Homework for
Introduction to Abstract Algebra II

Nicholas Camacho
Department of Mathematics
University of Iowa
Spring 2017

Most exercises are from
Abstract Algebra (3rd Edition) by Dummit & Foote.
For example, "4.2.8" means exercise 8
from section 4.2 in Dummit & Foote .
Beware: Some solutions may be incorrect!

In these exercises R is a ring with 1 and M is a left R-module.

Exercise 10.1.5. For any left ideal I of R define

$$IM = \left\{ \sum_{\text{finite}} a_i m_i \mid a_i \in I,\ m_i \in M \right\}$$

to be the collection of all finite sums of elements of the form am where $a \in I$ and $m \in M$. Prove that IM is a submodule of M.

Proof. Since $0_R \in I$ and $0_M \in M$, then $0_R 0_M = 0_M \in IM$ and so $IM \neq \varnothing$. Let $r \in R$ and let

$$x = \sum_{i=1}^{n} a_i m_i, \quad y = \sum_{j=1}^{m} a_j m_j$$

be elements of IM for $a_i, a_j \in I$ and $m_i, m_j \in M$ for all i and j. Then,

$$\begin{aligned} x + ry &= \sum_{i=1}^{n} a_i m_i + r \left(\sum_{j=1}^{m} a_j m_j \right) \\ &= \sum_{i=1}^{n} a_i m_i + \sum_{j=1}^{m} r a_j m_j \end{aligned}$$

is a finite sum of products of elements from I and M since $ra_j \in I$ and hence IM is a submodule of M. ☕

Exercise 10.1.9. If N is a submodule of M, the *annihilator of N in R* is defined to be $\{r \in R \mid rn = 0 \text{ for all } n \in N\}$. Prove that the annihilator of N is R is a 2-sided ideal of R.

Proof. Let $\mathrm{Ann}_R(N) = \{r \in R \mid rn = 0 \text{ for all } n \in N\}$. Notice that $\mathrm{Ann}_R(N) \neq \varnothing$ since $0_R n = 0_N$ for all $n \in N$. Let $x, y \in \mathrm{Ann}_R(N)$ and $n \in N$. Then $(x-y)n = xn - yn = 0_N - 0_N = 0_N$, where the first equality holds by a module axiom, and so $x - y \in \mathrm{Ann}_R(N)$. If $r \in R$ then $(rx)n = r(xn) = r(0_N) = 0_N$ where the first equality holds by a module axiom. We also have by a module axiom that $(xr)n = x(rn) = 0_N$, since $rn \in N$. Hence $xr, rx \in \mathrm{Ann}_R(N)$ and thus $\mathrm{Ann}_R(N)$ is an ideal of R. ☕

Exercise 10.1.10. If I is a right ideal of R, the *annihilator of I in M* is defined to be $\{m \in M \mid am = 0 \text{ for all } a \in I\}$. Prove that the annihilator of I in M is a submodule of M.

Proof. Let $\mathrm{Ann}_M(I) = \{m \in M \mid am = 0 \text{ for all } a \in I\}$. Since $a0_M = 0_M$ for all $a \in I$, then $0_M \in \mathrm{Ann}_M(I)$ and so $\mathrm{Ann}_M(I) \neq \varnothing$. Let $x, y \in \mathrm{Ann}_M(I)$, $r \in R$, and $a \in I$. Then $x + ry \in M$ and $ar \in I$, and so by module axioms

$$a(x + ry) = ax + a(ry) = 0 + (ar)y = 0 + 0 = 0.$$

So $x + ry \in \mathrm{Ann}_M(I)$ and thus $\mathrm{Ann}_M(I)$ is a submodule of M. ☕

Exercise 10.2.4. Let A be an $\mathbb{Z}$-module, let a be any element of A and let n be a positive integer. Prove that the map $\varphi_a : \mathbb{Z}/n\mathbb{Z} \to A$ given by $\varphi_a(\overline{k}) = ka$ is a well-defined $\mathbb{Z}$-module homomorphism if and only if $na = 0$.

Proof. ($\Rightarrow$) We have $na = \varphi_a(\overline{n}) = \varphi_a(\overline{0}) = 0a = 0$.

($\Leftarrow$) To show that φ_a is well defined, suppose $\overline{k} = \overline{\ell}$. Then $\overline{k} - \overline{\ell} = \overline{k-\ell} = \overline{0} = \overline{n}$ and so

$$ka - \ell a = (k-\ell)a = \varphi_a(\overline{k-\ell}) = \varphi_a(\overline{n}) = na = 0$$

and hence $\varphi_a(\overline{k}) = ka = \ell a = \varphi_a(\overline{\ell})$.

Let $\overline{k}, \overline{\ell} \in \mathbb{Z}/n\mathbb{Z}$ and $m \in \mathbb{Z}$. Then $m\overline{k} + \overline{\ell} = \overline{mk} + \overline{\ell} = \overline{mk+\ell}$ and so

$$\varphi_a(m\overline{k} + \overline{\ell}) = \varphi_a(\overline{mk+\ell}) = (mk+\ell)a = (mk)a + \ell a = m(ka) + \ell a = m\varphi_a(\overline{k}) + \varphi_a(\overline{\ell}).$$

☕

Prove that $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/n\mathbb{Z}, A) \cong A_n$ where $A_n = \{a \in A \mid na = 0\}$ (so A_n is the annihilator in A of the ideal (n) of $\mathbb{Z}$ — cf. Exercise 10.1.10)

Proof. Since A_n is the annihilator of the ideal (n) of $\mathbb{Z}$ in A, it is a $\mathbb{Z}$-submodule by Exercise 10.1.10. Moreover, by Proposition 2(2), $\mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/n\mathbb{Z}, A)$ is a $\mathbb{Z}$-module. So we define a map Φ of $\mathbb{Z}$–modules

$$\begin{aligned} \Phi : A_n &\to \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/n\mathbb{Z}, A) \\ a &\mapsto \varphi_a. \end{aligned}$$

and show that Φ is an isomorphism.

Let $x, y \in A_n$, $m \in \mathbb{Z}$, and $\overline{k} \in \mathbb{Z}/n\mathbb{Z}$. Then $mx + y \in A_n$, and by the previous proof $\varphi_{mx+y} \in \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/n\mathbb{Z}, A)$. Then

$$\varphi_{mx+y}(\overline{k}) = k(mx+y) = k(mx) + ky = (km)x + ky = m(kx) + ky = m\varphi_x(\overline{k}) + \varphi_y(\overline{k}).$$

So

$$\Phi(mx+y) = \varphi_{mx+y} = m\varphi_x + \varphi_y = m\Phi(x) + \Phi(y),$$

and so Φ is an $\mathbb{Z}$-module homomorphism.

Suppose $\varphi_x = \varphi_y$. Then

$$x = 1x = \varphi_x(\overline{1}) = \varphi_y(\overline{1}) = 1y = y,$$

and so Φ is injective. Let $\varphi \in \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/n\mathbb{Z}, A)$. Then $\varphi(\overline{1}) = a$ for some $a \in A$ and for $\overline{k} \in \mathbb{Z}/n\mathbb{Z}$

$$\varphi(\overline{k}) = \varphi(\underbrace{\overline{1} + \cdots + \overline{1}}_{k-summands}) = \underbrace{\varphi(\overline{1}) + \cdots + \varphi(\overline{1})}_{k-summands} = ka.$$

Hence $\Phi(a) = \varphi_a = \varphi$ and so Φ is surjective. ☕

Exercise 10.3.7. Let N be a submodule of M. Prove that if both M/N and N are finitely generated, then so is M.

Proof. By hypothesis, we have a finite subset $A = \{a_1, \dots, a_n\} \subseteq M$ for which

$$RA = Ra_1 + \cdots + Ra_n = N.$$

Similarly, we have a finite subset of distinct coset representatives $B = \{b_1, \dots, b_m\} \subseteq M$ where if $\overline{B} = \{b_1 + N, \dots, b_m + N\} = \{\overline{b_1}, \dots, \overline{b_m}\}$, then $\overline{B} \subseteq M/N$ and

$$R\overline{B} = R\overline{b_1} + \cdots + R\overline{b_m} = M/N.$$

We show that $M = R(A \cup B)$ and hence that M is finitely generated. Let $x \in M$. Then $x + N = \overline{x} \in M/N$ and for some $r_1, \dots r_m \in R$

$$\begin{aligned}\overline{x} &= r_1\overline{b_1} + \cdots + r_m\overline{b_m} \\ &= \overline{r_1 b_1} + \cdots + \overline{r_m b_m} \\ &= \overline{r_1 b_1 + \cdots + r_m b_m}.\end{aligned}$$

So,

$$x - (r_1 b_2 + \dots r_m b_m) = n \quad \text{for some } n \in N.$$

Now for some $s_1, \dots, s_n \in R$, $n = s_1 a_1 + \cdots + s_n a_n$ and so

$$x - (r_1 b_2 + \dots r_m b_m) = s_1 a_1 + \cdots + s_n a_n$$

which gives

$$x = s_1 a_1 + \cdots + s_n a_n + r_1 b_2 + \dots r_m b_m$$

and thus $x \in R(A \cup B)$ and $M \subseteq R(A \cup B)$. Conversely, since M is a left R–module and $A, B \subseteq M$, then $R(A \cup B) \subseteq M$. ☕

In these exercises R is a ring with 1 and M is a left R-module.

Exercise 10.1.8. An element m of the R–module M is called a *torsion element* if $rm = 0$ for some nonzero element $r \in R$. The set of torsion elements is denoted

$$\mathrm{Tor}(M) = \{m \in M \mid rm = 0 \text{ for some nonzero } r \in R\}.$$

(a) Prove that if R is an integral domain then $\mathrm{Tor}(M)$ is a submodule of M (called the *torsion* submodule of M).

Proof. Since $1_R 0_M = 0_M$ then $\mathrm{Tor}(M) \neq \varnothing$. Let $x, y \in \mathrm{Tor}(M)$ and $r \in R$. Then $r_1 x = 0_M$ and $r_2 y = 0_M$ for some $r_1, r_2 \in R - \{0_R\}$. Since R is an integral domain $r_1 r_2 \neq 0_R$. Then

$$(r_1 r_2)(x + ry) = (r_1 r_2)x + (r_1 r_2)ry = r_2(r_1 x) + r_1 r(r_2 y) = 0_M$$

and so $x + ry \in \mathrm{Tor}(M)$. ☕

(b) Give an example of a ring R and an R–module M such that $\mathrm{Tor}(M)$ is not a submodule. [Consider the torsion elements in the R–module R.]

Proof. Consider $R = M = \mathbb{Z}/4\mathbb{Z}$. Then $\overline{2} \cdot \overline{2} = \overline{0}$ so that $\overline{2} \in \mathrm{Tor}(R)$, but $\overline{1} \cdot \overline{2} = \overline{2} \neq 0$ and so $\mathrm{Tor}(R)$ is not closed under the action of rings elements and thus not a submodule. ☕

(c) If R has zero divisors show that every nonzero R–module has nonzero torsion elements.

Proof. Let M be a nonzero R–module. Suppose $r, s \in R - \{0_R\}$ for which $rs = 0$. Then if $m \in M - \{0\}$,

$$0_m = (rs)m = r(sm)$$

and so $sm \in \mathrm{Tor}(M)$. If $sm \neq 0_M$, we are done. If $sm = 0_M$, then m is a nonzero torsion element of M. ☕

Exercise 10.1.19. Let $F = \mathbb{R}$, let $V = \mathbb{R}^2$ and let T be the linear transformation from V to V which is projection onto the y-axis. Show that $V, 0$, the x-axis, and the y-axis are the only $F[x]$–submodules for this T.

Proof. Let X be the x-axis and Y be the y-axis. Then

$$\begin{aligned} T(X) = 0 \subset X, &\quad \text{and} \quad T(Y) = Y \subseteq Y, \\ T(\mathbb{R}^2) = Y \subset \mathbb{R}^2, &\quad \text{and} \quad T(0) = 0. \end{aligned}$$

and since $X, Y, \mathbb{R}^2$, and 0 are subspaces of $\mathbb{R}^2$, then they are $\mathbb{R}^2[x]$-submodules.

Now suppose that $(W, T|_W)$ is a $\mathbb{R}^2[x]$-submodule for T that is not X, Y or 0. Then there exists $(u, v) \in W$ so that $u \neq 0 \neq v$. Then any scalar multiple of (u, v) is in W so that the entire line L through the origin containing (u, v) is in W. Then $T(W) = Y$ and since $T(W) \subseteq W$, then $Y \subseteq W$. Given $(x, y) \in \mathbb{R}^2$, let $(x, b) \in L \subset W$ and $(0, y - b) \in Y \subset$ so that $(x, y) = (x, b) + (0, y - b) \in W$. So $W = \mathbb{R}^2$. So any T-stable subspace of $\mathbb{R}^2$, and hence any $\mathbb{R}^2[x]$-submodule is $X, Y, \mathbb{R}^2$, or 0. ☕

Exercise 10.2.9. Let R be a commutative ring. Prove that $\mathrm{Hom}_R(R, M)$ and M are isomorphic as left R–modules. [Show that each element of $\mathrm{Hom}_R(R, M)$ is determined by its value on the identity of R.]

Proof. Let $\varphi \in \mathrm{Hom}_R(R, M)$. Given $r \in R$,

$$\varphi(r) = \varphi(r \cdot 1_R) = r\varphi(1_R)$$

and so φ is completely determined by its value on 1_R. Define

$$\begin{aligned} \Phi : \mathrm{Hom}_R(R, M) &\to M \\ \varphi &\mapsto \varphi(1_R). \end{aligned}$$

If $\varphi, \psi \in \mathrm{Hom}_R(R, M)$

$$\Phi(\varphi + \psi) = (\varphi + \psi)(1_R) = \varphi(1_R) + \psi(1_R) = \Phi(\varphi) + \Phi(\psi)$$

and for $r \in R$

$$\Phi(r\varphi) = (r\varphi)(1_R) = r(\varphi(1_R)) = r\Phi(\varphi).$$

If $\varphi(1_R) = \psi(1_R)$ then $\varphi = \psi$ since elements of $\mathrm{Hom}_R(R, M)$ are completely determined by their value on 1_R and thus Φ is injective. Given $m \in M$, define $\varphi : R \to M$ by $r \mapsto rm$. Then for $r, s, t \in R$,

$$\varphi(r + st) = (r + st)m = rm + (st)m = rm + s(tm) = \varphi(r) + s\varphi(t)$$

and so $\varphi \in \mathrm{Hom}_R(R, M)$. Moreover, $\varphi(1_R) = 1_R m = m$ and so Φ is surjective. ☕

Exercise 10.2.12. Let I be a left ideal of R and let n be a positive integer. Prove

$$R^n/IR^n \cong \underbrace{R/IR \times \cdots \times R/IR}_{n \text{ times}}$$

where IR^n is defined as in Exercise 5 of section 1.

Proof. Define a map $\psi : R^n \to R/IR \times \cdots \times R/IR$ by $(r_1, \ldots, r_n) \mapsto (\overline{r_1}, \ldots, \overline{r_n})$. Given $(r_1, \ldots, r_n)$ and $(s_1, \ldots, s_n)$ in R^n and $t \in R$,

$$\begin{aligned} \psi((r_1, \ldots, r_n) + t(s_1, \ldots, s_n)) &= \psi(r_1 + ts_1, \ldots r_n + ts_n) \\ &= (\overline{r_1 + ts_1}, \ldots \overline{r_n + ts_n} = \overline{r_1}) \\ &= (\overline{r_1} + t\overline{s_1}, \ldots, \overline{r_n} + t\overline{s_n}) \\ &= (\overline{r_1}, \ldots, \overline{r_n}) + t(\overline{s_1, \ldots, \overline{r_n}}) \\ &= \psi(r_1, \ldots, r_n) + t\psi(s_1, \ldots, s_n), \end{aligned}$$

and so ψ is an R–module homomorphism. If $(\overline{r_1}, \ldots, \overline{r_n}) \in R/IR \times \cdots \times R/IR$, then $\varphi(r_1, \ldots, r_n) = (\overline{r_1}, \ldots, \overline{r_n})$ and so ψ is surjective.

If $(r_1, \ldots, r_n) \in \operatorname{Ker} \psi$ then $r_j \in IR$ for all j and so $(r_1, \ldots, r_n) \in IR^n$. Conversely if $a(r_1, \ldots, r_n) \in IR^n$ for $a \in I$ then

$$\psi(ar_1, \ldots, ar_n) = (\overline{ar_1}, \ldots, \overline{ar_n}) = (\overline{0}, \ldots, \overline{0}),$$

where the last equality holds since $ar_j \in IR$ for all j. So $a(r_1, \ldots, r_n) \in \operatorname{Ker} \psi$ and this extends to all elements of IR^n since ψ is linear. Therefore, $\operatorname{Ker} \psi = IR^n$ and by the First Isomorphism Theorem, $R^n/IR^n \cong R/IR \times \cdots \times R/IR$. ☕

Exercise 10.3.2. Assume R is commutative. Prove that $R^n \cong R^m$ is and only if $n = m$, i.e., two free R–modules of finite rank are isomorphic if and only if they have the same rank. [Apply Exercise 12 of Section 2 with I a maximal ideal of R. You may assume that if F is a field, then $F^n \cong F^m$ if and only if $n = m$, i.e., two finite dimensional vector spaces over F are isomorphic if and only if they have the same dimension — this will be proved later in Section 11.1.]

Proof. If $n = m$ then $R^n = R^m$. Let I be a maximal ideal of R. If $: R^n \to R^m$ is an R–module isomorphism, then for $a(r_1, \dots, r_n) \in IR^n$,

$$f(a(r_1, \dots, r_n)) = af(r_1, \dots, r_n) = a(s_1, \dots, s_m)$$

for some $(s_1, \dots, s_m) \in R^m$. Since f is linear, this extends to all finite sums of elements of the form $a(r_1, \dots, r_n)$ and so $f(IR^n) \subseteq IR^m$. Similarly, if $b(s_1, \dots, s_m) \in IR^m$, then $f^{-1}(b(s_1, \dots, s_m)) \in IR^n$ and so $f(IR^n) = IR^m$.

Define $\tilde{f} : R^n \to R^m/IR^m$ by $\tilde{f}(r_1, \dots, r_n) = \overline{f(r_1, \dots, r_n)}$. Then $\tilde{f}$ is an R–module epimorphism with $\operatorname{Ker} f = IR^n$ and we get an isomorphism $R^n/IR^n \cong R^m/IR^m$. Notice that $IR = I$. Let F be the field R/I. By Exercise 10.2.12,

$$F^n = \underbrace{F \times \cdots \times F}_{n \text{ times}} \cong R^n/IR^n \cong R^m/IR^m \cong \underbrace{F \times \cdots \times F}_{m \text{ times}} = F^m,$$

and by the hint, $n = m$.

In these exercises R is a ring with 1 and M is a left R-module.

Exercise 10.3.5. Let R be and Integral domain. Prove that every finitely generated torsion R–module has a nonzero annihilator. Give an example of a torsion R–module whose annihilator is the zero ideal.

Proof. Let M be a finitely generated torsion R–module with generating set $\{a_1, \dots, a_n\}$. For all a_i, there exists $s_i \in R - \{0_R\}$ such that $s_i a_i = 0$. Let $s = s_1 s_2 \cdots s_n$. Since R is an integral domain, $s \neq 0_R$. Now if $m = r_1 a_n + \cdots + r_n a_n \in M$, then

$$\begin{aligned} sm &= (s_1 \cdots s_n)(r_1 a_n + \cdots + r_n a_n) \\ &= (s_1 \cdots s_n) r_1 a_n + \cdots + (s_1 \cdots s_n) r_n a_n \\ &= (s_2 \cdots s_n)(s_1 r_1) a_n + \cdots + (s_1 \cdots s_{n-1})(s_n r_n) a_n = 0. \end{aligned}$$

So, we have $0_R \neq s \in \operatorname{Ann}_R(M)$.

Consider the internal direct sum

$$M = \bigoplus_{n \in \mathbb{Z}^+} \mathbb{Z}/n\mathbb{Z} = 0_{\mathbb{Z}} \oplus \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z} \oplus \cdots$$

Then E is a torsion $\mathbb{Z}$–module. To see this, let

$$m = (0_{\mathbb{Z}}, \dots, \overline{m_1}, \dots, \overline{m_2}, \dots, \overline{m_\ell}, \overline{0}, \dots)$$

be an element of M where we are assuming $\overline{m_\ell}$ is the last nonzero coordinate of m and $\overline{m_j} \neq \overline{0}$ for all $1 \leq j \leq \ell$. If we let $r = m_1 \cdot m_2 \cdots m_\ell$, then $rm = (0_{\mathbb{Z}}, \overline{0}, \overline{0}, \dots, \overline{0}, \dots)$.

Let $n \in \mathbb{Z}^+$, $r \in \mathbb{Z}$, and $\overline{j} \in \mathbb{Z}/n\mathbb{Z}$. If $r\overline{j} = \overline{rj} = \overline{0}$, then r is a multiple of n. Now if $r \in \operatorname{Ann}_{\mathbb{Z}}(M)$, this must be true for all n, which means r is an integer which is a multiple of every integer, which is not possible for any nonzero integer. Thus, $r = 0_{\mathbb{Z}} = \operatorname{Ann}_{\mathbb{Z}}(M)$. ☕

Exercise 10.3.15. An element $e \in R$ is called a *central idempotent* if $e^2 = e$ and $er = re$ for all $r \in R$. If e is a central idempotent in R, prove that $M = eM \oplus (1-e)M$.

Proof. If $n \in eM \cap (1-e)M$, then $em_1 = n = m_2 - em_2$ for some $m_1, m_2 \in M$. Then $m_2 = e(m_1 + m_2)$ and so

$$n = e(m_1 + m_2) - e[e(m_1 + m_2)] = e(m_1 + m_2) - e^2(m_1 + m_2) = 0_M,$$

which implies $eM \cap (1-e)M = \{0_M\}$.

Since $e0_M = 0_M$ and $(1-e)0_M = 0_M$, then the sets eM and $(1-e)M$ are nonempty. Let $em_1, em_2 \in eM$, $(1-e)m_3, (1-e)m_4 \in (1-e)M$ and $r \in R$. Then

$$em_1 + rem_2 = em_1 + erm_2 = e(m_1 + rm_2)$$

is in eM and

$$(1-e)m_3 + r(1-e)m_4 = (1-e)m_3 + (1-e)rm_4 = (1-e)(m_3 + rm_4)$$

is in $(1-e)M$. Thus eM and $(1-e)M$ are submodules of M. Since M is a module, then certainly $eM + (1-e)M \subseteq M$. Now if $m \in M$, then

$$m = em + m - em = em + (1-e)m$$

is an element of $eM + (1-e)M$. So $M = eM \oplus (1-e)M$. ☕

Exercise 10.4.16. Suppose R is a commutative ring and let I and J be ideals of R, so R/I and R/J are naturally R–modules.

Throughout, we use the following notation: $\overline{r} := r + I$, $\widetilde{r} := r + J$ and $\widehat{r} := r + (I + J)$ for all $r \in R$.

(a) Prove that every element of $R/I \otimes_R R/J$ can be written as a simple tensor of the form $\overline{1_r} \otimes \widetilde{r}$.

Proof. Let $n \in R/I \otimes R/J$. Then

$$\begin{aligned} n = \sum_{i=1}^{k} \overline{r_i} \otimes \tilde{s}_i &= \sum (\overline{1_R}) r_i \otimes \widetilde{s_i} \\ &= \sum (\overline{1_R}) \otimes r_i \widetilde{s_i} \\ &= \overline{1_R} \otimes \sum \widetilde{r_i s_i} \\ &= \overline{1_R} \otimes \widetilde{\sum r_i s_i}. \end{aligned}$$

☕

(b) Prove that there is an R–module isomorphism $R/I \otimes_R R/J \cong R/(I + J)$ mapping $\overline{r} \otimes \widetilde{r'}$ to $\widehat{rr'}$.

Proof. Call the map given above ψ. We first show that ψ is well-defined. In order to do this, we show that the corresponding map on the cartesian product

$$\varphi : R/I \times R/J \to R/(I + J), \quad (\overline{m}, \widetilde{n}) \mapsto \widehat{mn}$$

is R-balanced. Then by the universal property of the tensor product, there is a unique $\mathbb{Z}$–module homomorphism $\Phi : R/I \otimes_R R/J \to R/(I+J)$ such that $\Phi(\overline{m} \otimes \widetilde{n}) = \varphi(\overline{m}, \widetilde{n})$ for all $\overline{m} \in R/I$ and $\widetilde{n} \in R/J$. Then by uniqueness of this map, this means that then $\psi = \Phi$ is a well-defined $\mathbb{Z}$–module homomorphism. Let's begin:

Let $\overline{m_1}, \overline{m_2}, \overline{m} \in R/I, \widetilde{n_1}, \widetilde{n_2}, \widetilde{n} \in R/J$, and $r \in R$. Then we have

$$\begin{aligned} \varphi(\overline{m_1} + \overline{m_1}, \widetilde{n}) &= \varphi(\overline{m_1 + m_1}, \widetilde{n}) \\ &= \widehat{(m_1 + m_2)n} \\ &= \widehat{m_1 n} + \widehat{m_2 n} \\ &= \varphi(\overline{m_1}, \widetilde{n}) + \varphi(\overline{m_2}, \widetilde{n}), \\ \varphi(\overline{m}, \widetilde{n_1} + \widetilde{n_2}) &= \varphi(\overline{m}, \widetilde{n_1 + n_2}) \\ &= \widehat{m(n_1 + n_2)} \\ &= \widehat{mn_1} + \widehat{mn_2} \\ &= \varphi(\overline{m}, \widetilde{n_1}) + \varphi(\overline{m}, \widetilde{n_2}), \end{aligned}$$

and finally $\varphi(\overline{m}, \widetilde{rn}) = \widehat{mrn} = \varphi(\overline{mr}, \widetilde{n})$. So φ is R-balanced.

