Visual Abstract Algebra

Marcus Pivato

March 25, 2003

Contents

List of Figures

Part I

Groups

Chapter 1

Homomorphisms

1.1 Cosets and Coset Spaces

Prerequisites: Subgroups

If $\mathcal{A} < \mathcal{G}$ is a subgroup, and $g \in \mathcal{G}$, then the corresponding **left** and **right cosets** of $\mathcal{A}$ are the sets

$$g\mathcal{A} \ = \ \{ga \ ; \ a \in \mathcal{A}\}, \qquad \text{and} \qquad \mathcal{A}g \ = \ \{ag \ ; \ a \in \mathcal{A}\}.$$

Example 1:

(a) Let $\mathcal{G} = (\mathbb{Z}, +)$, and let $\mathcal{A} \ = \ 5\mathbb{Z} \ = \ \{5z \ ; \ z \in \mathbb{Z}\} \ = \ \{\ldots, -5, 0, 5, 10, 15, \ldots\}$. Then $3 + 5\mathbb{Z} \ = \ \{3 + 5z \ ; \ z \in \mathbb{Z}\} \ = \ \{\ldots, -2, 3, 8, 13, 18, \ldots\}$. Note that $\mathbb{Z}$ is abelian, so $\mathcal{A} + 3 = 3 + \mathcal{A}$.

(b) Let $\mathcal{G} = \mathbf{S}_3 = \Big\{e, \ (12), \ (13), \ (23), \ (123), \ (132)\Big\}$, and let $\mathcal{A} = \Big\{e, \ (12)\Big\}$. Then

$(123) \cdot \mathcal{A} \ = \ \Big\{(123), \ (123)(12)\Big\} \ = \ \Big\{(123), \ (13)\Big\}$, but

$\mathcal{A} \cdot (123) \ = \ \Big\{(123), \ (12)(123)\Big\} \ = \ \Big\{(123), \ (23)\Big\}$, so $(123) \cdot \mathcal{A} \ \neq \ \mathcal{A} \cdot (123)$.

(c) Again, let $\mathcal{G} = \mathbf{S}_3$, and let $\mathcal{A} = \mathbf{A}_3 = \Big\{e, \ (123), \ (132)\Big\}$. Then

$(12) \cdot \mathbf{A}_3 \ = \ \Big\{(12), \ (12)(123), \ (12)(132)\Big\} \ = \ \Big\{(12), \ (23), \ (13)\Big\}$.

Also, $(12) \cdot \mathbf{A}_3 \ = \ \Big\{(12), \ (123)(12), \ (132)(12)\Big\} \ = \ \Big\{(12), \ (13), \ (23)\Big\}$,

so that $(12)\mathbf{A}_3 = \mathbf{A}_3(12)$.

(d) Let $\mathcal{X}$ and $\mathcal{Y}$ be groups, and consider the product group

$$\mathcal{G} \ = \ \mathcal{X} \times \mathcal{Y} \ = \ \{(x, y) \ ; \ x \in \mathcal{X} \text{ and } \ y \in \mathcal{Y}\}.$$

Let $\mathcal{A} = \{(e_{\mathcal{X}}, y) \,;\, y \in \mathcal{Y}\} = \{e_{\mathcal{X}}\} \times \mathcal{Y}$. Then for any $(x, y_1) \in \mathcal{G}$,
$(x, y_1)\cdot\mathcal{A} = \{(x, y_1) \cdot (e_{\mathcal{X}}, y_2) \,;\, y_2 \in \mathcal{Y}\} = \{(x,\ y_1 y_2) \,;\, y \in \mathcal{Y}\} = \{(x, y) \,;\, y \in \mathcal{Y}\}$
$= \{x\} \times \mathcal{Y}$.

Lemma 2 *Let $\mathcal{G}$ be a group and let $\mathcal{A} < \mathcal{G}$ be a subgroup.*

(a) *For any $g \in \mathcal{G}$,* $\Big(g\mathcal{A} = \mathcal{A} \Big) \iff \Big(g \in \mathcal{A} \Big) \iff \Big(\mathcal{A}g = \mathcal{A} \Big)$.

(b) *For any $g, h \in \mathcal{G}$, the following are equivalent:*

$$\Big(g \in h\mathcal{A} \Big) \iff \Big(g\mathcal{A} = h\mathcal{A} \Big) \iff \Big(h \in g\mathcal{A} \Big) \iff \Big(h^{-1}g \in \mathcal{A} \Big) \iff \Big(g^{-1}h \in \mathcal{A} \Big)$$

Likewise, the following are equivalent:

$$\Big(g \in \mathcal{A}h \Big) \iff \Big(\mathcal{A}g = \mathcal{A}h \Big) \iff \Big(h \in \mathcal{A}g \Big) \iff \Big(hg^{-1} \in \mathcal{A} \Big) \iff \Big(gh^{-1} \in \mathcal{A} \Big)$$

Proof: **Exercise 1** □

Example 3:

(a) Let $\mathcal{G} = (\mathbb{Z}, +)$, and let $\mathcal{A} = 5\mathbb{Z}$ as in Example ⟨1a⟩. Then
$10 + 5\mathbb{Z} = \{10 + 5z \,;\, z \in \mathbb{Z}\} = \{\ldots, 5, 10, 15, 20, 25, \ldots\} = 5\mathbb{Z}$.

(b) Let $\mathcal{G} = \mathbf{S}_3$ and let $\mathcal{A} = \mathbf{A}_3 = \Big\{e,\ (123),\ (132)\Big\}$, as in Example ⟨1c⟩. Then
$(123) \cdot \mathbf{A}_3 = \Big\{(123),\ (132),\ e\Big\} = \mathbf{A}_3$.

(c) Let $\mathcal{X}$ and $\mathcal{Y}$ be groups; let $\mathcal{G} = \mathcal{X} \times \mathcal{Y}$ and let $\mathcal{A} = \{e_{\mathcal{X}}\} \times \mathcal{Y}$, as in Example ⟨1d⟩. Then for any $y \in \mathcal{Y}$, $(e_{\mathcal{G}}, y) \cdot \mathcal{A} = \{e_{\mathcal{X}}\} \times \mathcal{Y} = \mathcal{A}$.

The **left coset space** of $\mathcal{A}$ is the set of all its left cosets:

$$\mathcal{G}/\mathcal{A} = \{g\mathcal{A} \,;\, g \in \mathcal{G}\}.$$

Example 4:

(a) Let $\mathcal{G} = (\mathbb{Z}, +)$, and let $\mathcal{A} = 5\mathbb{Z}$ as in Example ⟨1a⟩. Then

$$\frac{\mathbb{Z}}{5\mathbb{Z}} = \{5\mathbb{Z},\ 1 + 5\mathbb{Z},\ 2 + 5\mathbb{Z},\ 3 + 5\mathbb{Z},\ 4 + 5\mathbb{Z}\}.$$

(b) Let $\mathcal{G} = \mathbf{S}_3$, and let $\mathcal{A} = \left\{e,\ (12)\right\}$, as in Example (1b). Then

$$\frac{\mathbf{S}_3}{\mathcal{A}} \quad = \quad \left\{\mathcal{A},\ \ (123)\mathcal{A},\ \ (23)\mathcal{A}\right\} \quad = \quad \left\{\{e,\ (12)\},\ \ \{(123),\ (13)\},\ \ \{(23),\ (132)\}\right\}.$$

(c) Let $\mathcal{G} = \mathbf{S}_3$, and let $\mathcal{A} = \mathbf{A}_3 = \left\{e,\ (123),\ (132)\right\}$, as in Example $\langle$1c$\rangle$. Then

$$\frac{\mathbf{S}_3}{\mathbf{A}_3} \quad = \quad \left\{\mathbf{A}_3,\ (12)\mathbf{A}_3\right\} \quad = \quad \left\{\{e,\ (123),\ (132)\},\ \ \{(12),\ (23),\ (13)\}\right\}.$$

(d) Let $\mathcal{X}$ and $\mathcal{Y}$ be groups; let $\mathcal{G} = \mathcal{X} \times \mathcal{Y}$, and let $\mathcal{A} = \{e_{\mathcal{X}}\} \times \mathcal{Y}$, as in Example $\langle$1d$\rangle$. Then $\dfrac{\mathcal{G}}{\mathcal{A}} = \{\mathcal{A}_x\ ;\ x \in \mathcal{X}\}$, where, for any fixed $x \in \mathcal{X}$, $\mathcal{A}_x = \{(x,y)\ ;\ y \in \mathcal{Y}\}$. ———

1.2 Lagrange's Theorem

Prerequisites: §1.1

Let $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_N \subset \mathcal{G}$ be a collection of subsets of $\mathcal{G}$. We say that $\mathcal{G}$ is a **disjoint union** of $\mathcal{S}_1, \ldots, \mathcal{S}_N$ if:

1. $\mathcal{G} = \mathcal{S}_1 \cup \mathcal{S}_2 \cup \ldots \cup \mathcal{S}_N$.

2. $\mathcal{S}_n \cap \mathcal{S}_m = \emptyset$, whenever $n \neq m$.

We then write: "$\mathcal{G} = \mathcal{S}_1 \sqcup \mathcal{S}_2 \sqcup \ldots \sqcup \mathcal{S}_N$", or "$\mathcal{G} = \displaystyle\bigsqcup_{n=1}^{N} \mathcal{S}_n$". We say that the collection $\mathfrak{S} = \{\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_N\}$ is a **partition** of $\mathcal{G}$.

More generally, let $\mathfrak{S}$ be any (possibly infinite) collection of subsets of $\mathcal{G}$. We say that $\mathcal{G}$ is a **disjoint union** of the elements in $\mathfrak{S}$ if:

1. $\mathcal{G} = \displaystyle\bigcup_{\mathcal{S} \in \mathfrak{S}} \mathcal{S}$.

2. $\mathcal{S} \cap \mathcal{S}' = \emptyset$, whenever $\mathcal{S}$ and $\mathcal{S}'$ are distinct elements of $\mathfrak{S}$.

We then write: "$\mathcal{G} = \displaystyle\bigsqcup_{\mathcal{S} \in \mathfrak{S}} \mathcal{S}$". We say that $\mathfrak{S}$ is a **partition** of $\mathcal{G}$.

Lemma 5 *Let $\mathcal{G}$ be a group and $\mathcal{A}$ a subgroup. Then $\mathcal{G}/\mathcal{A}$ is a partition of $\mathcal{G}$.*

In other words:

1. *Any two cosets of $\mathcal{A}$ are either identical or disjoint: for any $g_1, g_2 \in \mathcal{G}$, either $g_1\mathcal{A} = g_2\mathcal{A}$, or $(g_1\mathcal{A}) \cap (g_2\mathcal{A}) = \emptyset$.*

2. $\mathcal{G} \;=\; \bigsqcup_{(g\mathcal{A})\;\in\;\mathcal{G}/\mathcal{A}} (g\mathcal{A}).$

Proof: **Exercise 2** Hint: Apply Lemma 2(b) ______ □

Example 6:

(a) Let $\mathcal{G} = (\mathbb{Z}, +)$, and let $\mathcal{A} = 5\mathbb{Z}$ as in Example ⟨4a⟩. Then

$$\begin{aligned}\mathbb{Z} &= \{\ldots,-5,0,5,10,\ldots\} \sqcup \{\ldots,-4,1,6,11,\ldots\} \sqcup \{\ldots,-3,2,7,12,\ldots\} \\ &\qquad \sqcup \{\ldots,-2,3,8,13,\ldots\} \sqcup \{\ldots,-1,4,9,14,\ldots\} \\ &= (5\mathbb{Z}) \sqcup (1+5\mathbb{Z}) \sqcup (2+5\mathbb{Z}) \sqcup (3+5\mathbb{Z}) \sqcup (4+5\mathbb{Z}) \;=\; \bigsqcup_{(n+5\mathbb{Z})\;\in\;\mathbb{Z}/5\mathbb{Z}} (n+5\mathbb{Z}).\end{aligned}$$

(b) Let $\mathcal{G} = \mathbf{S}_3$, and let $\mathcal{A} = \Big\{e,\ (12)\Big\}$, as in Example (4b). Then

$$\begin{aligned}\mathbf{S}_3 &= \Big\{e,\,(12)\Big\} \sqcup \Big\{(123),\,(13)\Big\} \sqcup \Big\{(23),\,(132)\Big\} \;=\; \mathcal{A} \sqcup (123)\mathcal{A} \sqcup (23)\mathcal{A} \\ &= \bigsqcup_{(g\mathcal{A})\;\in\;\mathbf{S}_3/\mathcal{A}} (g\mathcal{A}).\end{aligned}$$

(c) Let $\mathcal{X}$ and $\mathcal{Y}$ be groups; let $\mathcal{G} = \mathcal{X} \times \mathcal{Y}$, and let $\mathcal{A} \;=\; \{e_{\mathcal{X}}\} \times \mathcal{Y}$, as in Example ⟨4d⟩. For any fixed $x \in \mathcal{X}$, let $\mathcal{A}_x = \{(x,y)\ ;\ y \in \mathcal{Y}\}$. Then $\mathcal{G} \;=\; \bigsqcup_{x\in\mathcal{X}} \mathcal{A}_x \;=\; \bigsqcup_{(x,y)\mathcal{A}\;\in\;\mathcal{G}/\mathcal{A}} (x,y)\mathcal{A}.$.

If $\mathcal{G}/\mathcal{A}$ is a finite set, then the **index** of $\mathcal{A}$ in $\mathcal{G}$ is the cardinality of $\mathcal{G}/\mathcal{A}$. It is sometimes denoted by "$|\mathcal{G} : \mathcal{A}|$".

Corollary 7 Lagrange's Theorem

Let $\mathcal{G}$ be a finite group and let $\mathcal{A}$ be any subgroup. Then:

1. $|\mathcal{G}| \;=\; |\mathcal{A}| \cdot \left|\dfrac{\mathcal{G}}{\mathcal{A}}\right|.$

2. In particular, $|\mathcal{A}|$ divides $|\mathcal{G}|$.

Proof:

Claim 1: *For any $g \in \mathcal{G}$,* $\quad |g\mathcal{A}| \;=\; |\mathcal{A}|$.

Proof: Define a map $\phi : \mathcal{A} \longrightarrow g\mathcal{A}$ by: $\phi(a) = ga$. It is **Exercise 3** to verify that ϕ is a bijection. Thus, $|g\mathcal{A}| \;=\; |\mathcal{A}|$. . . . . □ [Claim 1]

Now, Lemma 5 says that the cosets of $\mathcal{A}$ partition $\mathcal{G}$. Since $\mathcal{G}$ is finite, we know that there are a finite number of distinct cosets for $\mathcal{A}$ —let's say they are $(g_1\mathcal{A})$, $(g_2\mathcal{A})$, $\ldots$, $(g_N\mathcal{A})$, for some elements $g_1, g_2, \ldots, g_N \in \mathcal{G}$. Thus,

$$\begin{array}{rcl} \mathcal{G} & = & (g_1\mathcal{A}) \sqcup (g_2\mathcal{A}) \sqcup \ldots \sqcup (g_N\mathcal{A}), \\ \text{therefore,}\quad |\mathcal{G}| & = & |g_1\mathcal{A}| + |g_2\mathcal{A}| + \ldots + |g_N\mathcal{A}|. \\ & \overline{\overline{\text{Clm.1}}} & \underbrace{|\mathcal{A}| \quad + |\mathcal{A}| \quad + \ldots + |\mathcal{A}|}_{N} \quad = \quad N \cdot |\mathcal{A}|, \end{array}$$

where N is the number of distinct cosets—that is, $N = |\mathcal{G}/\mathcal{A}|$. □

1.3 Normal Subgroups & Quotient Groups

Prerequisites: §1.1; Homomorphisms

Let $\mathcal{A} < \mathcal{G}$ be a subgroup. We say that $\mathcal{A}$ is **normal** if $g\mathcal{A} = \mathcal{A}g$ for every $g \in \mathcal{G}$. We indicate this by writing "$\mathcal{A} \lhd \mathcal{G}$".

Example 8:

(a) Let $\mathcal{G} = \mathbb{Z}$ and let $\mathcal{A} = 5\mathbb{Z}$, as in Example 4a. Then $\mathcal{A}$ is normal in $\mathbb{Z}$, because for any $n \in \mathbb{Z}$, $\quad n + 5\mathbb{Z} \;=\; \{n + 5z\ ;\ z \in \mathbb{Z}\} \;=\; \{5z + n\ ;\ z \in \mathbb{Z}\} \;=\; 5\mathbb{Z} + n$.

(b) In general, if $\mathcal{G}$ is any *abelian* group, then *every* subgroup is normal (**Exercise 4**).

(c) Let $\mathcal{G} = \mathbf{S}_3$, and let $\mathcal{A} = \Big\{e,\ (12)\Big\}$, as in Example $\langle$1b$\rangle$. Then $\mathcal{A}$ is *not* normal in $\mathbf{S}_3$, because $(123) \cdot \mathcal{A} \;\neq\; \mathcal{A} \cdot (123)$.

(d) Let $\mathcal{G} = \mathbf{S}_3$, and let $\mathcal{A} = \mathbf{A}_3$, as in Example $\langle$4c$\rangle$. Then $\mathbf{A}_3$ is normal in $\mathbf{S}_3$, because the only nontrivial *left* coset of $\mathbf{A}_3$ is $(12)\mathbf{A}_3$, and the only nontrivial *right* coset of $\mathbf{A}_3$ is $\mathbf{A}_3(12)$, and we saw in Example (1c) that $(12)\mathbf{A}_3 = \mathbf{A}_3(12)$.

(e) Let $\mathcal{X}$ and $\mathcal{Y}$ be groups; let $\mathcal{G} = \mathcal{X} \times \mathcal{Y}$, and let $\mathcal{A} \;=\; \{e_{\mathcal{X}}\} \times \mathcal{Y}$, as in Example $\langle$4d$\rangle$. Then $\mathcal{A}$ is normal, because for any $(x, y) \in \mathcal{G}$, $\quad (x, y) \cdot \mathcal{A} \;=\; \mathcal{A}_x \;=\; \mathcal{A} \cdot (x, y)$.

If $\mathcal{A}, \mathcal{B} \subset \mathcal{G}$ are subsets, then their **product** is the set $\mathcal{AB} \;=\; \{ab\ ;\ a \in \mathcal{A} \text{ and } b \in \mathcal{B}\}$. In particular, if $(g\mathcal{A})$ and $(h\mathcal{A})$ are two cosets of $\mathcal{A}$, then their product is the set:

$$(g\mathcal{A})(h\mathcal{A}) \quad = \quad \{ga_1ha_2\ ;\ a_1, a_2 \in \mathcal{A}\}. \tag{1.1}$$

Example 9:

(a) Let $\mathcal{G} = (\mathbb{Z}, +)$ and let $\mathcal{A} = 5\mathbb{Z}$, as in Example (4a). Let $g = 1$ and $h = 2$. Then

$$\begin{aligned} 1 + 5\mathbb{Z} &= \{\ldots, -9, -4, 1, 6, 11, 16, \ldots\} \\ \text{and} \quad 2 + 5\mathbb{Z} &= \{\ldots, -8, -3, 2, 7, 12, 17, \ldots\}, \\ \text{so} \quad (1 + 5\mathbb{Z}) + (2 + 5\mathbb{Z}) &= \{\ldots, -7, -2, 3, 8, 13, 18, \ldots\} \quad = \quad (3 + 5\mathbb{Z}). \end{aligned}$$

(b) Let $\mathcal{X}$ and $\mathcal{Y}$ be groups; let $\mathcal{G} = \mathcal{X} \times \mathcal{Y}$, and let $\mathcal{A} \;=\; \{e_{\mathcal{X}}\} \times \mathcal{Y}$, as in Example ⟨4d⟩. Then for any (x_1, y_1) and (x_2, y_2) in $\mathcal{G}$,

$$\begin{aligned} (x_1, y_1) \cdot \mathcal{A} &= \{(x_1, y') \; ; \; y' \in \mathcal{Y}\}, \\ \text{and} \quad (x_2, y_2) \cdot \mathcal{A} &= \{(x_2, y'') \; ; \; y'' \in \mathcal{Y}\}, \\ \text{so that } (x_1, y_1) \cdot \mathcal{A} \cdot (x_2, y_2) \cdot \mathcal{A} &= \{(x_1, y') \cdot (x_2, y'') \; ; \; y' \in \mathcal{Y} \text{ and } \; y'' \in \mathcal{Y}\} \\ &= \left\{ \Big((x_1 x_2), \; (y' y'') \Big) \; ; \; y' \in \mathcal{Y} \text{ and } \; y'' \in \mathcal{Y} \right\} \\ &= \left\{ \Big((x_1 x_2), \; y \Big) \; ; \; y \in \mathcal{Y} \right\} \\ &= \big(x_1 x_2, \; e_{\mathcal{Y}} \big) \cdot \mathcal{A}, \end{aligned}$$

where $e_{\mathcal{Y}}$ is the identity element in $\mathcal{Y}$. Thus, the multiplication of cosets in $\mathcal{G}/\mathcal{A}$ 'mimics' the multiplication of elements in $\mathcal{X}$.

