Chapter 3. Fields

Algebraic and Transcendental Extensions

3.1 Definition: When we say that F and K are fields with $F \subseteq K$, we shall assume, unless otherwise stated, that F is using the same operation used in K, so that F is a subfield of K (that is K is an extension field of F). In this case, the field K is a vector space over the field F, and we define the **index** of K over F to be $[K:F] = \dim_F K$.

3.2 Example: $\mathbb{R}$ is a subfield of $\mathbb{C}$, and $[\mathbb{C}:\mathbb{R}] = 2$.

3.3 Exercise: Verify that

$$\begin{aligned}\mathbb{Q}[i] &= \left\{a+b\,i \,\middle|\, a,b\in\mathbb{Q}\right\} \\ \mathbb{Q}[\sqrt{2}\,] &= \left\{a+b\sqrt{2} \,\middle|\, a,b\in\mathbb{Q}\right\} \\ \mathbb{Q}[\sqrt{2},i] &= \left\{a+b\sqrt{2}+c\,i+d\sqrt{2}\,i \,\middle|\, a,b,c,d\in\mathbb{Q}\right\}\end{aligned}$$

are all fields with $\big[\mathbb{Q}[i]:\mathbb{Q}\big] = 2$, $\big[\mathbb{Q}[\sqrt{2}\,]:\mathbb{Q}\big] = 2$, and $\big[\mathbb{Q}[\sqrt{2},i]:\mathbb{Q}\big] = 4$.

3.4 Theorem: *Let F, K and L be fields with $F \subseteq K \subseteq L$. Then $[L:F] = [K:F][L:K]$. Indeed if U is a basis for K over F and V is a basis for L over K then*

$$W = \left\{uv \,\middle|\, u \in U, v \in V\right\}$$

is a basis for K over F.

Proof: Let U be a basis for K over F, and let V be a basis for L over K, and let $W = \left\{uv \,\middle|\, u\in U, v\in V\right\}$. In the case that U or V is infinite, recall that $|X|$ denotes the cardinality of a set X, recall that for two sets X and Y, we write $|X| = |Y|$ when there is a bijection $F: X \to Y$, and recall that $|X|\,|Y| = |X \times Y|$ (by definition). With this in mind, note that $|W| = |U|\,|V| = |U \times V|$ because the map $F: U \times V \to W$ given by $F(u,v) = uv$ is bijective: indeed it is clearly surjective, and it is injective because for $u_1, u_2 \in U$ and $v_1, v_2 \in V$, if $u_1 v_1 = u_2 v_2$ then since V is linearly independent we have $u_1 = u_2$, and since U is linearly independent so that $u_1 \neq 0$, $u_1 v_1 = u_1 v_2$ implies that $v_1 = v_2$.

Note that W spans L because given $w \in L$, since V spans L over K, we can choose $s_1, \cdots, s_n \in K$ such that $w = \sum_{j=1}^n s_j v_j$, then since U spans K over F, for each index j, we can choose $t_{j,1}, \cdots, t_{j,\ell_j}$, such that $s_j = \sum_{i=1}^{\ell_j} t_{j,i} u_i$, and then we have

$$w = \sum_{j=1}^n s_j v_j = \sum_{j=1}^n \big(\sum_{i=1}^{\ell_j} t_{j,i} u_i\big) v_j = \sum_{j=1}^n \sum_{i=1}^{\ell_j} t_{j,i} u_i v_j \,.$$

It remains to show that W is linearly independent. Suppose $\sum_{k=1}^m s_k w_k = 0$ where $w_1, \cdots, w_m$ are distinct elements in W and each $s_k \in F$. By the bijective correspondence $F: U \times V \to W$, the distinct elements $w_1, \cdots, w_m$ can be written as $u_{i,j} v_j$ with $1 \le j \le n$ and $1 \le i \le \ell_j$, such that $v_1, \cdots, v_n \in V$ are distinct and, for each index j, $u_{1,j}, \cdots, u_{\ell_j,j} \in U$ are distinct, and when $w_k = u_{i,j} v_j$, we write the corresponding coefficient as $s_k = t_{i,j}$. Then we have $0 = \sum_{k=1}^m s_k w_k = \sum_{j=1}^n \sum_{i=1}^{\ell_j} t_{i,j} u_{i,j} v_j$. Since V is linearly independent, we must have $\sum_{i=1}^{\ell_j} t_{i,j} u_{i,j} = 0$ for all j, and since U is linearly independent, we have $t_{i,j} = 0$ for all i,j, and hence $s_k = 0$ for all k.

3.5 Definition: When R and S are commutative rings with 1, where $R \subseteq S$ and U is a subset of S, the **subring of S generated by U over R**, denoted by $R[U]$, is the smallest subring of S which contains $R \cup U$. When $U = \{u_1, u_2, \cdots, u_n\}$ we write $R[U]$ as $R[u_1, u_2, \cdots, u_n]$, and we have

$$R[u_1, u_2, \cdots, u_n] = \left\{ f(u_1, u_2, \cdots, u_n) \,\middle|\, f \in R[x_1, x_2, \cdots, x_n] \right\}.$$

When $S = R[u_1, u_2, \cdots, u_n]$ for some $u_1, u_2, \cdots, u_n \in S$, we say that S is **finitely generated** as a ring over R.

When F and K are fields with $F \subseteq K$ and $U \subseteq K$, the **subfield of K generated by U over F**, denoted by $F(U)$, is the smallest subfield of K which contains $F \cup U$. When $U = \{u_1, u_2, \cdots, u_n\}$ we write $F(U)$ as $F(u_1, u_2, \cdots, u_n)$, and we have

$$F(u_1, \cdots, u_n) = \left\{ \frac{f(u_1, \cdots, u_n)}{g(u_1, \cdots, u_n)} \,\middle|\, f, g \in F[x_1, \cdots, x_n] \text{ and } g(u_1, \cdots, u_n) \neq 0 \right\}.$$

When $K = F(u_1, \cdots, u_n)$ for some $u_1, \cdots, u_n \in K$, we say that K is **finitely generated** as a field over F.

3.6 Definition: Let F and K be fields with $F \subseteq K$. For $a \in K$, we say that a is **algebraic** over F when there exists a polynomial $f(x) \in F[x]$ such that $f(a) = 0$ in K, otherwise we say that a is **transcendental** over F. We say that K is **algebraic** over F when every element $a \in K$ is algebraic over F, otherwise we say that K is **transcendental** over F.

3.7 Theorem: *Let F and K be fields with $F \subseteq K$ and let $a \in K$.*

(1) If a is transcendental over F then we have

$$F[a] \cong F[x] \quad \text{and} \quad F(a) \cong F(x).$$

In this case $[F(a) : F] = \infty$ and the set $\{1, a, a^2, \cdots\}$ is linearly independent over F.

(2) If a is algebraic over F then there is a unique monic irreducible polynomial $f(x) \in F[x]$ with $f(a) = 0$, the ideal generated by this polynomial in $F[x]$ is $\langle f \rangle = \left\{ g \in F[x] \,\middle|\, g(a) = 0 \right\}$ and we have

$$F(a) = F[a] \cong F[x]/\langle f \rangle .$$

For $n = \deg(f)$ the set $\{1, a, a^2, \cdots, a^{n-1}\}$ is a basis for $F(a)$ over F, and $[F(a) : F] = n$.

