent, but rather algebraically independent:

- Definition: Let K/F be a field extension. We say a subset S of K is algebraically dependent over F if there exists a finite subset $\{s_1, \dots, s_n\} \in S$ and a nonzero polynomial $p \in F[x_1, \dots, x_n]$ such that $p(s_1, \dots, s_n) = 0$. If there exists no such p for any finite subset of S, we say S is algebraically independent.
  - The general idea here is that a set of elements is algebraically dependent if they satisfy some algebraic (i.e., polynomial) relation over F.
  - Example: Over $\mathbb{Q}$, the set $\{\pi, \pi^2\}$ is algebraically dependent, since $p(x, y) = x^2 - y$ has $p(\pi, \pi^2) = 0$.
  - Example: Over $\mathbb{Q}$, the set $\{\sqrt[3]{2}\}$ is algebraically dependent, since $p(x) = x^3 - 2$ has $p(\sqrt[3]{2}) = 0$.
  - More generally, the set $\{\alpha\}$ is algebraically independent over F if and only if α is transcendental over F.
  - Example: Over $\mathbb{R}$, the set $\{x + y, x^2 + y^2\}$ is algebraically independent.
  - Example: Over $\mathbb{R}$, the set $\{x+y, x^2+y^2, x^3+y^3\}$ is algebraically dependent, since $p(a, b, c) = a^3-3ab+2c$ has $p(x + y, x^2 + y^2, x^3 + y^3) = 0$.
  - Example: If $x_1, \dots, x_n$ are indeterminates inside $F(x_1, \dots, x_n)$, the function field in n variables, then the set $\{x_1, \dots, x_n\}$ is algebraically independent over F.

- The notion of algebraic independence generalizes the notion of linear independence, and as such the two concepts are related in various ways.
  - It is easy to see that any subset of an algebraically independent set is algebraically independent, while any set containing an algebraically dependent set is algebraically dependent.
  - Also, we observe that linear dependence is a special type of algebraic dependence; namely, a set is linearly dependent precisely when it is algebraically dependent where the polynomial p is linear.
  - We have already defined the algebraic notion of the span of a set S: it is simply the subfield generated by S.
  - We might therefore hope to define a "transcendence basis" to be an algebraically independent set that generates the extension K/F.
  - Unfortunately, such a set need not exist: for example, $\mathbb{Q}(\sqrt{2})/\mathbb{Q}$ has no such set, because there are no transcendental elements at all.
  - The correct analogy is instead to observe that a basis for a vector space is a maximal linearly independent set:

- Definition: Let K/F be a field extension. A transcendence base for K/F is an algebraically independent subset S of K that is maximal in the set of all algebraically independent subsets of K.
  - Remark: The term "transcendence basis" is also used occasionally. We will prefer to use the word "base" to keep a distinction between a basis of a vector space and a transcendence base of a field extension.
  - By a straightforward Zorn's lemma argument, every extension has a transcendence base.
  - Example: The empty set $\emptyset$ is a transcendence base for $\mathbb{Q}(\sqrt{2})/\mathbb{Q}$. More generally, K/F is algebraic if and only if $\emptyset$ is a transcendence base.
  - Example: The set $\{x\}$ is a transcendence base for $F(x)$ over F.

- Here are some of the fundamental properties of transcendence bases, many of which are analogous to properties of vector spaces:

- Proposition (Transcendence Bases): Suppose K/F is a field extension and S is a subset of K.