Lemma 10 *Let $\mathcal{G}$ be a group.*

(a) *Subset multiplication in $\mathcal{G}$ is associative. That is, for any subsets $\mathcal{A}, \mathcal{B}, \mathcal{C} \subset \mathcal{G}$,*

$$\mathcal{A} \cdot (\mathcal{B} \cdot \mathcal{C}) \quad = \quad (\mathcal{A} \cdot \mathcal{B}) \cdot \mathcal{C}.$$

(b) *If $\mathcal{A} < \mathcal{G}$ is a subgroup of $\mathcal{G}$, then $\mathcal{A} \cdot \mathcal{A} = \mathcal{A}$.*

Proof: **Exercise 5** □

If $\Phi : \mathcal{G} \longrightarrow \mathcal{H}$ is a group homomorphism, recall that the **kernel** of Φ is the subgroup:

$$\ker(\Phi) \quad = \quad \{g \in \mathcal{G} \; ; \; \Phi(g) \; = \; e_{\mathcal{H}}\}.$$

Proposition 11 *Let $\mathcal{G}$ be a group, and let $\mathcal{A} < \mathcal{G}$ be a subgroup. The following are equivalent:*

(a) *$\mathcal{A} = \ker(\Phi)$ for some group homomorphism $\Phi : \mathcal{G} \longrightarrow \mathcal{H}$ (where $\mathcal{H}$ is some group).*

(b) *For any $g \in \mathcal{G}$, $\quad g\mathcal{A}g^{-1} \;=\; \mathcal{A}$.*

(c) *$\mathcal{A} \lhd \mathcal{G}$.*

(d) *The coset space $\mathcal{G}/\mathcal{A}$ is a* group *under the multiplication operation (1.1). Furthermore:*

1. *For any $g, h \in \mathcal{G}$, $(g\mathcal{A})(h\mathcal{A}) = (gh)\mathcal{A}$.*
2. *Define $\pi : \mathcal{G} \longrightarrow \mathcal{G}/\mathcal{A}$ by: $\pi(g) = g\mathcal{A}$. Then π is a* group epimorphism, *and $\ker(\pi) = \mathcal{A}$.*

Proof: **(a)**$\Longrightarrow$**(b)** Suppose $\mathcal{A} = \ker(\Phi)$, and let $g \in \mathcal{G}$. Then $g\mathcal{A}g^{-1} = \{gag^{-1} \; ; \; a \in \mathcal{A}\}$, and we want to show this set is just $\mathcal{A}$.

Claim 1: $g\mathcal{A}g^{-1} \subseteq \mathcal{A}$.

Proof: Let gag^{-1} be some element of $g\mathcal{A}g^{-1}$. Then $\Phi\left(gag^{-1}\right) = \Phi(g) \cdot \Phi(a) \cdot \Phi(g^{-1}) = \Phi(g) \cdot e_{\mathcal{H}} \cdot \Phi(g)^{-1} = \Phi(g) \cdot \Phi(g)^{-1} = e_{\mathcal{H}}$. Thus, $gag^{-1} \in \mathcal{A}$. .... □ [Claim 1]

Claim 2: $\mathcal{A} \subseteq g\mathcal{A}g^{-1}$.

Proof: It follows from Claim 1 that $g^{-1}\mathcal{A}g \subseteq \mathcal{A}$ (just use g^{-1} in place of g). Thus, $g(g^{-1}\mathcal{A}g)g^{-1} \subseteq g\mathcal{A}g^{-1}$. But $g(g^{-1}\mathcal{A}g)g^{-1} = (gg^{-1})\mathcal{A}(gg^{-1}) = e_{\mathcal{G}}\mathcal{A}e_{\mathcal{G}} = \mathcal{A}$. In other words, $\mathcal{A} \subseteq g\mathcal{A}g^{-1}$. .... □ [Claim 2]

It follows from Claims 1 and 2 that $g\mathcal{A}g^{-1} = \mathcal{A}$.

(b)$\Longrightarrow$**(c)** Let $g \in \mathcal{G}$. Then $g\mathcal{A} = g\mathcal{A}(gg^{-1}) = (g\mathcal{A}g^{-1})g \underset{\mathbf{(b)}}{=\!=} \mathcal{A}g$. This holds for any $g \in \mathcal{G}$, so $\mathcal{A} \lhd \mathcal{G}$.

(c)$\Longrightarrow$**(d)** We will prove this in several steps...

Claim 3: $(g\mathcal{A})(h\mathcal{A}) = (gh)\mathcal{A}$.

Proof: $(g\mathcal{A})(h\mathcal{A}) \underset{\text{Lem.10(a)}}{=\!=\!=} g(\mathcal{A}h)\mathcal{A} \underset{\mathcal{A}\lhd\mathcal{G}}{=\!=} g(h\mathcal{A})\mathcal{A} \underset{\text{Lem.10(a)}}{=\!=\!=} (gh)\mathcal{A}\cdot\mathcal{A} \underset{\text{Lem.10(b)}}{=\!=\!=} (gh)\mathcal{A}$. □ [Claim 3]

Claim 4: *$\mathcal{G}/\mathcal{A}$ is a group.*

Proof: It follows from Claim 3 that $\mathcal{G}/\mathcal{A}$ is closed under multiplication. By Lemma 10(a), this multiplication is associative. By applying Claim 3, it is easy to verify:

1. The identity element of $\mathcal{G}/\mathcal{A}$ is just the coset $e_{\mathcal{G}}\mathcal{A} = \mathcal{A}$.
2. For any $g \in \mathcal{G}$, the *inverse* of the coset $g\mathcal{A}$ is just the coset $g^{-1}\mathcal{A}$.

Thus, $\mathcal{G}/\mathcal{A}$ satisfies all the properties of a group .... □ [Claim 4]

Claim 5: *$\pi : \mathcal{G} \longrightarrow \mathcal{G}/\mathcal{A}$ is a homomorphism.*

Proof: $\pi(g \cdot h) = (gh) \cdot \mathcal{A} \underset{\text{Clm.3}}{=\!=} (g\mathcal{A}) \cdot (h\mathcal{A}) = \pi(g) \cdot \pi(h)$. .. □ [Claim 5]

Also, π is surjective: any element of $\mathcal{G}/\mathcal{A}$ is a coset of the form $g\mathcal{A}$ for some $g \in \mathcal{G}$, and thus, $g\mathcal{A} = \pi(g)$. Thus, π is an epimorphism.

Claim 6: $\ker(\pi) = \mathcal{A}$.

Proof: Let $g \in \mathcal{G}$. Then $\Big(g \in \ker(\pi) \Big) \iff \Big(\pi(g) = e_{\mathcal{G}/\mathcal{A}} \Big) \iff \Big(g\mathcal{A} \;=\; \mathcal{A} \Big)$ $\Leftarrow_{\text{Lem.2(a)}}\Rightarrow \Big(g \in \mathcal{A} \Big)$. . . . . □ `[Claim 6]`

(d)$\Longrightarrow$**(a)** This follows immediately: let $\mathcal{H} = \mathcal{G}/\mathcal{A}$ and let $\Phi = \pi$. □

The group $\mathcal{G}/\mathcal{A}$ is called the **quotient group**, and the epimorphism $\pi : \mathcal{G} \longrightarrow \mathcal{G}/\mathcal{A}$ is called the **projection map** or **quotient map**.

1.4 The Fundamental Isomorphism Theorem

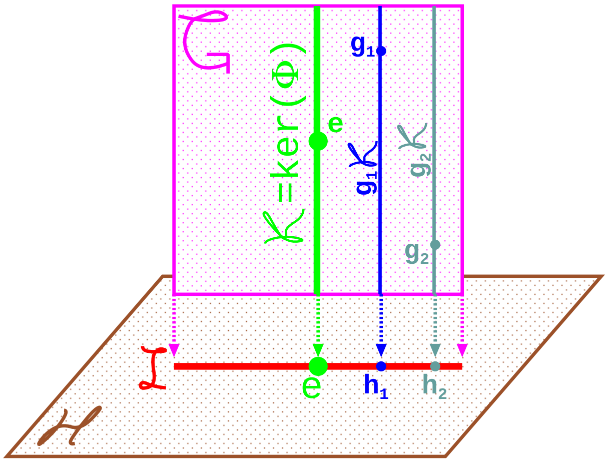
Prerequisites: §1.3

Theorem 12 Fundamental Isomorphism Theorem

Let $\mathcal{G}$ and $\mathcal{H}$ be groups, and let $\phi : \mathcal{G} \longrightarrow \mathcal{H}$ be a group homomorphism, with image $\mathcal{I} = \phi(\mathcal{G}) \subset \mathcal{H}$, and kernel $\mathcal{K}$. Then:

(a) $\mathcal{I} \;\cong\; \mathcal{G}/\mathcal{K}$.

(b) *For any $g \in \mathcal{G}$ with $h = \phi(g)$, the ϕ-preimage of h is the g-coset of $\mathcal{K}$. That is:* $\phi^{-1}\{h\} \;=\; g\mathcal{K}$. □



Example 13:

(a) Let $\mathcal{G} = \mathbb{Z}$ and $\mathcal{H} = \mathbb{Z}_{/3}$, and let $\Phi : \mathbb{Z} \ni n \mapsto [n]_3 \in \mathbb{Z}_{/3}$ (see Figure 1.1a). Then $\mathcal{K} = 3\mathbb{Z}$, so $\mathcal{I} = \mathbb{Z}_{/3} \;=\; \mathbb{Z}/(3\mathbb{Z}) \;=\; \mathcal{G}/\mathcal{K}$.

Observe: $\phi(5) \;=\; [5]_3 \;=\; [2]_3$, and thus, $\phi^{-1}\{[2]_3\} \;=\; 5 + 3\mathbb{Z} \;=\; \{\ldots, -1, 2, 5, 8, 11, \ldots\}$.

(b) Let $\mathcal{A}$ and $\mathcal{B}$ be groups, and let $\mathcal{G} = \mathcal{A} \times \mathcal{B}$ (see Figure 1.1b). Let $\Phi : \mathcal{G} \longrightarrow \mathcal{A}$ be the projection map; that is, $\Phi(a,b) = a$. Then

$$\ker(\Phi) \quad = \quad \{(e_{\mathcal{A}}, b)\ ;\ b \in \mathcal{B}\} \quad = \quad \{e_{\mathcal{A}}\} \times \mathcal{B}$$

is a subgroup isomorphic to $\mathcal{B}$. Thus, the Fundamental Isomorphism Theorem says:

$$\frac{\mathcal{A} \times \mathcal{B}}{\{e_{\mathcal{A}}\} \times \mathcal{B}} \quad = \quad \frac{\mathcal{G}}{\ker(\phi)} \quad \cong \quad \mathcal{A}.$$

For any fixed $(a,b) \in \mathcal{G}$, observe that $\phi(a,b) = a$. Thus,

$\phi^{-1}\{a\} \;=\; (a,b) \cdot \{(e_{\mathcal{A}}, b')\ ;\ b' \in \mathcal{B}\} \;=\; \{(a, b')\ ;\ b' \in \mathcal{B}\}$.

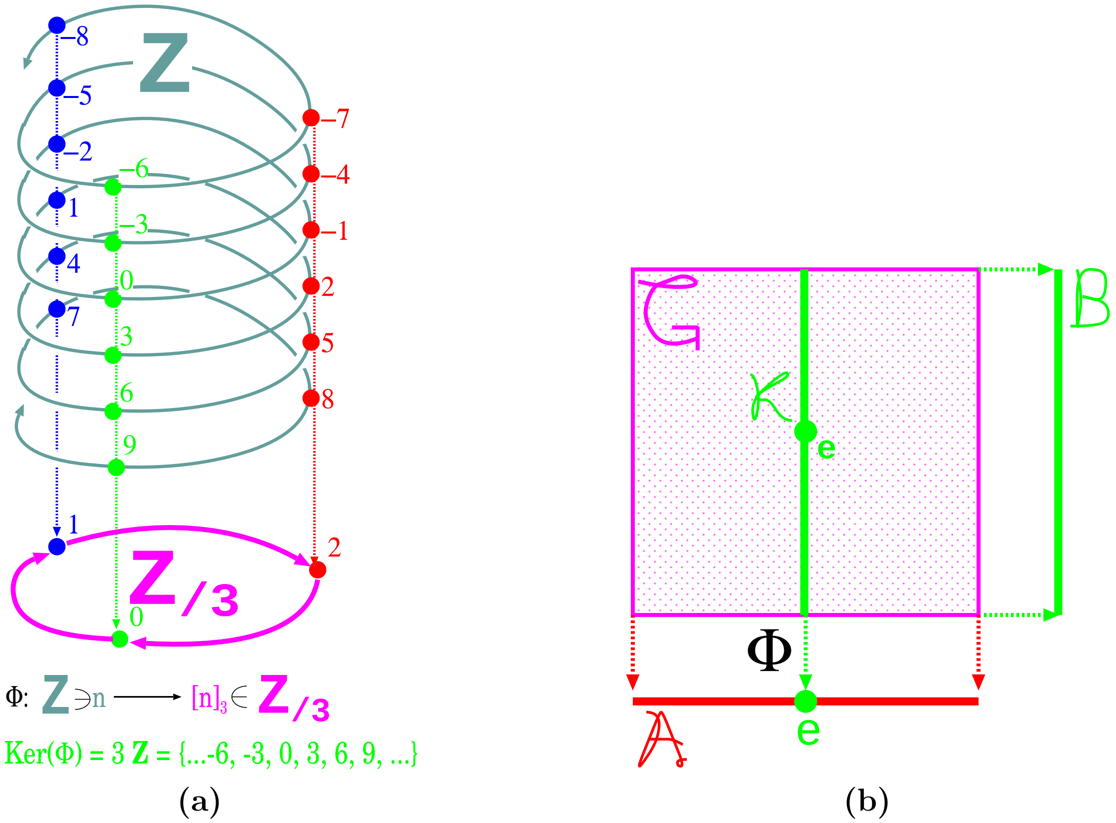


Figure 1.1: Examples (13a) and (13b)

(c) Let $\mathcal{G} = \mathbf{S}_3$, and let $\mathcal{A} = \mathbf{A}_3$, as in Example ⟨8d⟩. Then $\mathbf{A}_3$ is normal in $\mathbf{S}_3$, and $\mathbf{S}_3/\mathbf{A}_3 = \{\mathbf{A}_3,\ (12)\mathbf{A}_3\}$ is a two-element group, isomorphic to $(\mathbb{Z}_{/2}, +)$, via the map $\phi : \mathbb{Z}_{/2} \longrightarrow \mathbf{S}_3/\mathbf{A}_3$ defined: $\phi(\bar{0}) = \mathbf{A}_3$ and $\phi(\bar{1}) = (12)\mathbf{A}_3$.

Proof of Theorem 129: **(a)** We will build an explicit isomorphism between $\mathcal{I}$ and $\mathcal{G}/\mathcal{K}$.

Claim 1: *For any $g_1, g_2 \in \mathcal{G}$,* $\Big(g_1\mathcal{K} = g_2\mathcal{K} \Big) \iff \Big(\phi(g_1) = \phi(g_2) \Big)$.

Proof:
$$\Big(g_1\mathcal{K} = g_2\mathcal{K} \Big) \xLeftrightarrow{\text{Lem.2(b)}} \Big(g_2^{-1}g_1 \in \mathcal{K} \Big) \iff \Big(\phi(g_2^{-1}g_1) = e_{\mathcal{H}} \Big)$$
$$\iff \Big(\phi(g_2)^{-1} \cdot \phi(g_1) = e_{\mathcal{H}} \Big) \iff \Big(\phi(g_1) = \phi(g_2) \Big).$$
.... □ [Claim 1]

Define the map $\Psi : \mathcal{G}/\mathcal{K} \longrightarrow \mathcal{I}$ by: $\Psi(g\mathcal{K}) = \phi(g)$.

Claim 2: *Ψ is well-defined and injective.*

Proof: For any $g_1, g_2 \in \mathcal{G}$, $\Big(g_1\mathcal{K} = g_2\mathcal{K} \Big) \xLeftrightarrow{\text{Clm.1}} \Big(\phi(g_1) = \phi(g_2) \Big) \iff \Big(\Psi(g_1\mathcal{K}) = \Psi(g_2\mathcal{K}) \Big)$. .... □ [Claim 2]

Claim 3: *Ψ is surjective.*

Proof: By definition, any $i \in \mathcal{I}$ is the image of some $g \in \mathcal{G}$ —ie. $\phi(g) = i$. Thus, $\Psi(g\mathcal{K}) = i$. . . . . □ [Claim 3]

Claim 4: *Ψ is a homomorphism.*

Proof: For any cosets $g_1\mathcal{K}$ and $g_2\mathcal{K}$ in $\mathcal{G}/\mathcal{K}$,

$$\Psi\Big((g_1\mathcal{K})(g_2\mathcal{K})\Big) \underset{(*)}{=} \Psi\Big((g_1g_2)\mathcal{K}\Big) = \phi(g_1g_2) = \phi(g_1)\cdot\phi(g_2) = \Psi(g_1\mathcal{K})\cdot\Psi(g_2\mathcal{K}),$$

where $(*)$ follows from Proposition 11(d) part 1 . . . . □ [Claim 4]

(b) Let $g' \in \mathcal{G}$. Then: $\Big(g' \in g\mathcal{K} \Big) \Leftarrow_{\text{Lem2(b)}}\Rightarrow \Big(g'\mathcal{K} = g\mathcal{K} \Big) \Leftarrow_{\text{Clm.1}}\Rightarrow \Big(\phi(g') = \phi(g) = h \Big) \iff \Big(g' \in \phi^{-1}\{h\} \Big)$. □

Corollary 14 *Let $\phi : \mathcal{G} \longrightarrow \mathcal{H}$ be a group homomorphism. Then*

$$\Big(\phi \text{ is injective} \Big) \iff \Big(\ker \phi = \{e_{\mathcal{G}}\} \Big).$$

Proof: <u>**Exercise 6**</u> □

Chapter 2

The Isomorphism Theorems

2.1 The Diamond Isomorphism Theorem

Prerequisites: §1.4

Let $\mathcal{G}$ be a group and $\mathcal{B} < \mathcal{G}$ a subgroup. The **normalizer** of $\mathcal{B}$ in $\mathcal{G}$ is the subgroup

$$\mathcal{N}_{\mathcal{G}}^{rmlzr}(\mathcal{B}) \quad = \quad \{g \in \mathcal{G}\ ;\ g\mathcal{B} = \mathcal{B}g\}.$$

Thus, $\mathcal{N}_{\mathcal{G}}^{rmlzr}(\mathcal{B})$ is 'the set of all elements of $\mathcal{G}$ who think that $\mathcal{B}$ is normal'.