Note that in view of part (a), we can write an arbitrary element of $R/I \otimes_R R/J$ as a simple tensor of the form $\overline{1_R} \otimes \widetilde{r}$. Then ψ is injective because

$$\begin{aligned} \widehat{s} = \widehat{r} &\implies s - r \in I + J \\ &\implies \widehat{0_R} = \widehat{s - r} \\ &\implies 0_R - (s - r) \in I + J \\ &\implies r - s \in J \\ &\implies \widetilde{r} = \widetilde{s} \\ &\implies \overline{1_R} \otimes \widetilde{r} = \overline{1_R} \otimes \widetilde{s}. \end{aligned}$$

Certainly ψ is surjective: if $\widehat{r} \in R/(I+J)$, then $1_R \otimes \widetilde{r} \mapsto \widehat{r}$. Let $t \in R$. Then,

$$\begin{aligned} \overline{1_R} \otimes \widetilde{r} + t(\overline{1_R} \otimes \widetilde{s}) &= \overline{1_R} \otimes \widetilde{r} + t\overline{1_R} \otimes \widetilde{s} \\ &= \overline{1_R} \otimes \widetilde{r} + \overline{1_R}t \otimes \widetilde{s} \\ &= \overline{1_R} \otimes \widetilde{r} + \overline{1_R} \otimes \widetilde{ts} \\ &= \overline{1_R} \otimes \left(\widetilde{r} + \widetilde{ts}\right) \\ &= \overline{1_R} \otimes \left(\widetilde{r + ts}\right). \end{aligned}$$

So, we get $\widehat{r + ts} = \widehat{r} + \widehat{ts} = \widehat{r} + t\widehat{s}$, and we see that the map is an R–module homomorphism.

Exercise 10.5.2. Suppose that

$$\begin{tikzcd} A \arrow[r, "\psi"] \arrow[d, "\alpha"'] & B \arrow[r, "\varphi"] \arrow[d, "\beta"'] & C \arrow[r, "\eta"] \arrow[d, "\gamma"'] & D \arrow[d, "\delta"'] \\ A' \arrow[r, "\psi'"'] & B' \arrow[r, "\varphi'"'] & C' \arrow[r, "\eta'"'] & D' \end{tikzcd}$$

is a commutative diagram of groups, and that the rows are exact.

(a) Prove that if α is surjective, and β, δ are injective, then γ is injective.

Proof. Suppose $\gamma(c) = 1$. Then $\eta'(\gamma(c)) = 1$ and since $\eta' \circ \gamma = \delta \circ \eta$ then $\delta(\eta(c)) = 1$. Since δ is injective, $\eta(c) = 1$. Then $c \in \ker \eta = \operatorname{Im} \varphi$, and so there exists $b \in B$ such that $\varphi(b) = c$. Now,

$$\varphi'(\beta(b)) = \gamma(\varphi(b)) = \gamma(c) = 1,$$

and so $\beta(b) \in \ker \varphi' = \operatorname{Im} \psi'$. Thus there exists $a' \in A'$ such that $\psi'(a') = \beta(b)$. Since α is surjective there exists $a \in A$ such that $\alpha(a) = a'$. Then

$$\beta(b) = \psi'(\alpha(a)) = \beta(\psi(a)),$$

which implies $b = \psi(a)$ since β is injective. So $b \in \operatorname{Im} \psi = \ker \varphi$ and $c = \phi(b) = 1$.

(b) Prove that if δ is injective, and α, γ are surjective, then β is surjective.

Proof. Let $b' \in B'$. Since γ is surjective, then there exists $c \in C$ such that $\gamma(c) = \varphi'(b')$. So, $\gamma(c) \in \operatorname{Im}\varphi' = \ker \eta'$ and since $\eta' \circ \gamma = \delta \circ \eta$, then

$$1 = \eta'(\gamma(c)) = \delta(\eta(c))$$

which gives $\eta(c) = 1$ since δ is injective. Now, $c \in \ker \eta = \operatorname{Im}\varphi$, which means there exists $b_1 \in B$ such that $\varphi(b_1) = c$. Since $\gamma \circ \varphi = \varphi' \circ \beta$, then

$$\varphi'(b') = \gamma(c) = \gamma(\varphi(b_1)) = \varphi'(\beta(b_1)),$$

and so $b' - \beta(b_1) \in \ker \varphi' = \operatorname{Im}\psi$. So there exists $a' \in A'$ such that $\psi(a') = b' - \beta(b_1)$. Since α is surjective, there exists $a \in A$ so that $\alpha(a) = a'$. Since $\beta \circ \psi = \psi' \circ \alpha$, then

$$\beta(\psi(a)) = \psi'(\alpha(a)) = \psi'(a') - b' - \beta(b_1)$$

implies $b' = \beta(\psi(a) + b_1)$. Therefore, β is surjective.

In these exercises R is a ring with 1 and M is a left R-module.

Exercise 10.3.9. An R–module M is called *irreducible* if $M \neq 0$ and if 0 and M are the only submodules of M. Show that M is irreducible if and only if $M \neq 0$ and M is a cyclic module with any nonzero element as generator. Determine all the irreducible $\mathbb{Z}$–modules.

Proof. ($\Rightarrow$) Since M is irreducible, $M \neq 0$. Let $m \in N - \{0_N\}$, $r_1 m, r_2 m \in Rm$ and $s \in R$. Certainly $Rm \neq 0$ and

$$r_1 m + s(r_2 m) = (r_1 + sr_2)m \in Rm,$$

and so Rm is a submodule of M. Since M is irreducible, then $Rm = M$.

($\Leftarrow$) Let $N \subseteq M$ be a submodule. If $N = 0$ we are done. Suppose $N \neq 0$. Then for any $n \in N - \{0_N\}$, $Rn = M$ and since $Rn \subseteq N$, then $M = N$ and so M is irreducible.

Let M be an irreducible $\mathbb{Z}$–module, i.e. an abelian group that is cyclic. Notice that $M \neq \mathbb{Z}$, since $\mathbb{Z}$ has plenty of proper nontrivial subgroups. Thus M must have finite order. If M is to have no nontrivial proper subgroups, it must have order a prime number. Since all groups of prime order are cyclic, and all cyclic groups are abelian, then M must be a group of prime order. ☕

Exercise 10.3.11. Show that if M_1 and M_2 are irreducible R–modules then any nonzero R–module homomorphism from M_1 to M_2 is an isomorphism. Deduce that if M is irreducible then $\mathrm{End}_R(M)$ is a division ring (this result is called $Schur's Lemma$). [Consider the kernel and the image.]

Proof. If $\varphi : M_1 \to M_2$ is a nonzero homomorphism, then $\operatorname{Im}\varphi$ is a submodule of M_2: either 0_{M_2} or M_2. Since φ is nonzero, $\operatorname{Im}\varphi = M_2$. Similarly, $\ker\varphi$ is a submodule of M_1 and cannot be M_1 since φ is nonzero, and thus $\ker\varphi = 0_{M_1}$. So φ is an isomorphism.

Therefore, if M is irreducible and $\varphi \in \mathrm{End}_R(M)$ then φ^{-1} is also an element of $\mathrm{End}_R(M)$ and thus $\mathrm{End}_R(M)$ is a division ring. ☕

Exercise 10.3.12. Let R be a commutative ring and let A, B, and M be R–modules. Prove the following isomorphism of R–modules:

$$\mathrm{Hom}_R(A \times B, M) \cong \mathrm{Hom}_R(A, M) \times \mathrm{Hom}_R(B, M)$$

Proof. If $\varphi \in \mathrm{Hom}_R(A \times B, M)$, define $\varphi|_A : A \to M$ by $\varphi|_A(a) = \varphi(a, 0)$. It follows that $\varphi|_A \in \mathrm{Hom}_R(A, M)$ since φ is an $\mathbb{R}$ module homomorphism. Define $\varphi|_B : B \to M$ similarly: $\varphi|_B(b) = \varphi(0, b)$. Now define a map

$$\Phi : \mathrm{Hom}_R(A \times B, M) \to \mathrm{Hom}_R(A, M) \times \mathrm{Hom}_R(B, M),$$
$$\varphi \mapsto (\varphi|_A, \varphi|_B).$$

We first show that Φ is an R–module homomorphism: Let $\varphi_1, \varphi_2 \in \mathrm{Hom}_R(A \times B, M)$ and $r \in R$. Then

$$\begin{aligned}
\Phi(\varphi_1 + r\varphi_2) &= ((\varphi_1 + r\varphi_2)\,|_A, (\varphi_1 + r\varphi_2)\,|_B) \\
&= (\varphi_1|_A + r\,(\varphi_2|_A)\,, \varphi_1|_B + r\,(\varphi_2|_B)) \\
&= (\varphi_1|_A, \varphi_1|_B) + (r\,(\varphi_2|_A)\,, r\,(\varphi_2|_B)) \\
&= (\varphi_1|_A, \varphi_1|_B) + r\,((\varphi_2|_A)\,, (\varphi_2|_B)) \\
&= \Phi(\varphi_1) + r\Phi(\varphi_2).
\end{aligned}$$

If $\varphi \in \ker \Phi$, then for all $a \in A$ and for all $b \in B$

$$\begin{aligned} 0 &\equiv \varphi|_A(a) = \varphi(a, 0) \\ 0 &\equiv \varphi|_B(b) = \varphi(0, b), \end{aligned}$$

and so for all $a \in A$ and $b \in B$, we get

$$\varphi(a, b) = \varphi((a, 0) + (0, b)) = \varphi(a, 0) + \varphi(0, b) = 0.$$

Thus, $\varphi \equiv 0$ and so Φ is injective.

If $(\psi, \eta) \in \mathrm{Hom}_R(A, M) \times \mathrm{Hom}_R(B, M)$, define a map

$$\begin{aligned} \varphi : A \times B &\to M, \\ (a, b) &\mapsto \psi(a) + \eta(b). \end{aligned}$$

We check that $\varphi \in \mathrm{Hom}_R(A \times B, M)$: Let $r \in R$ and $(a_1, b_1), (a_2, b_2) \in A \times B$. Then

$$\begin{aligned} \varphi((a_1, b_1) + r(a_2, b_2)) &= \varphi(a_1 + ra_2, b_1 + rb_2) \\ &= \psi(a_1 + ra_2) + \eta(b_1 + rb_2) \\ &= \psi(a_1) + r\psi(a_2) + \eta(b_1) + r\eta(b_2) \\ &= [\psi(a_1) + \eta(b_1)] + r[\psi(a_2) + \eta(b_2)] \\ &= \varphi(a_1, b_1) + r\varphi(b_1, b_2). \end{aligned}$$

Notice that for all $a \in A$ and for all $b \in B$

$$\begin{aligned} \varphi|_A(a) &= \varphi(a, 0) = \psi(a) + \eta(0) = \psi(a), \text{ and} \\ \varphi|_B(b) &= \varphi(0, b) = \psi(0) + \eta(b) = \eta(b). \end{aligned}$$

So, $\Phi(\varphi) = (\psi, \eta)$ and hence Φ is surjective. ☕

Exercise 10.4.20. Let $I = (2, x)$ be the ideal generated by 2 and x in the ring $R = \mathbb{Z}[x]$. Show that the element $2 \otimes 2 + x \otimes x$ in $I \otimes_R I$ is not a simple tensor, i.e., cannot be written as $a \otimes b$ for some $a, b \in I$.

From Dr. Bleher: Prove that there is an R-balanced map $f : I \times I \to I^2$ defined by $f(p(x), q(x)) = p(x)q(x)$. Use this to obtain a well-defined $\mathbb{Z}$–module homomorphism $F : I \otimes_R I \to I^2$. Argue that F is surjective. Show that F is an R-module homomorphism by showing F preserves the R–module structure on $I \otimes_R I$ coming from the fact that R is commutative. Then use the map F to do #20.

Proof. Let $p_1(x), p_2(x), p(x), q_1(x)q_2(x), q(x) \in I$ and $r \in R$. Then

$$\begin{aligned} f(p_1(x) + p_2(x), q(x)) &= p_1(x)q(x) + p_2(x)q(x) = f(p_1(x), q(x)) + f(p_2(x), q(x)), \\ f(p(x), q_1(x) + q_2(x)) &= p(x)q_1(x) + p(x)q_2(x) = f(p(x), q_1(x)) + f(p(x), q_2(x)), \\ f(p(x), rq(x)) &= (p(x)r)q(x) = f(p(x)r, q(x)). \end{aligned}$$

So f is R-balanced and so we have a well-defined $\mathbb{Z}$–module homomorphism $F : I \otimes_R I \to I^2$. Given $\sum_{i=1}^n p_i(x)q_i(x) \in I^2$, we have

$$F\left(\sum_{i=1}^n p_i(x) \otimes q_i(x)\right) = \sum_{i=1}^n p_i(x)q_i(x),$$

and so F is surjective. Since R is commutative $I \otimes_R I$ is an R–module. Moreover, we know that F is additive since F is a $\mathbb{Z}$–module homomorphism. Now

$$\begin{aligned} F\left(r\sum_{i=1}^{n} p_i(x)\otimes q_i(x)\right) &= F\left(\sum_{i=1}^{n} rp_i(x)\otimes q_i(x)\right) \\ &= \sum_{i=1}^{n} F\left(rp_i(x)\otimes q_i(x)\right) \\ &= \sum_{i=1}^{n} rp_i(x)q_i(x) \\ &= r\sum_{i=1}^{n} p_i(x)q_i(x) \\ &= rF\left(\sum_{i=1}^{n} p_i(x)\otimes q_i(x)\right), \end{aligned}$$

and so F preserves the R–module structure on $I \otimes_R I$.

Suppose $2\otimes 2 + x\otimes x = p(x)\otimes q(x)$ for some $p(x), q(x) \in I$. Then

$$4 + x^2 = F(2\otimes 2) + F(x\otimes x) = F(p(x)\otimes q(x)) = p(x)q(x).$$

So if $p(x)$ were constant, then $p(x)$ would have to divide 1 and 4, and thus $p(x) = \pm 1$. But $\pm \notin I$ and so $p(x)$ and $q(x)$ are not constant. So, both $p(x)$ and $q(x)$ must be of the form $p(x) = x+n, q(x) = x+m$ for some even integers m, n. Then

$$x^2 + 4 = (x+n)(x+m) = x^2 + (m+n)x + nm,$$

which means $m+n = 0$ and so $m = -n$. We also get $4 = mn = -nn = -n^2$, a contradiction.

Exercise 11.1.6. Let V be a vector space of finite dimension. If φ is any linear transformation from V to V prove there is an integer m such that the intersection of the image of φ^m and the kernel of φ^m is $\{0_V\}$.

Proof. Let $i \in \mathbb{Z}^+$. Then if $k \in \operatorname{Ker} \varphi^i$,

$$\varphi^{i+1}(k) = \varphi(\varphi(k)) = \varphi(0_V) = 0_V$$

and so $k \in \operatorname{Ker} \varphi^{i+1}$. So,

$$\operatorname{Ker} \varphi \subseteq \operatorname{Ker} \varphi^2 \subseteq \cdots \subseteq \operatorname{Ker} \varphi^i \subseteq \operatorname{Ker} \varphi^{i+1} \subseteq \dots$$

is an ascending chain of subspaces of V. Since V is finite dimensional, the dimensions of this chain cannot strictly increase indefinitely. Thus there exists $m \in \mathbb{Z}^+$ such that $\operatorname{Ker} \varphi^i = \operatorname{Ker} \varphi^m$ for all $i \geq m$.

If $v \in \operatorname{Im} \varphi^m \cap \operatorname{Ker} \varphi^m$, then

$$\varphi^m(u) = v \quad \text{and} \quad \varphi^m(v) = 0_V$$

for some $u \in V$. Then

$$\varphi^{2m}(u) = \varphi^m(\varphi^m(u)) = \varphi^m(v) = 0_V,$$

and so $u \in \operatorname{Ker} \varphi^{2m} = \operatorname{Ker} \varphi^m$. Hence we get $0_V = \varphi^m(u) = v$. ☕

Exercise 11.1.8. Let V be a vector space over F and let φ be a linear transformation of the vector space V to itself. A nonzero element $v \in V$ satisfying $\varphi(v) = \lambda v$ for some $\lambda \in F$ is call an *eigenvector* of φ with *eigenvalue* λ. Prove that for any fixed $\lambda \in F$ the collection of eigenvectors of φ with eigenvalue λ together with 0 forms a subspace of V.

Proof. Let $E_\lambda = \{v \in V - \{0_V\} \mid \varphi(v) = \lambda v\} \cup \{0_V\}$. Since $0_V \in E_\lambda$, $E_\lambda \neq \varnothing$. If $E_\lambda = \{0_V\}$ we are done. Let $u, v \in E_\lambda$ so that at least one of the vectors u, v are nonzero, and let $\eta \in F$. Then

$$\varphi(u + \eta v) = \varphi(u) + \eta\varphi(v) = \lambda u + \eta(\lambda v) = \lambda u + \lambda(\eta v) = \lambda(u + \eta v),$$

and so $u + \eta v \in E_\lambda$, and hence E_λ is a subspace of V. ☕

Exercise 11.1.9. Let V be a vector space over F and let φ be a linear transformation of the vector space V to itself. Suppose for $i = 1, 2, \dots, k$ that $v_i \in V$ is an eigenvector for φ with eigenvalue $\lambda_i \in F$ and that all the eigenvalues λ_i are distinct. Prove that $v_1, v_2, \dots, v_k$ are linearly independent. [Use induction on k: write a linear dependence relation among the v_i and apply φ to get another linear dependence relation among the v_i involving the eigenvalues — now subtract a suitable multiple of the first linear relation to get a linear dependence relation on fewer elements.] Conclude that any linear transformation on an n-dimensional vector space has at most n distinct eigenvalues.

Proof. We proceed by induction on k. If $k = 1$, there's nothing to show. Suppose $v_1, \dots, v_m$ are linearly independent for some $1 \leq m \leq k - 1$. Suppose

$$0 = \alpha_1 v_1 + \dots \alpha_m v_m + \alpha_{m+1} v_{m+1}. \tag{1}$$

If we can show that $\alpha_{m+1} = 0$, then (1) would then give $\alpha_i = 0$ for all $1 \leq i \leq m + 1$ since $v_1, \dots, v_m$ are linearly independent. Suppose $\alpha_{m+1} \neq 0$. Then

$$v_{m+1} = -\frac{1}{\alpha_{m+1}}(\alpha_1 v_1 + \cdots + \alpha_m v_m) \tag{2}$$

Applying φ to (1) and substituting the value of v_{m+1} given in (2),

$$\begin{aligned}0=\varphi(0)&=\alpha_1\varphi(v_1)+\cdots+\alpha_m\varphi(v_m)+\alpha_{m+1}\varphi(v_{m+1})\\&=\alpha_1\lambda_1v_1+\cdots+\alpha_m\lambda_mv_m+\alpha_{m+1}\lambda_{m+1}v_{m+1}\\&=\alpha_1\lambda_1v_1+\cdots+\alpha_m\lambda_mv_m+\alpha_{m+1}\lambda_{m+1}\left(-\frac{1}{\alpha_{m+1}}(\alpha_1v_1+\cdots+\alpha_mv_m)\right)\\&=(\lambda_1-\lambda_{m+1})\alpha_1v_1+\cdots+(\lambda_m-\lambda_{m+1})\alpha_mv_m.\end{aligned}$$

Since $v_1,\dots,v_m$ are linearly independent, $(\lambda_i-\lambda_{m+1})\alpha_i=0$ for all $1\le i\le m$. Since the λ_j are all distinct for $1\le j\le m+1$, then $\lambda_i-\lambda_{m+1}\neq 0$ and so $\alpha_i=0$ for all $1\le i\le m$. Therefore, we have $0=\alpha_{m+1}v_{m+1}$ by (1), a contradiction since $v_{m+1}\neq 0$. Thus $\alpha_{m+1}=0$. ☕

Exercise 11.2.15. Prove that the row rank of two row equivalent matrices is the same. [It suffices to prove this for two matrices differing by an elementary row operation.]

Proof. Let A and B be two row equivalent matrices in $\mathrm{Mat}_{m\times n}(F)$ with row vectors $a_1,\dots,a_m$ and $b_1,\dots,b_m$. Notice that

$$\dim(\operatorname{span}\{a_1,\dots,a_m\})=\#\{\text{ linearly independent rows of } A\}=\text{ row rank of } A.$$

Similarly for the rows of B. If A and B differ by a interchange of rows, then $\{a_1,\dots,a_m\}=\{b_1,\dots,b_m\}$ and so trivially $\dim(\operatorname{span}\{a_1,\dots,a_m\})=\dim(\operatorname{span}\{b_1,\dots,b_m\})$.

Suppose A and B differ by the second elementary row operation. That is, suppose for some i, $a_i=\lambda b_i$ for some $\lambda\in F$ and $a_\ell=b_\ell$ for all $\ell\neq i$. Then

$$\begin{aligned}\dim(\operatorname{span}\{a_1,\dots,a_i,\dots,a_m\})&=\dim(\operatorname{span}\{a_1,\dots,a_{i-1},\lambda b_i,a_{i+1}\dots,a_m\})\\&=\dim(\operatorname{span}\{b_1,\dots,b_{i-1},\lambda b_i,b_{i+1}\dots,b_m\})\\&=\dim(\operatorname{span}\{b_1,\dots,b_m\})\end{aligned}$$

Now suppose A and B differ by the third elementary row operation. That is, suppose for some i, $a_i=b_i+\lambda b_j$ for some $\lambda\in F$ and $i\neq j$ and $a_\ell=b_\ell$ for all $\ell\neq i$. Then

$$\begin{aligned}\dim(\operatorname{span}\{a_1,\dots,a_i,\dots,a_m\})&=\dim(\operatorname{span}\{a_1,\dots,a_{i-1},b_i+\lambda b_j,a_{i+1}\dots,a_m\})\\&=\dim(\operatorname{span}\{b_1,\dots,b_{i-1},b_i+\lambda b_j,b_{i+1}\dots,b_m\})\\&=\dim(\operatorname{span}\{b_1,\dots,b_m\})\end{aligned}$$

☕

Exercise 11.3.4. If V is infinite dimensional with basis $\mathcal{A}$, prove that $\mathcal{A}^*=\{v^*\mid v\in\mathcal{A}\}$ does *not* span V^*.

Proof. Let $f\in V^*$ be defined by $f(e_\alpha)=1_F$ for all $e_\alpha\in A$. Suppose $f\in\operatorname{span}\mathcal{A}^*$. Then there exists $n\in\mathbb{Z}^+$, $c_1,\dots,c_n\in F$ and $v_1^*,\dots,v_n^*\in\mathcal{A}^*$ such that

$$f=\sum_{i=1}^n c_iv_i^*.$$

However, for $\alpha\notin\{1,\dots,n\}$,

$$1_F=f(e_\alpha)=\sum_{i=1}^n c_iv_i^*(e_\alpha)=0_F.$$

☕

Exercise 11.1.10. Prove that any vector space V has a basis.

Proof. Let $\mathcal{S} = \{J \subseteq V \mid J \text{ consists of linearly independent vectors}\}$ be partially ordered by inclusion. $\mathcal{S}$ is nonempty since $\{0_V\} \in \mathcal{S}$. Let $\mathcal{C}$ be a chain in $\mathcal{S}$. We claim

$$U = \bigcup_{J \in \mathcal{C}} J$$

is an upper bound for $\mathcal{C}$. Certainly $J \subseteq U$ for all $J \in \mathcal{C}$. It remains to show that $U \in \mathcal{S}$. Let $\{u_1, \ldots, u_n\}$ be a finite collection of vectors in U. For all i, there exists J_i containing u_i. Since $\mathcal{C}$ is a chain, there must be a J_k containing all J_i and therefore all $u_1, \ldots, u_n$. Hence, the $u_1, \ldots, u_n$ are linearly independent. Since this is true for all finite collections of vectors in U, then $U \in \mathcal{S}$. By Zorn's Lemma, $\mathcal{S}$ contains a maximal element, call it M.

We show $M = \{m_1, \ldots, m_\ell\}$ is a basis for V. We already know M is a linearly independent set. Suppose M does not span V. Then there exists $v \in V - \{0_V\}$ which is not a linear combination of the vectors in M. In other words, for all sets of scalars $a_1, \ldots, a_\ell \in F - \{0_F\}$,

$$v - \sum_{i=1}^{\ell} a_i m_i \neq 0_V.$$

Then $\{v\} \cup M$ is a linearly independent set in V, a contradiction since M is a maximal set of linearly independent vectors in V. Thus M is a basis for V. ☕

Exercise 2. Suppose V is a non-zero vector space over a field F and A is a subset of V that spans V. Prove that A contains a basis of V.

Proof. Let $\mathcal{S} = \{J \subseteq A \mid J \text{ consists of linearly independent vectors}\}$. The proof that $\mathcal{S}$ contains a maximal element M is identical to the previous problem.