Proof: To prove Part 1, suppose a is transcendental over F. The evaluation homomorphism $\phi : F[x] \to F[a]$, given by $\phi(f) = f(a)$, is clearly surjective $\left(\text{since } F[a] = \left\{ f(a) \,\middle|\, f \in F[x] \right\}\right)$, and it is injective because a is transcendental (so if $f(a) = 0$ then $f = 0$). Thus $F[a] \cong F[x]$.

The evaluation homomorphism $\phi : F(x) \to F(a)$, given by $\phi(f/g) = f(a)/g(a)$ where $f, g \in F[x]$ with $g \neq 0$, is well-defined because a is transcendental (so when $g \neq 0$, we have $g(a) \neq 0$). Also ϕ is clearly surjective $\left(\text{since } F(a) = \left\{ \frac{f(a)}{g(a)} \,\middle|\, f, g \in F[x], g(a) \neq 0 \right\}\right)$, and ϕ is injective, indeed every nonzero homomorphism from a field to a ring is injective because its kernel is an ideal, and the only ideals in a field are the trivial ideal and the entire field).

Finally, note that $\left\{1, a, a^2, \cdots\right\}$ is linearly independent (hence $[F(a) : F] = \infty$) since if we have $\sum_{i=0}^{n} c_i a^i = 0$ with $c_i \in F$, and we let $g(x) = \sum_{i=0}^{n} c_i x^i \in F[x]$, then we have $g(a) = 0$, hence $g = 0$ (since a is transcendental), and hence $c_i = 0$ for all i.

To prove Part 2, suppose that a is algebraic over F. The evaluation homomorphism $\phi : F[x] \to F[a]$, given by $\phi(f) = f(a)$, is clearly surjective $\big($since $F[a] = \big\{f(a) \,\big|\, f \in F[x]\big\}\big)$, and we have $\operatorname{Ker}\phi = \big\{g \in F[x] \,\big|\, g(a) = 0\big\}$. Note that $\operatorname{Ker}\phi \neq \{0\}$ because a is algebraic (so there exists $0 \neq g \in F[x]$ such that $g(a) = 0$). Since $F[x]$ is a Euclidean domain, with Euclidean norm $E(g) = \deg(g)$, we know that $F[x]$ is a principal ideal domain, and that $\operatorname{Ker}\phi = \langle g \rangle$ where g is a nonzero polynomial of smallest degree in $\operatorname{Ker}\phi$, that is a nonzero polynomial of smallest degree with $g(a) = 0$. Note that $\deg g \geq 1$ because $g \neq 0$ and if g was a nonzero constant polynomial, say $g(x) = c$ with $0 \neq 0 \in F$, then $g(a) = c \neq 0$. Also note that there is a unique monic polynomial f with $\langle f \rangle = \langle g \rangle$, namely $f = \frac{1}{c}\, g$ where c is the leading coefficient of g. Thus we have

$$\operatorname{Ker}\phi = \big\{g \in F[x] \,\big|\, g(a) = 0\big\} = \langle f \rangle$$

where f is the unique monic polynomial of minimal degree with $f(a) = 0$. By the First Isomorphism Theorem, since ϕ is surjective with $\operatorname{Ker}\phi = \langle f \rangle$, we have

$$F[a] \cong F[x]/\langle f \rangle.$$

Since $F[a]$ is a subring of a field, it is an integral domain, so the ideal $\langle f \rangle$ must be a prime ideal in $F[x]$, and hence f is a prime element in $F[x]$. Since $F[x]$ is a principal ideal domain, it follows that f is irreducible in $F[x]$, and the ideal $\langle f \rangle$ is maximal, and hence $F[x]/\langle f \rangle$ is a field. Thus the ring $F[a] \cong F[x]/\langle f \rangle$ is actually a field, and so we have

$$F(a) = F[a].$$

Note that f is the unique monic irreducible polynomial with $f(a) = 0$, since if g is another monic irreducible polynomial with $g(a) = 0$ then, since $g(a) = 0$ we have $g \in \operatorname{Ker}\phi = \langle f \rangle$ so that $f \big| g$, hence f and g are associates (since f and g are irreducible), and hence $f = g$ (since f and g are monic).

Let $n = \deg f$. We claim that $\{1, a, a^2, \cdots, a^{n-1}\}$ spans $F(a) = F[a]$. Let $u \in F[a]$, say $u = g(a)$ where $g \in F[x]$. By the Division Algorithm, we can choose $q, r \in F[x]$ with $\deg r < n$ such that $g = qf + r$, say $r(x) = c_0 + c_1 x + \cdots + c_{n-1}x^{n-1}$ with $c_i \in F$. Since $f(a) = 0$ we have $u = g(a) = q(a)f(a) + r(a) = r(a) = \sum_{i=0}^{n-1} c_i a^i \in \operatorname{Span}\{1, a, \cdots, a^{n-1}\}$. Thus $\big\{1, a, \cdots, a^{n-1}\big\}$ spans $F(a) = F[a]$ as claimed. We claim that $\big\{1, a, \cdots, a^{n-1}\big\}$ is linearly independent. Let $c_0, c_1, \cdots, c_{n-1} \in F$ and suppose that $\sum_{i=0}^{n-1} c_i a^i = 0$. Let $g(x) = \sum_{i=0}^{n-1} c_i x^i \in F[x]$. Since f is a nonzero polynomial of minimal degree with $f(a) = 0$, and g is a polynomial with $\deg g < \deg f$ and $g(a) = 0$, it follows that $g = 0$, and hence $c_i = 0$ for all i. Thus $\{1, a, \cdots, a^{n-1}\big\}$ is linearly independent, as claimed.

3.8 Definition: When F and K are fields with $F \subseteq K$, and $a \in K$ is algebraic over F, the unique monic irreducible polynomial $f(x) \in K[x]$ with $f(a) = 0$ in K is called the **minimal polynomial** of a over F.

3.9 Exercise: Find the minimal polynomial of $\sqrt{1 + \sqrt{3}}$, and of $\sqrt{3 + 2\sqrt{5}}$, over $\mathbb{Q}$.

3.10 Exercise: Let $\theta = 2\cos\frac{\pi}{9}$. Find $[\mathbb{Q}(\theta) : \mathbb{Q}]$.

3.11 Exercise: Let F and K be fields. Let $u \in K$ be transcendental over F. Note that $F(u^2) \subseteq F(u)$. Find the minimal polynomial of $u + 1$ over $F(u^2)$.

3.12 Exercise: Let $f(x) = x^3 - 2$. Note that the roots of f in $\mathbb{C}$ are $\sqrt[3]{2}$, $\sqrt[3]{2}\,\omega$ and $\sqrt[3]{2}\,\omega^2$, where $\omega = e^{i\,2\pi/3}$. Let $K = \mathbb{Q}\big[\sqrt[3]{2}, \sqrt[3]{2}\,\omega, \sqrt[3]{2}\,\omega^2\big]$ and $L = \mathbb{Q}\big[\sqrt[3]{2}, \omega\big]$. Show that $K = L$ and find $[K : \mathbb{Q}]$.