  1. If S is algebraically independent and $\alpha \in K$, then $S \cup \{\alpha\}$ is algebraically independent over F if and only if α is transcendental over $F(S)$.
     - Proof: This is the algebraic analogue of the statement that if S is linearly independent, then $S \cup \{\alpha\}$ is linearly independent if and only if α is not in the span of S.
     - Suppose $S \cup \{\alpha\}$ is algebraically dependent. Then there exists $s_i \in S$ and $p \in F[x]$ with $p(\alpha, s_1, \ldots, s_n) = 0$ and $p \neq 0$. View p as a polynomial in its first variable with coefficients in $F[s_1, \ldots, s_n]$: there must be at least one term involving α, as otherwise p would give an algebraic dependence in S. Then α is the root of a nonzero polynomial with coefficients in $F[s_1, \ldots, s_n] \subseteq F(s_1, \ldots, s_n) \subseteq F(S)$, so it is algebraic over $F(S)$.
     - Conversely, suppose that α is algebraic over $F(S)$. Then α is the root of some nonzero polynomial with coefficients in $F(S)$. Each coefficient of this polynomial is an element of $F(S)$; clearing denominators yields a nonzero polynomial p with coefficients in $F[s_1, \ldots, s_n]$ for the elements $s_i \in S$ that appear in these coefficients. This polynomial yields an algebraic dependence in $S \cup \{\alpha\}$.
  2. S is a transcendence base of K/F if and only if K is algebraic over $F(S)$.
     - Proof: This follows from (1) and the maximality of transcendence bases: if S is transcendence base if and only if no elements in K can be adjoined to S while preserving algebraic independence, and by (1) this is equivalent to saying that all elements in K are algebraic over $F(S)$.
  3. If T is a subset of K such that $K/F(T)$ is algebraic, then T contains a transcendence base of K/F.
     - Proof: Apply Zorn's lemma to the collection of all algebraically independent subsets of T, partially ordered by inclusion.
     - A maximal element M in this collection must then be a transcendence base for K/F: if $\beta \in K$ then β must be algebraic over $K/F(M)$ by the maximality of M, and then M is a transcendence base by (2).
  4. If T is an algebraically independent subset of K, then T can be extended to a transcendence base of K/F.
     - Proof: This is the analogue of the fact that every linearly independent subset can be extended to a basis, and the proof follows from a similar Zorn's lemma argument.
  5. If $S = \{s_1, \ldots, s_n\}$ is a transcendence base for K/F and $T = \{t_1, \ldots, t_m\}$ is any algebraically independent set, then there is a reordering of S, say $\{a_1, \ldots, a_n\}$, such that for each $1 \leq k \leq m$, the set $\{t_1, t_2, \ldots, t_k, a_{k+1}, \ldots, a_n\}$ is a transcendence base for K/F.
     - Proof: This is the analogue of the replacement theorem, and the proof proceeds inductively in essentially the same way. (We will omit the details.)
  6. If S is a (finite) transcendence base for K/F, then any subset T of K having larger cardinality than S must be algebraically dependent.
     - Proof: If $S = \{s_1, \ldots, s_n\}$ is finite, apply the replacement theorem (5) to S and T. At the end of the replacement, the result is that $\{t_1, \ldots, t_n\}$ is a transcendence base. But then by (2), any additional element of T would be algebraic over $\{t_1, \ldots, t_n\}$, contradicting the algebraic independence of T.
  7. Any two transcendence bases S and T for K/F have the same cardinality.
     - Proof: If the bases are infinite the result is immediate. If S has finite cardinality n, then the result follows by applying (6), since then T's cardinality m must satisfy $m \leq n$ (since T is algebraically independent and S is a transcendence base) and also $n \leq m$ (since S is algebraically independent and T is a transcendence base).

- The result of the last part of the proposition shows that any two transcendence bases have the same cardinality, and in analogy with the situation for vector spaces, this cardinality behaves somewhat like an extension degree:

- Definition: Let K/F be a field extension. The transcendence degree of K/F, denoted $\text{trdeg}(K/F)$, is the cardinality of any transcendence base of K/F.