To see that M is a basis for V, suppose $v \notin \operatorname{span} M$. Since $v \in \operatorname{span} A$, there exists $n \in \mathbb{Z}^+$, $a_1, \ldots, a_n \in A$, and $c_1, \ldots c_n \in F$ so that

$$v = \sum_{i=1}^{n} c_i a_i.$$

Notice that if $a_i \in \operatorname{span} M$ for all $i \in \{1, \ldots, n\}$, then v would be in the span of M. So, there exists $j \in \{1, \ldots, n\}$ such that $a_j \notin \operatorname{span} M$. Notice that $\{a_j\} \cup M$ is linearly independent; otherwise, we could write a_j as a linear combination of elements in M, (but $a_j \notin \operatorname{span} M$). So $\{a_j\} \cup M \subseteq A$ is linearly independent, a contradiction since M is maximal in A with respect to linear independence. Hence M is a basis for V. ☕

Exercise 11.2.6. Prove if $\varphi \in \mathrm{Hom}_F(F^n, F^m)$, and $\mathcal{B}, \mathcal{E}$ are the natural bases of F^n, F^m respectively, then the range of φ equals the span of the set of columns of $M_{\mathcal{B}}^{\mathcal{E}}(\varphi)$. Deduce that the rank of φ (as a linear transformation) equals the column rank of $M_{\mathcal{B}}^{\mathcal{E}}(\varphi)$.

Proof. Let $\mathcal{B} = \{b_1, \ldots, b_n\}$ and $\mathcal{E} = \{e_1, \ldots, e_m\}$ be the natural bases of F^n and F^m, respectively. For all $j \in \{1, \ldots, n\}$

$$\varphi(b_j) = \sum_{i=1}^{m} a_{ij} e_i$$

for some $a_{ij} \in F$. Notice that since $\mathcal{E}$ is a natural basis for F^m, the sum above is precisely the jth column vector of $M_{\mathcal{B}}^{\mathcal{E}}$; that is,

$$\varphi(b_j) = \sum_{i=1}^{m} a_{ij} e_i = \vec{a_j}, \tag{$*$}$$

where $\vec{a_j}$ denotes the jth column vector of $M_{\mathcal{B}}^{\mathcal{E}}$. If $v \in F^m$ is in the range of φ, there exists $u \in F^n$ such that $\varphi(u) = v$. Moreover, there exists $c_1, \ldots, c_n \in F$ such that

$$u = \sum_{j=1}^{n} c_j b_j.$$

So,

$$v = \varphi(u) = \sum_{j=1}^{n} c_j \varphi(b_j) = \sum_{j=1}^{n} c_j \sum_{i=1}^{m} a_{ij} e_i = \sum_{j=1}^{n} c_j \vec{a_j}$$

and hence v is in the span of the columns of $M_{\mathcal{B}}^{\mathcal{E}}$. Conversely, if w is in the span of the columns of $M_{\mathcal{B}}^{\mathcal{E}}$, then there exists $d_1, \ldots, d_n \in F$ such that

$$w = \sum_{j=1}^{n} d_j \vec{a_j}.$$

Letting $x = \sum_{j=1}^{n} d_j b_j \in F^n$, we get by $(*)$,

$$w = \sum_{j=1}^{n} d_j \vec{a_j} = \sum_{j=1}^{n} d_j \sum_{i=1}^{m} a_{ij} e_i = \sum_{j=1}^{n} d_j \varphi(b_j) = \varphi(x),$$

and so w is in the range of φ. Therefore we have set equality between the range of φ and the span of the columns of $M_{\mathcal{B}}^{\mathcal{E}}$. This gives

$$\dim(\mathrm{Im}\,\varphi) = \dim(\mathrm{span}\{\vec{a_1}, \ldots, \vec{a_n}\}),$$

i.e., the rank of φ equals the column rank of $M_{\mathcal{B}}^{\mathcal{E}}$. ☕

Exercise 11.2.11. Let φ be a linear transformation from the finite dimensional vector space V to itself such that $\varphi^2 = \varphi$.

A linear transformation φ satisfying $\varphi^2 = \varphi$ is called an *idempotent* linear transformation. This exercise proves that idempotent linear transformations are simply projections onto some subspace.

(a) Prove that $\operatorname{Im}\varphi \cap \operatorname{Ker}\varphi = 0$.

Proof. Let $v \in \operatorname{Im}\varphi \cap \operatorname{Ker}\varphi$. Then there exists $u \in V$ such that $\varphi(u) = v$ and $\varphi(v) = 0$. Then $v = \varphi(u) = \varphi^2(u) = \varphi(\varphi(u)) = \varphi(v) = 0$.

(b) Prove that $V = \operatorname{Im}\varphi \oplus \operatorname{Ker}\varphi$.

Proof. We know $\operatorname{Im}\varphi$ and $\operatorname{Ker}\varphi$ are subspaces of V. By part (a), their intersection is trivial. We also have $\operatorname{Im}\varphi + \operatorname{Ker}\varphi \subseteq V$. Moreover, if $v \in V$, then $v - \varphi(v) \in \operatorname{Ker}\varphi$ since

$$\varphi(v - \varphi(v)) = \varphi(v) - \varphi^2(v) = 0.$$

So $v = \varphi(v) + (v - \varphi(v)) \in \operatorname{Im}\varphi + \operatorname{Ker}\varphi$. Therefore, $V = \operatorname{Im}\varphi \oplus \operatorname{Ker}\varphi$.

(c) Prove that there is a basis of V such that the matrix of φ with respect to this basis is a diagonal matrix whose entries are all 0 or 1.

Proof. Let $\{v_1, \dots, v_k\}$ be a basis for $\operatorname{Ker}\varphi$. Extend this to a basis for V, $\mathcal{B} = \{v_1, \dots, v_k, \dots, v_n\}$. Notice that $\{v_{k+1}, \dots, v_n\} \subseteq \operatorname{Im}\varphi$ since $V = \operatorname{Im}\varphi \oplus \operatorname{Ker}\varphi$. Then $\varphi(v_i) = 0$ for all $i \in \{1, \dots, k\}$. So, the ith column in $M_{\mathcal{B}}^{\mathcal{B}}(\varphi)$ consists of all zeros for $i \in \{1, \dots, k\}$.

For all $j \in \{k+1, \dots, n\}$, there exists $u_j \in V$ such that $\varphi(u_j) = v_j$. So,

$$\varphi(v_j) = \varphi(\varphi(u_j)) = \varphi(u_j) = v_j.$$

Therefore, the jth column in $M_{\mathcal{B}}^{\mathcal{B}}(\varphi)$ has a 1 in the jth row and zeroes everywhere else. Thus we get

$$M_{\mathcal{B}}^{\mathcal{B}}(\varphi) = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 0 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$$

Exercise 11.2.13. Let V, W be vector spaces over F with dimensions n and m, respectively. Let $\varphi : V \to W$ be a linear transformation; let $\mathcal{B}_1, \mathcal{B}_2$ be bases for V and $\mathcal{E}_1, \mathcal{E}_2$ be bases for W. Define

$$A = M_{\mathcal{B}_1}^{\mathcal{E}_1}(\varphi), \quad B = M_{\mathcal{B}_2}^{\mathcal{E}_2}(\varphi), \quad P = M_{\mathcal{B}_2}^{\mathcal{B}_1}(\mathbb{1}_V), \quad \text{and} \quad Q = M_{\mathcal{E}_2}^{\mathcal{E}_1}(\mathbb{1}_W).$$

where $\mathbb{1}_V : V \to V$ and $\mathbb{1}_W : W \to W$ denote the identity maps on V and W, respectively. Prove that $Q^{-1} = M_{\mathcal{E}_1}^{\mathcal{E}_2}(\mathbb{1}_W)$ and that $Q^{-1}AP = B$, giving the general relation between matrices representing the same linear transformation but with respect to two different choices of bases.

Proof. We have

$$QM_{\mathcal{E}_1}^{\mathcal{E}_2}(\mathbb{1}_W) = M_{\mathcal{E}_2}^{\mathcal{E}_1}(\mathbb{1}_W)M_{\mathcal{E}_1}^{\mathcal{E}_2}(\mathbb{1}_W) = M_{\mathcal{E}_1}^{\mathcal{E}_1}(\mathbb{1}_W) = I,$$

and

$$M_{\mathcal{E}_1}^{\mathcal{E}_2}(\mathbb{1}_W)Q = M_{\mathcal{E}_1}^{\mathcal{E}_2}(\mathbb{1}_W)M_{\mathcal{E}_2}^{\mathcal{E}_1}(\mathbb{1}_W) = M_{\mathcal{E}_2}^{\mathcal{E}_2}(\mathbb{1}_W) = I,$$

and so $Q^{-1} = M_{\mathcal{E}_1}^{\mathcal{E}_2}(\mathbb{1}_W)$.

Moreover,

$$Q^{-1}AP = M_{\mathcal{E}_1}^{\mathcal{E}_2}(\mathbb{1}_W)M_{\mathcal{B}_1}^{\mathcal{E}_1}(\varphi)M_{\mathcal{B}_2}^{\mathcal{B}_1}(\mathbb{1}_V) = M_{\mathcal{B}_1}^{\mathcal{E}_2}(\mathbb{1}_W \circ \varphi)M_{\mathcal{B}_2}^{\mathcal{B}_1}(\mathbb{1}_V) = M_{\mathcal{B}_2}^{\mathcal{E}_2}(\varphi \circ \mathbb{1}_V) = B.$$

☕

Exercise 11.2.25. Let A be an $n \times n$ matrix.

(a) Show that A has an inverse matrix B with columns $B_1, \dots, B_n$ if and only if the system of equations:

$$AB_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad AB_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad AB_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

has solutions.

Proof. ($\Rightarrow$) If I_j denotes the jth column of the identity matrix I, then since $AB = I$, we get $AB_j = I_j$ for all $j \in \{1, \dots, n\}$. So the B_j are the solutions to the system of equations.

($\Leftarrow$) Let $B = (B_1, \dots, B_n)$. Since $AB_j = I_j$, it follows that $AB = I$. Let φ and ψ be the linear transformations (with respect to some chosen basis) of V associated with A and B, respectively. Since $AB = I$, then $\varphi \circ \psi = \mathbb{1}_V$ and so φ is surjective, i.e. $\dim \operatorname{Im} \varphi = n$. Then

$$n = \dim(V) = \dim \operatorname{Ker} \varphi + \dim \operatorname{Im} \varphi \implies \operatorname{Ker} \varphi = 0$$

So φ is bijective and thus A has a left inverse, say C. Then

$$C = CI = C(AB) = (CA)B = IB = B$$

and so $BA = I$. ☕

(b) Prove that A has an inverse if and only if A is row equivalent to the $n \times n$ identity matrix.

Proof. ($\Rightarrow$) Viewing A as a map $A : F^n \to F^n$, A is invertible. In particular, $\dim \operatorname{Ker} A = 0$. Since $n = \dim F^n = \dim \operatorname{Ker} A + \operatorname{Im} A$, then $\operatorname{Im} A = n$. So the (row) rank of A is n. Since the rank of a matrix is unaffected by row operations (cf. Exercise 11.2.15), then if we reduce A to its reduced row echelon form, call it A', then A' has rank n. Since the only matrix in reduced row echelon form of rank n is I, then $A' = I$. That is, A is row equivalent to I.

($\Leftarrow$) If A is row equivalent to the identity matrix, then there are a finite number of row operations which reduce A to I. For each elementary row operation, there is a corresponding elementary matrix P_i which, when multiplied by A has the same effect on A as that of an elementary row operation. Let $P_1, \ldots, P_k$ be the k elementary matrices corresponding to the k elementary row operations on A which reduce A to I. Then if $B = P_1 \cdots P_k$, we have $BA = I$. Using a similar argument as in part (a), we get $AB = I$. ☕

(c) Prove that A has an inverse B if and only if the augmented matrix $(A \mid I)$ can be row reduced to the augmented matrix $(I \mid B)$ where I is the $n \times n$ identity matrix.

Proof. This follows almost immediately from part (b). A is invertible if and only if A is row equivalent to the identity matrix, if and only if $(A \mid I)$ is row equivalent to $(I \mid C)$ for some matrix C. But notice that C was obtained by the elementary row operations on I which reduced A to I. Hence $CA = I$, and since inverses are unique, $B = C$. ☕

Exercise 11.4.4.

(a) (i) interchanging two rows changes the sign of the determinant.

(ii) adding a multiple of one row to another does not change the sign of the determinant.

(iii) multiplying any row by a nonzero element u from F multiplies the determinant by u.

Proof. We know that the determinant function is alternating on the columns of a matrix, and that $\det(A^T) = \det(A)$. If we interchange two rows of A, we interchange two columns of A^T. This will change the sign of $\det(A^T)$ by -1, and thus change the sign of $\det(A)$ by -1. This gives (i). Analogously, adding a multiple of one row to another row in A corresponds to adding a multiple of one column to another column in A^T. So if $A_1, \ldots, A_n$ are the columns of A^T (the rows of A), then

$$\begin{aligned}
\det(A_1, \ldots, A_i + \lambda A_j, \ldots, A_n) &= \det(A_1, \ldots, A_i, \ldots, A_n) + \det(0, \ldots, 0, \lambda A_j, 0, \ldots, 0) \\
&= \det(A_1, \ldots, A_i, \ldots, A_n) \\
&= \det(A^T) \\
&= \det(A).
\end{aligned}$$

This gives (ii). Finally for (iii), we have

$$\begin{aligned}\det(A_1, \dots, uA_i, \dots, A_n) &= \det(A_1, \dots, uA_i, \dots, A_n) \\ &= u\det(A_1, \dots, A_i, \dots, A_n) \\ &= u\det(A^T) \\ &= u\det(A).\end{aligned}$$

(b) Prove that $\det A$ is nonzero if and only if A is row equivalent to the $n \times n$ identity matrix. Suppose A can be row reduced to the identity matrix using a total of s row interchanges as in (i) and by multiplying the rows by nonzero elements $u_1, \dots, u_t$ as in (iii). Prove that $\det(A) = (-1)^s(u_1u_2 \dots u_t)^{-1}$.

Proof. ($\Rightarrow$) If the determinant of A is nonzero, then the columns of A are linearly independent. In particular, A has rank n. Then the kernel of A is trivial and thus A is invertible. By exercise 11.2.25, A is row equivalent to the identity matrix.

($\Leftarrow$) If A is row equivalent to the identity matrix, then the columns of A are linearly independent and so $\det(A) \neq 0$.

The last statement follows from the fact that the determinant function is multiplicative. By part (a), s row interchanges changes the value of the determinant of A by $(-1)^s$. Let P be the matrix obtained by performing s row interchanges in A. Then $\det(P) = (-1)^s \det(A)$. Let Q now be the matrix obtained by multiplying rows of P be $u_1, \dots, u_t$. In particular, $Q = I$. Then by part (a),

$$1 = \det(I) = \det(Q) = u_1 \cdots u_t \det(P) = u_1 \cdots u_t(-1)^s \det(A)$$

and so $\det(A) = (-1)^s(u_1 \cdots u_t)^{-1}$.

Exercise 11.3.3. Let S be any subset of V^* for some finite dimensional space V. Define $\text{Ann}(S) = \{v \in V \mid f(v) = 0 \text{ for all } f \in S\}$. ($\text{Ann}(S)$ is called the *annihilator* of S in V).

(a) Prove that $\text{Ann}(S)$ is a subspace of V.

(b) Let W_1 and W_2 be subspaces of V^*. Prove that $\text{Ann}(W_1+W_2) = \text{Ann}(W_1) \cap \text{Ann}(W_2)$ and $\text{Ann}(W_1 \cap W_2) = \text{Ann}(W_1) + \text{Ann}(W_2)$.

(c) Let W_1 and W_2 be subspaces of V^*. Prove that $W_1 = W_2$ if and only if $\text{Ann}(W_1) = \text{Ann}(W_2)$.

(d) Prove that the annihilator of S is the same as the annihilator of the subspace of V^* spanned by S.

(e) Assume V is finite dimensional with basis $v_1, \ldots, v_n$. Prove that if $S = \{v_1^*, \ldots, v_k^*\}$ for some $k \leq n$ then $\text{Ann}(S)$ is the subspace spanned by $\{v_{k+1}, \ldots, v_n\}$.

(f) Assume V is finite dimensional. Prove that if W^* is any subspace of V^* then $\dim \text{Ann}(W^*) = \dim V - \dim W^*$.

Proof. (a): $\text{Ann}(S)$ is nonempty since $f(0_V) = 0_V$ for all $f \in S$. If $u, w \in \text{Ann}(S), f \in S$, and $r \in F$, then $f(u + rw) = f(u) + rf(w) = 0$ and so $u + rw \in \text{Ann}(S)$. Hence $\text{Ann}(S)$ is a subspace of V.

(b): Let B be a basis for $W_1 \cap W_2$. Extend B to bases B_1, and B_2 for W_1, and W_2 respectively, with

$$B = \{x_1^*, \ldots, x_\ell^*\}, \quad B_1 - B = \{y_1^*, \ldots, y_{n_1}^*\}, \quad B_2 - B = \{z_1^*, \ldots, z_{n_2}^*\}.$$

We claim $B \cup (B_1 - B) \cup (B_2 - B) = B \cup B_1 \cup B_2$ is a linearly independent set. If there exists $r_i, s_j, t_k \in F$ so that

$$\sum_{i=1}^{\ell} r_i x_i^* + \sum_{j=1}^{n_1} s_j y_j^* + \sum_{k=1}^{n_2} t_k z_k^* = 0, \qquad \text{(Eqn 1)}$$

then

$$\sum_{i=1}^{\ell} r_i x_i^* + \sum_{j=1}^{n_1} s_j y_j^* = -\sum_{k=1}^{n_2} t_k z_k^* \in W_1 \cap W_2$$

which means

$$-\sum_{k=1}^{n_2} t_k z_k^* \in \text{span}\{x_1^*, \ldots, x_\ell^*\},$$

which is a contradiction unless all t_k are zero. Then (Eqn 1) becomes

$$\sum_{i=1}^{\ell} r_i x_i^* + \sum_{j=1}^{n_1} s_j y_j^* = 0,$$

and since the x_i* and y_j^* form a basis for W_2, all the r_i, s_j are zero.

Now, extend $B \cup B_1 \cup B_2$ to a basis for V^*, say

$$\mathfrak{B} = \{x_1^*, \ldots, x_\ell^*\} \cup \{y_1^*, \ldots, y_{n_1}^*\} \cup \{z_1^*, \ldots, z_{n_2}^*\} \cup \{f_1^*, \ldots, f_m^*\}.$$

Since V is finite dimensional, $V^{**} \cong V$, which means $\mathfrak{B}$ is dual to a basis

$$\{x_1, \dots, x_\ell\} \cup \{y_1, \dots, y_{n_1}\} \cup \{z_1, \dots, z_{n_2}\} \cup \{f_1, \dots, f_m\}$$

of V. Now let $v \in \text{Ann}(W_1 \cap W_2)$. Then $v \in V$ and so there exists $\alpha_i, \beta_j, \gamma_k, \delta_p \in F$ so that

$$v = \sum_{i=1}^{\ell} \alpha_i x_i + \sum_{j=1}^{n_1} \beta_j y_j + \sum_{k=1}^{n_2} \gamma_k z_k + \sum_{p=1}^{m} \delta_p f_p.$$

Since $x_{i_0}^*(v) = 0$ for all $i_0 \in \{1, \dots, \ell\}$, then

$$0 = x_{i_0}^*(v) = \sum_{i=1}^{\ell} \alpha_i x_{i_0}(x_i) = a_{i_0},$$

and so $a_i = 0$ for all $i \in \{1, \dots, \ell\}$. Thus,

$$v = \sum_{j=1}^{n_1} \beta_j y_j + \sum_{k=1}^{n_2} \gamma_k z_k + \sum_{p=1}^{m} \delta_p f_p.$$

Notice that $\sum_{j=1}^{n_1} \beta_j y_j \in \text{Ann}(W_2)$ since

$$z_{k_0}^* \left(\sum_{j=1}^{n_1} \beta_j y_j \right) = \sum_{j=1}^{n_1} \beta_j z_{k_0}^*(y_j) = 0,$$

and $\sum_{k=1}^{n_2} \gamma_k z_k + \sum_{p=1}^{m} \delta_p f_p \in \text{Ann}(W_1)$ since

$$y_{j_0}^* \left(\sum_{k=1}^{n_2} \gamma_k z_k + \sum_{p=1}^{m} \delta_p f_p \right) = \sum_{k=1}^{n_2} \gamma_k y_{j_0}^*(z_k) + \sum_{p=1}^{m} \delta_p y_{j_0}^*(f_p) = 0.$$

Thus $v \in \text{Ann}(W_1) + \text{Ann}(W_2)$.

Conversely, if $v = w + u \in \text{Ann}(W_1) + \text{Ann}(W_2)$, then for $f \in W_1 \cap W_2$

$$f(v) = f(w + u) = f(w) + f(v) = 0,$$

so $v \in \text{Ann}(W_1 \cap W_2)$.

(c):($\Rightarrow$) If W_1 and W_2 are the same then certainly so are their annihilators.

($\Leftarrow$) Let $\{v_1, \dots, v_k\}$ be a basis of $\text{Ann}(W_1) = \text{Ann}(W_2) \subseteq V$. Now extend this to a basis $\{v_1, \dots, v_k, \dots, v_n\}$ of V, with dual basis $\{v_1^*, \dots, v_n^*\}$ for V^*. For any $w^* \in W_1$, we can write

$$w^* = \sum_{i=1}^{n} a_i v_i^*.$$

Then for $\ell \in \{1 \dots k\}$, we have $w^*(v_\ell) = 0$ since $w^* \in W_1$ and $v_1, \dots, v_k \in \text{Ann}(W_1)$, and so

$$w^*(v_\ell) = \sum_{i=1}^{n} a_i v_i^*(v_\ell) = \sum_{i=k+1}^{n} a_i v_i^*(v_\ell).$$

So $W_1 \subseteq \text{span}(v_{k+1}^*, \dots, v_k^*)$ and similarly, $W_2 \subseteq \text{span}(v_{k+1}^*, \dots, v_k^*)$.

(d): If $v \in \operatorname{Ann}(\operatorname{span} S)$, then for any $s^* \in S \subseteq \operatorname{span} S$, we have $s^*(v) = 0$. Conversely, if $v \in \operatorname{Ann}(S)$, then for $s^* = \sum a_i s_i^* \in \operatorname{span} S$ for $s_i^* \in S$, we have

$$s^*(v) = \sum a_i s_i^*(v) = 0.$$

(e): Let $u \in \operatorname{Ann}(S)$, with $u = \sum_{i=1}^n a_i v_i$. For $j \in \{1, \dots, k\}$, $v_j^* \in S$, so

$$0 = v_j^*(u) = \sum_{i=1}^n a_i v_j(v_i) = a_j.$$

So $a_j = 0$ for all $j \in \{1, \dots, k\}$. Thus $u \in \operatorname{span}\{v_{k+1}, \dots, v_n\}$.

Conversely, let $u \in \operatorname{span}\{v_{k+1}, \dots, v_n\}$ with

$$u = \sum_{i=k+1}^n a_i v_i.$$

Then for $v_j^* \in S$, $1 \leq j \leq k$,

$$v_j^*(u) = \sum_{i=k+1}^n a_i v_j^*(v_i) = 0.$$

(f): Let $\{v_1^*, \dots, v_k^*\}$ be a basis for W^*. Extend to a basis $\mathfrak{B} = \{v_1^*, \dots, v_n^*\}$ for V^*. The dual basis of $\mathfrak{B}$ in $V^{**} \cong V$ is a basis $\{v_1, \dots, v_n\}$ for V. By part (e),

$$\dim \operatorname{Ann}(W^*) = \dim \operatorname{span}\{v_{k+1}, \dots, v_n\} = n - k = \dim(V) - \dim(W^*).$$

☕

Exercise 12.1.2. Let M be a module over the integral domain R.

(a) Suppose that M has rank n and that $x_1, x_2, \dots, x_n$ is any maximal set of linearly independent elements of M. Let $N = Rx_1 + \cdots + Rx_n$ be the submodule generated by $x_1, x_2, \dots, x_n$. Prove that N is isomorphic to R^n and that the quotient M/N is a torsion R–module (equivalently, the elements $x_1, \dots, x_n$ are linearly independent and for any $y \in M$ there is a nonzero element $r \in R$ such that ry can be written as a linear combination $r_1x_2 + \cdots + r_nx_n$ of the x_i).

Proof. Define a map $\varphi : N \to R^n$ by $r_1x_1 + \cdots + r_nx_n \mapsto (r_1, \dots, r_n)$. Let $s \in R$. Then

$$\begin{aligned}
\varphi((r_1x_1 + \cdots + r_nx_n) + s(t_1x_1 + \cdots + t_nx_n)) &= \varphi((r_1 + st_1)x_1 + \cdots + (r_n + st_n)x_n) \\
&= (r_1 + st_1, \dots, r_n + st_n) \\
&= (r_1, \dots, r_n) + s(t_1, \dots, t_n) \\
&= \varphi(r_1x_2 + \cdots + r_nx_n) \\
&\quad + s\varphi(t_1x_1 + \cdots + t_nx_n),
\end{aligned}$$

and so φ is an $R--$module homomorphism. If

$$\varphi(r_1x_1 + \cdots + r_nx_n) = (0, \dots, 0),$$

then $r_1, \dots, r_n = 0$ and hence $r_1x_1 + \cdots + r_nx_n = 0$. Thus φ is injective. If $(r_1, \dots, r_n) \in R^n$ then clearly $\varphi(r_1x_1 + \cdots + r_nx_n) = (r_1, \dots, r_n)$. Therefore φ is an isomorphism of R–modules.