3.13 Corollary: *Let F, K and L be fields with $F \subseteq K \subseteq L$ and let $a \in L$. Let $f(x) \in F[x]$ be the minimal polynomial for a over F and let $g(x) \in K[x]$ be the minimal polynomial for a over K. Then $g(x)\big|f(x)$ in $K[x]$.*

Proof: Since $f \in F[x]$ we also have $f \in K[x]$. Since $f \in K[x]$ with $f(a) = 0$, and g is the minimal polynomial of a over K, we have $f \in \langle g \rangle \subseteq K[x]$, and hence $g\big|f \in K[x]$.

3.14 Theorem: *Let F and K be fields with $F \subseteq K$. Then the following are equivalent:*
(1) $[K:F]$ is finite.
(2) K is algebraic and finitely generated as a field over F.
(3) There exist $a_1, \cdots, a_n \in K$, with each a_k algebraic over F, such that $K = F[a_1, \cdots, a_n]$.

Proof: Suppose that $[K : F]$ is finite. Note that K is algebraic over F since if we had an element $u \in K$ which was transcendental over F, then the set $\{1, u, u^2, \cdots\}$ would be linearly independent so that $[K:F] = \infty$. Also note that K is finitely generated because if $\{a_1, a_2, \cdots, a_n\}$ is a basis for K over F then we have $K = F[a_1, \cdots, a_n]$. Indeed given $u \in K$, we can write $u = \sum_{k=1}^{n} c_k a_k$ with each $c_k \in F$, and then for the polynomial $g(x_1, \cdots, x_n) = \sum_{i=1}^{n} c_i x_i \in F[x_1, \cdots, x_n]$, we have $u = g(a_1, \cdots, a_n) \in F[a_1, a_2, \cdots, a_n]$.

Suppose that K is algebraic and finitely generated over F. Since K is finitely generated over F we can choose $a_1, \cdots, a_n \in K$ such that $K = F(a_1, \cdots, a_n)$. Since K is algebraic over F, each a_k is algebraic over F, so we have $K = F(a_1, \cdots, a_n) = F[a_1, \cdots, a_n]$.

Suppose that $K = F[a_1, a_2, \cdots, a_n]$ with each $a_k \in K$ algebraic over F. Let $F_0 = F$ and $F_k = F[a_1, \cdots, a_k]$ for $1 \leq k \leq n$. Note that $F_n = K$ and $F_k = F_{k-1}[a_k]$ for $1 \leq k \leq n$. Since a_k is algebraic over F, it is also algebraic over F_{k-1}, so we have $\big[F_k : F_{k-1}\big] = d_k$ where d_k is the degree of the minimal polynomial of a_k over F_{k-1}. Since $F = F_0 \subseteq F_1 \subseteq \cdots \subseteq F_n = K$, by Theorem 3.4 we have

$$[K:F] = [F_n:F_0] = [F_n:F_{n-1}][F_{n-1}:F_{n-2}] \cdots [F_2:F_1][F_1:F_0] = d_n d_{n-1} \cdots d_2 d_1,$$

which is finite.

3.15 Corollary: *Let F, K and L be fields with $F \subseteq K \subseteq L$. If L is algebraic over K and K is algebraic over F then L is algebraic over F.*

Proof: Suppose that L is algebraic over K and K is algebraic over F. Let $u \in L$. Since u is algebraic over K we can choose $0 \neq g \in K[x]$ such that $g(u) = 0$, say $g(x) = \sum_{k=0}^{n} c_k x^k$ with each $c_k \in K$. Let $E = F[c_0, c_1, \cdots, c_n]$, and note that since each $c_i \in E$ we have $g \in E[x]$. Since $g \in E[x]$ and $g(u) = 0$, u is algebraic over E, and hence $[E(u) : E]$ is finite. Since $E = F[c_0, \cdots, c_n]$ with each c_k algebraic over F, the above theorrem implies that $[E:F]$ finite, and hence so is $[E(u):F] = [E(u):E][E:F]$. Since $[E(u):F]$ is finite, $E(u)$ is algebraic over F, so every element in $E(u)$ is algebraic over F, so in particular u is algebraic over F. Since $u \in L$ was arbitrary, L is algebraic over F, as required.

3.16 Corollary: *Let F be a subfield of K. Let $E = \big\{a \in K \,\big|\, a \text{ is algebraic over } F\big\}$. Then E is a field with $F \subseteq E \subseteq K$.*

Proof: Note that $F \subseteq E$ because every $a \in F$ is algebraic over F with minimal polynomial $x - a \in F[x]$. We claim that E is a subfield of K. Let $a, b \in E$. Then we have $a, b \in F[a, b]$, which is a field, and so $a + b$, $a - b$, ab and (if $b \neq 0$) $\frac{a}{b}$ all lie in $F[a, b]$. Since $F[a, b]$ is algebraic over F, each of the elements $a + b$, $a - b$, ab and (if $b \neq 0$) $\frac{a}{b}$ is algebraic over F, so they all lie in E. Thus E is a subfield of K, as claimed.

Geometric Constructions

3.17 Definition: Let S be a set in $\mathbb{R}^2$ which contains at least two points. A **line on** S is a line through any two distinct points in S, and a **circle on** S is a circle centred at one point in S which passes through another.

A point $p \in \mathbb{R}^2$ is **constructible in one step from** S when $p \in A \cap B$ for some $A \neq B$ where each of the sets A and B is either a line on S or a circle on S. We say that a point $p \in \mathbb{R}^2$ is **constructible from** S when there is a finite sequence of points $p_1, p_2, \cdots, p_n$ with $p_n = p$ such that each p_k is constructible in one step from $S \cup \{p_1, \cdots, p_{k-1}\}$. We say that a line L (or a circle C) is **constructible from** S when there is a finite set P of points, constructible from S, such that L is a line on $S \cup P$ (or C is a circle on $S \cup P$).

When a point (or line or circle) in $\mathbb{R}^2$ is constructible from the set $S_0 = \{(0,0),(1,0)\}$, we simply say that the point (or line or circle) is **constructible** (in $\mathbb{R}^2$). For $a \in \mathbb{R}$, we say that a is **constructible** (in $\mathbb{R}$) when $p = (a,0)$ is constructible in $\mathbb{R}^2$.

3.18 Note: Given two distinct points $a, b \in \mathbb{R}^2$, the perpendicular bisector of a and b is constructible from $\{a,b\}$ because it is the line through the two points of intersection of the circle C centred at a through b with the circle D centred at b through a.

Given two distinct points $a, b \in \mathbb{R}^2$ and a point $p \in \mathbb{R}^2$, let L be the line through a and b. Note that we can construct the line M through p perpendicular to L: indeed, if q is the point on L nearest to p then at least one of the two points a, b is not equal to q, say $a \neq q$, then the circle centred at p through a meets L at another point c, and the desired line M is the perpendicular bisector of a and c. It follows that we can also construct the line N through p parallel to L, which is the line through p perpendicular to M.

3.19 Note: From the set $S_0 = \{(0,0),(1,0)\}$ we can (of course) construct the x-axis, and we can construct the y-axis (since it is the line through $(0,0)$ perpendicular to the x-axis).

Note that for $a \in \mathbb{R}$, the point $(a,0)$ is constructible if and only if the point $(0,a)$ is constructible. Indeed when $a = 0$ we have $(a,0) = (0,0) = (0,a)$ and when $a \neq 0$, the circle centred at $(0,0)$ through $(a,0)$ intersects the y-axis at $(0,a)$ (and also at $(0,-a)$), and the circle centred at $(0,0)$ through $(0,a)$ intersects the x-axis at $(a,0)$.