Lemma 15 *Let $\mathcal{G}$ be a group with subgroup $\mathcal{B} < \mathcal{G}$. Then:*

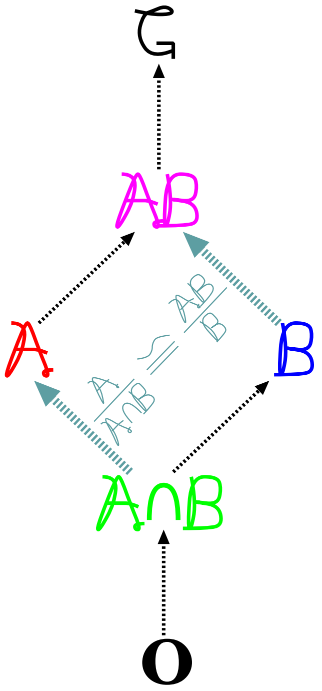
(a) $\mathcal{B} < \mathcal{N}_{\mathcal{G}}^{rmlzr}(\mathcal{B}) < \mathcal{G}$.

(b) $\Big(\mathcal{B} \lhd \mathcal{G} \Big) \iff \Big(\mathcal{N}_{\mathcal{G}}^{rmlzr}(\mathcal{B}) = \mathcal{G} \Big)$.

Proof: Exercise 7 □

Thus, the condition "$\mathcal{A} \subset \mathcal{N}_{\mathcal{G}}^{rmlzr}(\mathcal{B})$" is true automatically if $\mathcal{B}$ is a normal subgroup of $\mathcal{G}$.

Recall that $\mathcal{A} \cdot \mathcal{B} = \{a \cdot b\ ;\ a \in \mathcal{B},\ b \in \mathcal{A}\}$. In general, $\mathcal{AB}$ is not a subgroup of $\mathcal{G}$. But $\mathcal{AB}$ *is* a subgroup under the conditions of the following theorem...



Theorem 16 Diamond Isomorphism Theorem

Let $\mathcal{G}$ be a group, with subgroups $\mathcal{A}$ and $\mathcal{B}$. Suppose that $\mathcal{A} \subset \mathcal{N}^{rmlzr}_{\mathcal{G}}(\mathcal{B})$. Then:

(a) *$\mathcal{A} \cdot \mathcal{B}$ is a subgroup of $\mathcal{G}$.*

(b) *$\mathcal{B} \lhd (\mathcal{A} \cdot \mathcal{B})$.*

(c) *$(\mathcal{A} \cap \mathcal{B}) \lhd \mathcal{A}$.*

(d) *There is an isomorphism:* $\dfrac{\mathcal{A} \cdot \mathcal{B}}{\mathcal{B}} \cong \dfrac{\mathcal{A}}{\mathcal{A} \cap \mathcal{B}}$ *given by the map*

$$\begin{array}{rcl} \Phi : \dfrac{\mathcal{A} \cdot \mathcal{B}}{\mathcal{B}} & \longrightarrow & \dfrac{\mathcal{A}}{\mathcal{A} \cap \mathcal{B}} \\ (ab)\mathcal{B} & \mapsto & a \cdot (\mathcal{A} \cap \mathcal{B}) \end{array}$$

Proof: **(a)** Let $a_1 b_1$ and $a_2 b_2$ be elements of $\mathcal{AB}$. We must show that $a_1 b_1 (a_2 b_2)^{-1}$ is also in $\mathcal{AB}$. But

$$a_1 b_1 (a_2 b_2)^{-1} \quad = \quad a_1 b_1 b_2^{-1} a_2^{-1} \quad = \quad a_1 (b_1 b_2^{-1}) a_2^{-1} \underset{(*)}{=} a_1 a_2^{-1} b_3 \quad = \quad (a_1 a_2^{-1}) b_3 \quad \in \mathcal{AB}.$$

To see $(*)$, recall that $\mathcal{A} \subset \mathcal{N}^{rmlzr}_{\mathcal{G}}(\mathcal{B})$, so $\mathcal{B} a_2^{-1} = a_2^{-1} \mathcal{B}$. Thus $(b_1 b_2^{-1}) a_2^{-1} = a_2^{-1} b_3$ for some $b_3 \in \mathcal{B}$.

(b) and **(c)** **Exercise 8** .

(d) [In progress...]

Claim 1: *For any $a \in \mathcal{A}$ and $b \in \mathcal{B}$,* $(ab)\mathcal{B} = a\mathcal{B}$

□

Example 17:

(a) Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be groups, and let $\mathcal{G} = \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$, as shown in Figure 2.1. Define:

$$\begin{array}{rcccccccl} \mathcal{A} & = & \mathcal{X} & \times & \{e_y\} & \times & \mathcal{Z} & = & \big\{(x, e_y, z)\ ;\ x \in \mathcal{X},\ z \in \mathcal{Z}\big\}, \\ \text{and } \mathcal{B} & = & \{e_x\} & \times & \mathcal{Y} & \times & \mathcal{Z} & = & \big\{(e_x, y, z)\ ;\ y \in \mathcal{Y},\ z \in \mathcal{Z}\big\}. \end{array}$$

Then:

$$\begin{array}{rcccccccl} \mathcal{AB} & = & \mathcal{X} & \times & \mathcal{Y} & \times & \mathcal{Z} & = & \mathcal{G}, \\ \text{and } \mathcal{A} \cap \mathcal{B} & = & \{e_x\} & \times & \{e_y\} & \times & \mathcal{Z} & = & \big\{(e_x, e_y, z)\ ;\ z \in \mathcal{Z}\big\}. \end{array}$$

Thus, $\dfrac{\mathcal{AB}}{\mathcal{B}} = \dfrac{\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}}{\{e_x\} \times \mathcal{Y} \times \mathcal{Z}} \cong \mathcal{X} \cong \dfrac{\mathcal{X} \times \{e_y\} \times \mathcal{Z}}{\{e_x\} \times \{e_y\} \times \mathcal{Z}} = \dfrac{\mathcal{A}}{\mathcal{A} \cap \mathcal{B}}$. —

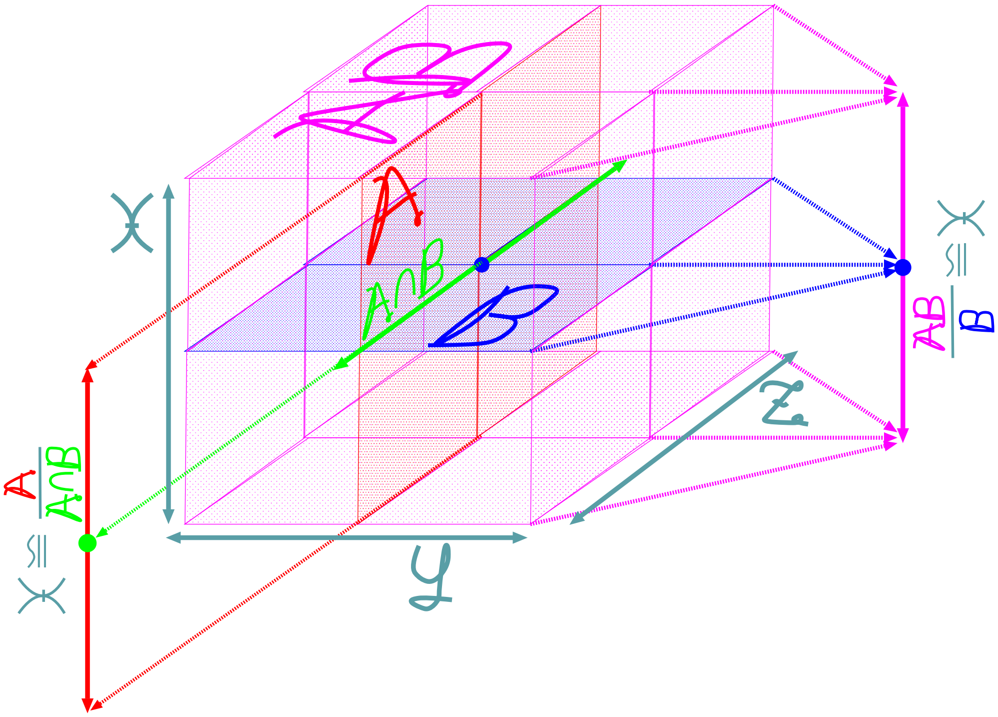


Figure 2.1: Example ⟨17a⟩.

2.2 The Chain Isomorphism Theorem

Prerequisites: §1.4

Let $\mathcal{G}$ be a group, with normal subgroup $\mathcal{A} \triangleleft \mathcal{G}$. Suppose $\mathcal{A} < \mathcal{B} < \mathcal{G}$. Then $\mathcal{A}$ is also a normal subgroup of $\mathcal{B}$, and the quotient group

$$\frac{\mathcal{B}}{\mathcal{A}} \quad = \quad \{b\mathcal{A}\ ;\ b \in \mathcal{B}\}$$

is a subset of the quotient group $\dfrac{\mathcal{G}}{\mathcal{A}} \ = \ \{g\mathcal{A}\ ;\ g \in \mathcal{G}\}$.

Theorem 18 Chain Isomorphism Theorem

Let $\mathcal{G}$ be a group, with normal subgroups $\mathcal{A} \triangleleft \mathcal{G}$ and $\mathcal{B} \triangleleft \mathcal{G}$. Suppose $\mathcal{A} < \mathcal{B}$. Then:

(a) $\dfrac{\mathcal{B}}{\mathcal{A}}$ *is a normal subgroup of* $\dfrac{\mathcal{G}}{\mathcal{A}}$.

(b) *There is an isomorphism* $\dfrac{(\mathcal{G}/\mathcal{A})}{(\mathcal{B}/\mathcal{A})} \quad \cong \quad \dfrac{\mathcal{G}}{\mathcal{B}}$.

(c) *Use 'bar' notation to denote elements of $\mathcal{G}/\mathcal{A}$. Thus, $g\mathcal{A} = \overline{g}$, $\mathcal{B}/\mathcal{A} = \overline{\mathcal{B}}$, $\mathcal{G}/\mathcal{A} = \overline{\mathcal{G}}$, and* $\dfrac{\mathcal{G}/\mathcal{A}}{\mathcal{B}/\mathcal{A}} \ = \ \overline{\mathcal{G}}/\overline{\mathcal{B}}$. *Then the isomorphism $\Phi : \overline{\mathcal{G}}/\overline{\mathcal{B}} \longrightarrow \mathcal{G}/\mathcal{B}$ is defined: $\Phi\left(\overline{g}\overline{\mathcal{B}}\right) \ = \ g\mathcal{B}$.*

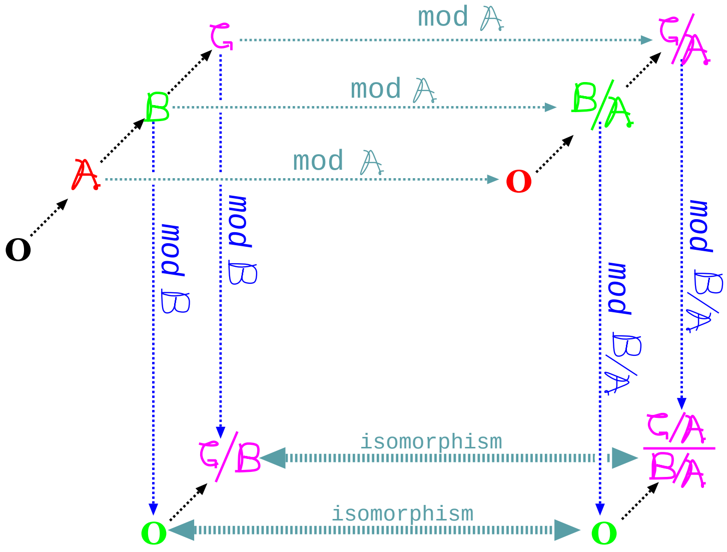


Figure 2.2: The Chain Isomorphism Theorem

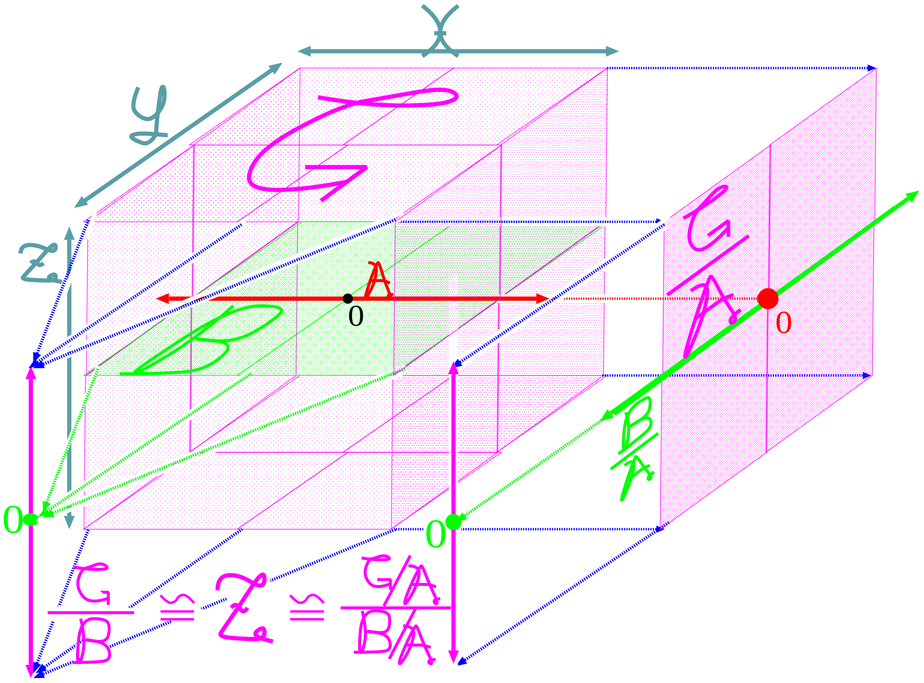


Figure 2.3: Example ⟨19a⟩.

Example 19:

(a) Let $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ be groups, and let $\mathcal{G} = \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. Define:

$$\begin{array}{rcccccccl} \mathcal{A} & = & \mathcal{X} & \times & \{e_{\mathcal{Y}}\} & \times & \{e_{\mathcal{Z}}\} & = & \big\{(x, e_{\mathcal{Y}}, e_{\mathcal{Z}}) \; ; \; x \in \mathcal{X}\big\}, \\ \text{and } \mathcal{B} & = & \mathcal{X} & \times & \mathcal{Y} & \times & \{e_{\mathcal{Z}}\} & = & \{(x, y, e_{\mathcal{Z}}) \; ; \; x \in \mathcal{Z}, \; y \in \mathcal{Y}\}. \end{array}$$

Then:

$$\frac{\mathcal{G}}{\mathcal{A}} \quad = \quad \frac{\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}}{\mathcal{X} \times \{e_{\mathcal{Y}}\} \times \{e_{\mathcal{Z}}\}} \quad \cong \quad \mathcal{Y} \quad \times \quad \mathcal{Z};$$

$$\frac{\mathcal{B}}{\mathcal{A}} \quad = \quad \frac{\mathcal{X} \times \mathcal{Y} \times \{e_{\mathcal{Z}}\}}{\mathcal{X} \times \{e_{\mathcal{Y}}\} \times \{e_{\mathcal{Z}}\}} \quad \cong \quad \mathcal{Y} \quad \times \quad \{e_{\mathcal{Z}}\};$$

$$\text{and } \frac{\mathcal{G}}{\mathcal{B}} \quad = \quad \frac{\mathcal{X} \times \mathcal{Y} \times \mathcal{Z}}{\mathcal{X} \times \mathcal{Y} \times \{e_{\mathcal{Z}}\}} \quad \cong \quad \mathcal{Z}.$$

Thus, $\dfrac{(\mathcal{G}/\mathcal{A})}{(\mathcal{B}/\mathcal{A})} \quad \cong \quad \dfrac{\mathcal{Y} \; \times \; \mathcal{Z}}{\mathcal{Y} \times \{e_{\mathcal{Z}}\}} \quad \cong \quad \mathcal{Z} \quad \cong \quad \dfrac{\mathcal{G}}{\mathcal{B}}.$

2.3 The Lattice Isomorphism Theorem

Prerequisites: §2.2

If $\mathcal{G}$ is a group, recall that the *subgroup lattice* of $\mathcal{G}$ is a directed graph $\mathfrak{L}(\mathcal{G})$, whose vertices are the subgroups of $\mathcal{G}$. We draw a (directed) edge from subgroup $\mathcal{A}$ to subgroup $\mathcal{B}$ if $\mathcal{A} < \mathcal{B}$ and there is no $\mathcal{C}$ such that $\mathcal{A} < \mathcal{C} < \mathcal{B}$. The graph is drawn on paper so that $\mathcal{A}$ appears *below* $\mathcal{B}$ on the page if and only if $\mathcal{A} < \mathcal{B}$.

This graph is called a *lattice* because it has two special properties:

1. For any subgroups $\mathcal{A}, \mathcal{B} < \mathcal{G}$, there is a *minimal* subgroup of $\mathcal{G}$ which contains both $\mathcal{A}$ and $\mathcal{B}$ —namely, the **join** of $\mathcal{A}$ and $\mathcal{B}$:

$$\langle \mathcal{A}, \mathcal{B} \rangle \quad = \quad \{a_1 b_1 a_2 b_2 \ldots a_n b_n \; ; \; n \in \mathbb{N}, \; a_1, a_2, \ldots, a_n \in \mathcal{A}, \; \text{and} \; b_1, b_2, \ldots, b_n \in \mathcal{B}\}.$$

2. For any subgroups $\mathcal{A}, \mathcal{B} < \mathcal{G}$, there is a *maximal* subgroup of $\mathcal{G}$ which is *contained in* both $\mathcal{A}$ and $\mathcal{B}$ —namely, their *intersection* $\mathcal{A} \cap \mathcal{B}$.

Let $\mathcal{N} \lhd \mathcal{G}$. We will adopt the 'bar' notation for objects in the quotient group $\mathcal{G}/\mathcal{N}$. Thus, $\overline{\mathcal{G}} = \mathcal{G}/\mathcal{N}$. If $g \in \mathcal{G}$, then $\overline{g} = g\mathcal{N} \in \overline{\mathcal{G}}$. If $\mathcal{N} < \mathcal{A} < \mathcal{G}$, then $\overline{\mathcal{A}} = \mathcal{A}/\mathcal{N} = \{a\mathcal{N} \; ; \; a \in \mathcal{A}\} \subset \overline{\mathcal{G}}$.