Let $y \in M - \{0_M\}$. Then the set $\{x_1, \dots, x_n, y\}$ is a linearly dependent set since M has rank n. In particular, there exists $r_1, \dots, r_{n+1} \in R$, not all zero so that

$$r_1x_1 + \cdots + r_nx_n + r_{n+1}y = 0.$$

If $r_{n+1} = 0$, then $r_1, \dots, r_n = 0$ since $x_1, \dots, x_n$ are linearly independent. So $r_{n+1} \neq 0$, and we have

$$-r_{n+1}y = r_1x_1 + \dots r_nx_n.$$

Thus $-r_{n+1}y \in N$ and hence M/N is a torsion R–module. ☕

(b) Prove conversely that if M contains a submodule N that is free of rank n (i.e., $N \cong R^n$) such that the quotient M/N is a torsion R–module then M has rank n.

Proof. Let $y_1, \dots, y_{n+1} \in M$ and let $\{a_1, \dots, a_n\}$ be an R-basis for N. Since M/N is torsion, there exists $r_1, \dots, r_{n+1} \in R - \{0_R\}$ such that $r_iy_i + N = N$, i.e., $r_iy_i \in N$ for all $1 \leq i \leq n + 1$. Since N is a free R–module, then any $n + 1$ elements in N are linearly dependent. So there exists $t_1, \dots, t_{n+1} \in R$, not all zero so that

$$t_1(r_1y_1) + \cdots + t_{n+1}(r_{n+1}y_{n+1}) = 0.$$

Letting $\alpha_i = t_ir_i$ for all $1 \leq i \leq n + 1$, we have

$$\alpha_1y_1 + \cdots + \alpha_{n+1}y_{n+1} = 0,$$

i.e., we have a linear dependence relationship for the y_i since $r_i \neq 0_R$ and at least one t_i is nonzero. Thus M has rank n. ☕

Exercise 12.1.3. Let R be an integral domain and let A and B be R–modules of ranks m and n, respectively. Prove that the rank of $A \oplus B$ is $m+n$.

Proof. By the previous exercise, A contains a submodule C (namely, the submodule generated by a maximal set of linearly independent elements in A) which is isomorphic to R^m so that $A/C \cong A/R^m$ is torsion. Similarly, B contains a submodule D so that $B/D \cong B/R^n$ is torsion. Notice that the map $a+b \mapsto (a+C)+(b+D)$ gives an isomorphism between $A \oplus B$ and $A/C + B/D$. The map is clearly surjective, and is a homomorphism since it is simply the natural projection in each coordinate. Moreover, $C+D$ is contained in the kernel of this map; and if $a+b \mapsto 0$, then $a \in C$ and $b \in D$, i.e., $a+b \in C+D$. Thus we get

$$A \oplus B/(C+D) \cong A/C + B/D.$$

In particular, $A \oplus B/(C+D)$ is torsion since both A/C and B/D are torsion. Moreover, $C+D \cong R^m + R^n \cong R^{m+n}$ is free of rank $m+n$. Therefore, $A \oplus B$ contains free submodule $C+D$ of rank $m+n$ so that $A \oplus B/(C+D)$ is torsion. Hence by the previous exercise, (part (b)), $A \oplus B$ has rank $m+n$. ☕

Exercise 12.1.4. Let R be an integral domain, let M be an R–module and let N be a submodule of M. Suppose M has rank n, N has rank r, and the quotient M/N has rank s. Prove that $n = r+s$.

Proof. Let $x_1, \ldots, x_s$ be elements of M such that $\overline{x_1}, \ldots, \overline{x_s}$ is a maximal set of linearly independent elements in M/N. Let $x_{s+1}, \ldots, x_{s+r}$ be a maximal set of linearly independent elements in N. Suppose

$$t_1 x_1 + \cdots + t_{s+r} x_{s+r} = 0_M \tag{$*$}$$

for some $t_1, \ldots, t_{r+s} \in R$. if $\pi : M \to M/N$ is the natural projection homomorphism, then applying π to $(*)$ gives

$$\begin{aligned}\overline{0_M} = \pi(t_1 x_1 + \cdots + t_{s+r} x_{s+r}) &= t_1 \pi(x_1) + \cdots + t_{s+r} \pi(x_{s+r}) \\ &= t_1 \pi(x_1) + \cdots + t_s \pi(x_s) \\ &= t_1 \overline{x_1} + \cdots + t_s \overline{x_s}.\end{aligned}$$

Then $t_1, \ldots, t_s$ are all 0 since $\overline{x_1}, \ldots, \overline{x_s}$ are linearly independent in M/N. Then $(*)$ becomes

$$t_{s+1} x_{s+1} + \cdots + t_{s+r} x_{s+r} = 0_M,$$

which means $t_{s+1}, \ldots, t_{s+r}$ are all 0 since $x_{s+1}, \ldots, x_{s+r}$ are linearly independent. Hence $x_1, \ldots, x_{s+r}$ are linearly independent in M.

Let $y \in M$. Then either $y \in M$ or $y \in M - N$. Consider the set $\{x_1, \ldots, x_{s+r}, y\}$. If this set is linearly independent, then if $y \in N$, the elements $x_{s+1}, \ldots, x_{s+r}, y$ are linearly independent in N, a contradiction. If $y \in M - N$, then $\overline{y} \neq \overline{0_M}$ and so the elements $\overline{x_1}, \ldots, \overline{x_s}, \overline{y}$ are linearly independent, a contradiction. Hence if we let $P = Rx_1 + \ldots Rx_{r+s}$, then M/P is a torsion R–module and P has rank $r+s$. Then be exercise 12.1.2 (b), M has rank $r+s$. ☕

Exercise 12.1.11. Let R be a P.I.D., let a be a nonzero element of R and let $M = R/(a)$. For any prime p of R prove that

$$p^{k-1}M/p^kM \cong \begin{cases} R/(p) & \text{if } k \leq n \\ 0 & \text{if } k > n, \end{cases}$$

where n is the power of p dividing a in R.

Proof. Suppose $k \leq n$ and define a map

$$\varphi : p^{k-1}(R/(a)) \to R/(p) \quad \text{by} \quad \overline{p^{k-1}r} = p^{k-1}r + (a) \mapsto r + (p).$$

Suppose $\overline{p^{k-1}r_1} = \overline{p^{k-1}r_2}$. Then

$$p^{k-1}(r_1 - r_2) \in (a) \subseteq (p^n) \subseteq (p) \implies p^{k-1}r_1 + (p) = p^{k-1}r_2 + (p)$$

and so $\varphi\left(\overline{p^{k-1}r_1}\right) = \varphi\left(\overline{p^{k-1}r_2}\right)$ and hence φ is well defined. It's clear that φ is surjective. Moreover, if $\varphi\left(\overline{p^{k-1}r}\right) = 0 + (p)$, then $r \in (p)$, and so $r = ps$ for some $s \in R$. Then

$$\overline{p^{k-1}r} = \overline{p^{k-1}ps} = \overline{p^k s} \in p^k(R/(a)).$$

Conversely if $\overline{p^k t} \in p^k(R/(a))$, then

$$\varphi\left(\overline{p^k t}\right) = \varphi\left(\overline{p^{k-1}pt}\right) = pt + (p) = 0 + (p),$$

and so $\ker \varphi = p^k(R/(a))$. Then by the First Isomorphism Theorem,

$$p^{k-1}(R/(a))/p^k(R/(a)) \cong R/(p).$$

Now we want to show that $p^m M = p^n M$ for all $m \geq n$. One inclusion is clear: If $p^m m_1 \in p^m M$, then

$$p^m m_1 = p^n p^{m-n} m_1 \in p^n M.$$

For the other inclusion, let $p^n x + (a) \in p^n M$. Notice that we can write $a = p^n b$ with $p \nmid b$. Then $\gcd(b, p^{m-n}) = 1$, and so there exists $r, s \in R$ for which $rb + sp^{m-n} = 1$. Then $x = xrb + xsp^{m-n}$, and so

$$\begin{aligned} p^n x + (a) &= p^n(xrb + xsp^{m-n}) + (a) \\ &= (p^n xrb + xsp^m) + (a) \\ &= (p^n xrb + (a)) + (xsp^m + (a)) \\ &= xsp^m + (a) \in p^m M. \end{aligned}$$

Therefore, when $k > n$, $k - 1 \geq n$, and so $p^{k-1}M/p^kM = p^nM/p^nM = 0$.

☕

Exercise 12.1.12. Let R be a P.I.D. and let p be a prime in R.

(a) Let M be a finitely generated torsion R–module. Use the previous exercise to prove that $p^{k-1}M/p^kM \cong F^{n_k}$ where F is the field $R/(p)$ and n_k is the number of elementary divisors of M which are powers p^α with $\alpha \geq k$.

Proof. By the Fundamental Theorem of Finitely Generated Modules over a P.I.D.,

$$M \cong R/(p_1^{\alpha_1}) \oplus \cdots \oplus R/(p_t^{\alpha_t})$$

where $p_1^{\alpha_1}, \ldots, p_t^{\alpha_t}$ are positive powers of primes in R. Define $M_i := R/(p_i^{\alpha_i})$ for all $1 \leq i \leq t$. Then

$$\begin{aligned} p^{k-1}M/p^kM &\cong p^{k-1}(M_1 \oplus \cdots \oplus M_t)\Big/p^k(M_1 \oplus \cdots \oplus M_t) \\ &\cong p^{k-1}M_1 \oplus \cdots \oplus p^{k-1}M_t \Big/ p^kM_1 \oplus \cdots \oplus p^kM_t \end{aligned}$$

By Exercise 12.1.7,

$$p^{k-1}M/p^kM \cong \left(p^{k-1}M_1\Big/p^kM_1\right) \oplus \cdots \oplus \left(p^{k-1}M_t\Big/p^kM_t\right).$$

Now, if there is an elementary divisor $p_i^{\alpha_i}$ of M which is associate to p^{α_i} and $k \leq \alpha_i$ then we get

$$p^{k-1}M/p^kM \cong R/(p) = F$$

by the previous exercise. On the other hand, if there is an elementary divisor $p_i^{\alpha_i}$ of M which is not associate to p^{α_i}, *or* which has power $\alpha_i < k$, then

$$p^{k-1}M_i/p^kM_i \cong 0.$$

Let n_k be the number of elementary divisors of M which are associate p^α with $\alpha \geq k$. Then

$$p^{k-1}M/p^kM \cong R/(p) \oplus \cdots \oplus R/(p) \cong F^{n_k}.$$

☕

(b) Suppose M_1 and M_2 are isomorphic finitely generated torsion R–modules. Use (a) to prove that, for every $k \geq 0$, M_1 and M_2 have the same number of elementary divisors p^α with $\alpha \geq k$. Prove that this implies M_1 and M_2 have the same set of elementary divisors.

Proof. Let $\varphi : M_1 \to M_2$ be an isomorphism between M_1 and M_2. Define a map

$$\psi : p^{k-1}M_1 \to p^{k-1}M_2/p^kM_2$$

by

$$p^{k-1}m \mapsto \overline{p^{k-1}\varphi(m)} = p^{k-1}\varphi(m) + p^kM_2.$$

Let $r \in R$ and $m, n \in M_1$. Then

$$\begin{aligned}\psi((p^{k-1}m) + r(p^{k-1}n)) &= \psi(p^{k-1}(m + rn)) \\ &= \overline{p^{k-1}\varphi(m + rn)} \\ &= \overline{p^{k-1}\varphi(m)} + r\overline{p^{k-1}\varphi(n)} \\ &= \psi(p^{k-1}m) + r\psi(p^{k-1}n).\end{aligned}$$

Hence ψ is an R–module homomorphism.

If $\overline{p^{k-1}m} \in p^{k-1}M_2/p^kM_2$, then there exists $n \in M_1$ such that $\varphi(n) = m$, and so

$$\psi(p^{k-1}n) = \overline{p^{k-1}\varphi(n)} = \overline{p^{k-1}m},$$

and thus ψ is surjective.

If $\psi(p^{k-1}m) = \overline{0}$, then $p^{k-1}\varphi(m) \in p^kM_2$ and so $p^{k-1}\varphi(m) = p^k\ell$ for some $\ell \in M_2$. Let $q \in M_1$ be such that $\varphi(q) = \ell$. Then

$$\begin{aligned}p^{k-1}\varphi(m) - p^k\ell = 0 &\iff p^{k-1}(\varphi(m) - p\ell) = 0 \\ &\iff p^{k-1}(\varphi(m) - p\varphi(q)) = 0 \\ &\iff \varphi(p^{k-1}(m - pq)) = 0 \\ &\iff p^{k-1}(m - pq) = 0 \\ &\iff p^{k-1}m = p^kq \in p^kM_1.\end{aligned}$$

Hence $\ker\psi = p^kM_1$. Therefore by the First Isomorphism Theorem,

$$p^{k-1}M_1/p^kM_1 \cong p^{k-1}M_2/p^kM_2.$$

Let n_k and number of elementary divisors of M_1 which are powers p^α with $\alpha \geq k$. Similarly let n'_k be this number for M_2. Then by part (a),

$$F^{n_k} \cong p^{k-1}M_1/p^kM_1 \cong p^{k-1}M_2/p^kM_2 \cong F^{n'_k},$$

where F is the field $R/(p)$. Then $n_k = n'_k$.

Since, n_k is number of elementary divisors of M_1 and M_2 which are associate to p^α, for $\alpha \geq k$, and n_{k+1} is number of elementary divisors of M_1 and M_2 which are associate to p^α, for $\alpha \geq k+1$, then $n_k - n_{k+1}$ is the number of elementary divisors of M_1 and M_2 which are associate to p^k. Since this is true for all k and for all primes $p \in R$, then M_1 and M_2 have the same elementary divisors.

☕

Exercise 12.1.16. Prove that M is finitely generated if and only if there is a surjective R-homomorphism $\varphi : R^n \to M$ for some integer n (this is true for any ring R).

Proof. ($\Rightarrow$) If $M = \{0\}$ then $\varphi : R^0 = \{0\} \to M$ is surjective. If $M \neq 0$, then let $M = Rx_1 + \cdots + Rx_n$ and define

$$\varphi : R^n \to M \quad \text{by} \quad (r_1, \ldots, r_n) \mapsto r_1x_1 + \ldots r_nx_n.$$

This map is certainly a surjective R–module homomorphism.

($\Leftarrow$) If $\varphi : R^n \to M$ is a surjective R–module homomorphism, then let e_i be the standard basis elements of R^n. Define $\varphi(e_i) = x_i$. If $m \in M$ there exists $r = \sum_{i=1}^n c_ie_i \in R^n$ such that

$$m = \varphi(r) = \sum_{i=1}^n c_i\varphi(e_i) = \sum_{i=1}^n c_ix_i.$$

Hence the set $\{x_1, \ldots, x_n\}$ generates M.

☕

Exercise 12.1.15. Prove that if R is a Neotherian ring then R^n is a Neotherian R–module.

Proof. We proceed by induction on n. When $n = 1$, we're done. Assume R^{n-1} is Neotherian for $n > 1$. Let $\{b_1, \dots, b_n\}$ be a basis for R^n. Let M be a submodule of R^n. Define

$$A = \{a_1 \mid (a_1, \dots, a_n) \in M\}.$$

Then A is nonempty since M is, and if $x, y \in A$ and $r \in R$, then there exists elements of M with x and y in their first coordinate: $(x, \dots), (y, \dots) \in M$. Then $(x, \dots) + r(y, \dots) = (x+ry, \dots)$, and so $x+ry \in A$, and thus A is a submodule of R. Since R is Neotherian, then A is finitely generated, say by $\{a_1, \dots, a_k\}$. For all $1 \leq i \leq k$ let a_i be the first coordinate of $m_i \in M$.

Now let $m \in M$ and a be the first coordinate of m. Then $a \in A$ and so

$$a = \sum_{i=1}^{k} r_i a_i$$

for some $r_1, \dots, r_k \in R$. Then

$$n := m - \sum_{i=1}^{k} r_i m_i$$

has first coordinate zero. So, $n \in R^{n-1} \cap M$ where we are viewing R^{n-1} as the set of elements in R^n whose first coordinate is zero. Then if $s, t \in R^{n-1} \cap M$ and $v \in R^{n-1}$, then clearly $s + vt \in inR^{n-1} \cap M$. So $R^{n-1} \cap M$ is a submodule of R^{n-1}, and by the induction hypothesis R^{n-1} is Neotherian and so $R^{n-1} \cap M$ is finitely generated, say by $\{n_1, \dots n_\ell\}$. Then we can write

$$n = \sum_{i=1}^{\ell} s_i n_i$$

for some $s_i \in R$. So,

$$m = n + \sum_{i=1}^{k} r_i m_i = \sum_{i=1}^{\ell} s_i n_i + \sum_{i=1}^{k} r_i m_i,$$

and hence $\{m_1, \dots, m_k, n_1, \dots, n_\ell\}$ generate M. Since M was arbitrary, every submodule of R^n is finitely generated and hence R^n is Neotherian. ☕

Exercise 12.1.19. By the previous two exercises, we may perform elementary row and column operations on a given relations matrix by choosing different generators for R^n and $\ker\phi$. If all relation matrices are the zero matrix then $\ker\varphi = 0$ and $M \cong R^n$. Otherwise let a_1 be the (nonzero) gcd of all the entries in a fixed initial relations matrix for M.

(a) Prove that by elementary row and column operations we may assume a_1 occurs in a relations matrix of the form

$$\begin{pmatrix} a_1 & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_2 & \dots & a_{mn} \end{pmatrix}$$

where a_1 divides a_{ij} for $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$.

(b) Prove that there is a relations matrix of the form

$$\begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

where a_1 divides all the entries.

Proof. Staring with the matrix in part (a), since $a_1 | a_{21}$, then there exists d_{21} such that $a_{21} = d_{21}a_1$. So we perform the following row operation to put a 0 in the $(2,1)$-entry of the matrix: $R_2 - d_{21}R_2 \to R_2$. Continuing to do this for each row, we perform the row operation $R_i - d_{i1}R_1 \to R_i$ for all $i \in \{1, \dots, m\}$ and obtain all zeroes in the first column (excluding the a_1 in the (1,1)-entry). Similarly, for each $j \in \{1, \dots, n\}$, there exists d_{1j} such that $a_{1j} = d_{1j}a_1$. Therefore, we perform the following column operation for each $j \in \{1, \dots, n\}$: $C_j - d_{1j}C_1 \to C_j$. By this we obtain the desired matrix. ☕

(c) Let a_2 be the gcd of all the entries excepts the element a_1 in the relations matrix in (b). Prove that there is a relations matrix of the form

$$\begin{pmatrix} a_1 & 0 & 0 & \dots & 0 \\ 0 & a_2 & 0 & \dots & 0 \\ 0 & 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & a_{m3} & \dots & a_{mn} \end{pmatrix}$$

where a_1 divides a_2 and a_2 divides all the other entries of the matrix.

Proof. Starting with the matrix obtained in part (b), we can apply part (b) again to obtain zeros in the second row and second column, except at the (2,2)-position. ☕

(d) Prove that there is a relations matrix of the form $\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$ where D is a diagonal matrix with nonzero entries $a_1, a_2, \dots, a_k$, $k \leq n$, satisfying

$$a_1|a_2|\dots|a_k.$$

Conclude that

$$M \cong R/(a_1) \oplus R/(a_2) \oplus \dots \oplus R/(a_k) \oplus R^{n-k}.$$

Proof. The matrix D is of the form

$$D = \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_k \end{pmatrix}.$$

By using part (c), we can obtain such a matrix D by induction. Moreover, if $\varphi : R^n \to M$ is a surjective R–module homomorphism we have $M \cong R^n / \ker \varphi$. From part (d), we have that

$$\ker \varphi = a_1 \oplus \dots a_k \oplus 0^{n-k}.$$

Then by Exercise 12.1.7,

$$M \cong R^n / \ker \varphi \cong R/(a_1) \oplus \dots \oplus R/(a_k) \oplus R^{n-k}/0^{n-k} \cong R/(a_1) \oplus \dots \oplus R/(a_k) \oplus R^{n-k}$$

☕

Exercise 12.2.14. Determine all possible rational canonical forms for a linear transformation with characteristic polynomial $x^2(x^2+1)^2$.

Proof. Let F be a field and T be a linear transformation over an F–module with $\chi_T = x^2(x^2+1)^2$. First suppose x^2+1 is irreducible over F. Since $m_T(x)$ divides $\chi_T(x)$ and must be divisible by all the factors appearing in $\chi_T(x)/m_T(x)$, we get the following possibilities for $m_T(x)$ and corresponding invariant factors:

(i) $m_T(x) = x^2(x^2+1)$. Invariant factors: $x^2+1 \mid x^2(x^2+1)^2 = x^4 + x$.

(ii) $m_T(x) = x^2(x^2+1)^2$. Invariant factors: $x^2(x^2+1) = x^6 + 2x^4 + x^2$.

(iii) $m_T(x) = x(x^2+1)$. Invariant factors: $x(x^2+1) \mid x(x^2+1)$.

(iv) $m_T(x) = x(x^2+1)^2$. Invariant factors: $x \mid x(x^2+1) = x^5 + 2x^3 + x$.

Then we get the corresponding rational canonical forms:

$$\text{(i)} \begin{pmatrix} 0 & -1 & & & & \\ 1 & 0 & & & & \\ & & 0 & 0 & 0 & 0 \\ & & 1 & 0 & 0 & 0 \\ & & 0 & 1 & 0 & -1 \\ & & 0 & 0 & 1 & 0 \end{pmatrix} \qquad \text{(ii)} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\text{(iii)} \begin{pmatrix} 0 & 0 & 0 & & & \\ 1 & 0 & -1 & & & \\ 0 & 1 & 0 & & & \\ & & & 0 & 0 & 0 \\ & & & 1 & 0 & -1 \\ & & & 0 & 1 & 0 \end{pmatrix} \qquad \text{(iv)} \begin{pmatrix} 0 & & & & & \\ & 0 & 0 & 0 & 0 & 0 \\ & 1 & 0 & 0 & 0 & -1 \\ & 0 & 1 & 0 & 0 & 0 \\ & 0 & 0 & 1 & 0 & -2 \\ & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Now suppose x^2+1 is irreducible in F with $x^2+1 = (x+\alpha)(x+\beta)$. By expanding and comparing coefficients, we find that $\beta = -\alpha$, i.e., $x^2+1 = (x+\alpha)(x-\alpha)$. Note that $\alpha^2 = -1$. This gives the following additional possibilities for $m_T(x)$:

(v) $m_T(x) = x^2(x+\alpha)^2(x-\alpha)$.
Invariant factors: $x-\alpha, \quad x^2(x+\alpha)^2(x-\alpha) = x^5 + \alpha x^4 + x^3 + \alpha x^2$.

(vi) $m_T(x) = x^2(x-\alpha)^2(x+\alpha)$.
Invariant factors: $x+\alpha, \quad x^2(x-\alpha)^2(x+\alpha) = x^5 - \alpha x^4 + x^3 - \alpha x^2$.

(vii) $m_T(x) = x(x+\alpha)^2(x-\alpha)$.
Invariant factors: $x(x-\alpha), \quad x(x+\alpha)^2(x-\alpha) = x^4 + \alpha x^3 + x^2 + \alpha x$.

(viii) $m_T(x) = x(x-\alpha)^2(x+\alpha)$.
Invariant factors: $x(x+\alpha), \quad x(x-\alpha)^2(x+\alpha) = x^4 - \alpha x^3 + x^2 - \alpha x$.

Then we get the corresponding rational canonical forms:

$$\text{(v)} \begin{pmatrix} \alpha & & & & & \\ & 0 & 0 & 0 & 0 & 0 \\ & 1 & 0 & 0 & 0 & 0 \\ & 0 & 1 & 0 & 0 & -\alpha \\ & 0 & 0 & 1 & 0 & -1 \\ & 0 & 0 & 0 & 1 & -\alpha \end{pmatrix} \qquad \text{(vi)} \begin{pmatrix} -\alpha & & & & & \\ & 0 & 0 & 0 & 0 & 0 \\ & 1 & 0 & 0 & 0 & 0 \\ & 0 & 1 & 0 & 0 & \alpha \\ & 0 & 0 & 1 & 0 & -1 \\ & 0 & 0 & 0 & 1 & \alpha \end{pmatrix}$$

$$\text{(vii)} \begin{pmatrix} 0 & 0 & & & & \\ 1 & \alpha & & & & \\ & & 0 & 0 & 0 & 0 \\ & & 1 & 0 & 0 & -\alpha \\ & & 0 & 1 & 0 & -1 \\ & & 0 & 0 & 1 & -\alpha \end{pmatrix} \qquad \text{(viii)} \begin{pmatrix} 0 & 0 & & & & \\ 1 & -\alpha & & & & \\ & & 0 & 0 & 0 & 0 \\ & & 1 & 0 & 0 & \alpha \\ & & 0 & 1 & 0 & -1 \\ & & 0 & 0 & 1 & \alpha \end{pmatrix}$$

Exercise 12.2.18. Let V be a finite dimensional vector space over $\mathbb{Q}$ and suppose T is a nonsingular linear transformation of V such that $T^{-1} = T^2 + T$. Prove that the dimension of V is divisible by 3. If the dimension of V is precisely 3, prove that all such linear transformations T are similar.

Proof. The given conditions give $I = T^3 + T^2$, i.e., $T^3 + T^2 - I = 0$ and so $m_T(x) = x^3 + x^2 - 1$, which is irreducible by the root test. Since m_T is irreducible, the invariant factors of T will be m_T itself, repeated say n times. Then $\chi_T = (m_T)^n$.The dimension of V is equal to the degree of χ_T, namely $3n$. Hence the dimension of V divides 3.