Note that for $a, b \in \mathbb{R}$ and $p = (a,b) \in \mathbb{R}^2$, the point p is constructible in $\mathbb{R}^2$ if and only if a and b are both constructible in $\mathbb{R}$. Indeed, if $p = (a,b)$ is constructible in $\mathbb{R}^2$, then $(a,0)$ is constructible (since it is the point of intersection of the x-axis with the line through (a,b) perpendicular to the x-axis) and $(0,b)$ is constructible (since it is the point of intersection of the y-axis with the line through (a,b) perpendicularr to the y-axis). And conversely, if $(a,0)$ and $(b,0)$, hence also $(0,b)$, are constructible, then so is (a,b) (since it is the point of intersection of the line through $(a,0)$ perpendicular to the x-axis with the line through $(0,b)$ perpendicular to the y-axis).

3.20 Theorem: (*Constructible Points*) *Let F be the set of all constructible real numbers.*
(1) F a field with $\mathbb{Q}\subseteq F\subseteq\mathbb{R}$.
(2) For $a\in\mathbb{R}$ we have $a\in F$ if and only if there is a tower of fields

$$\mathbb{Q}=F_0\subseteq F_1\subseteq\cdots\subseteq F_n$$

with $a\in F_n$ such that for $1\le k\le n$ we have $F_k=F_{k-1}\big[\sqrt{u_k}\,\big]$ for some $0\le u_k\in F_{k-1}$.
(3) For $a\in\mathbb{R}$, if $a\in F$ then $\big[\mathbb{Q}(a):\mathbb{Q}\big]$ is a power of 2.

Proof: To prove Part 1, let $a,b\in F$, and note that the points $(a,0)$, $(0,a)$, $(b,0)$, $(0,b)$ and (a,b) are constructible. If $b=0$ then we have $a\pm b=a\in F$. If $b\neq 0$ then the circle centred at $(a,0)$ through (a,b) meets the x-axis at $(a\pm b,0)$ so we have $a\pm b\in F$. We can construct the line L through $(0,a)$ and $(1,0)$ (L has equation $y=a-ax$), and we can construct the line M through $(b,0)$ parallel to L (M has equation $y=ab-ax$) and then the line M intersects the y-axis at $(0,ab)$, so we have $ab\in F$. Suppose $b\neq 0$. We can construct the line J through $(0,a)$ and $(b,0)$ (J has equation $y=a-\frac{a}{b}x$), and we can construct the line K through $(1,0)$ parallel to J (K has equation $y=\frac{a}{b}-\frac{a}{b}x$) and then the line K intersects the y-axis at $\left(0,\frac{a}{b}\right)$ so we have $\frac{a}{b}\in F$. Thus F is a field with $F\subseteq\mathbb{R}$, and note that every subfield of $\mathbb{R}$ contains $\mathbb{Q}$ so we have $\mathbb{Q}\subseteq F\subseteq\mathbb{R}$. .

To prove Part 2, let $a\in\mathbb{R}$. Suppose that $a\in F$, so $p=(a,0)$ is constructible. Choose points $p_1,\cdots,p_n\in\mathbb{R}^2$ with $p_n=p$ such that each p_k is constructible in one step from the set $S_0\cup\{p_1,p_2,\cdots,p_k\}$, where $S_0=\big\{(0,0),(1,0)\big\}$. Say $p_k=(a_k,b_k)$, let $F_0=\mathbb{Q}$, and let $F_k=\mathbb{Q}\big[a_1,b_1,a_2,b_2,\cdots,a_k,b_k\big]$ for $1\le k\le n$. Note that $F_k=F_{k-1}[a_k,b_k]$ and $a=a_n\in F_n$.

Fix k with $1\le k\le n$. We claim that $F_k=F_{k-1}\big[\sqrt{u_k}\,\big]$ for some $0\le u_k\in F_{k-1}$. Since p_k is constructible in one step from $S_0\cup\{p_1,\cdots,p_{k-1}\}$, we have $p\in A\cap B$ for some $A\neq B$ where each of the sets A and B is either a line or a circle on $S_0\cup\{p_1,\cdots,p_{k-1}\}$. Note that each of the points in $S_0\cup\{p_1,\cdots,p_{k-1}\}$ has coordinates which lie in F_{k-1}.

Any line through two distinct points in F_{k-1} has an equation of the form $Ax+By+C=0$ with $A,B,C\in F_{k-1}$. When two such lines are distinct and have a point of intersection, the point of intersection is unique, and its coordinates lie in F_{k-1}, so when p_k is the point of intersection of two such lines, we have $F_k=F_{k-1}[a_k,b_k]=F_{k-1}$, so we can take $u_k=0$.

Any circle centred at one point in F_{k-1} through another point in F_{k-1} has an equation of the form $x^2+y^2+ax+by+c=0$ with $a,b,c\in F_{k-1}$. A line $Ax+By+C=0$ (1) and a circle $x^2+y^2+ax+by+c=0$ (2) can have 0, 1 or 2 points of intersection, and we can find the points of intersection by solving the two equations. In the case $B\neq 0$, we can write (1) in the form $y=dx+e$ with $d,e\in F_{k-1}$, and we can put this into (2) to obtain a quadratic in x with coefficients in F_{k-1}. Let u_k be the discriminant of this quadratic. When the quadratic has a unique solution x, we have $u_k=0$, and $x\in F_{k-1}=F_{k-1}[0]$ so that $p_k=(a_k,b_k)$ with $a_k=x\in F_{k-1}$ and $b_k=da_k+e\in F_{k-1}$. When the quadratic has two distinct real solutions x_1,x_2, we have $u_k>0$, and the two solutions both lie in $F_{k-1}[\sqrt{u_k}]$, so that $p_k=(a_k,b_k)$ with $a_k=x_1$ or x_2 so $a_k\in F_{k-1}[\sqrt{u_k}]$ and $b_k=da_k+e\in F_{k-1}[\sqrt{u_k}]$.

Note that for two circles C and D with C given by $x^2+y^2+ax+by+c=0$ (1) and D given by $x^2+y^2+dx+ey+f=0$ (2) with $a,b,c,d,e,f\in F_{k-1}$, subtracting (2) from (1) gives $(a-d)x+(b-e)y+(c-f)=0$ so that the points of intersection of C and D are equal to the points of intersection of C and L where L is the line $(a-d)x+(c-e)y+(c-f)=0$ (noting that if $a=d$ and $c=e$ then C and D have the same centre so they are either equal or have no points of intersection).

We have proven that if $a \in F$ then there is a tower of fields $\mathbb{Q} = F_0 \subseteq F_1 \subseteq \cdots \subseteq F_n$ with $a \in F_n$ such that $F_k = F_{k-1}[\sqrt{u}_k]$ for some $0 \le u_k \in F_{k-1}$. Suppose, conversely, that we have such a tower of fields. By Part 1, we have $F_0 = \mathbb{Q} \subseteq F$. Fix $k \ge 1$ and suppose, inductively, that $F_{k-1} \subseteq F$. We have $0 \le u_k \in F_{k-1}$ so we can construct the point $(u_k, 0)$, and hence we can construct the points $(u_k \pm 1, 0)$. The circle C centred at $(0,0)$ through $(u_k + 1, 0)$ meets the line through $(u_k - 1, 0)$ perpendicular to the x-axis at the point $\big(u_k - 1, 2\sqrt{u_k}\big)$. Thus $2\sqrt{u_k} \in F$, and hence $\sqrt{u}_k \in F$. Since $F_{k-1} \subseteq F$ and $\sqrt{u_k} \in F$ we have $F_k = F_{k-1}[\sqrt{u}_k] \subseteq F$. By induction, $F_k \subseteq F$ for all k. In particular, we have $a \in F_n$ and $F_n \subseteq F$ so that $a \in F$. This completes the proof of Part 2.