Theorem 20 Lattice Isomorphism Theorem

Let $\mathcal{G}$ be a group and let $\mathcal{N} \lhd \mathcal{G}$ be a normal subgroup. Let $\overline{\mathcal{G}} = \mathcal{G}/\mathcal{N}$. Let $\mathfrak{L}(\overline{\mathcal{G}})$ be the subgroup lattice of $\overline{\mathcal{G}}$, and let $\mathfrak{L}_{\mathcal{N}}(\mathcal{G})$ be the 'fragment' of $\mathfrak{L}(\mathcal{G})$ consisting of all subgroups which contain $\mathcal{N}$. That is:

$$\mathfrak{L}_{\mathcal{N}}(\mathcal{G}) \quad = \quad \{\mathcal{A} < \mathcal{G} \; ; \; \mathcal{N} < \mathcal{A}\}.$$

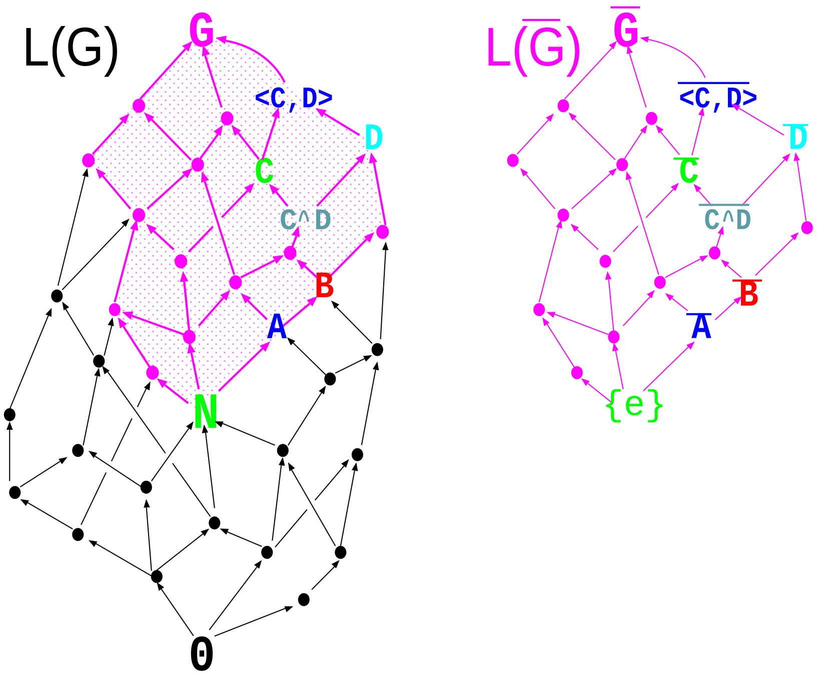


Figure 2.4: The Lattice Isomorphism Theorem

Then there is an order-preserving bijection from $\mathfrak{L}_{\mathcal{N}}(\mathcal{G})$ into $\mathfrak{L}(\overline{\mathcal{G}})$, given:

$$\mathfrak{L}_{\mathcal{N}}(\mathcal{G}) \ \ni \mathcal{A} \mapsto \overline{\mathcal{A}} \in \ \mathfrak{L}(\overline{\mathcal{G}}).$$

Furthermore, for any $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D} \in \mathfrak{L}_{\mathcal{N}}(\mathcal{G})$,

(a) $\Big(\mathcal{A} < \mathcal{B} \Big) \iff \Big(\overline{\mathcal{A}} < \overline{\mathcal{B}} \Big)$, *and in this case,* $|\mathcal{B} : \mathcal{A}| \ = \ |\overline{\mathcal{B}} : \overline{\mathcal{A}}|$.

(b) $\overline{\langle \mathcal{C}, \mathcal{D} \rangle} \ = \ \Big\langle \overline{\mathcal{C}}, \overline{\mathcal{D}} \Big\rangle$.

(c) $\overline{\mathcal{C} \cap \mathcal{D}} \ = \ \overline{\mathcal{C}} \cap \overline{\mathcal{D}}$.

(d) $\Big(\mathcal{A} \lhd \mathcal{G} \Big) \iff \Big(\overline{\mathcal{A}} \lhd \overline{\mathcal{G}} \Big)$, *and in this case,* $\mathcal{G}/\mathcal{A} \ \cong \ \overline{\mathcal{G}}/\overline{\mathcal{A}}$.

Proof: **(d)** just restates the Chain Isomorphism Theorem. The proofs of **(a,b,c)** are **Exercise 9** . □

Chapter 3

Free Abelian Groups

3.1 Rank and Linear Independence

Let $(\mathcal{A}, +)$ be an (additive) abelian group and let $\mathbf{a}_1, \ldots, \mathbf{a}_R \in \mathcal{A}$. We say that $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_R$ are **$\mathbb{Z}$-linearly independent** if there exist no $z_1, z_2, \ldots, z_R \in \mathbb{Z}$ (not all zero) such that $z_1\mathbf{a}_1 + z_2\mathbf{a}_2 + \ldots + z_R\mathbf{a}_R \;=\; 0$. This is a natural generalization of the notion of linear independence for vector spaces. The **rank** of $\mathcal{A}$ is the maximal cardinality of any linearly independent subset. In other words:

$$\Big(\mathsf{rank}\,(\mathcal{A}) = R \Big) \iff \left(\begin{array}{l} \text{1. There exists a linearly independent set } \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_R\} \subset \mathcal{A}. \\ \text{2. For any } \mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_R, \mathbf{b}_{R+1} \in \mathcal{A}, \text{ there exist } z_1, z_2, \ldots, z_R, z_{R+1} \in \mathbb{Z} \\ \quad \text{(not all zero) so that } z_1\mathbf{b}_1 + z_2\mathbf{b}_2 + \ldots + z_R\mathbf{b}_R + z_{R+1}\mathbf{b}_{R+1} \;=\; 0. \end{array} \right)$$

Think of rank as a notion of 'dimension' for abelian groups.

Example 21:

(a) $\mathsf{rank}\,(\mathbb{Z}, +) = 1$. First, we'll show that $\mathsf{rank}\,(\mathbb{Z}) \geq 1$. To see this, let $a \in \mathbb{Z}$ be nonzero. Then $z \cdot a \neq 0$ for all $z \in \mathbb{Z}$, so the set $\{a\}$ is $\mathbb{Z}$-linearly independent.

Next, we'll show that $\mathsf{rank}\,(\mathbb{Z}) \leq 1$. Suppose $a_1, a_2 \in \mathbb{Q}$ were nonzero. Let $z_1 = a_2$ and $z_2 = -a_1$. Then $z_1a_1 + z_2a_2 \;=\; a_2a_1 - a_1a_2 \;=\; 0$. So any set $\{a_1, a_2\}$ with two elements is linearly *dependent*.

(b) $\mathsf{rank}\,(\mathbb{Q}, +) = 1$.

First, we'll show that $\mathsf{rank}\,(\mathbb{Q}) \geq 1$. To see this, let $q \in \mathbb{Q}$ be nonzero. Then $z \cdot q \neq 0$ for all $z \in \mathbb{Z}$, so the set $\{q\}$ is $\mathbb{Z}$-linearly independent.

Next, we'll show that $\mathsf{rank}\,(\mathbb{Q}) \leq 1$. Suppose $q_1, q_2 \in \mathbb{Q}$ were nonzero; let $q_1 = \frac{r_1}{s_1}$ and $q_2 = \frac{r_2}{s_2}$. Let $z_1 = r_2s_1$ and $z_2 = -r_1s_2$. Then

$$z_1q_1 + z_2q_2 \quad = \quad r_2s_1\frac{r_1}{s_1} - r_1s_2\frac{r_2}{s_2} \quad = \quad r_2r_1 - r_1r_2 \quad = \quad 0.$$

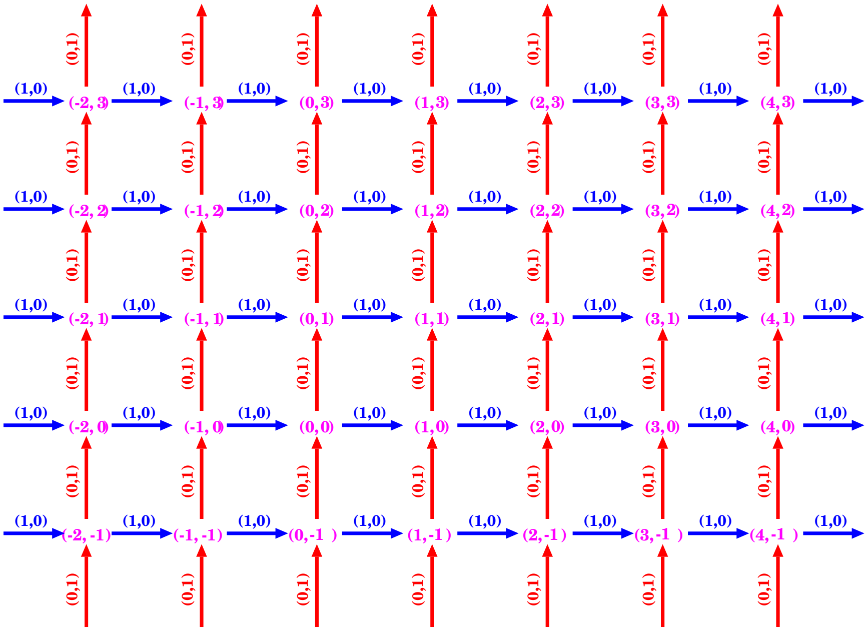


Figure 3.1: $\mathbb{Z}^2 = \mathbb{Z} \oplus \mathbb{Z}$ is the free group of rank 2.

(c) $\mathsf{rank}\left(\mathbb{Z}_{/n}\right) = 0$.

To see this, note that $n \cdot a = 0$ for any $a \in \mathbb{Z}_{/n}$. Thus, even a set like $\{a\}$, containing only *one* element, is not $\mathbb{Z}$-linearly independent.

(d) If $\mathcal{A}$ is any finite abelian group, then $\mathsf{rank}\,(\mathcal{A}) = 0$.

To see this, suppose $|\mathcal{A}| = n$. Then $n \cdot \mathbf{a} = 0$ for any $\mathbf{a} \in \mathcal{A}$. Thus, even a set like $\{\mathbf{a}\}$, containing only *one* element, is not $\mathbb{Z}$-linearly independent.

(e) $\mathsf{rank}\,(\mathbb{R}, +) = \infty$.

We need to construct an infinite, $\mathbb{Z}$-linearly independent subset $\{r_1, r_2, r_3, \ldots\} \subset \mathbb{R}$. This is **Exercise 10** .

3.2 Free Groups and Generators

Prerequisites: §3.1, §4.1

Let $R \in \mathbb{N}$. The **free abelian group** of **rank** R is the group

$$\mathbb{Z}^R \quad = \quad \{(z_1, z_2, \ldots, z_R)\ ;\ z_1, z_2, \ldots, z_R \in \mathbb{Z}\}, \qquad \text{(with componentwise addition)}$$

More generally, any abelian group is called **free** if it is isomorphic to $\mathbb{Z}^R$.

Lemma 22 *$\mathsf{rank}\left(\mathbb{Z}^R\right) = R$. Furthermore, if $\mathcal{B} < \mathbb{Z}^R$ is any subgroup, then $\mathsf{rank}\,(\mathcal{B}) \leq R$.*

Proof: I claim $\mathsf{rank}\left(\mathbb{Z}^R\right) \geq R$. To see this, let $\mathbf{a}_1 = (1,0,0,\ldots,0)$, $\mathbf{a}_2 = (0,1,0,\ldots,0), \ldots, \mathbf{a}_R = (0,0,\ldots,0,1)$. Then $\mathbf{a}_1,\ldots,\mathbf{a}_R$ are $\mathbb{Z}$-linearly independent (**Exercise 11**).

Claim 1: *If $\mathcal{B} < \mathbb{Z}^R$, then $\mathsf{rank}(\mathcal{B}) \leq R$.*

Proof: Suppose $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_R, \mathbf{b}_{R+1} \in \mathcal{B} \subset \mathbb{Z}^R$. Think of $\mathbb{Z}^R$ as a subset of $\mathbb{R}^R$. Then any set of $R+1$ vectors cannot be $\mathbb{R}$-linearly independent, so there are numbers $q_1,\ldots,q_{R+1}$ such that $q_1\mathbf{b}_1+\ldots+q_{R+1}\mathbf{b}_{R+1} = 0$. Since $\mathbf{b}_1,\ldots,\mathbf{b}_{R+1}$ have integer coefficients, we can assume that $q_1,\ldots,q_{R+1}$ are rational numbers; say $q_r = \frac{p_r}{m_r}$ for all r. Now, let $M = m_1 m_2 \ldots m_R$, and let $z_r = M \cdot q_r = m_1 m_2 \ldots m_{r-1}\, p_r\, m_{r+1} \ldots m_R$. Then $z_1,\ldots,z_{R+1}$ are integers, and $z_1\mathbf{b}_1 + \ldots + z_{R+1}\mathbf{b}_{R+1} = M \cdot (q_1\mathbf{b}_1 + \ldots + q_{R+1}\mathbf{b}_{R+1}) = M \cdot 0 = 0.$. □ `[Claim 1]`

Setting $\mathcal{B} = \mathbb{Z}^R$ in Claim 1 tells us that $\mathsf{rank}\left(\mathbb{Z}^R\right) \leq R$. Thus, $\mathsf{rank}\left(\mathbb{Z}^R\right) = R$. □

Lemma 23 *Let $\mathcal{A}$ be any abelian group. Then $\mathsf{rank}\left(\mathbb{Z}^R \oplus \mathcal{A}\right) = R + \mathsf{rank}(\mathcal{A})$.*

Proof: **Exercise 12** □

Example 24: $\mathsf{rank}\left(\mathbb{Z}^5 \oplus \mathbb{Z}_{/16} \oplus \mathbb{Z}_{/7}\right) = \mathsf{rank}\left(\mathbb{Z}^5\right) + \mathsf{rank}\left(\mathbb{Z}_{/16} \oplus \mathbb{Z}_{/7}\right) = 5 + 0 = 5.$

Lemma 25 *Let $\mathcal{A}$ and $\mathcal{B}$ be free abelian groups. Then:*

(a) *$\mathcal{A} \oplus \mathcal{B}$ is also a free abelian group.*

(b) $\mathsf{rank}(\mathcal{A} \oplus \mathcal{B}) = \mathsf{rank}(\mathcal{A}) + \mathsf{rank}(\mathcal{B})$.

Proof: Suppose $\mathsf{rank}(\mathcal{A}) = R$ and $\mathsf{rank}(\mathcal{B}) = S$. Thus, $\mathcal{A} \cong \mathbb{Z}^R$, and $\mathcal{B} \cong \mathbb{Z}^S$. Thus, $\mathcal{A} \oplus \mathcal{B} \cong \mathbb{Z}^R \oplus \mathbb{Z}^S = \mathbb{Z}^{R+S}$. □

Example 26: $\mathbb{Z}^3 \oplus \mathbb{Z}^5 = \mathbb{Z}^8$.

If $(\mathcal{A},+)$ is an (additive) abelian group, and $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_R \in \mathcal{A}$, then we say that $\mathcal{A}$ is **generated** by $\mathbf{g}_1,\ldots,\mathbf{g}_R$ if, for any $\mathbf{a} \in \mathcal{A}$, we can find integers $z_1, z_2, \ldots, z_R \in \mathbb{Z}$ so that

$$\mathbf{a} = z_1\mathbf{g}_1 + z_2\mathbf{g}_2 + \ldots + z_R\mathbf{g}_R. \tag{3.1}$$

In this case, we write "$\mathcal{A} = \langle \mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_R \rangle$", or, inspired by eqn. (3.1), we write:

$$\mathcal{A} = \mathbb{Z}\mathbf{g}_1 + \mathbb{Z}\mathbf{g}_2 + \ldots + \mathbb{Z}\mathbf{g}_R$$

If $\mathcal{A}$ has a finite generating set, we say that $\mathcal{A}$ is **finitely generated**.

Proposition 27 *Let $\mathcal{A}$ be an abelian group, and let $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_R \in \mathcal{A}$. The following are equivalent:*

(a) $\mathcal{A} = \mathbb{Z}\mathbf{g}_1 + \mathbb{Z}\mathbf{g}_2 + \ldots + \mathbb{Z}\mathbf{g}_R$, *and* $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_R$ *are* $\mathbb{Z}$*-linearly independent*

(b) $\mathcal{A} = \mathbb{Z}\mathbf{g}_1 \oplus \mathbb{Z}\mathbf{g}_2 \oplus \ldots \oplus \mathbb{Z}\mathbf{g}_R$ *(where* $\mathbb{Z}\mathbf{g}_r$ *is the cyclic subgroup generated by* $\mathbf{g}_r$*).*

(c) *For any* $\mathbf{a} \in \mathcal{A}$*, there are unique integers* $z_1, z_2, \ldots, z_R \in \mathbb{Z}$ *so that*
$$\mathbf{a} = z_1\mathbf{g}_1 + z_2\mathbf{g}_2 + \ldots + z_R\mathbf{g}_R.$$

(d) $\mathcal{A}$ *is isomorphic to* $\mathbb{Z}^R$*, via the mapping* $\phi : \mathbb{Z}^R \longrightarrow \mathcal{A}$ *defined:*
$$\phi(z_1, z_2, \ldots, z_R) = z_1\mathbf{g}_1 + z_2\mathbf{g}_2 + \ldots + z_R\mathbf{g}_R$$

Proof: (**Exercise 13**) □

If the conditions of Proposition 27 are satisfied, we say that $\mathcal{A}$ is the **free abelian group generated by** $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_R$, and we call $\{\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_R\}$ a **basis** for $\mathcal{A}$.

3.3 Universal Properties of Free Abelian Groups

Prerequisites: §3.2

Free abelian groups are called 'free' because they are the most 'structureless' of all abelian groups. It is thus very easy to construct epimorphisms from free abelian groups into any other abelian group. This endows the free groups with certain 'universal' properties....

Proposition 28 Universal Mapping Property of Free Abelian Groups

Let $\mathcal{A}$ be an abelian group, and let $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_R \in \mathcal{A}$. Define the function $\phi : \mathbb{Z}^R \longrightarrow \mathcal{A}$ by $\phi(z_1, z_2, \ldots, z_R) = z_1\mathbf{g}_1 + z_2\mathbf{g}_2 + \ldots + z_R\mathbf{g}_R$. Then:

(a) ϕ *is always a homomorphism (regardless of the choice of* $\mathbf{g}_1, \ldots, \mathbf{g}_R$*).*

(b) *Every homomorphism from* $\mathbb{Z}^R$ *into* $\mathcal{A}$ *has this form.*

(c) *If* $\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_R$ *generate* $\mathcal{A}$*, then* ϕ *is an epimorphism.*

Proof: (**Exercise 14**) □

Example 29:

(a) Let $\mathcal{A} = \mathbb{Z}_{/n}$, and recall that $\mathbb{Z}$ is the free abelian group of rank 1. Define $\phi : \mathbb{Z} \longrightarrow \mathbb{Z}_{/n}$ by $\phi(z) = z \cdot \overline{1} = \overline{z}$. Then ϕ is an epimorphism.

(b) Let $\mathcal{A}$ be any abelian group and let $\mathbf{a} \in \mathcal{A}$ be any element. Define $\phi : \mathbb{Z} \longrightarrow \mathcal{A}$ by $\phi(z) = z \cdot \mathbf{a}$. Then ϕ is a homomorphism, whose image is the cyclic subgroup $\langle \mathbf{a} \rangle$.

(c) Let $\mathcal{A} = \mathbb{Z}_{/5} \oplus \mathbb{Z}_{/7}$, and define $\phi : \mathbb{Z}^2 \longrightarrow \mathcal{A}$ by $\phi(z_1, z_2) = (z_1 \cdot [1]_5,\ z_2 \cdot [1]_7) = ([z_1]_5,\ [z_2]_7)$ (where $[n]_5$ is the congruence class of n, mod 5, etc.) Then ϕ is an epimorphism.

Corollary 30 Universal Covering Property of Free Abelian Groups

Any finitely generated abelian group is a quotient of a (finitely generated) free abelian group.

Proof: (**Exercise 15**) □

This is called the 'covering' property because any abelian group can be 'covered' by the projection of some free abelian group. Thus, to prove some result about *all* abelian groups, it is often sufficient to prove the result only for *free* abelian groups.

3.4 (∗) Homological Properties of Free Abelian Groups

Prerequisites: §3.3, Homomorphism Groups

The next two properties are not important to us at present, but arise frequently in homological algebra.

Corollary 31 Projective Property of Free Abelian Groups

Suppose $\mathcal{A}$ and $\mathcal{B}$ are abelian groups, and there are homomorphisms $\phi : \mathbb{Z}^R \longrightarrow \mathcal{A}$ and $\beta : \mathcal{B} \longrightarrow \mathcal{A}$. Then there is a homomorphism $\psi : \mathbb{Z}^R \longrightarrow \mathcal{B}$ so that $\beta \circ \psi = \phi$.

$$\begin{tikzcd}
\mathbb{Z}^R \arrow[rd, "\phi"'] & \mathcal{B} \arrow[d, "\beta"] & & \mathbb{Z}^R \arrow[r, "\psi"] \arrow[rd, "\phi"'] & \mathcal{B} \arrow[d, "\beta"] \\
& \mathcal{A} \arrow[rr, phantom, "\Longrightarrow"] & & & \mathcal{A}
\end{tikzcd}$$

In other words, given any diagram like the one on the left, we can always 'complete' it to get a commuting diagram like the one on the right.