If the dimension of V is 3, then and two linear transformations with minimal polynomial $x^3 + x^2 - 1$ will have the same invariant factors; namely, $x^3 + x^2 - 1$. Hence they will have the same rational canonical form and therefore be similar.

Exercise 12.3.16. Determine the Jordan canonical form for the matrix

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Proof. We start by finding the Smith Normal form for A. We obtain this by performing the following row/column operations:

$$\begin{aligned} \boldsymbol{-R_1 \to R_1};\ C_1 + xC_2 \to C-1;\ \boldsymbol{R_2 - x(x-1)R_1 \to R_2};\ C_2 - C_1 \to C_2; \\ C_3 - C_1 \to C_3;\ C_4 - C_1 \to C_4;\ C_4 - C_3 \to C_4;\ C_2 \leftrightarrow C_4; \\ \boldsymbol{R_3 + xR_2 \to R_3};\ \boldsymbol{R_4 - (x-1)R_2 \to R_4};\ C_3 + (x(x-1))C_2 \to C_3; \\ C_4 + (x-1)^2 C_2 \to C_4;\ C_4 - C_3 \to C_4;\ C_3 + xC_4 \to C_3;\ C_3 \leftrightarrow C_4; \\ \boldsymbol{R_3 + R_4 \to R_3};\ -C_3 \to C_3;\ -C_4 \to C_4. \end{aligned} \tag{$*$}$$

We get

$$xI - A = \begin{pmatrix} x-1 & -1 & -1 & -1 \\ 0 & x-1 & 0 & 1 \\ 0 & 0 & x-1 & -1 \\ 0 & 0 & 0 & x-1 \end{pmatrix} \xrightarrow{(*)} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & (x-1)^2 & 0 \\ 0 & 0 & 0 & (x-1)^2 \end{pmatrix}.$$

So the invariant factors of A are $a_1(x) = a_2(x) = (x-1)^2$. Now we find the matrix P' by performing on the identity matrix the column operations corresponding to the row operations used above. That is, we perform the following column operations on I:

$$\begin{aligned} (-1)C_1 \to C_1;\ C_1 + (A(A-I))C_2 \to C_1;\ C_2 - AC_3 \to C_3; \\ C_2 + (A-I)C_4 \to C_2;\ C_4 - C_3 \to C_4. \end{aligned} \tag{$**$}$$

$$I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{(**)} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix} = P'.$$

Now we find the matrix Q so that $Q^{-1}AQ$ is the Jordan canonical form of A. Let $C_i(P')$ denote the ith column of P'. The columns of Q will be given by:

Column 1: $(A-I)^{2-1}(C_3(P')) = (1\ 0\ 0\ 0)^T$, Column 2: $(A-I)^{2-2}(C_3(P')) = C_3(P')$

Column 3: $(A-I)^{2-1}(C_4(P')) = (0\ -1\ 1\ 0)^T$, Column 4: $(A-I)^{2-2}(C_4(P')) = C_4(P')$.

Therefore we get

$$Q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad Q^{-1}AQ = \begin{pmatrix} 1 & 1 & & \\ 0 & 1 & & \\ & & 1 & 1 \\ & & 0 & 1 \end{pmatrix}$$

☕

Exercise 12.3.21. Show that if $A^2 = A$ then A is similar to a diagonal matrix which only has 0's and 1's along the diagonal.

Proof. The minimal polynomial for A divides $x^2 - x$, and so the minimal polynomial will have distinct roots 0 and/or 1. Hence A is similar to a diagonal matrix (since a matrix whose minimal polynomial has distinct roots will be similar to a diagonal matrix). In particular, the diagonal will consist of the eigenvalues of A; namely, 0 and/or 1. ☕

Exercise 12.3.31. Let N be an $n \times n$ matrix with coefficients in the field F. The matrix N is said to be *nilpotent* if some power of N is the zero matrix, i.e., $N^k = 0$ for some k. Prove that any nilpotent matrix is similar to a block diagonal matrix whose blocks are matrices with 1's along the first superdiagonal and 0's elsewhere.

Proof. The minimal polynomial for N will divide the polynomial x^k, and hence the minimal polynomial will have all roots equal to 0. So the Jordan Normal form for N will have blocks with 1's along the first superdiagonal and 0's elsewhere. ☕

The following exercises outline the proof of Theorem 21:

Theorem (Theorem 21)**.** *Let A be an $n \times n$ matrix over a field F. Using the three elementary row and column operations, the $n \times n$ matrix $xI - A$ with entries from $F[x]$ can be put into the diagonal form, called the* Smith Normal Form for A

$$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & a_1(x) & & & \\ & & & & a_2(x) & & \\ & & & & & \ddots & \\ & & & & & & a_m(x) \end{pmatrix}$$

with monic nonzero elements $a_1(x), a_2(x), \ldots, a_m(x)$ of $F[x]$ with degrees at least one and satisfying $a_1(x) \mid a_2(x) \mid \ldots \mid a_m(x)$. The elements $a_1(x), a_2(x), \ldots, a_m(x)$ are the invariant factors of A.

Let V be an n-dimensional vector space with basis $v_1, v_2, \ldots, v_n$ and let T be the linear transformation of V defined by the matrix A and this choice of basis, i.e., T is the linear transformation with

$$T(v_j) = \sum_{i=1}^{n} a_{ij} v_i, \quad j = 1, 2, \ldots, n$$

where $A = (a_{ij})$. Let $F[x]^n$ be the free module of rank n over $F[x]$ and let $\xi_1, \xi_2, \ldots, \xi_n$ denote a basis. Then we have a natural surjective $F[x]$–module homomorphism

$$\varphi : F[x]^n \to V$$

defined by mapping ξ_i to v_i, $i = 1, 2, \ldots, n$. As indicated in the exercises of the previous section, the invariant factors for the $F[x]$–module V can be determined once we have determined a set of generators and the corresponding relations matrix for $\operatorname{Ker} \varphi$. Since by definition x acts on V by the linear transformation T, we have

$$x(v_j) = \sum_{i=1}^{n} a_{ij} v_i, \quad j = 1, 2, \ldots, n.$$

Exercise 12.2.22. Show that the elements

$$\nu_j = -a_{1j}\xi_i - \cdots - a_{j-1j}\xi_{j-1} + (x - a_{jj})\xi_j - a_{j+1j}\xi_{j+1} - \cdots - a_{nj}\xi_n$$

for $j = 1, 2, \ldots, n$ are elements of the kernel of φ.

Proof.

$$\begin{aligned} \varphi(\nu_j) &= -a_{1j}\varphi(\xi_1) - \cdots - a_{j-1j}\varphi(\xi_{j-1}) + (x - a_{jj})\varphi(\xi_j) - a_{j+1j}\varphi(\xi_{j+1}) - \cdots - a_{nj}\varphi(\xi_n) \\ &= -a_{1j}v_1 - \cdots - a_{j-1j}v_{j-1} + xv_j - a_{jj}v_j - a_{j+1j}v_{j+1} - \cdots - a_{nj}v_n \\ &= -a_{1j}v_1 - \cdots - a_{j-1j}v_{j-1} + \sum_{i=1}^{n} a_{ij}v_i - a_{jj}v_j - a_{j+1j}v_{j+1} - \cdots - a_{nj}v_n \\ &= 0. \end{aligned}$$

☕

Exercise 12.2.23.

(a) Show that $x\xi_j = \nu_j + f_j$ where $f_j \in F\xi_1 + \cdots + F\xi_n$ is an element of the F-vector space spanned by $\xi_1, \ldots, \xi_n$.

Proof.

$$\nu_j - x\xi_j = -a_{1j}\xi_1 - \cdots - a_{j-1j}\xi_{j-1} - a_{jj}\xi_j - a_{j+1j}\xi_{j+1} - \cdots - a_{nj}\xi_n \in F\xi_1 + \cdots + F\xi_n.$$

☕

(b) Show that

$$F[x]\xi_1 + \cdots + F[x]\xi_n = (F[x]\nu_1 + \cdots + F[x]\nu_n) + (F\xi_1 + \cdots + F\xi_n)$$

Proof. Notice that $F[x]\xi_1 + \cdots + F[x]\xi_n = F[x]^n$. Let $M := (F[x]\nu_1 + \cdots + F[x]\nu_n) + (F\xi_1 + \cdots + F\xi_n)$. We will show that M is a submodule of $F[x]^n$, and since M contains a basis of $F[x]^n$, it follows that $M = F[x]^n$.

Clearly $M \neq \varnothing$. Let $p_1, \ldots, p_n, q_1, \ldots, q_n \in F[x]$ and $a_1, \ldots, a_n, b_1, \ldots, b_n \in F$. Then

$$\left(\sum_{i=1}^n p_i\nu_i + \sum_{i=1}^n a_i\xi_i\right) + \left(\sum_{i=1}^n q_i\nu_i + \sum_{i=1}^n b_i\xi_i\right) = \sum_{i=1}^n (p_i + q_i)\nu_i + \sum_{i=1}^n (a_i + b_i)\xi_i$$

is an element of M. For $c \in F$,

$$c\left(\sum_{i=1}^n q_i\nu_i + \sum_{i=1}^n b_i\xi_i\right) = \sum_{i=1}^n (cq_i)\nu_i + \sum_{i=1}^n (cb_i)\xi_i$$

is an element of M. By part (a), we get

$$x\left(\sum_{i=1}^n b_i\xi_i\right) = \sum_{i=1}^n b_i(x\xi_i) = \sum_{i=1}^n b_i(\nu_i + f_i) = \sum_{i=1}^n b_i\nu_i + \sum_{i=1}^n b_i f_i$$

where $f_i \in F\xi_1 + \cdots + F\xi_n$ for all $1 \leq i \leq n$. So

$$\begin{aligned} x\left(\sum_{i=1}^n q_i\nu_i + \sum_{i=1}^n b_i\xi_i\right) &= x\left(\sum_{i=1}^n q_i\nu_i\right) + x\left(\sum_{i=1}^n b_i\xi_i\right) \\ &= \sum_{i=1}^n xq_i\nu_i + \left(\sum_{i=1}^n b_i\nu_i + \sum_{i=1}^n b_i f_i\right) \\ &= \sum_{i=1}^n (xq_i + b_i)\nu_i + \sum_{i=1}^n b_i f_i \end{aligned}$$

is an element of M. Therefore M is a submodule of $F[x]^n$. ☕

Exercise 12.2.24. Show that $\nu_1, \nu_2, \dots, \nu_n$ generate the kernel of φ.

Proof. By the previous exercise, an element in $F[x]^n$ can be written as the sum of an element in the $F[x]$–module generated by $v_1, \dots, v_n$ with an element of the form $b_1\xi_1 + \dots b_n\xi_n$ where the b_i are elements of F. Let

$$\kappa = \sum_{i=1}^{n} q_i \nu_i + \sum_{i=1}^{n} b_i \xi_i$$

be an element of $F[x]^n$. Recall from Exercise 12.2.22 that $v_j \in \operatorname{Ker}\varphi$ for all $1 \le j \le n$. Then

$$\begin{aligned} \kappa \in \operatorname{Ker}\varphi &\iff 0_V = \varphi(\kappa) = \sum_{i=1}^{n} q_1 \varphi(\nu_n) + \sum_{i=1}^{n} b_i \varphi(\xi_i) = \sum_{i=1}^{n} b_i v_i \\ &\iff b_i = 0_F \text{ for all } 1 \le i \le n \\ &\iff \kappa = \sum_{i=1}^{n} q_i \nu_i \end{aligned}$$

and hence $\nu_1, \nu_2, \dots, \nu_n$ generate the the kernel of φ. ☕

Exercise 12.2.25. Show that the generators $\nu_1, \nu_2, \dots, \nu_n$ of $\operatorname{Ker}\varphi$ have corresponding relations matrix

$$\begin{pmatrix} x-a_{11} & -a_{21} & \dots & -a_{n1} \\ -a_{12} & x-a_{22} & \dots & -a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{1n} & -a_{2n} & \dots & x-a_{nn} \end{pmatrix} = xI - A^t$$

where A^t is the transpose of A. Conclude that Theorem 21 and the algorithm for determining the invariant factors of A follows by Exercises 16 and 19 in the previous section (note that the row and column operations necessary to diagonalize this relations matrix are the column and row operations necessary to diagonalize the matrix in Theorem 21, which explains why the invariant factor algorithm keeps track of the *row* operations used).

Proof. Based on the form of ν_j given in Exercise 12.2.22, the jth row of the relations matrix (which corresponds the generator ν_j of $\operatorname{Ker}\varphi$) is by definition

$$\begin{pmatrix} -a_{1j} & \dots & -a_{j-1j} & x-a_{jj} & -a_{j+1j} & -a_{nj} \end{pmatrix},$$

which gives the desired relations matrix. Now, by exercise 19 of the previous section, we can diagonalize the relations matrix $xI - A^t$ by using row and column operations on to get a matrix of the form

$$\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \tag{$*$}$$

where D is a diagonal matrix with nonzero entries $a_1(x), a_2(x), \dots, a_k(x)$ and we have the divisibility conditions $a_1(x)|a_2(x)|\dots|a_k(x)$. Additionally, we have that

$$(V, T) \cong F[x]/(a_1(x)) \oplus F[x]/(a_2(x)) \oplus \dots \oplus F[x]/(a_k(x)) \oplus F[x]^{n-k}.$$

Since V is finite dimensional, then (V, T) is torsion, i.e., $k = n$. Therefore the matrix in $(*)$ is just D.

Let α_i be the leading coefficient for $a_i(x)$ for all $1 \le i \le n$. In the case that $a_i(x)$ is constant, let α_i be the constant of $a_i(x)$. To obtain monic polynomials, we multiply row i (or column i) of D by α_i^{-1} for all $1 \le i \le n$. Due to the divisibility conditions, any constant polynomials will appear in the beginning of the list $a_1(x), a_2(x), \dots, a_n(x)$. So, we obtain a matrix of the form desired in Theorem 21, which proves the theorem.

Now, the row and column operations used to obtain D are those which are described in (a) and (b) of the first step of the Invariant Factor Decomposition Algorithm. Also, multiplying the rows (or columns) of D by units as we did in the proof of Theorem 21 corresponds to part (c) of the algorithm.

Moreover, parts (a) and (b) of step 2 in the algorithm correspond directly to exercises 17 and 18 of section 12.1, which say that interchanging generators and multiplying one generator by a multiple of another does not alter the relations matrix. Per exercises 17 and 18 of the previous section, multiplying the ith row by a unit corresponds to changing the ith generator; so, part (c) of step 2 corresponds to this action. ☕

Exercise 12.3.26. Determine the Jordan canonical form for the $n \times n$ matrix over $\mathbb{F}_p$ whose entries are all equal to 1.

Proof. The minimal polynomial for such a matrix A will be $m_A(x) = x^2 - nx$ since no linear polynomial will send A to zero, and

$$\begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}^2 - n \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix} = \begin{pmatrix} n & \cdots & n \\ \vdots & \ddots & \vdots \\ n & \cdots & n \end{pmatrix} - \begin{pmatrix} n & \cdots & n \\ \vdots & \ddots & \vdots \\ n & \cdots & n \end{pmatrix} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}.$$

Notice that x and $x - n$ are the only divisors of $x^2 - xn = x(x-n)$. Because of the divisibility condition on the invariant factors, and since the product of the invariant factors must have degree equal to n, we get the following possibilities for the invariant factors of A:

$$\underbrace{x, x, \ldots, x}_{n-2 \text{ terms}}, x^2 - nx, \quad \text{or} \quad \underbrace{x-n, x-n, \ldots, x-n}_{n-2 \text{ terms}}, x^2 - nx.$$

However, the latter set of invariant factors would yield a characteristic polynomial so that the geometric multiplicity of the eigenvalue n, namely n, would exceed the algebraic multiplicity, namely $n - 1$. This contradicts a basic linear algebra fact that the geometric multiplicity is bounded above by the algebraic multiplicity. Hence the invariant factors of A are those in the first list.

Now to determine the Jordan Canonical form, we must consider two cases: when $p \nmid n$ (hence $n \neq 0_{\mathbb{F}_p}$), and when $p|n$ (hence $n = 0_{\mathbb{F}_p}$). If $p \nmid n$, the minimal polynomial has distinct roots 0 and n, the Jordan canonical form will be a diagonal matrix, with the roots of the invariant factors along the diagonal. If $p|n$, the minimal polynomial will not have distinct roots, and hence have Jordan blocks of size 1 for the invariant factors, except for the Jordan block corresponding to the minimal polynomial, which will be a block of size 2. That is, with respect to these cases, the possible Jordan canonical form for A are

$$\begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & 0 & \\ & & & n \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & 0 & 1 \\ & & 0 & 0 \end{pmatrix}.$$

☕

Exercise 13.2.8. Let F be a field of characteristic $\neq 2$. let D_1 and D_2 be elements of F, neither of which is a square in F. Prove that $F(\sqrt{D_1}, \sqrt{D_2})$ is of degree 4 over F if D_1D_2 is not a square in F and is of degree 2 over F otherwise. When $F(\sqrt{D_1}, \sqrt{D_2})$ is of degree 4 over F the field is called a *biquadratic extension of* F.

Proof. We have the tower of fields $F \subseteq F\left(\sqrt{D_1}\right) \subseteq F\left(\sqrt{D_1}, \sqrt{D_2}\right)$. Hence

$$\left[F\left(\sqrt{D_1}, \sqrt{D_2}\right) : F\right] = \left[F\left(\sqrt{D_1}, \sqrt{D_2}\right) : F\left(\sqrt{D_1}\right)\right]\left[F\left(\sqrt{D_1}\right) : F\right].$$

Notice that $\sqrt{D_1}$ is a root of $x^2 - D_1$ over F, and so $m_{\sqrt{D_1},F}(x)$ divides 2. We show that $x^2 - D_1$ is irreducible over F and hence is the minimal polynomial of $\sqrt{D_1}$ over F so that $\left[F\left(\sqrt{D_1}\right) : F\right] = 2$. Suppose $x^2 - D_1 = (x+\alpha)(x+\beta) = x^2 + (\alpha+\beta)x + \alpha\beta \in F[x]$. Comparing coefficients, we get that $\alpha = -\beta$ and $-D_1 = \alpha\beta = -\beta^2 \implies D_1 = \beta^2$, which is a contradiction since D_1 is not a square in F.

Next, we show that the degree of the minimal polynomial of $\sqrt{D_2}$ over $F\left(\sqrt{D_1}\right)$ (and hence $\left[F\left(\sqrt{D_1}, \sqrt{D_2}\right) : F\left(\sqrt{D_1}\right)\right]$) equals 2 if D_1D_2 is not a square in F, and equals 1 if D_1D_2 is a square in F. The desired result will follow.

Since $\sqrt{D_2}$ is a root of $x^2 - D_2$ over $F\left(\sqrt{D_1}\right)$, then degree of the minimal polynomial of $\sqrt{D_2}$ over $F\left(\sqrt{D_1}\right)$ is 2 or 1. Suppose

$$x^2 - D_1 = (x+\alpha)(x+\beta) = x^2 + (\alpha+\beta)x + \alpha\beta$$

for $\alpha, \beta \in F\left(\sqrt{D_1}\right)$. Comparing coefficients, we get $D_2 = \beta^2$. By Corollary 7 (§13.1, D& F), we have $F\left(\sqrt{D_1}\right) = \{a + b\sqrt{D_1} \mid a, b \in F\}$. So $\beta = a + b\sqrt{D_1}$ for some $a, b \in F$. Then

$$D_2 = \beta^2 = a^2 + 2ab\sqrt{D_1} + b^2D_1 \implies D_2 - a^2 - b^2D_1 = 2ab\sqrt{D_1}.$$

Comparing the coefficients of $\sqrt{D_1}$, we have that $2ab = 0$. Since F has characteristic $\neq 2$, this reduces to $ab = 0$, and so $a = 0$ or $b = 0$. If $b = 0$, then $D_2 = a^2$, a contradiction since D_2 is not a square in F. So $a = 0$ and hence $D_2 = b^2D_1$, which gives $D_1D_2 = (bD_1)^2$.

If D_1D_2 is not a square in F, this gives a contradiction, showing that $x^2 - D_2$ is irreducible over $F(\sqrt{D_1})$ and hence the minimal polynomial of $\sqrt{D_2}$ over $F(\sqrt{D_1})$. Otherwise, $x^2 - D_1$ is reducible and hence the minimal polynomial of $\sqrt{D_2}$ over $F(\sqrt{D_1})$ has degree 1.

Exercise 13.2.14. Prove that if $[F(\alpha) : F]$ is odd then $F(\alpha) = F(\alpha^2)$.

Proof. We prove the contrapositive statement. Suppose $F(\alpha) \neq F(\alpha^2)$. We have the tower of fields $F \subseteq F(\alpha^2) \subsetneq F(\alpha)$, and in particular $\alpha \notin F(\alpha^2)$. Notice that α is a root of the polynomial $x^2 - \alpha^2 \in F(\alpha^2)[x]$. Hence $\deg(m_{\alpha,F(\alpha^2)}(x))$ divides 2. Moreover, $\deg(m_{\alpha,F(\alpha^2)}(x)) \neq 1$ since the only possible linear polynomial in $F(\alpha^2)[x]$ of which α could be a root would be $x - \alpha$, but $\alpha \notin F(\alpha^2)$. So $[F(\alpha) : F]$ is even since

$$[F(\alpha) : F] = [F(\alpha^2) : F][F(\alpha) : F(\alpha^2)] = [F(\alpha^2) : F] \cdot 2.$$

Exercise 13.2.16. Let K/F be an algebraic extension and let R be a *ring* contained in K and containing F. Show that R is a subfield of K containing F.

Proof. Since $1_K = 1_F \in F \subseteq R$, then R has a 1. Let $\alpha \in R - \{0_R\}$. Consider the subring $F[\alpha] \subseteq R$. Since α is algebraic over F, then $F[\alpha] = F(\alpha)$. Since $F(\alpha)$ is a field, then $\alpha^{-1} \in F(\alpha) \subseteq R$. For any $\alpha, \beta \in R$, $\alpha\beta = \beta\alpha \in K$. Since R is closed, $\alpha\beta = \beta\alpha \in R$. Hence R is a commutative division ring with 1, i.e., a field.

Exercise 12.3.30. Let λ be an eigenvalue of the linear transformation T on the finite dimensional vector space V over the field F. Let $r_k = \dim_F(T-\lambda)^k V$ be the rank of the linear transformation $(T-\lambda)^k$ on V. For any $k \geq 1$, prove that $r_{k-1} - 2r_k + r_{k+1}$ is the number of Jordan blocks of T corresponding to λ of size k [use Exercise 12 in Section 1].

Proof. By Exercise 12 in Section 1, using $p(x) = x - \lambda$ (which is irreducible in the UFD $F[x]$ and hence prime), $R = F[x]$, and $M = (V, T)$ (which is a torsion $F[x]$–module since V is finite dimensional), we have that

$$(x-\lambda)^{k-1}(V,T)/(x-\lambda)^k(V,T) \cong (F[x]/(x-\lambda))^{n_k},$$

where n_k is the number of elementary divisors of (V, T) which are of the form $(x-\lambda)^\alpha$ for $\alpha \geq k$. This gives

$$\dim_F(T-\lambda)^{k-1}V - \dim_F(T-\lambda)^k V = \dim_F (F[x]/(T-\lambda))^{n_k} = n_k,$$

i.e., $r_{k-1} - r_k = n_k$. Define $\#J_{\lambda,k} :=$ the number of Jordan Blocks of T corresponding to λ of size k. Notice that $\#J_{\lambda,k}$ corresponds to the number of elementary divisors of (V, T) which are of the form $(x-\lambda)^k$. Hence $\#J_{\lambda,k} = n_k - n_{k+1}$, which gives

$$\#J_{\lambda,k} = n_k - n_{k+1} = r_{k-1} - r_k - (r_k - r_{k+1}) = r_{k-1} - 2r_k + r_{k+1}.$$

Exercise 13.2.9. Let F be a field of characteristic $\neq 2$. Let a, b be elements of the field F with b not a square in F. Prove that a necessary and sufficient condition for $\sqrt{a+\sqrt{b}} = \sqrt{m} + \sqrt{n}$ for some m and n in F is that $a^2 - b$ is a square in F. Use this to determine when the field $\mathbb{Q}(\sqrt{a+\sqrt{b}})$, $(a, b \in \mathbb{Q})$ is biquadratic over $\mathbb{Q}$.

Proof. ($\Rightarrow$) First suppose $\sqrt{a+\sqrt{b}} = \sqrt{m} + \sqrt{n}$ for some m and n in F. Squaring both sides we obtain $a + \sqrt{b} = m + 2\sqrt{mn} + n$. We claim that $a = m + n$ and $\sqrt{b} = 2\sqrt{mn}$. Once this is verified, it follows that $a^2 - b = (m-n)^2$ and hence is a square in F.

By assumption, $\sqrt{b} \notin F$ and hence $a + \sqrt{b} \notin F$. Note that $\sqrt{mn} \notin F$ since otherwise, $a + \sqrt{b} = m + 2\sqrt{mn} + n$ is an element of F, a contradiction.

Certainly $m + n \neq \sqrt{b}$. Suppose $\sqrt{b} = c + 2\sqrt{mn}$ for c equal to m, n or $m + n$. This implies $b = c^2 + 4c\sqrt{mn} + 4mn \notin F$, since $\sqrt{mn} \notin F$, which is a contradiction. The claim now follows.

($\Leftarrow$) Now suppose $a^2 - b := d^2$ is a square in F. Then $m := \frac{a+d}{2}$ and $n := \frac{a-d}{2} \in F$ give $m + 2\sqrt{mn} + n = a + \sqrt{b}$, i.e., $\sqrt{a+\sqrt{b}} = \sqrt{m} + \sqrt{n}$.