Let $a \in F$. By Part 2, we can choose a tower of fields $\mathbb{Q} = F_0 \subseteq F_1 \subseteq \cdots \subseteq F_n$ with $a \in F_n$ where $F_k = F_{k-1}[\sqrt{u}_k]$ with $0 \le u_k \in F_{k-1}$. Since $\sqrt{u_k}$ is a root of $f(x) = x^2 - u_k \in F_{k-1}[x]$, the minimal polynomial of $\sqrt{u}_k$ over F_{k-1} is of degree 1 or 2, so that we have $[F_k : F_{k-1}] \in \{1, 2\}$. Since $[F_n : \mathbb{Q}] = [F_n, F_0] = \prod_{k=1}^{k} [F_k : F_{k-1}]$ with each $[F_k : F_{k-1}] \in \{1, 2\}$, we see that $[F_n : \mathbb{Q}]$ is a power of 2, say $[F_n : \mathbb{Q}] = 2^\ell$. Since $a \in F_n$ we have $2^\ell = [F_n : \mathbb{Q}] = [F_n : \mathbb{Q}(a)][\mathbb{Q}(a) : \mathbb{Q}]$, and hence $[\mathbb{Q}(a) : \mathbb{Q}]$ divides 2^ℓ, and so $[\mathbb{Q}(a) : \mathbb{Q}]$ is a power of 2.

3.21 Example: The real number $\sqrt[3]{2}$ is not constructible (so we cannot construct the vertices of a square of area 2).

3.22 Exercise: Show that, for a real number θ, we can construct $\cos\theta$ if and only if we can construct $\sin\theta$ if and only if we can construct $\cos 2\theta$ and/or $\sin 2\theta$.

3.23 Exercise: Show that we can construct the real number $\cos\frac{\pi}{5}$ (so we can construct the vertices of a regular pentagon or decagon), but we cannot construct the real number $\cos\frac{\pi}{9}$ (so we cannot construct the vertices of a regular 9-gon or 18-gon).

3.24 Example: It has been shown that the real numbers e and π are transcendental, so we cannot construct them (and hence we cannot construct a circle of area 1).

Splitting Fields

3.25 Definition: Let F be a subfield of K and let $f \in F[x]$ be a polynomial with $n = \deg f \in \mathbb{Z}^+$. We say that f **splits** over K (or that f splits in $K[x]$) when f factors as a product of linear factors in $K[x]$, that is when $f(x) = c\,(x-a_1)(x-a_2)\cdots(x-a_n) \in K[x]$ for some $0 \neq c \in F$ and some $a_1, a_2, \cdots, a_n \in K$ (not necessarily distinct). In this case the field $F[a_1, a_2, \cdots, a_n] \subseteq K$ is called the **splitting field** of f over F in K.

3.26 Theorem: (*Kronecker's Theorem*) *Let F be a field and let $f \in F[x]$ be a nonconstant polynomial. Then there exists an extension field K of F which contains a root of f.*

Proof: Let $g \in F[x]$ be an irreducible factor of f. Note that it suffices to construct an extension field K of F which contains a root of g (since if $f = gh$ and $g(a) = 0$ then $f(a) = g(a)h(a) = 0$ so that a is a root of f). Let $L = F[x]/\langle g\rangle$. Since g is irreducible, the ideal $\langle g\rangle$ is maximal, so L is a field. Note that the homomorphism $\phi : F \to L$ given by $\phi(c) = c + \langle g\rangle$ is injective: indeed if $c, d \in F$ and $c + \langle g\rangle = d + \langle g\rangle$ then we have $d - c \in \langle g\rangle$ and hence $d - c = 0$ (because 0 is the only constant polynomial in $\langle g\rangle$). Let $E = \phi(F)$ and note that $\phi : F \to E$ is an isomorphism. Extend ϕ to the isomorphism $\phi : F[x] \to E[y]$ given by $\phi\big(\sum_{k=0}^n c_k x^k\big)(y) = \sum_{k=0}^n (c_k + \langle g\rangle) y^k$. Note that the element $x + \langle g\rangle \in L$ is a root of $\phi(g)$ because if $g(x) = \sum a_k x^k$ then we have

$$\phi(g)(x + \langle g\rangle) = \sum_{k=0}^{n} (a_k + \langle g\rangle)\big(x + \langle g\rangle\big)^k = \Big(\sum_{k=0}^{n} a_k x^k\Big) + \langle g\rangle = g + \langle g\rangle = 0 + \langle g\rangle,$$

which is the zero element in L.

If we are willing to identify F with E (by identifying $c \in F$ with $\phi(c) = c + \langle g\rangle \in E$) and to identify $f \in F[x]$ with $\phi(f) \in E[y]$, then we can take $K = L$ and we have obtained a field extension of F which contains a root of f. If we insist on constructing a field K which actually contains F as a subfield, we can do so as follows. Choose any set A which is disjoint from F and has the same cardinality as $L \setminus E$, let $\theta : A \to L \setminus E$ be any bijection, and let $K = F \cup A$. Let $\psi : K \to L$ be the bijection given by $\psi(u) = \phi(u)$ when $u \in F$ and $\psi(u) = \theta(u)$ when $u \in A$. Use ψ to pull the field operations from L back to K by defining $u + v = \psi^{-1}\big(\psi(u) + \psi(v)\big)$ and $u \cdot v = \psi^{-1}\big(\psi(u) \cdot \psi(v)\big)$ for $u, v \in K$. Using these operations on K, K is an extension field of F, and $\psi : K \to L$ is an isomorphism with $\psi(u) = \phi(u)$ for all $u \in F$, and the element $a = \psi^{-1}\big(x + \langle g\rangle\big) \in K$ is a root of f.

3.27 Corollary: *Let F be a field and let $f \in F[x]$ be a nonconstant polynomial. Then there exists a splitting field for f over F.*

Proof: We apply Kronecker's Theorem repeatedly: Let g_1 be an irreducible factor of f in $F[x]$. Let K_1 be an extension field of F which contains a root a_1 of g_1. Let $F_1 = F[a_1] \subseteq K_1$. Note that $f(a_1) = 0$ in F_1 so that $(x - a_1)\big|f(x)$ in $F_1[x]$, say $f(x) = (x - a_1)f_2(x)$. If $\deg f = 1$ so that $f_2(x)$ is a constant polynomial, we are done. Otherwise repeat the argument. Let g_2 be an irreducible factor of f_2 in $F_1[x]$. Let K_2 be an extension field of F_1 which contains a root a_2 of g_2 and let $F_2 = F_1[a_2] = F[a_1, a_2]$. Note that $f_2(a_2) = 0$, say $f_2(x) = (x - a_2)f_3(x)$ and note that we have $f(x) = (x - a_1)(x - a_2)f_3(x) \in F_2[x]$. If $\deg f = 2$ we are done and, if not, we repeat.