Proof: Proposition 28(c) says we can find some elements $\mathbf{a}_1, \ldots, \mathbf{a}_R \in \mathcal{A}$ so that ϕ has the form: $\phi(z_1, \ldots, z_R) = z_1\mathbf{a}_1 + \ldots + z_R\mathbf{a}_R$. Now, $\beta : \mathcal{B} \longrightarrow \mathcal{A}$ is a surjection, so for each r, find some $\mathbf{b}_r \in \mathcal{B}$ with $\beta(\mathbf{b}_r) = \mathbf{a}_r$. Now, define the map $\psi : \mathbb{Z}^R \longrightarrow \mathcal{B}$ by $\psi(z_1, \ldots, z_R) = z_1\mathbf{b}_1 + \ldots + z_R\mathbf{b}_R$. Proposition 28(a) says ψ is a homomorphism. To see that $\beta \circ \psi = \phi$, observe that $\beta \circ \psi(z_1, \ldots, z_R) = \beta(z_1\mathbf{b}_1 + \ldots + z_R\mathbf{b}_R) = z_1\beta(\mathbf{b}_1) + \ldots + z_R\beta(\mathbf{b}_R) = z_1\mathbf{a}_1 + \ldots + z_R\mathbf{a}_R = \phi(z_1, \ldots, z_R)$. □

Proposition 32 Adjoint Property

Let $\mathcal{A}$ be an abelian group, and let $\mathsf{Hom}\left(\mathbb{Z}^R, \mathcal{A}\right)$ *be the group of all homomorphisms from* $\mathbb{Z}^R$ *into* $\mathcal{A}$. *Then* $\mathsf{Hom}\left(\mathbb{Z}^R, \mathcal{A}\right)$ *is isomorphic to* $\mathcal{A}^R$, *via the map*

$$\Phi : \mathcal{A}^R \longrightarrow \mathsf{Hom}\left(\mathbb{Z}^R, \mathcal{A}\right)$$

sending $(\mathbf{a}_1, \ldots, \mathbf{a}_R)$ *to the morphism* $\alpha : \mathbb{Z}^R \longrightarrow \mathcal{A}$ *such that* $\alpha(z_1, \ldots, z_R) = z_1\mathbf{a}_1 + \ldots + z_R\mathbf{a}_R$.

Proof: Φ *is a homomorphism:* Let $\mathbf{a} = (\mathbf{a}_1, \ldots, \mathbf{a}_R) \in \mathcal{A}^R$ and $\mathbf{b} = (\mathbf{b}_1, \ldots, \mathbf{b}_R) \in \mathcal{A}^R$. Let $\alpha = \Phi(\mathbf{a})$ and $\beta = \Phi(\mathbf{b})$. Thus, $\alpha, \beta : \mathbb{Z}^R \longrightarrow \mathcal{A}$ are homomomorphisms, such that $\alpha(z_1, \ldots, z_R) = z_1\mathbf{a}_1 + \ldots + z_R\mathbf{a}_R$, and $\beta(z_1, \ldots, z_R) = z_1\mathbf{b}_1 + \ldots + z_R\mathbf{b}_R$. We want to show that $\Phi(\mathbf{a}+\mathbf{b}) = \alpha+\beta$. But $\mathbf{a}+\mathbf{b} = (\mathbf{a}_1+\mathbf{b}_1, \ldots, \mathbf{a}_R+\mathbf{b}_R)$, so $\Phi(\mathbf{a}+\mathbf{b})$ is the homomorphism $\chi(z_1, \ldots, z_R) = z_1(\mathbf{a}_1+\mathbf{b}_1)+\ldots+z_R(\mathbf{a}_R+\mathbf{b}_R) = (z_1\mathbf{a}_1+\ldots+z_R\mathbf{a}_R)+(z_1\mathbf{b}_1+\ldots+z_R\mathbf{b}_R) = \alpha(z_1, \ldots, z_R) + \beta(z_1, \ldots, z_R)$.

Φ *is surjective:* This is just Proposition 28(b).

Φ *is injective:* This is **Exercise 16** . □

3.5 Subgroups of Free Abelian Groups

Prerequisites: §3.2

Any subgroup of a free abelian group is also free. Furthermore, there is a 'common basis' for both the group and its subgroup...

Proposition 33 *Let $\mathcal{A}$ be a free abelian group, and let $\mathcal{B} < \mathcal{A}$ be any subgroup. Then:*

(a) *$\mathcal{B}$ is also a free abelian group, of rank $S \leq R$.*

(b) *There exists a basis $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_R\}$ for $\mathcal{A}$, and numbers $m_1, m_2, \ldots, m_S \in \mathbb{N}$, so that, if $\mathbf{b}_s = m_s\mathbf{a}_s$ for all $s \in [1..S]$, then $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_S\}$ is a basis for $\mathcal{B}$.*

Example 34:

(a) As shown in Figure 3.2, Let $\mathcal{A} = \mathbb{Z}^2$, and let $\mathcal{B}$ be the subgroup generated by $\mathbf{x} = (2, 0)$ and $\mathbf{y} = (1, 2)$. We can see that $\mathcal{B} = (\mathbb{Z}\mathbf{x}) \oplus (\mathbb{Z}\mathbf{y})$, so $\mathcal{B}$ is a free abelian group. However, we want to find a 'common basis' for $\mathcal{A}$ and $\mathcal{B}$.

As shown in Figure 3.3, let $\mathbf{a}_1 = (1, 2)$ and $\mathbf{a}_2 = (0, 1)$. Let $m_1 = 1$ and $m_2 = 4$, so that $\mathbf{b}_1 = m_1\mathbf{a}_1 = (1, 2)$ and $\mathbf{b}_2 = m_2\mathbf{a}_2 = (0, 4)$. Then $\mathcal{A} = (\mathbb{Z}\mathbf{a}_1) \oplus (\mathbb{Z}\mathbf{a}_2)$ and $\mathcal{B} = (\mathbb{Z}\mathbf{b}_1) \oplus (\mathbb{Z}\mathbf{b}_2)$.

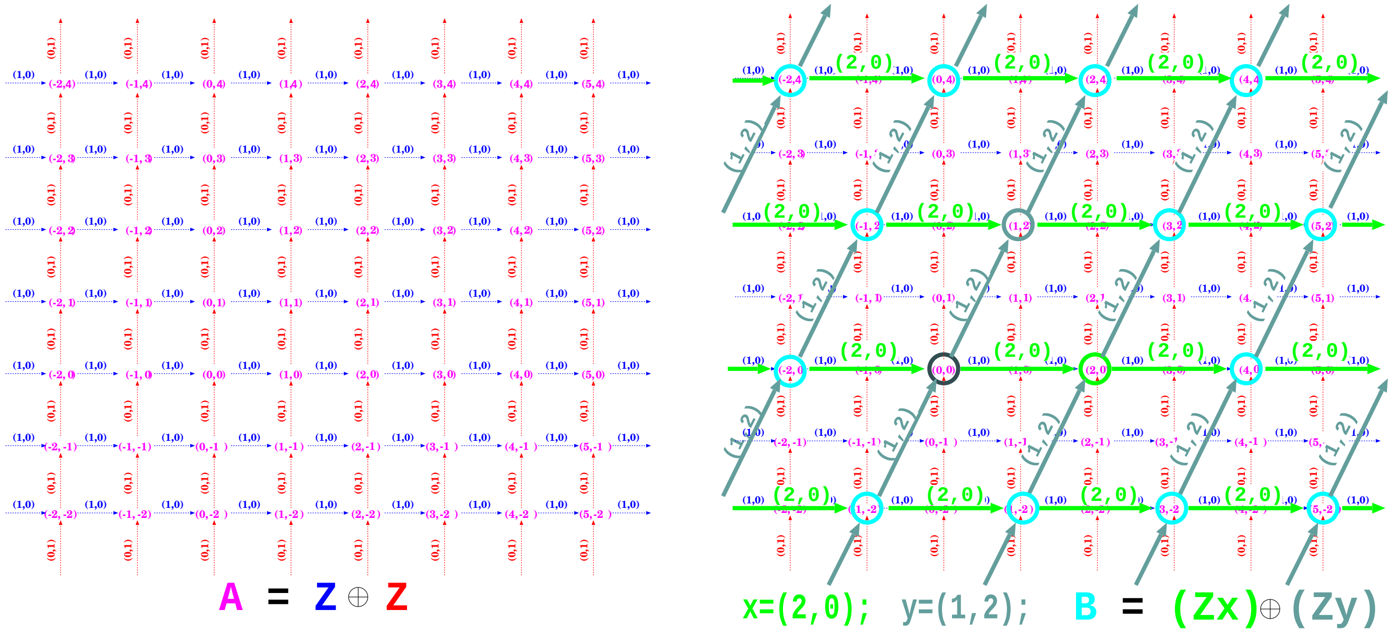


Figure 3.2: Example ⟨34a⟩: $\mathcal{A} = \mathbb{Z} \oplus \mathbb{Z}$, and $\mathcal{B}$ is the subgroup generated by $\mathbf{x} = (2, 0)$ and $\mathbf{y} = (1, 2)$.

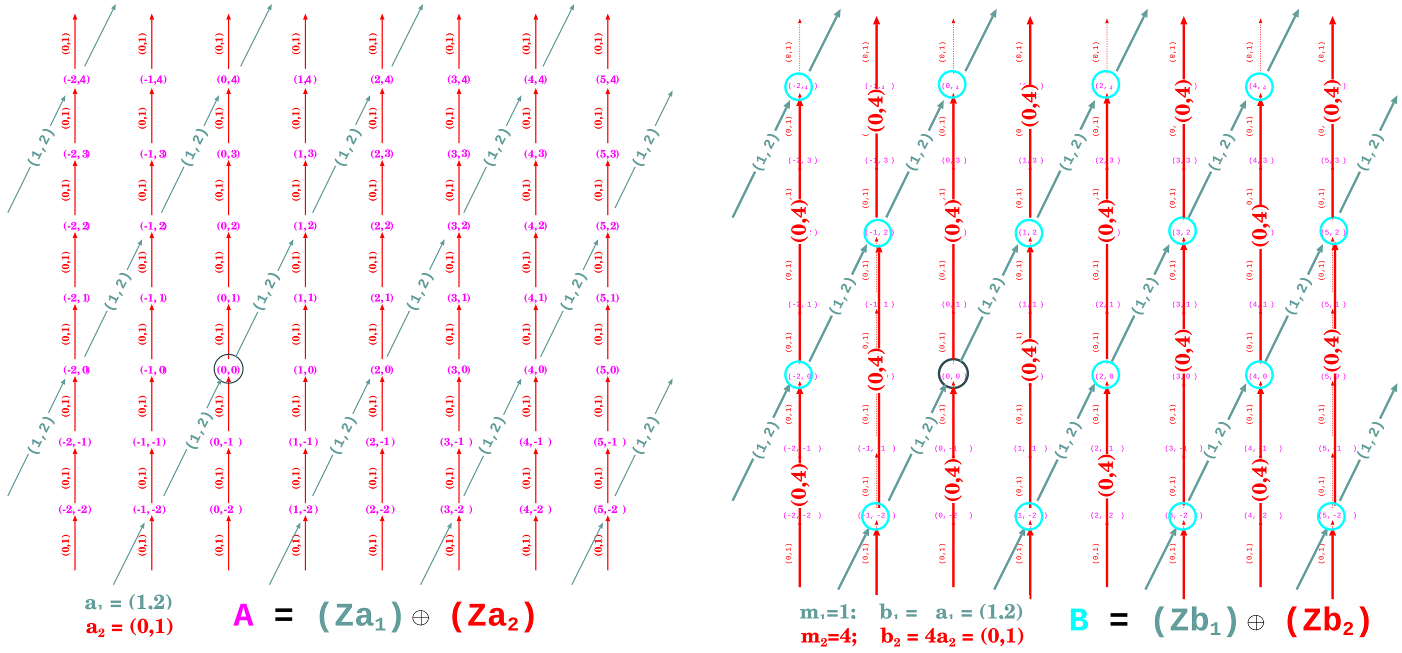


Figure 3.3: Example ⟨34a⟩: $\mathcal{A} = (\mathbb{Z}\mathbf{a}_1) \oplus (\mathbb{Z}\mathbf{a}_2)$, where $\mathbf{a}_1 = (1, 2)$ and $\mathbf{a}_2 = (0, 1)$. Let $m_1 = 1$ and $m_2 = 4$, so that $\mathbf{b}_1 = m_1\mathbf{a}_1 = (1, 2)$ and $\mathbf{b}_2 = m_2\mathbf{a}_2 = (0, 4)$. Then $\mathcal{B} = (\mathbb{Z}\mathbf{b}_1) \oplus (\mathbb{Z}\mathbf{b}_2)$.

(b) Let $\mathcal{A} = \mathbb{Z}^3$, and let $\mathcal{B} = \{(5y, 7z, 0) \; ; \; y, z \in \mathbb{Z}\}$. In this case, $R = 3$ and $S = 2$. Define:

$$\begin{array}{rclcrclcrcl} \mathbf{a}_1 & = & (1,0,0) & \text{and} & m_1 & = & 5, & \text{so that} & \mathbf{b}_1 & = & (5,0,0); \\ \mathbf{a}_2 & = & (0,1,0) & \text{and} & m_2 & = & 7, & \text{so that} & \mathbf{b}_2 & = & (0,7,0); \\ \mathbf{a}_3 & = & (0,0,1). \end{array}$$

Then $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ is a basis for $\mathcal{A}$, and $\{\mathbf{b}_1, \mathbf{b}_2\}$ is a basis for $\mathcal{B}$.

Proof of Proposition 33: We will prove **(a)** by induction on R, and prove **(b)** by induction on S. We will use the following fact:

Claim 8: *There exists elements $\mathbf{a}_1 \in \mathcal{B}$ and $\mathbf{b}_1 \in \mathcal{B}$, and a number $m_1 \in \mathbb{N}$ so that $\mathbf{b}_1 = m_1 \mathbf{a}_1$, such that*

$$\mathcal{A} \;=\; (\mathbb{Z}\mathbf{a}_1) \oplus \widetilde{\mathcal{A}} \qquad \textit{and} \qquad \mathcal{B} \;=\; (\mathbb{Z}\mathbf{b}_1) \oplus \widetilde{\mathcal{B}},$$

where $\widetilde{\mathcal{A}} < \mathcal{A}$ is a subgroup of rank $R-1$, and $\widetilde{\mathcal{B}} < \widetilde{\mathcal{A}}$ is a subgroup of rank $S-1$.

....we will then apply the induction hypotheses to $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{B}}$.

Claim 8 will follow from Claims 1 through 7 below. Without loss of generality, suppose $\mathcal{A} = \mathbb{Z}^R$ (we know that $\mathcal{A}$ is always isomorphic to $\mathbb{Z}^R$, so this is okay). For all $r \in [1..R]$, let $\mathbf{pr}_r : \mathcal{A} \longrightarrow \mathbb{Z}$ be projection into the rth coordinate —ie. $\mathbf{pr}_r(z_1, \ldots, z_R) \;=\; z_r$.

Let $\phi : \mathcal{A} \longrightarrow \mathbb{Z}$ be any group homomorphism. Then $\phi(\mathcal{B}) < \mathbb{Z}$ is a subgroup, so there is some $m_\phi \in \mathbb{Z}$ such that $\phi(\mathcal{B}) \;=\; m_\phi \mathbb{Z}$. Define:

$$\mathbb{M} \quad = \quad \{m_\phi \; ; \; \phi : \mathcal{A} \longrightarrow \mathbb{Z} \text{ any homomorphism}\}.$$

Claim 1: $\mathbb{M} \neq \{0\}$.

Proof: We must show that there is some homomorphism $\phi : \mathcal{A} \longrightarrow \mathbb{Z}$ so that $\phi(\mathcal{B}) \neq \{0\}$. Recall that $\mathcal{A} = \mathbb{Z}^R$, and $\mathbf{pr}_r : \mathcal{A} \longrightarrow \mathbb{Z}$ is projection onto the rth coordinate. Given any $\mathbf{a} \in \mathcal{A}$, there must be some $r \in [1..R]$ so that $\mathbf{pr}_r(\mathbf{a}) \neq 0$ —otherwise $\mathbf{a} = (0, 0, \ldots, 0)$. In particular, for any $\mathbf{b} \in \mathcal{B}$, there is some $r \in [1..R]$ so that $\mathbf{pr}_r(\mathbf{b}) \neq 0$. Thus, $\mathbf{pr}_r(\mathcal{B}) \neq \{0\}$. □ [Claim 1]

Now, let m_1 be the *minimal nonzero element* in $\mathbb{M}$. Let $\phi_1 : \mathcal{A} \longrightarrow \mathbb{Z}$ be a homomorphism such that $m_{(\phi_1)} = m_1$, and let $\mathbf{b}_1 \in \mathcal{B}$ be an element such that

$$\phi_1(\mathbf{b}_1) \quad = \quad m_1. \qquad (*)$$

Claim 2: *Let $\psi : \mathcal{A} \longrightarrow \mathbb{Z}$ be any homomorphism. Then m_1 divides $\psi(\mathbf{b}_1)$.*

Proof: Let $n = \psi(\mathbf{b}_1)$, and let $d = \gcd(m_1, n)$. Thus $d = z_1 m_1 + zn$ for some integers $z_1, z \in \mathbb{Z}$. Define the homomorphism $\delta : \mathcal{A} \longrightarrow \mathbb{Z}$ by: $\delta(\mathbf{a}) \;=\; z_1\phi_1(\mathbf{a}) + z\psi(\mathbf{a})$. Then

$$\delta(\mathbf{b}_1) \quad = \quad z_1\phi_1(\mathbf{b}_1) + z\psi(\mathbf{b}_1) \quad \underset{\text{by } (*)}{=\!=} \quad z_1 m_1 + zn \quad = \quad d.$$

Thus, $d \in \delta(\mathcal{B})$, so m_δ must divide d, so $m_\delta \leq d$. But d divides m_1, so $d \leq m_1$. Thus, $m_\delta \leq m_1$. But m_1 is the *minimal* nonzero element in $\mathbb{M}$, so $m_1 \leq m_\delta$. Therefore, $m_\delta = m_1$, which means $d = m_1$, which means m_1 divides n. .... □ [Claim 2]

Recall that $\mathcal{A} = \mathbb{Z}^R$, and $\mathbf{pr}_r : \mathcal{A} \longrightarrow \mathbb{Z}$ is projection onto the rth coordinate. Claim 2 says m_1 divides $\mathbf{pr}_r(\mathbf{b}_1)$ for all $r \in [1..R]$. In other words, $\mathbf{b}_1 = (b_1, b_2, \ldots, b_R)$, where $b_1 = m_1 a_1, \quad b_2 = m_2 a_2, \ldots, \; b_R = m_R a_R$, for some $a_1, \ldots, a_R \in \mathbb{Z}$. Thus, $\mathbf{b}_1 = m_1 \mathbf{a}_1$, where $\mathbf{a}_1 = (a_1, a_2, \ldots, a_R)$.

Claim 3: $\phi_1(\mathbf{a}_1) = 1$.

Proof: $m_1 = \phi_1(\mathbf{b}_1) = \phi_1(m_1\mathbf{a}_1) = m_1\phi_1(\mathbf{a}_1)$. Thus, $\phi(\mathbf{a}_1) = 1$. ...... □ [Claim 3]

Now, let $\widetilde{\mathcal{A}} = \ker(\phi_1)$, and let $\widetilde{\mathcal{B}} = \mathbf{K} \cap \mathcal{B}$.