- The key property of transcendence degree is that it is additive in towers:
- Theorem (Transcendence in Towers): If $L/K/F$ is a tower of extensions, then $\mathrm{trdeg}(L/F) = \mathrm{trdeg}(L/K) + \mathrm{trdeg}(K/F)$.
  - The idea here is quite simple: we want to show that the union of transcendence bases for K/F and L/K gives a transcendence base for L/F.
  - Proof: First suppose that both $\mathrm{trdeg}(K/F)$ and $\mathrm{trdeg}(L/K)$ are finite, and let $S = \{s_1, \ldots, s_n\}$ and $T = \{t_1, \ldots, t_m\}$ be transcendence bases for K/F and L/K. Then $S \cap T = \emptyset$ since each t_i is transcendental over K.
  - Furthermore, K is algebraic over $F(S)$, so $K(T)$ is algebraic over $F(T)(S) = F(S \cup T)$ by our results on algebraic extensions.
  - Then since L is algebraic over $K(T)$, we deduce that L is algebraic over $F(S \cup T)$, also by our results on algebraic extensions.
  - Thus, by property (3) above, $S \cup T$ contains a transcendence base of L/F.
  - Finally, we claim $S \cup T$ is algebraically independent over F, so suppose that $p(s_1, \ldots, s_n, t_1, \ldots, t_m) = 0$ for some $p \in F[x_1, \ldots, x_n, y_1, \ldots, y_m]$.
  - Separate monomial terms to write $p(s_1, \ldots, s_n, t_1, \ldots, t_m) = 0$ as a sum $\sum f_i(s_1, \ldots, s_n) g_i(t_1, \ldots, t_m) = 0$ with $f_i \in F[x_1, \ldots, x_n]$ and $g_i \in F[y_1, \ldots, y_m]$.
  - Now, since T is algebraically independent over $F(S) \subseteq K$, all of the $f_i(s_1, \ldots, s_n)$ must be zero (as elements of K). But since S is algebraically independent over F, that means all of the polynomials $f_i(x_1, \ldots, x_n)$ must be zero (as polynomials).
  - This means p is the zero polynomial, and so $S \cup T$ is algebraically independent.
- Fields that are generated by a transcendence base are particularly convenient:
- Definition: The extension K/F is purely transcendental if $K = F(S)$ for some transcendence base S of K/F.
  - Equivalently, K/F is purely transcendental when it is generated (as a field extension) by an algebraically independent set.
  - If $S = \{s_1, \ldots, s_n\}$, then the purely transcendental extension $K = F(S)$ is ring-isomorphic to the function field $F(x_1, \ldots, x_n)$ in n variables: it is not hard to check that the map sending s_i to x_i is an isomorphism.
  - If K/F has transcendence degree 1 or 2 and E/F is an intermediate extension, then in fact E is also purely transcendental: the degree-1 case is a theorem of Lüroth, while the degree-2 case is a theorem of Castelnuovo. In higher degrees, there do exist extensions that are not purely transcendental, but it is not easy to verify this fact.
- Since any extension K/F has a transcendence base S, property (2) of transcendence bases implies that K/F is an algebraic extension of the purely transcendental extension $F(S)/F$.
  - This shows that any field extension can be written as an algebraic extension of a purely transcendental extension.
  - One might wonder whether it is possible to reverse the order and put the algebraic piece first: the answer turns out to be no, for reasons related to algebraic geometry.
  - For example, if F is algebraically closed (e.g., $\mathbb{C}$) any example of a transcendental extension that is not purely transcendental cannot have the order reversed, since there are no algebraic extensions of $\mathbb{C}$.
  - One example of such a field is the elliptic function field $\mathbb{C}(t, \sqrt{t^3 + t})$, which arises as the function field of the elliptic curve $y^2 = x^3 + x$; the relationship between these two follows from the fact that $\mathbb{C}(t, \sqrt{t^3 + t}) \cong \mathbb{C}[x, y]/(y^2 - x^3 - x)$.
  - We have barely scratched the surface of what can be said here, but as a closing remark we note that much of elementary algebraic geometry is concerned with understanding these connections between algebraic properties of function fields and geometric properties of varieties.

Well, you're at the end of my handout. Hope it was helpful.