3.28 Theorem: Let F *be a subfield of* K, *let* E *be a subfield of* L, *let* $\phi : F \to E$ *be an isomorphism of fields, and extend* ϕ *to obtain an isomorphism of rings* $\phi : F[x] \to E[x]$ *given by* $\phi\big(\sum_{k=0}^{n} c_k x^k\big) = \sum_{k=0}^{n} \phi(c_k) x^k$. *Let* $f \in F[x]$ *be an irreducible polynomial, and let* $g = \phi(f) \in E[x]$. *Let* $a \in K$ *be a root of* f *in* K *and let* $b \in L$ *be a root of* g *in* L. *Then the isomorphism* $\phi : F \to E$ *extends uniquely to an isomorphism* $\phi : F[a] \to E[b]$ *such that* $\phi(a) = b$.

Proof: Let $n = \deg f$. Note that $g = \phi(f)$ is irreducible with $\deg g = n$. By Theorem 3.7, the set $A = \{1, a, a^2, \cdots, a^{n-1}\}$ is a basis for $F[a]$ over F, and the set $B = \{1, b, b^2, \cdots, b^{n-1}\}$ is a basis for $E[b]$ over E. The desired extension $\phi : F[a] \to E[b]$ must be given by $\phi\big(\sum_{k=0}^{n-1} c_k a^k\big) = \sum_{k=0}^{n-1} \phi(c_k) b^k$. This map is bijective since A and B are bases for $F[a]$ ov er F and $E[b]$ over E, and it is easy to see that this map is a ring homomorphism, so it is an isomorphism.

3.29 Corollary: Let F *be a field and let* $f \in F[x]$ *be a nonconstant polynomial. Let* K *and* L *be splitting fields of* f *over* F. *Then there is an isomorphism of fields* $\phi : K \to L$ *with* $\phi(x) = x$ *for every* $x \in F$.

Proof: Let K and L be splitting fields for f over F , with $K = F[a_1, \cdots, a_\ell]$ where $a_1, \cdots, a_\ell$ are the distinct roots of f in K, and $L = F[b_1, \cdots, b_m]$ where $b_1, \cdots, b_m$ are the distinct roots of f in L. If f splits over F, then $K = L = F$ and we are done. Suppose f does not split over F. Let g_1 be an irreducible factor of f in $F[x]$ with $g_1(a_1) = 0$ in K. Let $h_1 = g_1$ and reorder the roots $b_1, \cdots, b_m$, if necessary, so that b_1 is a root of h_1 in L. By Theorem 3.28, we can extend the identity map $I : F \to F$ to a field isomorphism $\phi : F[a_1] \to F[b_1]$ with $\phi(a_1) = b_1$. We also write $\phi : F[a_1][x] \to F[b_1][x]$ to denote the associated ring isomorphism. Note that f splits over $F[a_1]$ if and only if f splits over $F[b_1]$ and, in this case, we have $K = f[a_1]$ and $L = F[b_1]$ we are done. Suppose that f does not split over $F[a_1]$ or over $F[b_1]$. Let g_2 be an irreducible factor of f in $F[a_1][x]$ with $g_2(a_2) = 0$. Let $h_2 = \phi(g_2)$ and note that h_2 is an irreducible factor of f in $F[b_1][x]$. Also note that h_2 has a root in L which is distinct from b_1 (if b_1 was the only root of h_2 in L then h_2 would split over $F[b_1]$, hence g_1 would split over $F[a_1]$ with a_1 as its only root). Reorder the roots $b_2, \cdots, b_m$, if necessary, so that b_2 is a root of h_2 in L. By Theorem 3.28, we can extend the isomorphism $\phi : F[a_1] \to F[b_1]$ to an isomorphism $\phi : F[a_1, a_2] \to F[b_1, b_2]$ with $\phi(a_1) = b_1$ and $\phi(a_2) = b_2$. We also write $\phi : F[a_1, a_2][x] \to F[b_1, b_2][x]$ to denote the associated ring isomorphism. Note that f splits over $F[a_1, a_2]$ if and only if f splits over $F[b_1, b_2]$ and, in this case, we have $K = F[a_1, a_2]$ and $L = F[b_1, b_2]$ and we are done. Otherwise, we repeat the above procedure until f splits.

3.30 Exercise: For each of the following polynomials $f \in \mathbb{Q}[x]$, find the splitting field K of f over $\mathbb{Q}$ in $\mathbb{C}$, and find $[K : \mathbb{Q}]$.

(a) $f(x) = x^4 - 4$. (b) $f(x) = x^4 - 2$. (c) $f(x) = x^5 - 1$.

3.31 Exercise: Let $f(x) = x^2 + x + 1 \in \mathbb{Z}_2[x]$ and note that f is irreducible (because it has no roots in $\mathbb{Z}_2$). Let K be a splitting field for f over $\mathbb{Z}_2$, and let $a \in K$ be a root of f in K. Note that $[K : \mathbb{Z}_2] = 2$ and $\{1, a\}$ is a basis for K over $\mathbb{Z}_2$, and so we have $K = \{s \cdot 1 + t \cdot a \,|\, s, t \in \mathbb{Z}_2\} = \{0\,,\,1\,,\,a\,,\,a+1\}$. Determine the addition and multiplication tables in K.

Multiple Roots

3.32 Definition: Let F be a field. For $f(x)=\sum\limits_{k=0}^{n}a_kx^k\in F[x]$, we define the (formal) **derivative** of f to be

$$\tfrac{d}{dx}f(x)=f'(x)=\sum_{k=0}^{n}k\,a_kx^{k-1}.$$

3.33 Theorem: *Let F be a field. For all $f,g\in F[x]$ and $c\in F$, we have*

$$(cf)'=cf'\ ,\ (f+g)'=f'+g'\ ,\ (fg)'=f'g+fg'\ \text{ and }\ (f\circ g)'=(f'\circ g)\,g'.$$

Proof: We leave the proof that $(cf)'=cf'$ and $(f+g)'=f'+g'$ as an exercise. Let $f(x)=\sum\limits_{k=0}^{n}a_kx^k$ and $g(x)=\sum\limits_{\ell=0}^{m}b_\ell x^\ell$. We have

$$\begin{aligned}(fg)'(x)&=\tfrac{d}{dx}\Big(\sum_{k=0}^{n}\sum_{\ell=0}^{m}a_kb_\ell x^{k+\ell}\Big)=\sum_{k=0}^{n}\sum_{\ell=0}^{m}(k+\ell)a_kb_\ell x^{k+\ell-1},\text{ and}\\ (f'g+fg')(x)&=\Big(\sum_{k=0}^{n}k\,a_kx^{k-1}\Big)\Big(\sum_{\ell=0}^{m}b_\ell x^\ell\Big)+\Big(\sum_{k=0}^{n}a_kx^k\Big)\Big(\sum_{\ell=0}^{m}\ell\,b_\ell x^{\ell-1}\Big)\\ &=\sum_{k=0}^{n}\sum_{\ell=0}^{m}k\,a_kb_\ell x^{k+\ell-1}+\sum_{k=0}^{n}\sum_{\ell=0}^{m}\ell\,a_kb_\ell x^{k+\ell-1}\\ &=\sum_{k=0}^{n}\sum_{\ell=0}^{m}(k+\ell)a_kb_\ell x^{k+\ell-1}=(fg)'(x).\end{aligned}$$