Claim 4: $\mathcal{A} \;=\; (\mathbb{Z}\mathbf{a}_1) \oplus \widetilde{\mathcal{A}}$.

Proof:

Claim 4.1: $\mathcal{A} \;=\; (\mathbb{Z}\mathbf{a}_1) + \widetilde{\mathcal{A}}$.

Proof: Let $\mathbf{a} \in \mathcal{A}$, and let $z = \phi_1(\mathbf{a})$. Let $\widetilde{\mathbf{a}} = \mathbf{a} - z\mathbf{a}_1$. I claim $\widetilde{\mathbf{a}} \in \widetilde{\mathcal{A}}$. To see this, note that $\phi_1(\widetilde{\mathbf{a}}) \;=\; \phi_1(\mathbf{a}) - z\cdot\phi_1(\mathbf{a}_1) \;\underset{\text{Clm 3}}{=\!=}\; z - z\cdot 1 \;=\; z - z = 0$.
Thus, $\mathbf{a} \;=\; z\mathbf{a}_1 + \widetilde{\mathbf{a}} \;\in\; (\mathbb{Z}\mathbf{a}_1) + \widetilde{\mathcal{A}}$. .... □ [Claim 4.1]

Claim 4.2: $(\mathbb{Z}\mathbf{a}_1) \cap \widetilde{\mathcal{A}} \;=\; \{0\}$.

Proof: Any element of $\mathbb{Z}\mathbf{a}_1$ has the form $z\mathbf{a}_1$ for some $z \in \mathbb{Z}$. If $z\mathbf{a}_1 \in \widetilde{\mathcal{A}}$, this means that $\phi_1(z\mathbf{a}_1) = 0$. But $\phi_1(z\mathbf{a}_1) \;=\; z\phi_1(\mathbf{a}_1) \;\underset{\text{Clm 3}}{=\!=}\; z\cdot 1 \;=\; z$. Thus, $z = 0$. □ [Claim 4.2]

Claims 4.1 and 4.2 imply that $\mathcal{A} \;=\; (\mathbb{Z}\mathbf{a}_1) \oplus \widetilde{\mathcal{A}}$. .... □ [Claim 4]

Claim 5: $\mathcal{B} \;=\; (\mathbb{Z}\mathbf{b}_1) \oplus \widetilde{\mathcal{B}}$.

Proof:

Claim 5.1: $\mathcal{B} \;=\; (\mathbb{Z}\mathbf{b}_1) + \widetilde{\mathcal{B}}$.

Proof: Let $\mathbf{b} \in \mathcal{B}$, and let $z = \phi_1(\mathbf{b})$. Let $\widetilde{\mathbf{b}} = \mathbf{b} - z\mathbf{a}_1$. As in Claim 4.1, $\widetilde{\mathbf{b}} \in \widetilde{\mathcal{A}}$.
I claim that $z\mathbf{a}_1 \in \mathbb{Z}\mathbf{b}$. To see this, recall that $z = \phi_1(\mathbf{b}) \in \phi_1(\mathcal{B}) \;=\; m_1\mathbb{Z}$ by definition of m_1. Thus, m_1 must divide z. Thus, $z = ym_1$ for some $y \in \mathbb{Z}$, so that $z\mathbf{a}_1 = ym_1\mathbf{a}_1 = y\mathbf{b}_1 \in \mathbb{Z}\mathbf{b}$.
I also claim that $\widetilde{\mathbf{b}} \in \widetilde{\mathcal{B}}$. We know that $\widetilde{\mathbf{b}} \in \widetilde{\mathcal{A}}$. Observe that $\widetilde{\mathbf{b}} \;=\; \mathbf{b} - z\mathbf{a}_1 \;=\; \mathbf{b} - y\mathbf{b}_1$ is a difference of two elements in $\mathcal{B}$, so $\widetilde{\mathbf{b}} \in \mathcal{B}$ also. Thus, $\widetilde{\mathbf{b}} \in \widetilde{\mathcal{A}} \cap \mathcal{B} = \widetilde{\mathcal{B}}$.
Thus, $\mathbf{a} \;=\; y\mathbf{b}_1 + \widetilde{\mathbf{b}} \;\in\; (\mathbb{Z}\mathbf{b}_1) + \widetilde{\mathcal{B}}$. .... □ [Claim 5.1]

Claim 4.1 implies that $(\mathbb{Z}\mathbf{b}_1) \cap \widetilde{\mathcal{B}} \;=\; \{0\}$. We conclude that $\mathcal{B} \;=\; (\mathbb{Z}\mathbf{b}_1) \oplus \widetilde{\mathcal{B}}$. □ [Claim 5]

Claim 6: $\mathsf{rank}\left(\widetilde{\mathcal{A}}\right) = R - 1$.

Proof: Claim 4 says that $\mathcal{A} = (\mathbb{Z}\mathbf{a}_1) \oplus \widetilde{\mathcal{A}}$. Thus,

$$R \;=\; \mathsf{rank}\,(\mathcal{A}) \;\underset{\text{Lem. 23}}{=\!=}\; \mathsf{rank}\,(\mathbb{Z}\mathbf{a}_1) + \mathsf{rank}\left(\widetilde{\mathcal{A}}\right) \;=\; 1 + \mathsf{rank}\left(\widetilde{\mathcal{A}}\right).$$

Thus $\mathsf{rank}\left(\widetilde{\mathcal{A}}\right) = R - 1$. .... □ [Claim 6]

Claim 7: $\mathsf{rank}\left(\widetilde{\mathcal{B}}\right) = S - 1$.

Proof: Claim 5 says $\mathcal{B} = (\mathbb{Z}\mathbf{b}_1) \oplus \widetilde{\mathcal{B}}$. Now proceed exactly as with Claim 6. □ `[Claim 7]`

Claim 8 follows from Claims 1 through 7. Now to prove the theorem....

Proof of **(a)***:* (by induction on S)

Case ($S = 0$) Suppose $\mathsf{rank}(\mathcal{B}) = 0$. I claim $\mathcal{B} = \{0\}$. To see this, suppose $\mathbf{b} \in \mathcal{B}$ was nonzero. Then for any $z \in \mathbb{Z}$, $z\mathbf{b} \neq 0$, so the set $\{\mathbf{b}\}$ is linearly independent, hence $\mathsf{rank}(\mathcal{B}) \geq 1$, a contradiction.

Induction on S: Suppose part **(a)** is true for all subgroups of $\mathcal{A}$ of rank $S - 1$. Then in particular, part **(a)** holds for all $\widetilde{\mathcal{B}}$; hence, $\widetilde{\mathcal{B}}$ is a free abelian group. Since $\mathcal{B} = (\mathbb{Z}\mathbf{b}_1) \oplus \widetilde{\mathcal{B}}$, Lemma 25(a) says that $\mathcal{B}$ is also a free abelian group.

We conclude that all subgroups of $\mathcal{A}$ are free.

Proof of **(b)***:* (by induction on R)

Case ($R = 1$) In this case, $\mathcal{A} \cong \mathbb{Z}$. Thus, $\mathcal{B} < \mathbb{Z}$, so we know there is some $m_1 \in \mathbb{N}$ so that $\mathcal{B} = m_1\mathbb{Z}$. So, let $\mathbf{a}_1 = 1$ and $\mathbf{b}_1 = m_1 \cdot \mathbf{a}_1 = m_1$. Then $\mathbb{Z}$ is generated by $\mathbf{a}_1$, and $\mathcal{B}$ is generated by $\mathbf{b}_1$.

Induction on R: Suppose part **(b)** is true for *all* subgroups of *all* free abelian groups of rank $R - 1$. Consider $\widetilde{\mathcal{A}}$. By Claim 6, we know that $\mathsf{rank}\left(\widetilde{\mathcal{A}}\right) = R - 1$. By part **(a)** of the theorem (already proved), we know that $\widetilde{\mathcal{A}}$ must be a free group. Also, $\widetilde{\mathcal{B}} < \widetilde{\mathcal{A}}$, and $\mathsf{rank}\left(\widetilde{\mathcal{B}}\right) = S - 1$ (Claim 7) so by induction hypothesis, there is some basis $\{\mathbf{a}_2, \ldots, \mathbf{a}_R\}$ for $\widetilde{\mathcal{A}}$, and numbers $m_2, \ldots, m_S \in \mathbb{N}$, so that, if $\mathbf{b}_s = m_s\mathbf{a}_s$ for all $s \in [2..S]$, then $\{\mathbf{b}_2, \ldots, \mathbf{b}_S\}$ is a basis for $\widetilde{\mathcal{B}}$.

Since $\mathcal{A} = (\mathbb{Z}\mathbf{a}_1) \oplus \widetilde{\mathcal{A}}$ (by Claim 4), it follows that $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_R\}$ is a basis for $\mathcal{A}$. Since $\mathcal{B} = (\mathbb{Z}\mathbf{b}_1) \oplus \widetilde{\mathcal{B}}$ (by Claim 5), it follows that $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_S\}$ is a basis for $\mathcal{B}$. ______□

Chapter 4

The Structure Theory of Abelian Groups

Group structure theory concerns the decomposition of groups into 'elementary components'. For example, the Jordan-Hölder theorem says that any group $\mathcal{G}$ has a *composition series*

$$\{e\} = \mathcal{N}_0 \lhd \mathcal{N}_1 \lhd \ldots \lhd \mathcal{N}_K = \mathcal{G}$$

where the groups $\mathcal{S}_k = \mathcal{N}_k/\mathcal{N}_{k-1}$ are simple for all k.

In this section, we will show that any finitely generated *abelian* group can be written as a *direct sum* of cyclic groups.

4.1 Direct Products

Let $(\mathcal{A}, \star)$, $(\mathcal{B}, *)$, and $(\mathcal{C}, \diamond)$ be three groups. The **direct product** of $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ is the group

$$\mathcal{A} \times \mathcal{B} \times \mathcal{C} \;=\; \{(a, b, c)\ ;\ a \in \mathcal{A},\ b \in \mathcal{B},\ c \in \mathcal{C}\}$$

with the multiplication operation:

$$(a_1, b_1, c_1) \cdot (a_2, b_2, c_2) \quad = \quad (a_1 \star a_2,\ b_1 * b_2,\ c_1 \diamond c_2)$$

We have defined this for *three* groups, but the same construction works for any number of groups. The direct product of $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_N$ is denoted by "$\mathcal{G}_1 \times \mathcal{G}_2 \times \ldots \times \mathcal{G}_N$" or "$\displaystyle\prod_{n=1}^{N} \mathcal{G}_n$".

Example 35: Suppose $\mathcal{A} = \mathcal{B} = \mathcal{C} = (\mathbb{R}, +)$. Then $\mathcal{A} \times \mathcal{B} \times \mathcal{C} = \mathbb{R} \times \mathbb{R} \times \mathbb{R} \;=\; \mathbb{R}^3$ is just three-dimensional Euclidean space, with the usual vector addition.

Proposition 36 $\quad \left|\mathcal{G}_1 \times \mathcal{G}_2 \times \ldots \times \mathcal{G}_N\right| \quad = \quad |\mathcal{G}_1| \cdot |\mathcal{G}_2| \cdots |\mathcal{G}_N|.$

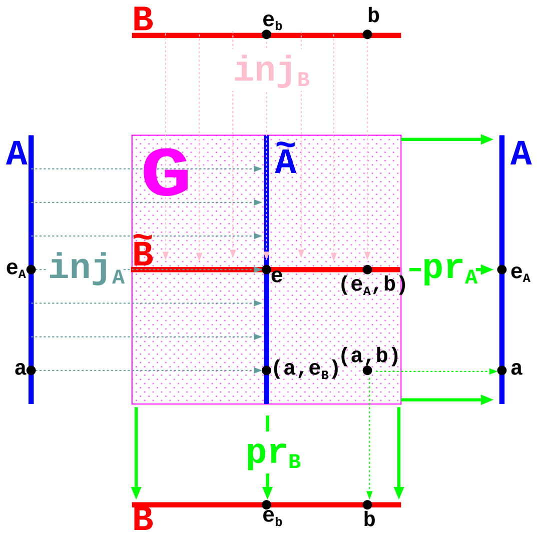


Figure 4.1: The product group $\mathcal{G} = \mathcal{A} \times \mathcal{B}$, with injection and projection maps.

Proof: (Exercise 17) □

Example 37: $\left|\mathbb{Z}_{/3} \times \mathbb{Z}_{/5} \times \mathbb{Z}_{/7}\right| = \left|\mathbb{Z}_{/3}\right| \cdot \left|\mathbb{Z}_{/5}\right| \cdot \left|\mathbb{Z}_{/7}\right| = 3 \cdot 5 \cdot 7 = 75.$

Lemma 38 *If $(\mathcal{A}, +)$, $(\mathcal{B}, +)$, and $(\mathcal{C}, +)$ are abelian groups, then their product $\mathcal{A} \times \mathcal{B} \times \mathcal{C}$ is also abelian.*

Proof: (Exercise 18) □

(This theorem likewise generalizes to a product of any number of groups)

The product of finitely many (additive) *abelian* groups is usually called a **direct sum**, and indicated with the notation "$\mathcal{A} \oplus \mathcal{B} \oplus \mathcal{C}$."

Proposition 39 (Properties of Product Groups)

Let $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ be groups, and let $\mathcal{G} = \mathcal{A} \times \mathcal{B} \times \mathcal{C}$ be their product.

(a) *The identity element of $\mathcal{G}$ is just $(e_{\mathcal{A}}, e_{\mathcal{B}}, e_{\mathcal{C}})$.*

(b) *For any $(a, b, c) \in \mathcal{G}$, $\quad (a, b, c)^{-1} = (a^{-1}, b^{-1}, c^{-1})$.*

(c) *Define injection maps:*

$$\begin{array}{rlrl} \mathsf{inj}_{\mathcal{A}} : \mathcal{A} \longrightarrow \mathcal{G} & \text{by} \quad \mathsf{inj}_{\mathcal{A}}(a) & = (a, e_{\mathcal{B}}, e_{\mathcal{C}}) & \text{for all } a \in \mathcal{A}; \\ \mathsf{inj}_{\mathcal{B}} : \mathcal{A} \longrightarrow \mathcal{G} & \text{by} \quad \mathsf{inj}_{\mathcal{B}}(b) & = (e_{\mathcal{A}}, b, e_{\mathcal{C}}) & \text{for all } b \in \mathcal{B}; \\ \text{and} \quad \mathsf{inj}_{\mathcal{A}} : \mathcal{A} \longrightarrow \mathcal{G} & \text{by} \quad \mathsf{inj}_{\mathcal{C}}(c) & = (e_{\mathcal{A}}, e_{\mathcal{B}}, c) & \text{for all } c \in \mathcal{C} \end{array}$$

Then $\mathsf{inj}_{\mathcal{A}}$, $\mathsf{inj}_{\mathcal{B}}$, and $\mathsf{inj}_{\mathcal{C}}$ are monomorphisms.

(d) *Define*

$$\begin{array}{rcccccc} \widetilde{\mathcal{A}} & = & \mathsf{image}\,[\mathsf{inj}_{\mathcal{A}}] & = & \{(a, e_{\mathcal{B}}, e_{\mathcal{C}})\ ;\ a \in \mathcal{A}\} & \subset & \mathcal{G}; \\ \widetilde{\mathcal{B}} & = & \mathsf{image}\,[\mathsf{inj}_{\mathcal{B}}] & = & \{(e_{\mathcal{A}}, b, e_{\mathcal{C}})\ ;\ b \in \mathcal{B}\} & \subset & \mathcal{G}; \\ \text{and } \widetilde{\mathcal{C}} & = & \mathsf{image}\,[\mathsf{inj}_{\mathcal{C}}] & = & \{(e_{\mathcal{A}}, e_{\mathcal{B}}, c)\ ;\ c \in \mathcal{B}\} & \subset & \mathcal{G} \end{array}$$

Then $\widetilde{\mathcal{A}}$, $\widetilde{\mathcal{B}}$, and $\widetilde{\mathbf{C}}$ are normal subgroups of $\mathcal{G}$

(e) *$\widetilde{\mathcal{A}}$ is isomorphic to $\mathcal{A}$ (via $\mathsf{inj}_{\mathcal{A}}$). Likewise, $\widetilde{\mathcal{B}} \cong \mathcal{B}$ and $\widetilde{\mathbf{C}} \cong \mathcal{C}$.*

(f) *$\mathcal{G} = \widetilde{\mathcal{A}} \cdot \widetilde{\mathcal{B}} \cdot \widetilde{\mathbf{C}}$. In other words, for any $g \in \mathcal{G}$, there are elements $\widetilde{a} \in \widetilde{\mathcal{A}}$, $\widetilde{b} \in \widetilde{\mathcal{B}}$, and $\widetilde{c} \in \widetilde{\mathbf{C}}$ so that $g = \widetilde{a} \cdot \widetilde{b} \cdot \widetilde{c}$.*

(g) *The elements of $\widetilde{\mathcal{A}}$, $\widetilde{\mathcal{B}}$, and $\widetilde{\mathbf{C}}$ commute with each other. In other words, for any $\widetilde{a} \in \widetilde{\mathcal{A}}$, $\widetilde{b} \in \widetilde{\mathcal{B}}$, and $\widetilde{c} \in \widetilde{\mathbf{C}}$,*

$$\widetilde{a} \cdot \widetilde{b} = \widetilde{b} \cdot \widetilde{a}, \qquad \widetilde{a} \cdot \widetilde{c} = \widetilde{c} \cdot \widetilde{a}, \qquad \text{and} \qquad \widetilde{b} \cdot \widetilde{c} = \widetilde{c} \cdot \widetilde{b},$$

(h) *$\widetilde{\mathcal{A}} \cap \widetilde{\mathcal{B}} = \{e\}$, $\widetilde{\mathcal{A}} \cap \widetilde{\mathbf{C}} = \{e\}$, and $\widetilde{\mathcal{B}} \cap \widetilde{\mathbf{C}} = \{e\}$.*

(i) *Define projection maps:*

$$\begin{array}{rccc} \mathbf{pr}_{\mathcal{A}} : \mathcal{G} \longrightarrow \mathcal{A} & \text{by} \quad \mathbf{pr}_{\mathcal{A}}(a, b, c) & = & a; \\ \mathbf{pr}_{\mathcal{B}} : \mathcal{G} \longrightarrow \mathcal{B} & \text{by} \quad \mathbf{pr}_{\mathcal{B}}(a, b, c) & = & b; \\ \text{and } \mathbf{pr}_{\mathcal{C}} : \mathcal{G} \longrightarrow \mathcal{C} & \text{by} \quad \mathbf{pr}_{\mathcal{C}}(a, b, c) & = & c \end{array}$$

Then $\mathbf{pr}_{\mathcal{A}}$, $\mathbf{pr}_{\mathcal{B}}$, and $\mathbf{pr}_{\mathcal{C}}$ are epimorphisms.