Now, we claim that $\mathbb{Q}(\sqrt{m} + \sqrt{n}) = \mathbb{Q}(\sqrt{m}, \sqrt{n})$. The inclusion " $\subseteq$ " is clear: If $f(x, y) = x + y \in \mathbb{Q}[x, y]$, then $f(\sqrt{m}, \sqrt{n}) = \sqrt{m} + \sqrt{n} \in \mathbb{Q}(\sqrt{m}, \sqrt{n})$. For the other inclusion, let

$$g(x) = x^3 - x(3m + n)(2n - 2m)^{-1} \quad \text{and} \quad h(x) = x^3 - x(m + 3n)(2m - 2n)^{-1}$$

be elements of $\mathbb{Q}[x]$. Then $\sqrt{m} = g(\sqrt{m} + \sqrt{n})$ and $\sqrt{n} = h(\sqrt{m} + \sqrt{n})$ are in $\mathbb{Q}(\sqrt{m} + \sqrt{n})$, which gives " $\supseteq$ ".

Now, $a^2 - b$ is a square in F if and only if $\mathbb{Q}(\sqrt{a+\sqrt{b}}) = \mathbb{Q}(\sqrt{m} + \sqrt{n}) = \mathbb{Q}(\sqrt{m}, \sqrt{n})$. By Exercise 13.2.8, $\mathbb{Q}(\sqrt{m}, \sqrt{n})$ is biquadratic if and only if m and n are not squares in $\mathbb{Q}$ and

$$mn = \frac{a^2 - d^2}{4} = \frac{b}{4}$$

is not a square in F. Since b is not a square in F, neither is $b/4$.

Exercise 13.2.15. A field F is said to be *formally real* if -1 is not expressible as a sum of squares in F. Let F be a formally real field, let $f(x) \in F[x]$ be an irreducible polynomial of odd degree and let α be a root of $f(x)$. Prove that $F(\alpha)$ is also formally real. [Pick α a counterexample of minimal degree. Show that $-1 + f(x)g(x) = (p_1(x))^2 + \cdots + (p_m(x))^2$ for some $p_i(x), g(x) \in F[x]$ where $g(x)$ has odd degree $< \deg f$. Show that some root β of g has odd degree over F and $F(\beta)$ is not formally real, violating the minimality of α.

Proof. Pick a root α of $f(x)$ of minimal degree so that $F(\alpha)$ is not formally real. Then $F(\alpha) \cong F[x]/(f(x))$. Since $F(\alpha)$ is not formally real, we can write $-1 \sum_{i=1}^{m} \lambda_i^2$ for some $\lambda_i \in F(\alpha)$. Let $\overline{p_i(x)}$ be the image of λ_i under the above isomorphism where $\overline{p_i(x)} = p_i(x) + (f(x))$, and the $p_i(x)$'s have strictly smaller degree than $f(x)$. Then

$$\overline{-1} = \sum_{i=1}^{m} \overline{p_i(x)}^2, \tag{$*$}$$

i.e., $-1 - \sum \overline{p_i(x)}^2 = f(x)g(x)$ for some $g(x)$ in $F[x]$. So,

$$-1 + f(x)g(x) = \sum_{i=1}^{m} p_i(x)^2.$$

Since the RHS is a polynomial of even degree then $f(x)g(x)$ has even degree; and since $f(x)$ has odd degree, so does $g(x)$. Moreover, the RHS has degree less than $2\deg f(x)$, which gives that the degree of $g(x)$ is less than that of $f(x)$.

Now since the degree of $g(x)$ is the sum of degrees of its irreducible factors, there is an irreducible factor $h(x)$ of $g(x)$ which has odd degree. If β is a root of $h(x)$, then $F(\beta) \cong F[x]/(h(x))$, and the equation in $(*)$ is still true in $F[x]/(h(x))$. Hence $F(\beta)$ is not formally real and β has degree equal to $\deg h(x)$, which is strictly less than $\deg f(x)$ which is the degree of α. This is a contradiction to the minimality of the degree of α. ☕

Exercise 13.2.17. Let $f(x)$ be an irreducible polynomial of degree n over a field F. Let $g(x)$ be any polynomial if $F[x]$. Prove that every irreducible factor of the composite polynomial $f(g(x))$ has degree divisible by n.

Proof. Let $p(x) \in F[x]$ be an irreducible factor of $f(g(x))$ of degree m. If α is a root of $p(x)$, then $f(g(\alpha)) = 0$, i.e., $g(\alpha)$ is a root of $f(x)$. Since f is irreducible, the degree of $g(\alpha)$ over F is n; that is, $[F(g(\alpha)) : F] = n$.

Now, since $F \subseteq F(g(\alpha))$, then $p(x) \in F(g(\alpha))[x]$. So α is algebraic over $F(g(\alpha))$, and hence $F(g(\alpha), \alpha) = F(\alpha)$ is a finite extension over F, say with index $[F(\alpha) : F] = \ell$. Then we have the tower of fields $F \subseteq F(g(\alpha)) \subseteq F(\alpha)$, which gives

$$m = [F(\alpha) : F] = [F(g(\alpha)) : F][F(\alpha) : F(g(\alpha))] = n \cdot \ell$$

and hence m divides n. ☕

Exercise 13.2.19. Let K be an extension of F of degree n.

(a) For any $\alpha \in K$ prove that α acting by left multiplication on K is an F-linear transformation of K.

(b) Prove that K is isomorphic to a subfield of the ring $n \times n$ matrices over F, so the ring of $n \times n$ matrices over F contains an isomorphic copy of *every* extension of F of degree $\leq n$.

Proof. For $\alpha \in K$, define $\phi_\alpha : K \to K$ by $\phi_\alpha(k) = \alpha k$. Then for $k, \ell \in K$ and $\lambda \in F$,

$$\phi_\alpha(k + \lambda\ell) = \alpha(k + \lambda\ell) = \alpha k + \alpha\lambda\ell = \alpha k + \lambda\alpha\ell = \phi_\alpha(k) + \lambda\phi_\alpha(\ell),$$

and hence ϕ_α is an F-linear transformation of K. For $\alpha, \beta \in K$ we have the following properties: $\phi_{\alpha+\beta} = \phi_\alpha + \phi_\beta$ and $\phi_{\alpha\beta} = \phi_\alpha \circ \phi_\beta$. Now, pick a basis $\mathcal{E}$ for K over F and define $\Phi : K \to \mathrm{Mat}_{n\times n}(F)$ by $\alpha \mapsto M_{\mathcal{E}}^{\mathcal{E}}(\phi_\alpha)$. Then for $\alpha, \beta \in K$,

$$M_{\mathcal{E}}^{\mathcal{E}}(\phi_{\alpha+\beta}) = M_{\mathcal{E}}^{\mathcal{E}}(\phi_\alpha + \phi_\beta) = M_{\mathcal{E}}^{\mathcal{E}}(\phi_\alpha) + M_{\mathcal{E}}^{\mathcal{E}}(\phi_\beta), \quad M_{\mathcal{E}}^{\mathcal{E}}(\phi_{\alpha\beta}) = M_{\mathcal{E}}^{\mathcal{E}}(\phi_\alpha \circ \phi_\beta) = M_{\mathcal{E}}^{\mathcal{E}}(\phi_\alpha) M_{\mathcal{E}}^{\mathcal{E}}(\phi_\beta).$$

So Φ is a field homomorphism and $\Phi(K)$ is a subfield of $\mathrm{Mat}_{n\times n}(F)$. Then

$$(a_{ij}) := M_{\mathcal{E}}^{\mathcal{E}}(\phi_\alpha) = M_{\mathcal{E}}^{\mathcal{E}}(\phi_\beta) =: (b_{ij}) \implies a_{ij} = b_{ij} \text{ for all } i, j \in \{1, \ldots, n\},$$

and since linear transformations are determined by their action on basis elements, $\phi_\alpha = \phi_\beta$, i.e., $\alpha = \beta$. So Φ is injective and hence $K \cong \Phi(K)$ as fields. ☕

Exercise 13.2.21. Let $K = \mathbb{Q}(\sqrt{D})$ for some squarefree integer D. Let $\alpha = a + b\sqrt{D}$ be an element of K. Use the basis $1, \sqrt{D}$ for K as a vector space over $\mathbb{Q}$ and show that the matrix of the linear transformation "multiplication by α " on K considered in the previous exercises has the matrix $\left(\begin{smallmatrix} a & bD \\ b & a \end{smallmatrix}\right)$. Prove directly that the map $a + b\sqrt{D} \mapsto \left(\begin{smallmatrix} a & bD \\ b & a \end{smallmatrix}\right)$ is an isomorphism of the field K with a subfield of the ring $\mathrm{Mat}_{2\times 2}(\mathbb{Q})$.

Proof. Let ϕ_α be left multiplication by $\alpha = a + b\sqrt{D}$. Let $\mathcal{E} = \{1, \sqrt{D}\}$ be a basis for K. Then $\phi_\alpha(1) = \alpha(1) = a + b\sqrt{D}$ and $\phi_\alpha(\sqrt{D}) = \alpha(\sqrt{D}) = bD + a\sqrt{D}$ gives

$$M_{\mathcal{E}}^{\mathcal{E}}(\phi_\alpha) = \begin{pmatrix} a & bD \\ b & a \end{pmatrix}.$$

The map$a + b\sqrt{D} \mapsto \left(\begin{smallmatrix} a & bD \\ b & a \end{smallmatrix}\right)$ is a homomorphism since

$$\begin{aligned}(a + b\sqrt{D}) + (a' + b'\sqrt{D}) = a + a' + (b + b')\sqrt{D} &\mapsto \begin{pmatrix} a + a' & bD + b'D \\ b + b' & a' + b' \end{pmatrix} \\ &= \begin{pmatrix} a & bD \\ b & a \end{pmatrix} + \begin{pmatrix} a' & b'D \\ b' & a' \end{pmatrix}\end{aligned}$$

and

$$\begin{aligned}(a + b\sqrt{D})(a' + b'\sqrt{D}) = (aa' + bb'D + (ab' + a'b)\sqrt{D}) &\mapsto \begin{pmatrix} aa' + bb'D & ab'D + a'bD \\ a'b + ab' & bb'D + aa' \end{pmatrix} \\ &= \begin{pmatrix} a & bD \\ b & a \end{pmatrix} \begin{pmatrix} a' & b'D \\ b' & a' \end{pmatrix}\end{aligned}$$

The image of this map is therefore a subfield of $\mathrm{Mat}_{2\times 2}(\mathbb{Q})$. If $\begin{pmatrix} a & bD \\ b & a \end{pmatrix} = \begin{pmatrix} a' & b'D \\ b' & a' \end{pmatrix}$ then $a = a'$, $b = b'$ and hence $a + b\sqrt{D} = a' + b'\sqrt{D}$, i.e., the map is injective. This shows K is isomorphic to a subfield of $\mathrm{Mat}_{2\times 2}(\mathbb{Q})$. ☕

Exercise 13.4.2. Determine the splitting field and its degree over $\mathbb{Q}$ for $x^4 + 2$.

Proof. The roots of $x^4 + 2$ are:

$$(-2)^{1/4}, (-2)^{1/4}\xi, (-2)^{1/4}\xi^2, (-2)^{1/4}\xi^3$$

where $\xi = e^{2\pi i/4} = e^{\pi i/2} = i$. Hence our list becomes

$$(-2)^{1/4}, (-2)^{1/4}i, -(-2)^{1/4}, (-2)^{1/4}(-i).$$

So our splitting field is $K = \mathbb{Q}((-2)^{1/4}, (-2)^{1/4}i, -(-2)^{1/4}, (-2)^{1/4}(-i))$. But

$$(-2)^{1/4}i, -(-2)^{1/4}, (-2)^{1/4}(-i)$$

are all elements of $\mathbb{Q}((-2)^{1/4}, i)$, and hence $K = \mathbb{Q}((-2)^{1/4}, i)$. Moreover, since $(-2)^{1/4}$ is a root of $x^4 + 2$, which is irreducible over $\mathbb{Q}$, and i is a root of $x^2 + 1$, which is irreducible over $\mathbb{Q}((-2)^{1/4})$, the degree of K over $\mathbb{Q}$ is

$$[K : \mathbb{Q}] = [K : \mathbb{Q}((-2)^{1/4}][\mathbb{Q}((-2)^{1/4}) : \mathbb{Q}] = 2 \cdot 4 = 8.$$

Exercise 13.4.3. Determine the splitting field and its degree over $\mathbb{Q}$ for $x^4 + x^2 + 1$.

Proof. Using the quadratic equation to find x^2, we find that

$$x^2 = \frac{-1 \pm \sqrt{-3}}{2} \implies x = \pm\frac{(-2 \pm 2\sqrt{3}i)^{1/2}}{2} = \pm\frac{(4e^{i\theta})^{1/2}}{2} = \pm e^{i\theta/2},$$

where $\theta = 2\pi/3$ and $4\pi/3$. So we have the splitting field

$$K = \mathbb{Q}(\pm e^{i\pi/3}, \pm e^{i2\pi/3}) = \mathbb{Q}(e^{i\pi/3}, e^{i2\pi/6}) = \mathbb{Q}(e^{i\pi/3}),$$

where the last equality follows from the fact that $(e^{i\pi/3})^2 = e^{i2\pi/3}$. Since $e^{i\pi/3}$ is a a root of $x^2 - x + 1$, which is irreducible over $\mathbb{Q}$, we have $[K : \mathbb{Q}] = 2$.

Exercise 13.2.18. Let k be a field and let $k(x)$ be the field of rational functions in x with coefficients from k. Let $t \in k(x)$ be the rational function $\frac{P(x)}{Q(x)}$ with relatively prime polynomials $P(x), Q(x) \in k[x]$ with $Q(x) \neq 0$. Then $k(x)$ is an extension of $k(t)$ and to compute its degree it is necessary to compute the minimal polynomial with coefficients in $k(t)$ satisfied by x.

(a) Show that the polynomial $P(X) - tQ(X)$ in the variable X and coefficients in $k(t)$ is irreducible over $k(t)$ and has x as a root.

Proof. Let $f(X) = P(X) - tQ(X)$. Then $f(x) = P(x) - \frac{P(x)}{Q(x)}(Q(x)) = 0$ and hence x is a root of $f(X)$. Consider the ring $k[t]$ consisting of all polynomials in the variable t with coefficients in k. The fraction field of $k[t]$ consists of all *rational* expressions of polynomials in the variable t with coefficients in k . But this is precisely the definition for $k(t)$, i.e., $k(t)$ is the fraction field of $k[t]$.

Now since $P(x)$ and $Q(x)$ are rational functions with coefficients in k, then $P(x) - tQ(x)$ is a rational function in the variable t with coefficients in k, i.e., $f(X) = P(X) - tQ(X) \in (k(t))[X]$. Since P and Q are relatively prime, then by Gauss' Lemma, $f(X)$ is irreducible in $(k[t])[X]$ if and only if it is irreducible in $(k(t))[X]$. But notice that $(k[t])[X] = (k[X])[t]$, and hence we need to show that $f(X)$ is irreducible in $(k[X])[t]$. But as a polynomial in t with coefficients in $k[X]$, $f(X)$ is linear, and hence irreducible. ☕

(b) Show that the degree of $P(X) - tQ(X)$ as a polynomial in X with coefficients in $k(t)$ is the maximum of the degrees of $P(x)$ and $Q(x)$.

Proof. Suppose $P(X) = a_nX^n + \cdots + a_0$ and $Q(X) = b_mX^m + \cdots + b_0$ for $a_i, b_j \in k$ where $a_n, b_m \neq 0$, then $P(X) - tQ(X)$ will be

$$a_nX^n + \cdots + a_0 - tb_mX^m - \cdots - tb_0$$

If $m \neq n$, then the result is clear. If $m = n$, then we have leading coefficient $a_n - tb_n$. We just need to make sure $a_n - tb_n \neq 0$ and the result follows. Suppose $a_n - tb_n = 0$, then $a_n = tb_n$. If $b_n = 0$, then $a_n = 0$, a contradiction. Otherwise, $t = a_n/b_m$, but this means $t \in k$, a contradiction. ☕

(c) Show that $[k(x) : k(t)] = \left[k(x) : k\left(\frac{P(x)}{Q(x)}\right)\right] = \max(\deg P(x), \deg(Q(x))$.

Proof. This follows immediately from parts (a) and (b), since the degree of the extension $k(x)$ over $k(t)$ is the degree of the minimal polynomial of x over $k(t)$. ☕

Exercise 13.2.20. Show that if the matrix of the linear transformation "multiplication by α" considered in the previous exercises is A then α is a root of χ_A. This gives an effective procedure for determining an equation of degree n satisfied by and element α in an extension of F of degree n. Use this procedure to obtain the monic polynomial of degree 3 satisfied by $\sqrt[3]{2}$ and by $1+\sqrt[3]{2}+\sqrt[3]{4}$.

Proof. If ϕ_α is the linear transformation "multiplication by α", notice that $\phi_\alpha - \alpha\mathbb{1}_K \equiv 0$. Hence for some basis $\mathcal{E}$ of K over F, $M_{\mathcal{E}}^{\mathcal{E}}(\phi_\alpha) = A$

$$\det(A-\alpha I) = \det\left(M_{\mathcal{E}}^{\mathcal{E}}(\phi_\alpha) - M_{\mathcal{E}}^{\mathcal{E}}(\alpha\mathbb{1}_K)\right) = \det\left(M_{\mathcal{E}}^{\mathcal{E}}(\phi_\alpha - \alpha\mathbb{1}_K)\right) = \det(0) = 0,$$

and hence $\chi_A(\alpha) = 0$.

The basis $\{1, \sqrt[3]{2}, \sqrt[3]{4}\}$ generates $\mathbb{Q}(\sqrt[3]{2})$. Multiplication by $\alpha = \sqrt[3]{2}$ is given by

$$1 \mapsto \sqrt[3]{2}, \quad \sqrt[3]{2} \mapsto \sqrt[3]{4}, \quad \sqrt[3]{4} \mapsto 2.$$

So, the matrix corresponding to this linear transformation is $\begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$, which has characteristic polynomial $x^3 - 2$.

The basis $\{1, \sqrt[3]{2}, \sqrt[3]{4}\}$ generates $\mathbb{Q}(1+\sqrt[3]{2}+\sqrt[3]{4})$. Multiplication by $\alpha = 1+\sqrt[3]{2}+\sqrt[3]{4}$ is given by

$$1 \mapsto 1+\sqrt[3]{2}+\sqrt[3]{4}, \quad \sqrt[3]{2} \mapsto 2+\sqrt[3]{2}+\sqrt[3]{4}, \quad \sqrt[3]{4} \mapsto 2+2\sqrt[3]{2}+\sqrt[3]{4}$$

So, the matrix corresponding to this linear transformation is $\begin{pmatrix} 1 & 2 & 2 \\ 1 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}$, which has characteristic polynomial $x^3 - 3x^2 - 3x - 1$.

☕

Exercise 13.4.6. Let K_1 and K_2 be finite extensions of F contained in the field K, and assume both are splitting fields over F.

(a) Prove that their composite K_1K_2 is a splitting field over F.

Proof. The assumption that K_1 and K_2 are splitting fields means that each of them is a splitting field of a family of polynomials in $F[x]$ of degrees ≥ 1. Since we assume K_1 and K_2 to be finite over F, we can take these families to be finite. Hence, by taking the product of the polynomials in each of these families we get that K_1 (respectively K_2) is the splitting field of a single polynomial $f_1(x)$ (respectively $f_2(x)$) in $F[x]$ of degree ≥ 1.

Now, since K_1K_2 contains both K_1 and K_2, it contains all the roots of $f_1(x)$ and $f_2(x)$, and so there exists a splitting field $L \subseteq K_1K_2$ of the family of polynomials $\{f_1(x), f_2(x)\}$. Since L is the contains the roots of $f_1(x)$ (respectively $f_2(x)$), then $K_1 \subseteq L$ (respectively $K_1 \subseteq L$). Since K_1K_2 is the *smallest* subfield of K containing both K_1 and K_2 then $K_1K_2 \subseteq L$, and hence $K_1K_2 = L$. ☕

(b) Prove that $K_1 \cap K_2$ is a splitting field over F.

Proof. Let $p(x) \in F[x]$ be an irreducible polynomial which has a root in $K_1 \cap K_2$. Then $p(x)$ splits into linear factors in both $K_1[x]$ and $K_2[x]$, and hence K_1 and K_2 are splitting fields of $p(x)$. Now since K_1 and K_2 are contained in K, then the factorizations of $p(x)$ in $K_1[x]$ and $K_2[x]$ are contained in $K[x]$. But since $K[x]$ is a UFD, these factorizations of $p(x)$ differ by at most a unit in K. Hence the linear factors of $p(x)$ in $K_1[x]$ and $K_2[x]$ are the same, i.e., $p(x)$ splits into linear factors in $(K_1 \cap K_2)[x]$, and hence $K_1 \cap K_2$ is a splitting field over F. ☕

Exercise 1. Let a be a real number such that $a^4 = 7$. Let $\mathbb{Q}$ be the field of rational numbers, and let i be a square root of -1. Show that $\mathbb{Q}(ia^2)$ is normal over $\mathbb{Q}$. Show that $\mathbb{Q}(a+ia)$ is normal over $\mathbb{Q}(ia^2)$. Show that $\mathbb{Q}(a+ia)$ is not normal over $\mathbb{Q}$.

Proof. We show that $\mathbb{Q}(ia^2)$ is a splitting field over $\mathbb{Q}$ and hence normal over $\mathbb{Q}$. Consider the polynomial $x^2+7 \in \mathbb{Q}[x]$. The roots of this polynomial are $\pm ia^2$ and hence has splitting field $\mathbb{Q}(\pm ia^2) = \mathbb{Q}(ia^2)$.

Similarly, we show that $\mathbb{Q}(a+ia)$ is a splitting field over $\mathbb{Q}(ia^2)$. The polynomial $x^2 - 2ia^2 = (x-(a+ia))(x+(a+ia)) \in \mathbb{Q}(ia^2)[x]$ has roots $\pm(a+ia)$, and hence has splitting field $\mathbb{Q}(ia^2)(\pm(a+ia)) = \mathbb{Q}(ia^2, a+ia)$. Now since $ia^2 = (1/2)(a+ai)^2$, then $\mathbb{Q}(ia^2, a+ia) = \mathbb{Q}(a+ia)$.

Notice that $f(x) := x^4 + 28$ is irreducible over $\mathbb{Q}$ by Eisenstein's Criterion. Moreover $\pm a \pm ia$ are the roots of this polynomial. However, $a - ia \notin \mathbb{Q}(a+ia)$. To see this, suppose otherwise. Since $f(x)$ is the minimal polynomial then $a - ia$ will have the form

$$\begin{aligned} a - ia &= \alpha + \beta(a+ia) + \gamma(a+ia)^2 + \delta(a+ia)^3 \\ &= \alpha + \beta(a+ia) + \gamma(2ia^2) + \delta(2ia^3 - 2a^3). \end{aligned}$$

Comparing the coefficients of i, we get

$$a = \beta a + 2\gamma a^2 + 2\delta a^3.$$

However, this would give that a is a root of the polynomial $g(x) := 2\delta x^2 + 2\gamma x + (\beta - 1) \in \mathbb{Q}[x]$, which means $g(x)$ is divisible by $f(x)$. But this cannot happen since $\deg g(x) < \deg f(x)$. So $\mathbb{Q}(a+ia)$ contains the root $a+ia$ of the irreducible polynomial $f(x) \in \mathbb{Q}[x]$ but not all of its roots. Hence $\mathbb{Q}(a+ia)$ is not a normal extension of $\mathbb{Q}$. ☕

Exercise 13.5.3. Prove that d divides n if and only if $x^d - 1$ divides $x^n - 1$. [Note that if $n = qd + r$ then $x^n - 1 = (x^{qd+r} - x^r) + (x^r - 1)$.]

Proof. We have the following formula (which I found online because I could not, for the life of me, figure out this problem):

$$x^{dq} - 1 = (x^d - 1)\left(\sum_{k=1}^{q-1} \left(x^d\right)^k\right) \tag{$\mathfrak{D}$}$$

($\Rightarrow$) Supposing d divides n, we have $n = dq$ for some q. So $x^n - 1 = x^{dq} - 1$ and the desired result follows from ($\mathfrak{D}$).

($\Leftarrow$) Suppose now that $x^d - 1$ divides $x^n - 1$. By the division algorithm, there exists q, r with $0 \le r < d$ so that $n = qd + r$. Then using the hint and applying ($\mathfrak{D}$) to , we have

$$\begin{aligned} x^n - 1 = (x^{dq}x^r - x^r) + (x^r - 1) &= (x^r)(x^{dq} - 1) + (x^r - 1) \\ &= (x^r)(x^d - 1)\left(\sum_{k=1}^{q-1} \left(x^d\right)^k\right) + (x^r - 1). \end{aligned}$$

Now $x^d - 1$ divides the first term above. However, if $0 < r < d$, then $x^d - 1$ does not divide the second term since the degree of the degree of $x^d - 1$ is greater than $x^r - 1$. Hence $r = 0$ and so $n = qd$, i.e., d divides n. ☕

Exercise 13.5.4. Let $a > 1$ be an integer. Prove for any positive integers n, d that d divides n if and only if $a^d - 1$ divides $a^n - 1$. Conclude in particular that $\mathbb{F}_{p^d} \subseteq \mathbb{F}_{p^n}$ if and only if d divides n.

Proof. ($\Rightarrow$) This direction follows immediately from the previous problem.