This proves the Product Rule. To prove the Chain Rule, we first claim that

$$\tfrac{d}{dx}g(x)^k=k\,g(x)^{k-1}g'(x)$$

for all $k\geq 0$. This clearly holds when $k=0$. Suppose, inductively, it holds for some $k\geq 0$. Then, by the Product Rule, we have

$$\begin{aligned}\tfrac{d}{dx}g(x)^{k+1}&=\tfrac{d}{dx}\big(g(x)^kg(x)\big)=\tfrac{d}{dx}g(x)^k\cdot g(x)+g(x)^kg'(x)\\ &=k\,g(x)^{k-1}g'(x)\cdot g(x)+g(x)^kg'(x)=(k+1)g(x)^kg'(x).\end{aligned}$$

By induction, we have $\frac{d}{dx}g(x)^k=k\,g(x)^{k-1}g'(x)$ for all $k\geq 0$, as claimed. Thus

$$\begin{aligned}(f\circ g)'(x)&=\tfrac{d}{dx}\sum_{k=0}^{n}a_kg(x)^k=\sum_{k=0}^{n}a_k\tfrac{d}{dx}g(x)^k=\sum_{k=0}^{n}a_k\cdot k\,g(x)^{k-1}g'(x)\\ &=\Big(\sum_{k=0}^{n}k\,a_k\,g(x)^{k-1}\Big)g'(x)=(f'\circ g)(x)\cdot g'(x).\end{aligned}$$

3.34 Theorem: *Let F be a field and let $f \in F[x]$ be a non-constant polynomial. Then f has no repeated roots in its splitting field if and only if* $\gcd(f, f') = 1$.

Proof: Let K be a splitting field for f over F. Note that we can consider f to lie in $F[x]$ or in $K[x]$. When we calculate $f'(x)$, the coefficients of f' lie in F, so the derivative of f in $F[x]$ is equal to the derivative of f in $K[x]$. The same holds for $\gcd(f, f')$: when we calculate $\gcd(f, f')$ using the Euclidean Algorithm, at each step in our calculation the coefficients of all the polynomials lie in F, and in particular, all of the coefficients of $\gcd(f, f')$ lie in F, so the greatest common divisor of f and f' in $K[x]$ is equal to the greatest common divisor of f and f' in $F[x]$.

Suppose f has a repeated root in K, say $f(x) = (x-a)^2 g(x) \in K[x]$. Then we have $f'(x) = 2(x-a)g(x) + (x-a)^2 g'(x) \in K[x]$. Thus $(x-a)$ divides both $f(x)$ and $f'(x)$ so that $(x-a)$ divides $\gcd(f, f')$ in $K[x]$. Thus the degree of $\gcd(f, f')$ is at least 2 (in $K[x]$ and in $F[x]$), and so $\gcd(f, f') \neq 1$. .

Suppose that f has no repeated roots in K, say $f(x) = c(x-a_1)(x-a_2)\cdots(x-a_n) \in K[x]$ with the elements $a_k \in K$ all distinct. Then we have $f'(x) = c\sum_{k=1}^{n}\prod_{i\neq k}(x-a_i)$. For each k we have $f'(a_k) = c\prod_{i\neq k}(a_k - a_i) \neq 0$, so none of the linear polynomials $(x-a_i)$ divides f' in $K[x]$. Thus we have $\gcd(f, f') = 1$ in $K[x]$, hence also in $F[x]$.

3.35 Corollary: *Let F and let $f \in F[x]$ be irreducible. Then f has a repeated root in its splitting field if and only if $f' = 0$. So when* $\mathrm{char}(F) = 0$, *f has no repeated roots, and when* $\mathrm{char}(F) = p$ *with p prime, f has a repeated root if and only if f is of the form $g(x^p)$ for some $g \in F[x]$.*

Solution: If $f' = 0$ then $\gcd(f, f') = f' \neq 1$, so f has a repeated root in its splitting field. Suppose that $f' \neq 0$. Let g be a common factor f and f' in $F[x]$. Since $g \big| f'$ and $f' \neq 0$ we have $\deg(g) \leq \deg(f') < \deg(f)$. Since $g \big| f$ and f is irreducible, either g is a unit or g is an associate of f, so either $\deg(g) = 0$ or $\deg(g) = \deg(f)$. Thus we must have $\deg(g) = 0$. So the common divisors of f and f' in $F[x]$ are the non-zero constant polynomials, so we have $\gcd(f, f') = 1$, hence f has no repeated roots.

Finally, note that when $\mathrm{char}(F) = 0$ we have $f' \neq 0$ (indeed $\deg(f') = \deg(f) - 1$) and when $\mathrm{char}(F) = p$ we have $f' = 0$ if and only if f is of the form $f(x) = \sum_{k=0}^{n} a_k x^{kp}$ for some $a_k \in F$, if and only if f is of the form $f(x) = g(x^p)$ for some $g \in F[x]$.

3.36 Example: Consider the polynomial $f(x) = x^2 + u \in \mathbb{Z}_2(u)[x]$ where u is a variable symbol, so that $\mathbb{Z}_2(u) = \left\{ \frac{f(u)}{g(u)} \,\middle|\, f(u), g(u) \in \mathbb{Z}_2[u], g(u) \neq 0 \right\}$, where $\mathbb{Z}_2[u]$ is the ring of polynomials over $\mathbb{Z}_2$ in the variable u. Note that f has no roots in $\mathbb{Z}_2(u)$ because if we had $\left(\frac{f(u)}{g(u)}\right)^2 + u = 0 \in \mathbb{Z}_2(u)$ then we would have $f(u)^2 = -u\,g(u)^2$ in $\mathbb{Z}_2[u]$, but this is not possible since the polynomial $f(u)^2$ has even degree but the polynomial $u\,g(u)^2$ has odd degree. Since $f(x)$ has degree 2 and has no roots in $\mathbb{Z}_2(u)$, it is irreducible in $\mathbb{Z}_2(u)[x]$. But since $f'(x) = 2x = 0 \in \mathbb{Z}_2(u)[x]$, it follows (from the above corollary) that f has a repeated root in its splitting field. Indeed we do not need to rely on the above corollary as it is easy to check that if K is a splitting field of f, and $a \in K$ is a root of f in K, then we have $a^2 = u \in K$ and we have $(x-a)^2 = x^2 - 2a\,x + a^2 = x^2 + a^2 = x^2 + u = f(x) \in K[x]$.