(j) *$\mathbf{pr}_{\mathcal{A}} \circ \mathsf{inj}_{\mathcal{A}} = \mathbf{Id}_{\mathcal{A}}$, $\mathbf{pr}_{\mathcal{B}} \circ \mathsf{inj}_{\mathcal{B}} = \mathbf{Id}_{\mathcal{B}}$, and $\mathbf{pr}_{\mathcal{C}} \circ \mathsf{inj}_{\mathcal{C}} = \mathbf{Id}_{\mathcal{C}}$.*

Proof: (**Exercise 19**) □

We have stated Proposition 39 for a product of *three* groups for the sake of simplicity, but the obvious generalization holds for any number of groups. The conditions of this propositon actually *characterize* direct products, as follows:

Proposition 40 *Let $\mathcal{G}$ be a group, with subgroups $\mathcal{A}, \mathcal{B} < \mathcal{G}$, such that $\mathcal{G} = \mathcal{A} \cdot \mathcal{B}$. The following are equivalent:*

(a) *$\mathcal{G}$ is isomorphic to $\mathcal{A} \times \mathcal{B}$, via the mapping:*

$$\phi : \mathcal{A} \times \mathcal{B} \ni (a, b) \ \mapsto \ a \cdot b \in \mathcal{G}$$

(b) *$\mathcal{A}\cap\mathcal{B} \ = \ \{e\}$, and for every $a\in\mathcal{A}$ and $b\in\mathcal{B}$, $\quad ab \ = \ ba$.*

(c) *$\mathcal{A}\cap\mathcal{B} \ = \ \{e\}$, and $\mathcal{A}\lhd\mathcal{G}$ and $\mathcal{B}\lhd\mathcal{G}$.*

Proof: The assertions (**a**)$\Longrightarrow$(**b**) and (**a**)$\Longrightarrow$(**c**) follow from Proposition 39(d,g,h).

(**b**)$\Longrightarrow$(**a**) We must show ϕ is a homomorphism and bijective.

ϕ is a homomorphism: $\phi\Big((a_1,b_1)\cdot(a_2,b_2)\Big) \underset{\text{(DP)}}{=\!=} \phi\Big((a_1a_2),\ (b_1b_2)\Big) \underset{(\text{D}\phi)}{=\!=} (a_1a_2)\cdot(b_1b_2) \underset{\text{(C)}}{=} (a_1b_1)\cdot(a_2b_2) \underset{(\text{D}\phi)}{=\!=} \phi(a_1,b_1)\cdot\phi(a_2,b_2)$. Here, equalities $(D\phi)$ are by definition of ϕ; (DP) is by definition of the product group, and (C) is because **(b)** says a_2 and b_1 commute.

ϕ is surjective: By hypothesis, for any $g\in\mathcal{G}$ we can find $a\in\mathcal{A}$ and $b\in\mathcal{B}$ so that $g=a\cdot b$. But then $g=\phi(a,b)$.

ϕ is injective: We'll show $\ker(\phi)=\{e\}$. Suppose $\phi(a,b)=e$. Thus, $a\cdot b=e$. Thus, $b=a^{-1}$, so $b\in\mathcal{A}$. Thus, $b\in\mathcal{A}\cap\mathcal{B}$, so $b=e$. Thus, $a=e$. Thus, $(a,b)=(e,e)$.

(**c**)$\Longrightarrow$(**b**) Let $a\in\mathcal{A}$ and $b\in\mathcal{B}$. Then:

$$\begin{array}{rcccccc}
\mathcal{A}\lhd\mathcal{G},\ \text{so} & & b & a^{-1} & b^{-1} & \in & \mathcal{A}, \\
\text{thus,} & a & b & a^{-1} & b^{-1} & \in & \mathcal{A}.
\end{array}
\qquad\qquad
\begin{array}{rcccccc}
\mathcal{B}\lhd\mathcal{G},\ \text{so} & a & b & a^{-1} & & \in & \mathcal{B}, \\
\text{thus,} & a & b & a^{-1} & b^{-1} & \in & \mathcal{B};
\end{array}$$

Thus, $aba^{-1}b^{-1}\in\mathcal{A}\cap\mathcal{B}=\{e\}$, so we conclude that $aba^{-1}b^{-1}=e$. Thus, $ab \ = \ \left(a^{-1}b^{-1}\right)^{-1} \ = \ \left(b^{-1}\right)^{-1}\left(a^{-1}\right)^{-1} \ = \ ba$. □

When the conditions of Proposition 41 are satisfied, we say that $\mathcal{G}$ is an **internal direct product** of $\mathcal{A}$ and $\mathcal{B}$. We then 'abuse notation' by writing, '$\mathcal{G} = \mathcal{A}\times\mathcal{B}$'. In terms of the notation of Proposition 39, we are implicitly identifying $\mathcal{A}$ with $\widetilde{\mathcal{A}}$ and $\mathcal{B}$ with $\widetilde{\mathcal{B}}$.

If $\mathcal{G}$ is *abelian*, then the conditions in parts (b) and (c) of Proposition 41 become trivial...

Corollary 41 (Interpretation for Abelian Groups)

Let $(\mathcal{G},+)$ be an (additive) abelian group, with subgroups $\mathcal{A},\mathcal{B}<\mathcal{G}$. Then the following are equivalent:

(a) *$\mathcal{A}\cap\mathcal{B}=\{0\}$, and $\mathcal{G}=\mathcal{A}+\mathcal{B}$ (that is, for any $g\in\mathcal{G}$ there are $a\in\mathcal{A}$ and $b\in\mathcal{B}$ so that $g=a+b$).*

(b) *$\mathcal{G}\cong\mathcal{A}\times\mathcal{B}$, via the mapping $\phi:\mathcal{A}\times\mathcal{B}\ni(a,b) \ \mapsto \ (a+b)\in\mathcal{G}$.* □

In the abelian cas, we say that $\mathcal{G}$ is an **internal direct sum** of $\mathcal{A}$ and $\mathcal{B}$. We again abuse notation by writing, '$\mathcal{G}=\mathcal{A}\oplus\mathcal{B}$'.

4.2 The Chinese Remainder Theorem

Prerequisites: §4.1

According to ancient legend from the War of the Three Kingdoms[1],

> *The redoubtable Shu general Zhu Geliang wanted to rapidly count his troops before entering a great battle against the Wei. He estimated that there were less than* 15000 *troops, and* $15000 < 17017 = 7 \times 11 \times 13 \times 17$. *So, first he had the troops line up in rows of* 7*, and found that there were* 6 *left over. Next, he lined them up in rows of* 11 *each, and found there were* 7 *left over. Next, in rows of* 13*, there were* 5 *left over. Finally, in rows of* 17*, there were* 2 *left over. He concluded that there were exactly* 14384 *troops.*

Theorem 42 (Chinese Remainder Theorem)

(a) *Let* $n_1, n_2, \ldots, n_K \in \mathbb{N}$ *be any collection of pairwise relatively prime numbers (ie.* $\gcd(n_j, n_k) = 1$ *whenever* $j \neq k$*), and let* $n = n_1 \cdot n_2 \cdots n_K$*. Then*

$$\mathbb{Z}_{/n} \quad \cong \quad \mathbb{Z}_{/n_1} \oplus \mathbb{Z}_{/n_2} \oplus \ldots \oplus \mathbb{Z}_{/n_k}.$$

(b) *To be specific, define* $\phi : \mathbb{Z}_{/n} \longrightarrow \mathbb{Z}_{/n_1} \oplus \mathbb{Z}_{/n_2} \oplus \ldots \oplus \mathbb{Z}_{/n_k}$ *by:*

$$\phi\left([z]_n\right) \quad = \quad \left([z]_{n_1},\ [z]_{n_2},\ \ldots [z]_{n_K}\right),$$

(where $[z]_n$ *is the congruence class of* z*, mod* n*, etc.). Then* ϕ *is an isomorphism.*

(c) *In particular, suppose* $n \in \mathbb{N}$ *has prime factorization* $n = p_1^{\nu_1} \cdot p_2^{\nu_2} \cdots p_K^{\nu_K}$*. Then*

$$\mathbb{Z}_{/n} \quad \cong \quad \mathbb{Z}_{/\left(p_1^{\nu_1}\right)} \oplus \mathbb{Z}_{/\left(p_2^{\nu_2}\right)} \oplus \ldots \oplus \mathbb{Z}_{/\left(p_K^{\nu_K}\right)}$$

Proof: **(c)** is a special case of **(a)** which follows from **(b)**, which is **Exercise 20** . □

Example 43:

(a) $12 = 4 \times 3$ and 4 is relatively prime to 3. Therefore, $\mathbb{Z}_{/12} \cong \mathbb{Z}_{/4} \oplus \mathbb{Z}_{/3}$. (Figure 4.2). Part **(b)** provides a specific isomorphism $\phi : \mathbb{Z}_{/12} \longrightarrow \mathbb{Z}_{/4} \oplus \mathbb{Z}_{/3}$, given by:

$$\phi\left([z]_{12}\right) \quad = \quad \left([z]_4,\ [z]_3\right).$$

For example, $\phi\left([5]_{12}\right)\left([5]_4,\ [5]_3\right) = \left([1]_{14},\ [2]_3\right)$.

[1] Actually, I made this up. But Zhu Geliang (C.E. 181-234) was a famously brilliant Shu military leader, inventor, and mathematician.

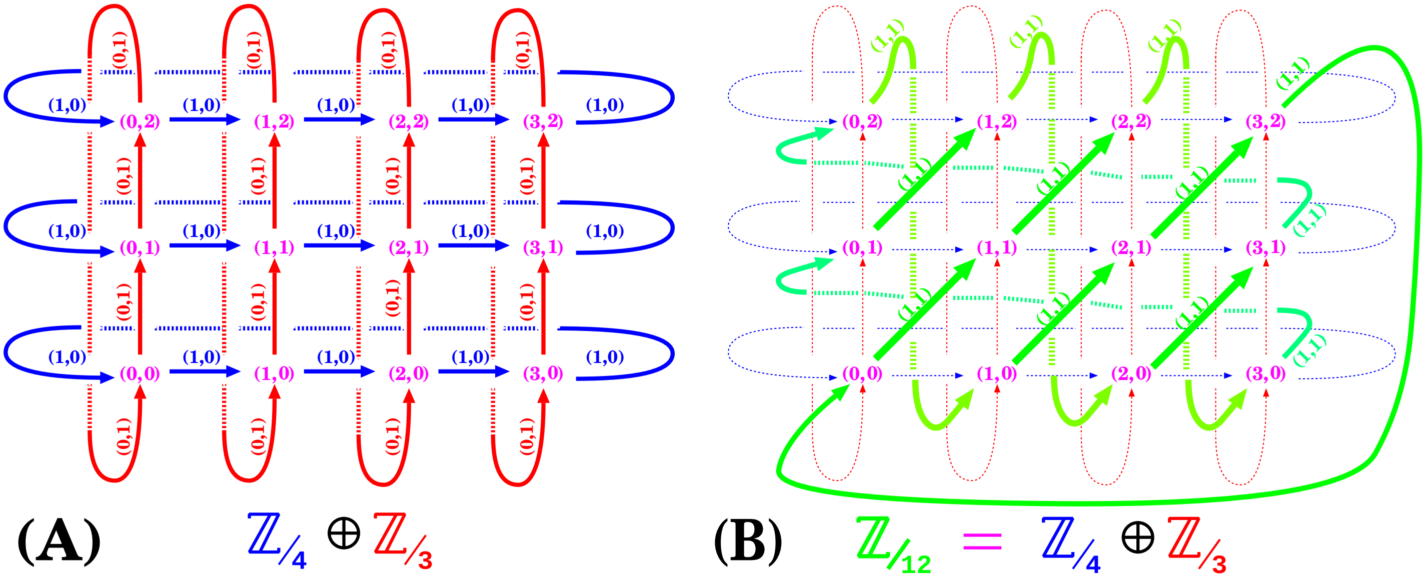

(0,1)
(1,0)
(0,2)
(1,2)
(2,2)
(3,2)
(0,1)
(1,1)
(2,1)
(3,1)
(0,0)
(1,0)
(2,0)
(3,0)
(1,1)
(A)
$\mathbb{Z}_{/4} \oplus \mathbb{Z}_{/3}$
(B)
$\mathbb{Z}_{/12} = \mathbb{Z}_{/4} \oplus \mathbb{Z}_{/3}$

Figure 4.2: Example ⟨43a⟩.

(b) $210 = 14 \times 15$, and 14 is relatively prime to 15. Thus, part **(a)** of the Chinese Remainder Theorem says that $\mathbb{Z}_{/210} \cong \mathbb{Z}_{/14} \oplus \mathbb{Z}_{/15}$. Part **(b)** provides a specific isomorphism $\phi : \mathbb{Z}_{/210} \longrightarrow \mathbb{Z}_{/14} \oplus \mathbb{Z}_{/15}$, given by: $\phi([z]_{210}) = ([z]_{14}, [z]_{15})$.

Part **(c)** of the Chinese Remainder Theorem says that *any finite cyclic group can be written as a direct sum of prime power cyclic groups*. This is a special case of the *Fundamental Theorem of Finitely Generated Abelian Groups*, which we will see in §4.3.

Example 44:

(a) $12 = 4 \times 3 = 2^2 \times 3$. Therefore, $\mathbb{Z}_{/12} \cong \mathbb{Z}_{/4} \oplus \mathbb{Z}_{/3}$.

(b) $210 = 2 \times 3 \times 5 \times 7$. Therefore, $\mathbb{Z}_{/210} \cong \mathbb{Z}_{/2} \oplus \mathbb{Z}_{/3} \oplus \mathbb{Z}_{/5} \oplus \mathbb{Z}_{/7}$.

(c) $720 = 16 \times 9 \times 5 = 2^4 \times 3^2 \times 5$. Therefore, $\mathbb{Z}_{/720} \cong \mathbb{Z}_{/16} \oplus \mathbb{Z}_{/9} \oplus \mathbb{Z}_{/5}$.

Important: The Chinese Remainder Theorem does *not* say that $\mathbb{Z}_{/nm} \cong \mathbb{Z}_{/n} \oplus \mathbb{Z}_{/m}$ when n and m are *not* relatively prime. For example, although $12 = 2 \times 6$, it is *not* true that $\mathbb{Z}_{/12} = \mathbb{Z}_{/2} \oplus \mathbb{Z}_{/6}$. Likewise, it is *not* true that $\mathbb{Z}_{/8} = \mathbb{Z}_{/2} \oplus \mathbb{Z}_{/2} \oplus \mathbb{Z}_{/2}$.

The story of Zhu Geliang comes from the following application of Theorem 42(b):

Corollary 45 *Let $n_1, n_2, \ldots, n_K \in \mathbb{N}$ be pairwise relatively prime and let $n = n_1 \cdot n_2 \cdots n_K$.*

Given any numbers $z_1 \in [0..n_1)$, $z_2 \in [0..n_2)$, $\ldots$, $z_K \in [0..n_K)$, there is a unique $z \in [0..n)$ such that $z \equiv z_1 \pmod{n_1}$, $z \equiv z_2 \pmod{n_2}$, $\ldots$, $z \equiv z_K \pmod{n_K}$. □

Example 46: In the story of Zhu Geliang, $N = 17017 = 7 \times 11 \times 13 \times 17$, $z_1 = 6$, $z_2 = 7$, $z_3 = 5$, and $z_4 = 2$. Thus, Zhu knows there is a unique $z \in [0...17016]$ so that

$$z \equiv 6 \pmod{7}, \quad z \equiv 7 \pmod{11}, \quad z \equiv 5 \pmod{13}, \text{ and } \; z \equiv 2 \pmod{17}.$$

To be precise, $z = 14384$.

In Example ⟨46⟩, how does Zhu Geliang know that $z = 14384$? The Chinese Remainder Theorem states that the solution z exists, but it does not say how to compute it. To compute z, we need to invert the isomorphism ϕ in Theorem 42(b).

Proposition 47 Inversion Formula

Let $N = n_1 \cdot n_2 \cdots n_K$ as in Theorem 42. Define:

$$\begin{array}{rclcl c rcll}
a_1 & = & N/n_1 & = & n_2 \cdot n_3 \cdot \ldots \cdot n_K & \text{and} & b_1 & = & a_1^{-1} & \pmod{n_1} \\
a_2 & = & N/n_2 & = & n_1 \cdot n_3 \cdot \ldots \cdot n_K & \text{and} & b_2 & = & a_2^{-1} & \pmod{n_2} \\
 & \vdots & & & & & & & & \\
a_K & = & N/n_K & = & n_1 \cdot n_2 \cdot \ldots \cdot n_{K-1} & \text{and} & b_K & = & a_K^{-1} & \pmod{n_K}
\end{array}$$

(By this, I mean that $a_1 b_1 \equiv 1 \pmod{n_1}$, $a_2 b_2 \equiv 1 \pmod{n_2}$, etc.)

Now, define $e_1 = a_1 b_1$, $e_2 = a_2 b_2$, $\ldots$, $e_K = a_K b_K$. Finally, let Now define $z = z_1 e_1 + z_2 e_2 + \ldots z_K e_K$. Then $z \equiv z_1 \pmod{n_1}$, $z \equiv z_2 \pmod{n_2}$, $\ldots$, $z \equiv z_K \pmod{n_K}$.

Proof: By construction:

$$\begin{array}{rcll rcll rcll c rcll}
e_1 & \equiv & 1 & \pmod{n_1}; & e_1 & \equiv & 0 & \pmod{n_2} & e_1 & \equiv & 0 & \pmod{n_3} & \ldots & e_1 & \equiv & 0 & \pmod{n_K} \\
e_2 & \equiv & 0 & \pmod{n_1}; & e_2 & \equiv & 1 & \pmod{n_2} & e_2 & \equiv & 0 & \pmod{n_3} & \ldots & e_2 & \equiv & 0 & \pmod{n_K} \\
e_3 & \equiv & 0 & \pmod{n_1}; & e_3 & \equiv & 0 & \pmod{n_2} & e_3 & \equiv & 1 & \pmod{n_3} & \ldots & e_3 & \equiv & 0 & \pmod{n_K} \\
 & \vdots & & & & \vdots & & & & \vdots & & & \ddots & & \vdots & & \\
e_K & \equiv & 0 & \pmod{n_1}; & e_K & \equiv & 0 & \pmod{n_2} & e_K & \equiv & 0 & \pmod{n_3} & \ldots & e_K & \equiv & 1 & \pmod{n_K}
\end{array}$$

Thus,

$$\begin{array}{rcccccccccl}
z_1 e_1 + z_2 e_2 + \ldots z_K e_K & \equiv & z_1 \cdot 1 & + & z_2 \cdot 0 & + & \ldots & + & z_K \cdot 0 & \equiv & z_1 \quad \pmod{n_1} \\
z_1 e_1 + z_2 e_2 + \ldots z_K e_K & \equiv & z_1 \cdot 0 & + & z_2 \cdot 1 & + & \ldots & + & z_K \cdot 0 & \equiv & z_1 \quad \pmod{n_2} \\
 & \vdots & & & & & & & & & \\
z_1 e_1 + z_2 e_2 + \ldots z_K e_K & \equiv & z_1 \cdot 0 & + & z_2 \cdot 0 & + & \ldots & + & z_K \cdot 1 & \equiv & z_1 \quad \pmod{n_K}
\end{array}$$

□

The Chinese Remainder Theorem is really a theorem about *abelian* groups[2], but it has the following partial generalization to nonabelian groups:

[2] Later we will also see versions of it for *rings* and for *modules*.

Proposition 48 *Let $\mathcal{G}$ be a (possibly nonabelian) group, with $|\mathcal{G}| = N = n_1 \cdot n_2 \cdots n_K$, where $n_1, n_2, \ldots, n_K$ are pairwise relatively prime. Suppose there are normal subgroups $\mathcal{N}_1, \mathcal{N}_2, \ldots, \mathcal{N}_K \lhd \mathcal{G}$ so that $|\mathcal{N}_k| = n_k$. Then $\mathcal{G} \cong \mathcal{N}_1 \times \mathcal{N}_2 \times \ldots \times \mathcal{N}_K$.*

Proof: (by induction)

Base Case ($K = 2$) We will apply Proposition 41(c).

Claim 1: $\mathcal{N}_1 \cap \mathcal{N}_2 = \{e\}$.

Proof: If $g \in \mathcal{N}_1$, then $|g|$ must divide $|\mathcal{N}_1| = n_1$ Likewise, if $g \in \mathcal{N}_2$, then $|g|$ divides n_2. Thus, $|g|$ divides $\gcd(n_1, n_2) = 1$. Thus, $|g| = 1$, so $g = e$. .... □ `[Claim 1]`

Claim 2: $\mathcal{G} = \mathcal{N}_1 \cdot \mathcal{N}_2$.