($\Leftarrow$) This follows *almost* immediately from the previous problem, except now in the last step we argue that if $0 < r < d$, then $a^d - 1$ does not divide the second term since $a^d - 1 > a^r - 1$. Hence $r = 0$ and so $n = qd$, i.e., d divides n.

Note that $|(\mathbb{F}_{p^n})^\times| = p^n - 1$ and so if $\alpha \in \mathbb{F}_{p^n}$, then $\alpha^{p^n-1} = 1$. Now, if d divides n then $p^d - 1$ divides $p^n - 1$, say $p^n - 1 = (p^d - 1)q$. So if $\beta \in \mathbb{F}_{p^d}$ then $\beta^{p^n-1} = (\beta^{p^d-1})^q = 1$, and so $\beta \in \mathbb{F}_{p^n}$.

Conversely, suppose $\mathbb{F}_{p^d} \subseteq \mathbb{F}_{p^n}$. Now, $x^{p^k} - x$ is separable since its derivative is $p^k x^{p^k-1} - 1 = -1$. So, since every root of $x^{p^d} - x$ is also root of $x^{p^n} - x$, then $x^{p^d} - x$ divides $x^{p^n} - x$. Then

$$x^{p^d} - x \Big| x^{p^n} - x \implies x^{p^d-1} - 1 \Big| x^{p^n-1} - 1 \overset{\text{Exer.5.3}}{\implies} p^d - 1 \Big| p^n - 1 \overset{\text{Exer.5.4}}{\implies} d|n.$$

Exercise 13.5.5. For any prime p and any nonzero $a \in \mathbb{F}_p$ prove that $x^p - x + a$ is irreducible and separable over $\mathbb{F}_p$. [For the irreducibility: One approach — prove first that if α is a root then $\alpha + 1$ is also a root. Another approach — suppose it's reducible and compute derivatives.]

Proof. If α is a root of $f(x) := x^p - x + a$, then $(\alpha+1)^p - (\alpha+1) + a = \alpha^p + 1 - \alpha - 1 + a = 0$, and so $\alpha + 1$ is also root of $f(x)$. Continuing inductively

$$\alpha + \underbrace{1 + 1 + \cdots + 1}_{k \text{ summands}}$$

is a root of $f(x)$ for all $0 \leq k \leq p-1$. Hence $\{\alpha + \beta\}_{\beta \in \mathbb{F}_p}$ are the p distinct roots of $f(x)$. So $\mathbb{F}_p(\alpha)$ is a splitting field for $f(x)$, and in particular $f(x)$ is separable since it has p distinct roots in $\mathbb{F}_p(\alpha)$. Now if $\alpha \in \mathbb{F}_p$ then $0 = \alpha + (-\alpha)$ is a root of $f(x)$, which is a contradiction since $a \neq 0_{\mathbb{F}_p}$. So $\alpha \notin \mathbb{F}_p$ and so none of the roots of $f(x)$ lie in $\mathbb{F}_p$.

Before moving on to show that $f(x)$ is irreducible, we prove the following by induction: For $n \geq 2$, the product $(x - a_1)(x - a_2) \dots (x - a_n)$ will have $-\sum_{i=1}^{n} a_i$ as the coefficient on the x^{n-1} term. We have

$$(x - a_1)(x - a_2) = x^2 - (a_1 + a_2)x + a_1 a_2.$$

Now suppose for induction that $(x - a_1)(x - a_2) \dots (x - a_{n-1})$ has coefficient $-\sum_{i=1}^{n-1} a_i$ on the x^{n-2}. Then

$$\begin{aligned}
(x - a_1) \dots (x - a_{n-1})(x - a_n) &= \left(x^{n-1} - \left(\sum_{i=1}^{n-1} a_i \right) x^{n-2} + \dots \right) (x - a_n) \\
&= \left(x^n - \left(\sum_{i=1}^{n} a_i \right) x^{n-1} + \dots \right) \\
&\qquad + \left(-a_n x^{n-1} - \left(a_n + \sum_{i=1}^{n-1} a_i \right) x^{n-2} + \dots \right) \\
&= \left(x^n - \left(\sum_{i=1}^{n} a_i \right) x^n + \dots \right).
\end{aligned}$$

This completes the induction. Now, any proper factor of $f(x)$ will be of the form

$$\prod_{\substack{\beta_i \in S, \\ S \subsetneq \mathbb{F}_p}} (x - \alpha + \beta_i)$$

in $\mathbb{F}_p(\alpha)[x]$. Suppose $|S| = k$. Then

$$\prod_{\substack{\beta_i \in S, \\ S \subsetneq \mathbb{F}_p}} (x - \alpha + \beta_i) = x^k - \left(\sum_{i=1}^{k} (\alpha + \beta_i) \right) x^{k-1} + \dots = x^k - \left(\sum_{i=1}^{k} \beta_i + k\alpha \right) x^{k-1} + \dots$$

However, $k\alpha \notin \mathbb{F}_p$, and so no proper factor of $f(x)$ is in $\mathbb{F}_p[x]$. Hence $f(x)$ is irreducible over $\mathbb{F}_p$. ☕

Exercise 1. Let F be a field, let $f(x) \in F[x]$ be of degree $n \geq 1$, and let K be a splitting field of $f(x)$ over F. Prove that $[K : F]$ divides $n!$. *Hint:* Use induction on n and distinguish between the cases when $f(x)$ is irreducible, resp. reducible, in $F[x]$. It may be helpful to remember that if $n_1 + n_2 = n$ for positive integers n_1 and n_2, then the product $n_1! \cdot n_2!$ divides $n!$.

Proof. When $n = 1$, $f(x)$ is irreducible and $K = F$, and so certainly $1 = [K : F]$ divides $1!$.

Suppose for induction that for $g(x) \in F[x]$ of degree less than n, if L is a splitting field of $g(x)$ over F, then $[L : F]$ divides $(\deg g(x))!$. Let $f(x) \in F[x]$ be of degree n, and let K be its splitting field over F. First suppose that $f(x)$ is reducible. Let $p(x)$ an irreducible factor of $f(x)$ and let L be the splitting field of $p(x)$ over F. Let K be the splitting field of $f(x)/p(x)$ over L. Then K is a splitting field for $f(x)$ over F. By the induction hypothesis,

$$[L : F] \text{ divides } (\deg p(x))! \quad \text{and} \quad [K : L] \text{ divides } (\deg f(x)/p(x))!.$$

So,

$$[K : F] = [L : F][K : L] \text{ divides } (\deg p(x))!(\deg f(x)/p(x))!. \tag{☽}$$

By the hint, $\deg f(x)/p(x) + \deg p(x) = \deg f(x)$ implies that $(\deg f(x)/p(x))!(\deg p(x))!$ divides $(\deg f(x))!$. So by (☽), $[K : F]$ divides $(\deg f(x))!$.

Now suppose $f(x)$ is irreducible, and let α be a root of $f(x)$ in K. Let $L = F(\alpha)$. Then $[F(\alpha) : F] = \deg f(x) = n$. By the induction hypothesis, $[K : L]$ divides $\deg(f(x)/x - \alpha)! = (n-1)!$. Hence $[K : F] = [K : F(\alpha)][F(\alpha) : F]$ divides $(n-1)! \cdot n = n! = (\deg f(x))!$. ☕

Exercise 13.5.7. Suppose K is a field of characteristic p which is not a perfect field: $K \neq K^p$. Prove there exists irreducible inseparable polynomials over K. Conclude that there exists inseparable finite extensions of K.

Proof. Since $K \neq K^p$ and $K^p \subsetneq K$, let $a \in K \setminus K^p$. Then define $f(x) := x^p - a$, and let α be a root of $f(x)$. This implies $\alpha^p = a$. Hence $f(x) = x^p - \alpha^p = (x - \alpha)^p$ and so $f(x)$ is inseparable since it has a repeated root. Now, let $g(x)$ be an irreducible factor of $f(x)$. Then $g(x)$ will have the form $g(x) = (x - \alpha)^k$ for some $1 \leq k \leq p$. If $k = 1$, then $\alpha \in K$, a contradiction. So, $1 < k \leq p$. Now since

$$g(x) = x^k - k\alpha x^{k-1} + \cdots + (-\alpha)^k,$$

we have that $-k\alpha \in K$. If $k \neq p$, then $k = 1+1+\cdots+1 \in K$, and so $\alpha \in K$, a contradiction. Hence $k = p$ and so $f(x) = g(x)$. ☕

Exercise 13.5.11. Suppose $K[x]$ is a polynomial ring over the field K and F is a subfield of K. If F is a perfect and $f(x) \in F[x]$ has no repeated irreducible factors in $F[x]$, prove that $f(x)$ has no repeated irreducible factors in $K[x]$.

Proof. Without loss of generality, suppose $f(x)$ is monic. Then let $f(x) = f_1(x)f_2(x)\cdots f_n(x)$ for distinct, monic, and irreducible polynomials $\{f_i\}$. We show that $f(x)$ has no repeated roots, and hence cannot have a repeated irreducible factor in $K[x]$. (If $f(x)$ has a repeated irreducible factor in $K[x]$, then $f(x)$ has a repeated root.)

Since F is perfect, each f_i is separable and so each f_i has no repeated roots. Therefore, $f(x)$ has a repeated root if and only if two of its irreduble factors share a root. Suppose α is a root of f_i and f_j for $i \neq j$. Then f_i and f_j are minimal polynomials for α, and by uniqueness of the minimal polynomial, $f_i = f_j$, a contradiction. ☕

Exercise 14.1.10. Let K be an extension of the field F. Let $\varphi : K \to K'$ be an isomorphism of K with a field K' which maps F to the subfield F' of K'. Prove that the map $\sigma \mapsto \varphi\sigma\varphi^{-1}$ defines a group isomorphism $\mathrm{Aut}(K/F) \xrightarrow{\sim} \mathrm{Aut}(K'/F')$.

Proof. Note that for $\sigma \in \mathrm{Aut}(K/F)$, $\varphi(\sigma(\varphi^{-1}(F'))) = \varphi(\sigma(F)) = \varphi(F) = F'$ and so $\varphi\sigma\varphi \in\in \mathrm{Aut}(K'/F')$. Define a map $\Phi : \mathrm{Aut}(K/F) \to \mathrm{Aut}(K'/F')$ by $\sigma \mapsto \varphi\sigma\varphi^{-1}$. If $\sigma, \gamma \in \mathrm{Aut}(K/F)$, then

$$\Phi(\sigma\gamma) = \varphi\sigma\gamma\varphi^{-1} = (\varphi\sigma\varphi^{-1})(\varphi\gamma\varphi^{-1}) = \Phi(\sigma)\Phi(\gamma),$$

and so Φ is a group homomorphism. If $\varphi\sigma\varphi^{-1} = \varphi\gamma\varphi^{-1}$ then $\sigma = \gamma$ and so Φ is injective. Finally, if $\tau \in \mathrm{Aut}(K'/F')$ then $\varphi^{-1}\tau\varphi \in \mathrm{Aut}(K/F)$ since $\varphi^{-1}\tau\varphi : K \to K$ is an isomorphism (as the composition of isomorphisms) and $\varphi^{-1}(\tau(\varphi(F))) = \varphi^{-1}(\tau(F')) = \varphi^{-1}(F') = F$. Then

$$\Phi(\varphi^{-1}\tau\varphi) = \varphi\varphi^{-1}\tau\varphi\varphi^{-1} = \tau,$$

and so Φ is surjective. ☕

Exercise 14.2.14. Show that $\mathbb{Q}(\sqrt{2+\sqrt{2}})$ is a cyclic quartic field i.e. is a Galois extension of degree 4 with cyclic Galois group.

Proof. Let $K = \mathbb{Q}(\sqrt{2+\sqrt{2}})$. Notice that $f(x) = x^4 - 4x + 2$ is an irreducible polynomial over $\mathbb{Q}$ by Eisenstein's Criterion. The roots of $f(x)$ are $\pm\sqrt{2+\sqrt{2}}$ and $\pm\sqrt{2-\sqrt{2}}$. Hence $f(x)$ is separable since it has no repeated roots and $\mathbb{Q}\left(\sqrt{2+\sqrt{2}}, \sqrt{2-\sqrt{2}}\right)$ is a splitting field of $f(x)$ over $\mathbb{Q}$. Now, $\sqrt{2} = \left(\sqrt{2+\sqrt{2}}\right)^2 - 2 \in K$, which gives

$$\sqrt{2-\sqrt{2}} = \frac{\sqrt{2-\sqrt{2}}\sqrt{2+\sqrt{2}}}{\sqrt{2+\sqrt{2}}} = \frac{\sqrt{2}}{\sqrt{2+\sqrt{2}}} \in K.$$

So $K = \mathbb{Q}\left(\sqrt{2+\sqrt{2}}, \sqrt{2-\sqrt{2}}\right)$, and so K is a splitting field for the separable polynomial $f(x)$ over $\mathbb{Q}$, and hence $K/\mathbb{Q}$ is Galois with $|\,\mathrm{Gal}(K/\mathbb{Q})| = [K : \mathbb{Q}] = 4$. So $G = \mathrm{Gal}(K/\mathbb{Q})$ is a subgroup of S_4 of order 4. There exists $\sigma \in G$ such that $\sigma\left(\sqrt{2-\sqrt{2}}\right) = -\sqrt{2+\sqrt{2}}$. Now,

$$\sigma(\sqrt{2}) = \sigma\left(-\left(\sqrt{2-\sqrt{2}}\right)^2 + 2\right) = -\sigma\left(\sqrt{2-\sqrt{2}}\right)^2 + \sigma(2) = -\left(2+\sqrt{2}\right) + 2 = -\sqrt{2}.$$

This gives

$$\sigma\left(-\sqrt{2+\sqrt{2}}\right) = \sigma\left(-\frac{\sqrt{2+\sqrt{2}}\sqrt{2-\sqrt{2}}}{\sqrt{2-\sqrt{2}}}\right) = \frac{-\sigma\left(\sqrt{2}\right)}{\sigma\left(\sqrt{2-\sqrt{2}}\right)} = -\sqrt{2-\sqrt{2}}.$$

Finally, we have

$$\sigma\left(-\sqrt{2-\sqrt{2}}\right) = -\sigma\left(\sqrt{2-\sqrt{2}}\right) = \sqrt{2+\sqrt{2}}.$$

Therefore,

$$\sigma = \left(\sqrt{2-\sqrt{2}}, -\sqrt{2+\sqrt{2}}, -\sqrt{2-\sqrt{2}}, \sqrt{2+\sqrt{2}}\right),$$

and so $|\,\langle\sigma\rangle\,| = 4$, which means $G = \langle\sigma\rangle$. ☕

Exercise 14.2.15. (Biquadratic Extensions) Let F be a field of characteristic $\neq 2$.

(a) If $K = F(\sqrt{D_1}, \sqrt{D_2})$ where $D_1, D_2 \in F$ have the property that none of D_1, D_2, or D_1D_2 is a square in F, prove that K/F is a Galois extension with $\mathrm{Gal}(K/F)$ isomorphic to the Klein 4-group.

Proof. Let

$$f(x) = \left(x - \sqrt{D_1}\right)\left(x + \sqrt{D_1}\right)\left(x - \sqrt{D_2}\right)\left(x + \sqrt{D_2}\right) \in F[x].$$

Then $f(x)$ is separable since it has no multiple roots and moreover, $f(x)$ has splitting field K. Hence K/F is Galois since K is the splitting field of a separable polynomial in $F[x]$. By Exercise 13.2.8, $[K : F] = 4$ and so $G = \mathrm{Gal}(K/F)$ has order 4.

Notice that $f(x) = (x^2 - D_1)(x^2 - D_2)$, and both factors of $f(x)$ are irreducible since they do not contain a root in $F[x]$. Since elements of G permute the roots of the irreducible factors of $f(x)$, we get the following (nonidentity) elements of G:

$$\sigma = \left(\sqrt{D_1}, -\sqrt{D_1}\right), \tau = \left(\sqrt{D_2}, -\sqrt{D_2}\right), \text{ and } \sigma\tau = \left(\sqrt{D_1}, -\sqrt{D_1}\right)\left(\sqrt{D_2}, -\sqrt{D_2}\right).$$

Then $|\sigma| = |\tau| = |\sigma\tau| = 2$, and hence $G \cong V_4$.

(b) Conversely, suppose K/F is a Galois extension with $\mathrm{Gal}(K/F)$ isomorphic to the Klein 4-group. Prove that $K = F(\sqrt{D_1}, \sqrt{D_2})$ where $D_1, D_2 \in F$ have the property that none of D_1, D_2, or D_1D_2 is a square in F.

Proof. Suppose $G = \mathrm{Gal}(K/F) = \{1, \sigma, \tau, \sigma\tau\} \cong V_4$. By the Fundamental Theorem of Galois Theory, we have corresponding lattices:

$$\begin{tikzcd} & 1 \arrow[dl, dash, "2"'] \arrow[d, dash, "2"] \arrow[dr, dash, "2"] & \\ \langle\sigma\rangle \arrow[dr, dash, "2"'] & \langle\tau\rangle \arrow[d, dash, "2"] & \langle\sigma\tau\rangle \arrow[dl, dash, "2"] \\ & G & \end{tikzcd} \quad \widehat{=} \quad \begin{tikzcd} & K \arrow[dl, dash, "2"'] \arrow[d, dash, "2"] \arrow[dr, dash, "2"] & \\ E_1 \arrow[dr, dash, "2"'] & E_2 \arrow[d, dash, "2"] & E_3 \arrow[dl, dash, "2"] \\ & F & \end{tikzcd}$$

where E_1, E_2, and E_3 are intermediate fields of K/F.

Since $\langle\sigma\rangle$ and $\langle\tau\rangle$ are normal subgroups of G, then E_1/F and E_2/F are Galois (again by the Fundamental Theorem). In particular, E_1/F and E_2/F are finite separable and hence $E_1 = F(\alpha)$ and $E_2 = F(\beta)$ for some α, β algebraic over F by the Primitive Element Theorem. By the example on page 522 (on quadratic extensions of fields of characteristic $\neq 2$), we must have $\alpha = \sqrt{D_1}$ and $\beta = \sqrt{D_2}$ for $D_1, D_2 \in F$ which are not squares in F. If D_1D_2 was a square in F then $\sqrt{D_1D_2} \in F$ and so

$$\sqrt{D_2} = \frac{\sqrt{D_1}\sqrt{D_2}}{\sqrt{D_1}} \in F\left(\sqrt{D_1}\right),$$

and similarly $\sqrt{D_1} \in F\left(\sqrt{D_2}\right)$. So $F\left(\sqrt{D_1}\right) = F\left(\sqrt{D_2}\right)$, a contradiction since the fixed fields of $\langle\sigma\rangle$ and $\langle\tau\rangle$ are unique.

Now, $E_1E_2 = F(\sqrt{D_1}, \sqrt{D_2}) \subseteq K$ and $[E_1E_2 : F] = 4$ by Exercise 13.2.8. Since $[K : F] = 4$, then $F(\sqrt{D_1}, \sqrt{D_2}) = K$.

Exercise 14.2.4. Let p be a prime. Determine the elements of the Galois group of $x^p - 2$.

Proof. Let $\sqrt[p]{2} \in \mathbb{R}$. The roots of $f(x)$ are $\sqrt[p]{2}$ and $\xi^i \sqrt[p]{2}$ for $1 \leq i \leq p-1$ where $\xi = e^{(2\pi i)/p}$. So $K = \mathbb{Q}(\sqrt[p]{2}, \xi)$ is a splitting field for $f(x)$ over $\mathbb{Q}$. We have the diagram

$$\begin{tikzcd}
 & K = \mathbb{Q}(\sqrt[p]{2}), \xi) \arrow[dl, dash] \arrow[dr, dash] & \\
\mathbb{Q}(\sqrt[p]{2}) \arrow[dr, dash, "p"'] & & \mathbb{Q}(\xi) \arrow[dl, dash, "p-1"] \\
 & \mathbb{Q} &
\end{tikzcd}$$

Since p and $p-1$ are coprime, $[K : F] = p(p-1)$, and so $G = \mathrm{Gal}(K/F) = p(p-1)$. Let $\sigma \in G$ and suppose that $\sigma(\sqrt[p]{2}) = \xi^a \sqrt[p]{2}$ and $\sigma(\xi \sqrt[p]{2}) = \xi^b \sqrt[p]{2}$. Then

$$\sigma(\xi) = \sigma\left(\frac{\xi \sqrt[p]{2}}{\sqrt[p]{2}}\right) = \xi^{b-a} = \xi^i \qquad \text{for some } i \in \{1, \ldots, p-1\}.$$

Hence the elements of G are given by

$$\sigma_{i,j} = \begin{cases} \xi \longmapsto \xi^i, & 1 \leq i \leq p-1, \\ \sqrt[p]{2} \longmapsto \xi^j \sqrt[p]{2}, & 0 \leq j \leq p. \end{cases}$$

☕

Exercise 14.2.5. Prove that the Galois group of $x^p - 2$ for p a prime is isomorphic to the group of matrices $\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix}$ where $a, b \in \mathbb{F}_p$, $a \neq 0$.

Proof. Using the notation in the previous exercise, define a map $\sigma_{i,j} \mapsto \begin{pmatrix} i & j \\ 0 & 1 \end{pmatrix}$. Let $\sigma_{i,j}, \sigma_{k,\ell} \in G$. Then

$$\sigma_{i,j}\sigma_{k,\ell} = \begin{cases} \xi \longmapsto \xi^{ik} \\ \sqrt[p]{2} \longmapsto \xi^{i\ell+j} \sqrt[p]{2}, \end{cases}$$

and $\sigma_{i,j}\sigma_{k,\ell} \mapsto \begin{pmatrix} ik & i\ell + j \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} i & j \\ 0 & 1 \end{pmatrix}\begin{pmatrix} k & \ell \\ 0 & 1 \end{pmatrix}$. So the map is a homomorphism, and if $\sigma_{i,j}$ maps to the identity matrix, then $i = 1$ and $j = 0$, i.e., $\sigma_{i,j}$ is the identity permutation. Finally, the map is surjective by definition of $\sigma_{i,j}$. ☕

Exercise 14.2.12. Determine the Galois group of the splitting field over $\mathbb{Q}$ of $x^4 - 14x^2 + 9$.

Proof. Using the quadratic formula $x^2 = 7 + 2\sqrt{10}$ and so the roots of $f(x) = x^4 - 14x^2 + 9$ are $\pm\sqrt{7 \pm 2\sqrt{10}}$. Let $\alpha = \sqrt{7 + 2\sqrt{10}}$ and $\beta = \sqrt{7 - 2\sqrt{10}}$. Since

$$f(x) = (x - \alpha)(x + \alpha)(x - \beta)(x + \beta) = \left(x^2 - \alpha^2\right)\left(x^2 - \beta^2\right),$$

then $f(x)$ is irreducible over $\mathbb{Q}$ because none of its quadratic factors lie in $\mathbb{Q}[x]$. (Since $\mathbb{Q}[x]$ is a UFD, this factorization into quadratic factors is unique). Now, $f(x)$ is separable since it has no repeated roots, and so $K = \mathbb{Q}(\alpha, \beta)$ is Galois over $\mathbb{Q}$. Notice that

$$\sqrt{7 - 2\sqrt{10}} = \frac{\sqrt{7 - 2\sqrt{10}}\sqrt{7 + 2\sqrt{10}}}{\sqrt{7 + 2\sqrt{10}}} = \frac{3}{\sqrt{7 + 2\sqrt{10}}} \in \mathbb{Q}(\alpha),$$

and so $K = \mathbb{Q}(\alpha)$. Now $G = \operatorname{Gal}(K/F)$ has order 4. If $\sigma \in G$ and $\sigma(\alpha) = \beta$, then $\sigma(\beta) = \sigma(3/\alpha) = 3/\beta = \alpha$. If $\tau \in G$ and $\tau(\alpha) = -\beta$, then $\sigma(-\beta) = \sigma(-3/\alpha) = -3/-\beta = \alpha$. Then $\sigma\tau(\alpha) = \sigma(-\beta) = -\alpha$ and $\sigma\tau(-\alpha) = \sigma(\beta) = \alpha$. Hence the (nonidentity) elements of G are

$$\sigma = (\alpha, \beta), \tau = (\alpha, -\beta), \text{ and } \sigma\tau = (\alpha, -\alpha).$$

Then $|\sigma| = |\tau| = |\sigma\tau| = 2$, and hence $G \cong V_4$. ☕

Exercise 14.2.16.

(a) Prove that $x^4 - 2x^2 - 2$ is irreducible over $\mathbb{Q}$.
Solution: Eisenstien.

(b) Show the roots of this quartic are

$$\alpha_1 = \sqrt{1 + \sqrt{3}} \qquad \alpha_3 = -\sqrt{1 + \sqrt{3}}$$
$$\alpha_2 = \sqrt{1 - \sqrt{3}} \qquad \alpha_4 = -\sqrt{1 - \sqrt{3}}$$

Solution:
By the quadratic formula $x^2 = \frac{2 \pm \sqrt{12}}{2} = 1 \pm \sqrt{3}$, and so the roots are as above.

(c) Let $K_1 = \mathbb{Q}(\alpha_1)$ and $K_2 = \mathbb{Q}(\alpha_2)$. Show that $K_1 \neq K_2$, and $K_1 \cap K_2 = \mathbb{Q}(\sqrt{3}) = F$.