Finite Fields

3.37 Definition: When X is a subset of Y, and $\phi : X \to Y$ is any function, the **fixed point set** of ϕ is the set

$$\mathrm{Fix}(\phi) = \big\{x\in X \,\big|\, \phi(x)=x\big\}.$$

When R is a subring of S and $\phi : R \to S$ is a ring homomorphism, note that $\mathrm{Fix}(\phi)$ is a subring of R: indeed if $a, b \in \mathrm{Fix}(\phi)$ so that $\phi(a) = a$ and $\phi(b) = b$, then we have $\phi(a + b) = \phi(a) + \phi(b) = a + b$ and $\phi(ab) = \phi(a)\phi(b) = ab$ so that $a + b \in \mathrm{Fix}(\phi)$ and $ab \in \mathrm{Fix}(\phi)$. When F is a subfield of K and $\phi : F \to K$ is a non-zero homomorphism, the fixed point set $\mathrm{Fix}(\phi)$ is a subfield of F: indeed recall (or verify) that $\phi(1) = 1$, so when $0 \neq a \in \mathrm{Fix}(\phi)$ so that $\phi(a) = a$, then we have $\phi(a) \cdot \phi\big(\frac{1}{a}\big) = \phi\big(a \cdot \frac{1}{a}\big) = \phi(1) = 1$ and hence $\phi\big(\frac{1}{a}\big) = \frac{1}{\phi(a)} = \frac{1}{a}$ so that $\frac{1}{a} \in \mathrm{Fix}(\phi)$.

3.38 Definition: When R is a commutative ring with 1 with prime characteristic p, the **Frobenius map** is the map $\phi : F \to F$ given by $\phi(x) = x^p$. Note that ϕ is a ring homomorphism because $\phi(xy) = (xy)^p = x^p y^p$ and $\phi(x + y)^p = \sum \binom{p}{k} x^k y^{p-k} = x^p + y^p$. When R is an integral domain, this map ϕ is injective since $x^p = 0 \Longrightarrow x = 0$. When F is a finite field, every injective map from F to F is bijective so, in particular, the Frobenius map ϕ is bijective. Also note that ϕ fixes every element in the prime subfield E because $E \cong \mathbb{Z}_p$ and for every $x \in \mathbb{Z}_p$ we have $x^p = x$ (by Fermat's Little Theorem), so $E \subseteq \mathrm{Fix}(\phi)$.

3.39 Theorem: I*f F is a finite field then the multiplicative group of units F^* is cyclic.*

Proof: Let F be a finite field. We claim that the multiplicative group of units F^* is cyclic. By the Classification of Finite Abelian groups, we have $F^* \cong C_{n_1} \times C_{n_2} \times \cdots \times C_{n_\ell}$ where C_n is the standard multiplicative cyclic group of order n (the group of n^{th} roots of unity in $\mathbb{C}^*$) and $n_k \big| n_{k+1}$. Note that every $a \in F^*$ satisfies $a^{n_\ell} = 1$, so every $a \in F^*$ is a root of $f(x) = x^{n_\ell} - 1 \in F[x]$. Since F is a field, f has at most n_ℓ roots, so we must have $\ell = 1$ and $n_\ell = |F^*|$ so that F^* is cyclic, as claimed.

3.40 Theorem: (*The Classification of Finite Fields*)

(1) If F is a finite field then $|F| = p^n$ for some prime number $p \in \mathbb{Z}^+$ and some $n \in \mathbb{Z}^+$.
(2) For every prime $p \in \mathbb{Z}^+$ and every $n \in \mathbb{Z}^+$ there is, up to isomorphism, a unique field F with $|F| = p^n$. This field F is the splitting field of $f(x) = x^{p^n} - x$ over the prime subfield E. Indeed, f has p^n distinct roots in F, and F is equal to the set of roots of f.

Proof: Recall that every finite field has prime characteristic, and in a field of prime characteristic p, the prime subfield is isomorphic to $\mathbb{Z}_p$ $\big($the prime subfield is the field $E = \big\{0 \cdot 1, 1 \cdot 1, 2 \cdot 1, \cdots, (p-1) \cdot 1\big\} \cong \mathbb{Z}_p\big)$. To prove Part 1, let F be a finite field. Let $p = \mathrm{char} F$, and let E be the prime subfield of F. Let $n = [F : E] = \dim_E F$. Note that n is finite (since any basis for F is a subset of F, which must be finite). Let $\{u_1, u_2, \cdots, u_n\}$ be a basis for F over E. Since each element in F can be expressed uniquely as a linear combination $\sum_{k=1}^n t_k u_k$, and we have p choices for each of the n elements $t_k \in E$, it follows that $|F| = p^n$.

To prove Part 2, let $p \in \mathbb{Z}^+$ be prime and let $n \in \mathbb{Z}^+$. First let us prove that there exists a field F with $|F| = p^n$. Let F be the splitting field of $f(x) = x^{p^n} - x$ over $\mathbb{Z}_p$. Note that since $f'(x) = p^n x^{p^n - 1} - 1 = -1$ we have $\gcd(f, f') = 1$ so that f has p^n distinct roots in F (so F has at least p^n elements). Note that since the Frobenius map $\phi(x) = x^p$ is an automorphism of F which fixes elements in E, so is the map $\psi = \phi^n$ given by $\psi(x) = x^{p^n}$. For $a \in F$, note that a is a root of f if and only if $a^{p^n} = a$ if and only if $\psi(a) = a$, so the set of roots of f in F is equal to $\text{Fix}(\psi) = \left\{a \in F \,\middle|\, \psi(a) = a\right\}$, which is a subfield of F. Since F is the splitting field of f over E, which is the smallest subfield of F which contains all the roots of f, and since the set of all roots of f is a field, it follows that F is equal to the set of all the roots of f, and so F has exactly p^n elements.

Now let us prove uniqueness. Suppose that F is any field with $|F| = p^n$. Let E be the prime subfield and note that $E \cong \mathbb{Z}_p$. Since F^* is cyclic of order $p^n - 1$, every $a \in F^*$ satisfies $a^{p^n - 1} = 1$, and hence every $a \in F$ (including $a = 0$) satisfies $a^{p^n} = a$. Thus every element in F is a root of $f(x) = x^{p^n} - x \in E[x]$. Since f has at most p^n roots in F, it follows that the the roots of f are distinct in F, and the elements in F are equal to the roots of f, hence F is the splitting field of f over E.

3.41 Corollary: If *F is a finite field and $n \in \mathbb{Z}^+$ then there exists an extension field K of F with $[K : F] = n$. This extension field K is unique up to isomorphism and it is of the form $K = F(a)$ for some $a \in K$.*

3.42 Corollary: If *F is a finite field and $n \in \mathbb{Z}^+$ then there exists an irreducible polynomial $f \in F[x]$ of degree n.*

3.43 Corollary: Let *K be a finite field with $|K| = p^n$ where $p \in \mathbb{Z}^+$ is prime and $n \in \mathbb{Z}^+$. If F is a subfield of K then $|F| = p^d$ for some divisor d of n. Conversely, for every divisor d of n, there is exactly one subfield F of K with $|K| = p^d$. This subfield F is the splitting field of (and the set of roots of) $f(x) = x^{p^d} - x$ over the prime subfield E in K.*

3.44 Corollary: Let *$p \in \mathbb{Z}^+$ be prime and let $n \in \mathbb{Z}^+$. In the ring $\mathbb{Z}_p[x]$, the polynomial $f(x) = x^{p^n} - x$ is equal to the product of all the distinct monic irreducible polynomials in $\mathbb{Z}_p[x]$ whose degree divides n.*

3.45 Example: In $\mathbb{Z}_2[x]$, the irreducible polynomials of degree 1 are x and $x+1$, and the irreducible polynomials of degree 3 are $x^3 + x + 1$ and $x^3 + x^2 + 1$, and we have

$$x^{2^3} - x = x^8 + x = x(x+1)(x^3 + x + 1)(x^3 + x^2 + 1).$$