Proof. Since $K_1 \subseteq \mathbb{R}$ but $K_2 \subsetneq \mathbb{R}$, then $K_1 \neq K_2$. since $K_1 \cap K_2 \subsetneq K_2$, we have the tower

$$\begin{tikzcd}
K_2 \arrow[d, dash, ">1"] \\
K_1 \cap K_2 \arrow[d, dash, "\geq 1"] \\
\mathbb{Q}(\sqrt{3}) \arrow[d, dash, "2"] \\
\mathbb{Q} \arrow[uuu, dash, bend left, "4"]
\end{tikzcd}$$

Hence we must have $[K_2 : K_1 \cap K_2] = 2$, which gives $[K_1 \cap K_2 : \mathbb{Q}(\sqrt{3})] = 1$, i.e., $K_1 \cap K_2 = \mathbb{Q}(\sqrt{3})$.

(d) Prove that K_1, K_2 and K_1K_2 are Galois over F with $\mathrm{Gal}(K_1K_2/F)$ the Klein 4-group. Write out the elements of $\mathrm{Gal}(K_1K_2/F)$ explicitly. Determine all the subgroups of the Galois group and give their corresponding fixed subfields of K_1K_2 containing F.

Proof. Let $f(x) = x^4 - 2x^2 - 2 = (x^2 - \alpha_1^2)(x^2 - \alpha_2^2)$. Then $f(x)$ and each of its quadratic factors are separable. Hence K_1, K_2, and K_1K_2 are Galois over $\mathbb{Q}$ since they are the splitting fields of $(x^2 - \alpha_1^2)$, $(x^2 - \alpha_2^2)$, and $f(x)$, respectively.

Define $\sigma = (\alpha_1, \alpha_3), \tau = (\alpha_2, \alpha_3)$. Then $G = \mathrm{Gal}(K_1K_2/F) = \{1, \sigma, \tau, \sigma\tau\}$. We then get the corresponding lattices

$$\begin{tikzcd}
 & 1 \arrow[dl, dash] \arrow[d, dash] \arrow[dr, dash] & \\
\langle\sigma\rangle \arrow[dr, dash] & \langle\tau\rangle \arrow[d, dash] & \langle\sigma\tau\rangle \arrow[dl, dash] \\
 & G &
\end{tikzcd}
\quad \widehat{=} \quad
\begin{tikzcd}
 & K_1K_2 \arrow[dl, dash] \arrow[d, dash] \arrow[dr, dash] & \\
K_2 \arrow[dr, dash] & K_1 \arrow[d, dash] & F(\sqrt{-2}) \arrow[dl, dash] \\
 & F &
\end{tikzcd}$$

We have $\langle\sigma\tau\rangle \,\widehat{=} F(\sqrt{-2})$ since $\sigma\tau(\sqrt{-2}) = \sigma\tau(\alpha_1\alpha_2) = (-\alpha_1)(-\alpha_2) = \alpha_1\alpha_2 = \sqrt{-2}$.

(e) Prove that the splitting field of $x^4 - 2x^2 - 2$ over $\mathbb{Q}$ is of degree 8 with dihedral Galois group.

Proof. Since $[K_1K_2 : \mathbb{Q}] = [K_1K_2 : F][F : \mathbb{Q}] = 4 \cdot 2$, then $H = \mathrm{Gal}(K_1K_2/\mathbb{Q})$ has order 8. Then $H \leq S_4$ and $H \cong D_8$, since the only subgroup of S_4 of order 8 is D_8.

Exercise 14.2.17. Let K/F be any finite extension and let $\alpha \in K$. Let L be a Galois extension of F containing K and let $H \leq \mathrm{Gal}(L/F)$ be the subgroup corresponding to K. Define the *norm* of α from K to F to be

$$N_{K/F}(\alpha) = \prod_{\sigma} \sigma(\alpha),$$

where the product is taken over all the embeddings of K into an algebraic closure of F (so over a set of coset representatives for H in $\mathrm{Gal}(L/F)$ by the Fundamental Theorem of Galois Theory). This is a product of Galois conjugates of α. In particular, if K/F is Galois this is $\prod_{\sigma \in \mathrm{Gal}(K/F)} \sigma(\alpha)$.

(a) Prove that $N_{K/F}(\alpha) \in F$.

Proof. Without loss of generality, we assume that a fixed algebraic closure $\overline{F}$ of F contains L. Then if $\tau \in \mathrm{Gal}(L/F)$ and $\sigma : K \to \overline{F}$ is an embedding, $\tau\sigma : K \to L \subseteq \overline{F}$ is an embedding of K into an an algebraic closure of F. So

$$\tau\left(N_{K/F}(\alpha)\right) = \tau\left(\prod_{\sigma} \sigma(\alpha)\right) = \prod_{\sigma} \tau\sigma(\alpha) = N_{K/F}(\alpha),$$

and hence $N_{K/F}(\alpha) \in F$. ☕

(b) Prove that $N_{K/F}(\alpha\beta) = N_{K/F}(\alpha)N_{K/F}(\beta)$, so that the norm is a multiplicative map from K to F.

Proof.

$$N_{K/F}(\alpha\beta) = \prod_{\sigma} \sigma(\alpha)\sigma(\beta) = \prod_{\sigma} \sigma(\alpha) \prod_{\sigma} \sigma(\beta) = N_{K/F}(\alpha)N_{K/F}(\beta).$$

☕

(c) Let $K = F(\sqrt{D})$ be a quadratic extension of F. Show that $N_{K/F}(a+b\sqrt{D}) = a^2 - Db^2$.

Proof. Since $p(x) = m_{\sqrt{D},F}(x) = x^2 - D$, then $[K : F] = 2$. Also, $\gcd(p(x), p'(x)) = 1$ and so $p(x)$ is separable. Hence K is Galois over F since it is a splitting field for $p(x)$ over F. Then $G = \mathrm{Gal}(K/F) = \{\mathrm{id}, \sigma\}$ where $\sigma \widehat{=} (\sqrt{D}, -\sqrt{D})$. Then

$$N_{K/F}(a + b\sqrt{D}) = \prod_{\sigma \in G} \sigma(a + b\sqrt{D}) = (a + b\sqrt{D})(a - b\sqrt{D}) = a^2 - Db^2.$$

☕

(d) Let $m_\alpha(x) = x^d + a_{d-1}x^{d-1} + \cdots + a_1 x + a_0 \in F[x]$ be the minimal polynomial for $\alpha \in K$ over F. Let $n = [K : F]$. Prove that d divides n, that there are d distinct Galois conjugates of α which are all repeated n/d times in the product above and conclude that $N_{K/F}(\alpha) = (-1)^n a_0^{n/d}$.

Proof. We have the tower $F \subseteq F(\alpha) \subseteq K \subseteq L$. Then $d = [F(\alpha) : F]$ divides $[K : F] = n$ by multiplicativity of degrees. Now, since L/F is Galois, it is in particular separable, and hence K/F is separable. So $m_\alpha(x)$ has d distinct roots, and so there are d distinct Galois conjugates of α. Let $\sigma_1, \ldots, \sigma_d \in \mathrm{Gal}(L/F)$ be such that $\sigma_1(\alpha), \ldots, \sigma_d(\alpha)$ are the distinct roots of $m_\alpha(x)$. By restricting each σ_i to $F(\alpha)$, we consider each as an embedding $\sigma_i : F(\alpha) \to \overline{L}$.

We now argue that $K/F(\alpha)$ is separable. Then for each i, we know that the number of distinct ways to extend σ_i to embeddings of K into $\overline{L}$ is $[K : F(\alpha)] = n/d$. To that end, notice that since L/F is separable then if $\beta \in K$, $m_{\beta,F}(x)$ is separable. Then $m_{\beta,F(\alpha)}(x)$ divides the $m_{\beta,F}(x)$, and so any root of the former must also be a root of the later. So $m_{\beta,F(a)}(x)$ must have distinct roots, because otherwise, a repeated root of $m_{\beta,F(\alpha)}(x)$ would be a repeated root of $m_{\beta,F}(x)$, a contradiction. Hence $K/F(a)$ is separable.

Now for all $1 \le i \le d$, let $\tau_{i,1}, \ldots, \tau_{i,n/d} : K \to \overline{L}$ be distinct embeddings extending σ_i. Hence each Galois conjugate is repeated n/d times in the product above. Note that $\tau_{i,j}(\alpha) = \sigma_i(\alpha)$ for all $1 \le i \le d$ and for all $1 \le j \le n/d$. Then, we can write $m_\alpha(x) = \prod_{i=1}^d (x - \sigma_i(\alpha))$, which means $a_0 = (-1)^d \prod_{i=1}^d \sigma_i(\alpha)$. So $\prod_{i=1}^d \sigma_i(\alpha) = (-1)^d a_0$ and thus

$$
\begin{aligned}
N_{K/F}(\alpha) = \prod_{i=1}^{d}\prod_{j=1}^{n/d} \tau_{i,j}(\alpha) = \prod_{i=1}^{d} (\sigma_i(\alpha))^{n/d} &= \left(\prod_{i=1}^{d} \sigma_i(\alpha)\right)^{n/d} \\
&= \left((-1)^d a_0\right)^{n/d} \\
&= (-1)^n a_0^{n/d}.
\end{aligned}
$$

☕

Exercise 14.3.8. Determine the splitting field of the polynomial $x^p - x - a$ over $\mathbb{F}_p$ where $a \neq 0$, $a \in \mathbb{F}_p$. Show explicitly that the Galois group is cyclic.

Proof. Let $f(x) = x^p - x - a$, and suppose α is a root of $f(x)$. Then $f(\alpha+1) = (\alpha+1)^p - (\alpha+1) - a = \alpha^p + 1 - \alpha - 1 - a = 0$. So $\alpha + 1$ is a root of $f(x)$. Continuing inductively, we get that

$$\alpha + \underbrace{1 + \cdots + 1}_{k-summands}$$

is a root of $f(x)$ for all $0 \le k \le p-1$. Hence $\{\alpha + \beta\}_{\beta \in \mathbb{F}_p}$ are the p distinct roots of $f(x)$. So the splitting field $\mathbb{F}_p(\alpha)$ of $f(x)$ over $\mathbb{F}_p$ is Galois over $\mathbb{F}_p$. Then there exists $\sigma \in G = \mathrm{Gal}(\mathbb{F}_p(\alpha), \mathbb{F}_p)$ such that $\sigma(\alpha) = \alpha + 1$. If $\tau \in G$ then $\tau(\alpha) = \tau + \beta$ for some $\beta \in \mathbb{F}_p$. Hence

$$\sigma^\beta(\alpha) = \sigma^{\beta-1}(\alpha+1) = \sigma^{\beta-2}(\alpha+2) = \cdots = \sigma^2(\alpha+\beta-2) = \sigma(\alpha+\beta-1) = \alpha+\beta,$$

and so $\sigma^\beta = \tau$. So $\langle \sigma \rangle = G$. ☕

Exercise 14.5.7. Show that complex conjugation restricts to the automorphism $\sigma_{-1} \in \mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q})$ of the cyclotomic field of n^{th} roots of unity. Show that the field $K^+ = \mathbb{Q}(\zeta_n + \zeta_n^{-1})$ is the subfield of real elements in $K = \mathbb{Q}(\zeta_n)$, called the *maximal real subfield of* K.

Proof. Let $\tau : \mathbb{C} \to \mathbb{C}$ be complex conjugation, and let $\zeta_n = e^{2\pi i/n}$. Then

$$\tau(\zeta_n) = e^{-2\pi i/n} = \zeta_n^{-1} = \sigma_{-1}(\zeta_n).$$

Notice that $K^{\langle \sigma_{-1} \rangle} = K \cap \mathbb{R}$, i.e., the fixed field of $\langle \sigma_{-1} \rangle$ in K is the maximal real subfield of K. Since $\sigma_{-1}(\zeta_n + \zeta_n^{-1}) = \sigma_{-1}(\zeta_n) + \sigma_{-1}(\zeta_n^{-1}) = \zeta_n^{-1} + \zeta_n$, then $K^+ \subseteq K^{\langle \sigma_{-1} \rangle}$. So we have the tower

$$\mathbb{Q} \subset K^+ \subset K^{\langle \sigma_{-1} \rangle} \subset K$$

Since $|\langle \sigma_{-1} \rangle| = 2$, then by the Fundamental Theorem of Galois Theory, $[K : K^{\langle \sigma_{-1} \rangle}] = 2$. Notice that $K = K^+(\zeta_n)$, and that $m_{\zeta_n, K^+}(x) = x^2 - (\zeta_n + \zeta_n^{-1})x + 1 = (x - \zeta_n)(x - \zeta_n^{-1})$. So $[K : K^+] = 2$. This forces $[K^{\langle \sigma_{-1} \rangle} : K^+] = 1$, i.e., $K^+ = K^{\langle \sigma_{-1} \rangle} = K \cap \mathbb{R}$. ☕

Exercise 14.5.8. Let $K_n = \mathbb{Q}(\zeta_{2^{n+2}})$ be the cyclotomic field of 2^{n+2}-th roots of unity, $n \geq 0$. Set $\alpha_n = (\zeta_{2^{n+2}} + \zeta_{2^{n+2}}^{-1})$ and $K_n^+ = \mathbb{Q}(\alpha_n)$, the maximal real subfield of K_n.

(a) Show that for all $n \geq 0$, $[K_n : \mathbb{Q}] = 2^{n+1}, [K_n : K_n^+] = 2, [K_n^+ : \mathbb{Q}] = 2^n$, and $[K_{n+1}^+ : K_n^+] = 2$.

Proof. In the first case, $[K_n : \mathbb{Q}] = \deg \Phi_{2^{n+2}}(x) = \varphi(2^{n+2}) = 2^{n+1}(2-1) = 2^{n+1}$, where $\Phi_{2^{n+2}}$ is the 2^{n+2} cyclotomic polynomial and φ is the Euler phi function. Then the minimal polynomial of $\zeta_{2^{n+2}}$ over K_n^+ is $(x - \zeta_{2^{n+2}})(x - \zeta_{2^{n+2}}^{-1})$, which gives $[K_n : K_n^+] = 2$.

Then $[K_n^+ : \mathbb{Q}] = [K_n^+ : \mathbb{Q}]/[K_n : K_n^+] = 2^{n+1}/2 = 2^n$. And finally $[K_{n+1}^+ : K_n^+] = [K_{n+1}^+ : \mathbb{Q}]/[K_n^+ : \mathbb{Q}] = 2^{n+1}/2^n = 2$. ☕

(b) Determine the quadratic equation satisfied by $\zeta_{2^{n+2}}$ over K_n^+ in terms of α_n.

Proof.

$$\begin{aligned}(x - \zeta_{2^{n+2}})(x - \zeta_{2^{n+2}}^{-1}) &= x^2 - (\zeta_{2^{n+2}} + \zeta_{2^{n+2}}^{-1})x + \zeta_{2^{n+2}}\zeta_{2^{n+2}}^{-1} \\ &= x^2 - \alpha_n + (\alpha_n - \zeta_{2^{n+2}}^{-1})(\alpha_n - \zeta_{2^{n+2}})\end{aligned}$$

☕

(c) Show that for $n \geq 0$, $\alpha_{n+1}^2 = 2 + \alpha_n$ and hence show that

$$\alpha_n = \pm\sqrt{2 \pm \sqrt{2 \pm \sqrt{\cdots \pm \sqrt{2}}}} \quad (n \text{ times})$$

giving an explicit formula for the (constructable) 2^{n+2}-th roots of unity.

Proof. We have

$$\alpha_{n+1}^2 = \zeta_{2^{n+3}}^2 + 2\zeta_{2^{n+3}}\zeta_{2^{n+3}}^{-1} + (\zeta_{2^{n+3}}^{-1})^2 = \zeta_{2^{n+2}} + 2(1) + \zeta_{2^{n+2}}^{-1} = 2 + \alpha_n,$$

which gives $\alpha_n = \pm\sqrt{2 + \alpha_{n-1}} = \pm\sqrt{2 \pm \sqrt{2 \pm \alpha_{n-2}}} = \pm\sqrt{2 \pm \sqrt{2 \pm \sqrt{\cdots \pm \sqrt{2}}}}$. ☕

Exercise 14.3.11. Prove that $x^{p^n} - x + 1$ is irreducible over $\mathbb{F}_p$ only when $n = 1$ or $n = p = 2$.

Proof. Suppose $f(x)$ is irreducible over $\mathbb{F}_p$. Let α be a root of $f(x) = x^{p^n} - x + 1$ inside a fixed algebraic closure $\overline{\mathbb{F}}_p$ of $\mathbb{F}_p$. Then for any $\beta \in \mathbb{F}_p$

$$f(\alpha+\beta) = (\alpha+\beta)^{p^n} - (\alpha+\beta) + 1 = \alpha^{p^n} + \beta^{p^n} - \alpha - \beta + 1 = \underbrace{(\alpha^{p^n} - \alpha + 1)}_{\substack{=0 \text{ since } \alpha \\ \text{is a root of } f(x)}} + \underbrace{(\beta^{p^n} - \beta)}_{\substack{=0 \text{ since} \\ \beta^{p^n} = \beta \ \forall \beta \in \mathbb{F}_{p^n}}} = 0.$$

Hence $\{\alpha + \beta\}_{\beta \in F_{p^n}}$ are the p^n roots of the irreducible polynomial $f(x)$, which gives

$$p^n = [\mathbb{F}_p(\alpha) : \mathbb{F}_p] = [\mathbb{F}_p(\alpha + \beta) : \mathbb{F}_p].$$

Now, know that for all $r \in \mathbb{Z}^+$, $\overline{\mathbb{F}}_p$ contains a *unique* subfield of order p^r. So, since

$$[\mathbb{F}_{p^{p^n}} : \mathbb{F}_p] = p^n \quad \text{and} \quad \mathbb{F}_p(\alpha + \beta), \mathbb{F}_p(\alpha) \subset \overline{\mathbb{F}}_p,$$

then

$$\mathbb{F}_p(\alpha + \beta) = \mathbb{F}_{p^{p^n}} = \mathbb{F}_p(\alpha)$$

in $\overline{\mathbb{F}}_p$. Now, let $\beta \in \mathbb{F}_{p^n}$. Then $\alpha + \beta, \alpha \in \mathbb{F}(\alpha)$, and so $\beta = \alpha + \beta - \alpha \in \mathbb{F}(\alpha)$, hence $\mathbb{F}_{p^n} \subseteq \mathbb{F}_p(\alpha)$.

Then $\mathbb{F}_p(\alpha)/\mathbb{F}_{p^n}$ is cyclic Galois (all the Galois groups of finite fields over subfields are cyclic). So let $\langle \sigma \rangle = G = \text{Gal}(\mathbb{F}_p(\alpha)/\mathbb{F}_{p^n})$. Then $\sigma(\alpha) = \alpha + \beta$ for some $\beta \in \mathbb{F}_{p^n}$. Since $\beta \in \mathbb{F}_{p^n}$ then $\sigma(\beta) = \beta$. So

$$\begin{aligned}
\sigma(\alpha) &= \alpha + \beta \\
\sigma(\alpha + \beta) &= \sigma(\alpha) + \sigma(\beta) = \alpha + 2\beta \\
\sigma(\alpha + 2\beta) &= \sigma(\alpha) + 2\sigma(\beta) = \alpha + 3\beta \\
&\vdots \\
\sigma(\alpha + (p-1)\beta) &= \sigma(\alpha) + (p-1)\sigma(\beta) = \alpha + p\beta = \alpha.
\end{aligned}$$

Therefore,

$$\sigma \mathrel{\widehat{=}} (\alpha, \alpha + \beta, \alpha + 2\beta, \dots, \alpha + (p-1)\beta),$$

and hence $p = |\langle \sigma \rangle| = |G| = [\mathbb{F}_p(\alpha) : \mathbb{F}_{p^n}]$. So we have the tower

$$\begin{tikzcd}
\mathbb{F}_p(\alpha) \arrow[d, dash, "p"] \arrow[dd, dash, bend right=60, "p^n"'] \\
\mathbb{F}_{p^n} \arrow[d, dash, "n"] \\
\mathbb{F}_p
\end{tikzcd}$$

So $pn = p^n$. If $n = 1$ then the equality holds. If $n = 2$, then $2 = p^{2-1} = p$. If $n \geq 3$, the equation $n = p^{n-1}$ has no solution. ☕

Exercise 14.4.4. Let K/F be a finite Galois extension, and $\overline{K}$ be an algebraic closure of K. Let $f(x) \in F[x]$ be separable and irreducible over F, with splitting field L inside $\overline{K}$. Let α be a root of $f(x)$ inside L. Show that $f(x)$ factors in $K[x]$ into a product of m irreducible polynomials each of degree d over K, where $m = [F(\alpha) \cap K : F]$ and $d = [K(\alpha) : K]$.

Proof. If $f(x) = p_1(x)p_2(x) \cdots p_m(x)$ for irreducibles $p_i(x) \in K[x]$, then since $f(x)$ splits completely into linear factors in $L[x]$, the coefficients of $p_i(x)$ are sums of products of the roots of $f(x)$, and so the $p_i(x)$ all lie in $L[x]$. Hence this factorization of $f(x)$ in $K[x]$ is the same as that in $(L \cap K)[x]$.

Suppose β_i is a root of $p_i(x)$ for some i. If $\sigma \in H := \mathrm{Gal}(L/L \cap K)$, then $\sigma(\beta_i)$ is also a root of $p_i(x)$, since the elements of H permute the roots of the irreducible factors of $f(x)$. Hence the orbit $\mathcal{O}_{\beta_i}$ precisely contains the roots of $p_i(x)$. So we get a correspondence $\{p_i(x)\} \leftrightarrow \{H_{\beta_i}\}$. Since H acts transitively on the roots of f, then by Exercise 9 of Section 4.1, that the H_{β_i} each have the same cardinality, i.e., the degrees of all the $p_i(x)$ are all the same.

If α is a root of $f(x)$ then without loss of generality, suppose α is a root of $p_1(x)$. Also suppose $\deg p_1(x) = d$. Then $[K(\alpha) : K] = d$ since $p_1(x)$ is irreducible. Hence all factors $p_i(x)$ have degree d.

Since $p_1(x) \in K[x]$ is of degree K and has $\alpha \in L$ as a root, then $[K(\alpha) : K] = d$. Since $\alpha \in L$ is a root of the irreducible polynomial $f(x)$ over F, then $[F(\alpha) : F] = md$. So we have the tower

$$\begin{tikzcd}
 & & K(\alpha) = K \cap F(\alpha) \arrow[dll, no head] \arrow[drr, no head, "d"] & & \\
F(\alpha) \arrow[drr, no head] \arrow[ddrr, no head, "md"'] & & & & K \arrow[dll, no head] \\
 & & K \cap F(\alpha) \arrow[d, no head] & & \\
 & & F & &
\end{tikzcd}$$

So we have by the formula given in Corollary 20, (page 592, D& F),

$$[K \cap F(\alpha)] = \frac{[K : F][F(\alpha) : F]}{[K(\alpha) : F]} = \frac{[K : F]md}{[K : F][K(\alpha) : K]} = \frac{md}{d} = m.$$

Exercise 14.7.4. Let $K = \mathbb{Q}(\sqrt[n]{a})$, where $a \in \mathbb{Q}$, $a > 0$ and suppose $[K : \mathbb{Q}] = n$ (i.e., $x^n - a$ is irreducible). Let E be an subfield of K and let $[E : \mathbb{Q}] = d$. Prove that $E = \mathbb{Q}(\sqrt[d]{a})$. [Consider $N_{K/E}(\sqrt[n]{a}) \in E$.]

Proof. The elements of $\mathrm{Gal}(K/E)$ are $\{\sigma_i\}$ where $\sigma_i(\sqrt[n]{a}) = \zeta^i \sqrt[n]{a}$ for all $0 \le i \le (n/d) - 1$. So we have

$$N_{K/E}(\sqrt[n]{a}) = \prod_{i=0}^{(n/d)-1} \sigma_i(\sqrt[n]{a}) = \zeta^{\sum i}(\sqrt[n]{a})^{n/d} = \zeta^{\sum i} \sqrt[d]{a} \in E.$$

Now since $E \subset K \subset \mathbb{R}$, then $\zeta^{\sum i} = \pm 1$. So $\mathbb{Q}(\sqrt[d]{a}) \subseteq E$, and we have the tower

$$\begin{tikzcd}
K \arrow[d, dash, "n/d"] \\
E \arrow[d, dash, "1"] \arrow[dd, dash, bend right=60, "d"'] \\
\mathbb{Q}(\sqrt[d]{a}) \arrow[d, dash, "d"] \\
\mathbb{Q}
\end{tikzcd}$$

Hence $E = \mathbb{Q}(\sqrt[d]{a})$. Now suppose α is an arbitrary root of $x^n - a$ with $K = \mathbb{Q}(\alpha)$, $E \subseteq K$, and $[E : \mathbb{Q}] = d$. We have an isomorphism $\mathbb{Q}(\sqrt[n]{a}) \cong \mathbb{Q}(\alpha)$ given by $\sqrt[n]{a} \mapsto \alpha$. This induces an isomorphism $\mathbb{Q}(\sqrt[d]{a}) \cong \mathbb{Q}(\alpha^{n/d})$ given by $\sqrt[d]{a} \mapsto \alpha^{n/d}$. So $E = \mathbb{Q}(\alpha^{n/d})$.

$$\begin{tikzcd}
\mathbb{Q}(\sqrt[n]{a}) \arrow[r, "\cong"] \arrow[d, dash] & \mathbb{Q}(\alpha) \arrow[d, dash] \\
\mathbb{Q}(\sqrt[d]{a}) \arrow[r, "\cong"] \arrow[d, dash, "d"] & \mathbb{Q}(\alpha^{n/d}) \arrow[d, dash, "d"] \\
\mathbb{Q} & \mathbb{Q}
\end{tikzcd}$$