Proof: Clearly, $\mathcal{N}_1 \cdot \mathcal{N}_2$ is a subset $\mathcal{G}$. We claim that $\mathcal{N}_1 \cdot \mathcal{N}_2$ has the same cardinality as $\mathcal{G}$. To see this, observe that

$$\left|\mathcal{N}_1 \cdot \mathcal{N}_2\right| \underset{(*)}{=\!=} \frac{|\mathcal{N}_1| \cdot |\mathcal{N}_2|}{|\mathcal{N}_1 \cap \mathcal{N}_2} \underset{\text{Claim 1}}{=\!=\!=} \frac{n_1 \cdot n_2}{1} = n = |\mathcal{G}|.$$

where $(*)$ follows from Prop 13, §3.2, p.94 in Dummit & Foote. .... □ `[Claim 2]`

Now combine Claims 1 and 2 with Proposition 41(c).

Induction: Observe that, since $\mathcal{N}_2, \mathcal{N}_3, \ldots, \mathcal{N}_K \lhd \mathcal{G}$, the product $\mathcal{M} = \mathcal{N}_2 \cdot \mathcal{N}_3 \cdots \mathcal{N}_K$ is a subgroup of $\mathcal{G}$. By induction hypothesis,

$$\mathcal{M} \cong \mathcal{N}_2 \times \mathcal{N}_3 \times \ldots \times \mathcal{N}_K. \tag{4.1}$$

Now, let $m = |\mathcal{M}| = n_2 \cdot n_3 \cdots n_K$. Then m is relatively prime to n_1 (because each of n_2, $n_3, \ldots, n_K$ is relatively prime to n_1). Thus, by applying the **Base Case** (with $\mathcal{M}$ playing the role of $\mathcal{N}_2$) we have: $\mathcal{G} \cong \mathcal{N}_1 \times \mathcal{M} \underset{\text{eqn.(4.1)}}{=\!=\!=} \mathcal{N}_1 \times \mathcal{N}_2 \times \mathcal{N}_3 \times \ldots \times \mathcal{N}_K$. ________□

Corollary 49 *Let $\mathcal{G}$ be a (possibly nonabelian) group. Suppose $|\mathcal{G}| = n$, and let n have prime factorization: $n = p_1^{\nu_1} p_2^{\nu_2} \ldots p_K^{\nu_K}$. Suppose there are subgroups[3] $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_K < \mathcal{G}$ so that $|\mathcal{P}_k| = p_k^{\nu_k}$.*

If $\mathcal{P}_1, \ldots, \mathcal{P}_K$ are normal subgroups of $\mathcal{G}$, then $\mathcal{G} \cong \mathcal{P}_1 \times \mathcal{P}_2 \times \ldots \times \mathcal{P}_K$.

Proof: Apply Proposition 48, with $n_1 = p_1^{\nu_1}$, $n_2 = p_2^{\nu_2}$, $\ldots, n_K = p_K^{\nu_K}$ ____________□

[3] These are called **Sylow** subgroups.

Another proof of part (c) of the Chinese Remainder Theorem: Let $\mathcal{G} = \mathbb{Z}_{/N}$, and for all $k \in [1..K]$, let $m_k = n/p_k^{\nu_k}$. Let $\overline{m}_k \in \mathcal{G}$ be the associated congruence class.

Claim 1: $|\overline{m}_k| = p_k^{\nu_k}$.

Proof: (**Exercise 21**) . . . . □ [Claim 1]

Let $\mathcal{P}_k = \langle \overline{m}_k \rangle$ be the subgroup generated by $\overline{m}_k$. It follows from Claim 1 that $|\mathcal{P}_k| = p_k^{\nu_k}$. Thus, Corollary 49 says that $\mathcal{G} = \mathcal{P}_1 \oplus \mathcal{P}_2 \oplus \ldots \oplus \mathcal{P}_K$. □

4.3 The Fundamental Theorem of Finitely Generated Abelian Groups

Prerequisites: §4.2, §3.5

Theorem 50 *Let $\mathcal{A}$ be a finitely generated abelian group. Then there are unqiue prime powers $p_1^{\alpha_1}, p_2^{\alpha_2}, \ldots, p_n^{\alpha_N}$ (where $p_1, p_2, \ldots, p_N$ are prime numbers, not necessarily distinct, and $\alpha_1, \ldots, \alpha_N$ are natural numbers), and a unique integer $R \geq 0$ so that*

$$\mathcal{A} \quad \cong \quad \mathbb{Z}^R \oplus \mathbb{Z}_{/(p_1^{\alpha_1})} \oplus \mathbb{Z}_{/(p_2^{\alpha_2})} \oplus \ldots \oplus \mathbb{Z}_{/(p_N^{\alpha_N})}.$$

Example 51: The following groups typify the Fundamental Theorem:

$\mathbb{Z}^{15}$; $\quad \mathbb{Z}_{/2} \oplus \mathbb{Z}_{/2} \oplus \mathbb{Z}_{/2}$; $\quad \mathbb{Z}_{/2} \oplus \mathbb{Z}_{/4} \oplus \mathbb{Z}_{/8}$; $\quad \mathbb{Z}^5 \oplus \mathbb{Z}_{/2} \oplus \mathbb{Z}_{/4} \oplus \mathbb{Z}_{/3} \oplus \mathbb{Z}_{/81} \oplus \mathbb{Z}_{/25} \oplus \mathbb{Z}_{/49} \oplus \mathbb{Z}_{/17}$.

On the other hand, the abelian group $(\mathbb{Q}, +)$ is *not* finitely generated, and therefore does not admit a decomposition of this kind. Instead,

$$\mathbb{Q} \quad = \quad \mathbb{Q}_2 + \mathbb{Q}_3 + \mathbb{Q}_5 + \mathbb{Q}_7 + \mathbb{Q}_{11} + \mathbb{Q}_{13} + \ldots$$

In other words, $\mathbb{Q}$ is a sum of the p-adic rational numbers, for all prime p. Note that this sum is *not* direct, because $\mathbb{Q}_2 \cap \mathbb{Q}_3 \cap \mathbb{Q}_5 \cap \mathbb{Q}_7 \cap \mathbb{Q}_{11} \cap \mathbb{Q}_{13} \cap \ldots \quad = \quad \mathbb{Z}$.

To obtain something more like a direct sum, let $\widetilde{\mathbb{Q}} = \mathbb{Q}/\mathbb{Z}$. If you think of the unit circle as $\mathbb{R}/\mathbb{Z}$, then you can imagine $\widetilde{\mathbb{Q}}$ as the subgroup of 'rational angles' on the circle. Likewise, let $\widetilde{\mathbb{Q}}_p = \mathbb{Q}_p/\mathbb{Z}$ for all prime p. Then

$$\widetilde{\mathbb{Q}} \quad = \quad \widetilde{\mathbb{Q}}_2 \oplus \widetilde{\mathbb{Q}}_3 \oplus \widetilde{\mathbb{Q}}_5 \oplus \widetilde{\mathbb{Q}}_7 \oplus \widetilde{\mathbb{Q}}_{11} \oplus \widetilde{\mathbb{Q}}_{13} \oplus \ldots$$

and $\mathbb{Q} = \widetilde{\mathbb{Q}} \rtimes \mathbb{Z}$ is a *semidirect product* of $\widetilde{\mathbb{Q}}$ with $\mathbb{Z}$.

Proof of the Fundamental Theorem: By the Universal Covering Property (Corollary 30 on page 23), find a free abelian group $\mathcal{F}$ and an epimorphism $\phi : \mathcal{F} \longrightarrow \mathcal{A}$. Let $\mathcal{K} = \ker(\phi)$; thus,

$$\mathcal{A} \quad \cong \quad \mathcal{F}/\mathcal{K}. \tag{4.2}$$

By Proposition 33 on page 24, $\mathcal{K}$ is also a free abelian group, and there is a basis $\{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_R\}$ for $\mathcal{F}$, and numbers $m_1, m_2, \ldots, m_S \in \mathbb{N}$ so that $\{(m_1\mathbf{a}_1),\ (m_2\mathbf{a}_2),\ \ldots, (m_S\mathbf{a}_S)\}$ is a basis for $\mathcal{K}$.

$$\begin{array}{rrcllllllllll}
\text{In other words,} & \mathcal{F} & = & \mathbb{Z}\mathbf{a}_1 & \oplus & \mathbb{Z}\mathbf{a}_2 & \oplus\ldots\oplus & \mathbb{Z}\mathbf{a}_S & \oplus & \mathbb{Z}\mathbf{a}_{S+1} & \oplus\ldots\oplus & \mathbb{Z}\mathbf{a}_R, \\
\text{and} & \mathcal{K} & = & \mathbb{Z}m_1\mathbf{a}_1 & \oplus & \mathbb{Z}m_2\mathbf{a}_2 & \oplus\ldots\oplus & \mathbb{Z}m_S\mathbf{a}_S & \oplus & \{0\} & \oplus\ldots\oplus & \{0\}. \\
\text{Thus,} & \mathcal{F}/\mathcal{K} & \cong & \mathbb{Z}/m_1\mathbb{Z} & \oplus & \mathbb{Z}/m_2\mathbb{Z} & \oplus\ldots\oplus & \mathbb{Z}/m_S\mathbb{Z} & \oplus & \mathbb{Z} & \oplus\ldots\oplus & \mathbb{Z} \\
 & & \cong & \mathbb{Z}_{/m_1} & \oplus & \mathbb{Z}_{/m_2} & \oplus\ldots\oplus & \mathbb{Z}_{/m_S} & \oplus & \mathbb{Z} & \oplus\ldots\oplus & \mathbb{Z}
\end{array}$$

In other words, if $R' = R - S$, then

$$\mathcal{F}/\mathcal{K} \quad \cong \quad \mathbb{Z}_{/m_1} \oplus \mathbb{Z}_{/m_2} \oplus \ldots \oplus \mathbb{Z}_{/m_S} \oplus \mathbb{Z}^{R'}. \tag{4.3}$$

Now, suppose that m_1 has prime factorization $m_1 = p_1^{\alpha_1} \cdot p_2^{\alpha_2} \cdots p_i^{\alpha_i}$. Then the part (c) of the Chinese Remainder Theorem (Theorem 42 on page 33) says

$$\mathbb{Z}_{/m_1} \quad \cong \quad \mathbb{Z}_{/\left(p_1^{\alpha_1}\right)} \oplus \mathbb{Z}_{/\left(p_2^{\alpha_2}\right)} \oplus \ldots \oplus \mathbb{Z}_{/\left(p_i^{\alpha_i}\right)}.$$

Likewise, if $m_2 = p_{i+1}^{\alpha_{i+1}} \cdot p_{i+2}^{\alpha_{i+2}} \cdots p_j^{\alpha_j}$. then $\mathbb{Z}_{/m_1} \ \cong\ \mathbb{Z}_{/\left(p_{i+1}^{\alpha_{i+1}}\right)} \oplus \mathbb{Z}_{/\left(p_{i+2}^{\alpha_{i+2}}\right)} \oplus \ldots \oplus \mathbb{Z}_{/\left(p_j^{\alpha_j}\right)}$. Proceeding this way, suppose that $m_1, \ldots, m_S$ collectively have prime factorizations:

$$\begin{array}{rcl}
m_1 & = & p_1^{\alpha_1} \cdot p_2^{\alpha_2} \cdots p_i^{\alpha_i} \\
m_2 & = & p_{i+1}^{\alpha_{i+1}} \cdot p_{i+2}^{\alpha_{i+2}} \cdots p_j^{\alpha_j} \\
 & \vdots & \\
m_S & = & p_k^{\alpha_k} \cdot p_{k+1}^{\alpha_{k+1}} \cdots p_N^{\alpha_N}
\end{array}$$

(where the prime number sets $\{p_1, \ldots, p_i\}$, $\{p_{i+1}, \ldots, p_j\}$, ..., $\{p_k, \ldots, p_N\}$ are not necessarily disjoint)

Then:

$$\begin{aligned}
\mathbb{Z}_{/m_1} \oplus \mathbb{Z}_{/m_2} \oplus \ldots \oplus \mathbb{Z}_{/m_S} \quad = \quad & \left(\mathbb{Z}_{/\left(p_1^{\alpha_1}\right)} \oplus \ldots \oplus \mathbb{Z}_{/\left(p_i^{\alpha_i}\right)}\right) \oplus \left(\mathbb{Z}_{/\left(p_{i+1}^{\alpha_{i+1}}\right)} \oplus \ldots \oplus \mathbb{Z}_{/\left(p_j^{\alpha_j}\right)}\right) \oplus \ldots \\
& \ldots \oplus \left(\mathbb{Z}_{/\left(p_k^{\alpha_k}\right)} \oplus \ldots \oplus \mathbb{Z}_{/\left(p_N^{\alpha_N}\right)}\right).
\end{aligned} \tag{4.4}$$

Now, combine equations (4.2,4.3,4.4) to conclude: $\mathcal{A} \ \cong\ \mathbb{Z}_{/\left(p_1^{\alpha_1}\right)} \oplus \ldots \oplus \mathbb{Z}_{/\left(p_N^{\alpha_N}\right)} \oplus \mathbb{Z}^{R'}$. □

Example: To illustrate the reasoning of the proof, suppose that $\mathcal{F} = \mathbb{Z}^3$, so that $R = 3$. Suppose $S = 2$, and that

$$\begin{array}{rclllcl}
\mathbf{a}_1 & = & (1,0,0) & \text{and} \quad m_1 = 60, & \text{so that} & m_1\mathbf{a}_1 & = & (60,\ 0,\ 0); \\
\mathbf{a}_2 & = & (0,1,0) & \text{and} \quad m_2 = 80, & \text{so that} & m_2\mathbf{a}_2 & = & (0,\ 80,\ 0); \\
\mathbf{a}_3 & = & (0,0,1). & & & & &
\end{array}$$

Thus, $\mathcal{K} \quad = \quad \{(60z_1,\ 80z_2,\ 0)\ ;\ z_1, z_2 \in \mathbb{Z}\} \quad = \quad (60\mathbb{Z}) \oplus (80\mathbb{Z}) \oplus \{0\}$. Thus,

$$\mathcal{A} \quad \cong \quad \mathcal{F}/\mathcal{K} \quad = \quad \frac{\mathbb{Z} \quad \oplus \quad \mathbb{Z} \quad \oplus \mathbb{Z}}{(60\mathbb{Z}) \oplus (80\mathbb{Z}) \oplus \{0\}} \quad \cong \quad \mathbb{Z}_{/60} \ \oplus\ \mathbb{Z}_{/80} \ \oplus\ \mathbb{Z}.$$

Now, we apply the Chinese Remainder Theorem:

$$\begin{array}{rclclcrcll}
60 & = & 4 \times 3 \times 5 & = & 2^2 \times 3 \times 5, & \text{so} & \mathbb{Z}_{/60} & \cong & \mathbb{Z}_{/3} \oplus \mathbb{Z}_{/4} \oplus \mathbb{Z}_{/5}; & \text{ie. } \{p_1, p_2, p_3\} = \{2,3,5\}. \\
80 & = & 16 \times 5 & = & 2^4 \times 5, & \text{so} & \mathbb{Z}_{/80} & \cong & \mathbb{Z}_{/16} \oplus \mathbb{Z}_{/5}; & \text{ie. } \{p_4, p_5\} = \{2,5\}.
\end{array}$$

Observe that the set $\{2,3,5\}$ is not disjoint from the set $\{2,5\}$, but we can still proceed to conclude that:

$$\mathcal{A} \quad \cong \quad \mathbb{Z} \ \oplus\ \mathbb{Z}_{/3} \oplus \mathbb{Z}_{/4} \oplus \mathbb{Z}_{/5} \quad \oplus \quad \mathbb{Z}_{/16} \oplus \mathbb{Z}_{/5}.$$

Chapter 5

The Structure Theory of Nonabelian Groups

5.1 Three Examples

Prerequisites: §1.4

Example (A) The arithmetic quotient equation

$$\frac{6}{3} \quad = \quad 2 \tag{5.1}$$

is equivalent to the factorization:

$$6 \quad = \quad 2 \cdot 3. \tag{5.2}$$

Can we factor the numbers 2 and 3 any further? No, because they are prime. The equation $6 = 2 \times 3$ is a *prime factorization* of 6. Prime numbers are the 'elementary components' of number theory: any number can be written in a unique way as a product of prime numbers.

The idea of *group structure theory* is to 'factor' groups into 'elementary components' in a similar fashion.

Example (B) Consider the cyclic group $\mathbb{Z}_{/6} = \{\bar{0},\ \bar{1},\ \bar{2},\ \bar{3},\ \bar{4},\ \bar{5}\}$, and the subgroup $\mathcal{A} = \{\bar{0},\ \bar{2},\ \bar{4}\}$. It is easy to check that $\mathcal{A} \lhd \mathbb{Z}_{/6}$, and that

$$\frac{\mathbb{Z}_{/6}}{\mathcal{A}} \quad \cong \quad \mathbb{Z}_{/2}. \tag{5.3}$$

Observe that $\mathcal{A} \cong \mathbb{Z}_{/3}$. Thus, we suspect

$$\mathbb{Z}_{/6} \quad \cong \quad \mathbb{Z}_{/3} \times \mathbb{Z}_{/2}, \tag{5.4}$$

and indeed, this is true. To see this, define $\phi : \mathbb{Z}_{/3} \times \mathbb{Z}_{/2} \longrightarrow \mathbb{Z}_{/6}$ by $\phi\left([n]_3,\ [m]_2\right) \ = \ \overline{2n+3m}$. (**Exercise 22** Check that ϕ is an isomorphism. Here, $[n]_3$ is the mod-3 congruence class of n, etc.)

Thus, the quotient equation (5.3) entails a *factorization* (5.4) of the group $\mathbb{Z}_{/6}$, just as the quotient equation (5.1) entailed a factorization (5.2) of the integer 6.

Can we break down $\mathbb{Z}_{/3}$ any further? No, because $\mathbb{Z}_{/3}$ contains no normal subgroups, so it cannot be 'factored' in this way. Likewise, $\mathbb{Z}_{/2}$ contains no normal subgroups. Thus, $\mathbb{Z}_{/3}$ and $\mathbb{Z}_{/2}$ are analogous to prime numbers. We say that $\mathbb{Z}_{/3}$ and $\mathbb{Z}_{/2}$ are **simple**.

Example (C) Now, consider the symmetric group $\mathbf{S}_3 \ = \ \Big\{e,\ (12),\ (13),\ (23),\ (123),\ (132)\Big\}$, and the subgroup $\mathbf{A}_3 = \Big\{e,\ (123),\ (132)\Big\}$. We saw in Example ⟨8d⟩ (page 7) that $\mathbf{A}_3 \lhd \mathbf{S}_3$, and we saw in Example ⟨13c⟩ (page 11) that

$$\frac{\mathbf{S}_3}{\mathbf{A}_3} \quad \cong \quad \mathbb{Z}_{/2}. \tag{5.5}$$

Observe that $\mathbf{A}_3 = \Big\{e,\ (123),\ (132)\Big\}$ is isomorphic to $\mathbb{Z}_{/3}$. Thus, we suspect $\mathbf{S}_3 \ \cong \ \mathbb{Z}_{/3} \times \mathbb{Z}_{/2}$. However, this is *false*, because $\mathbb{Z}_{/3} \times \mathbb{Z}_{/2}$ is abelian, but $\mathbf{S}_3$ is *not* abelian. So if $\mathbf{S}_3$ is a 'product' of $\mathbb{Z}_{/3}$ and $\mathbb{Z}_{/2}$, it is only a product in a metaphorical sense[1]. but we can still write:

$$\mathbf{S}_3 \quad \cong \quad \mathbb{Z}_{/3} \rtimes \mathbb{Z}_{/2}, \tag{5.6}$$

as long as we interpret the symbol "$\rtimes$" in the appropriate way. Thus, the quotient equation (5.5) entails a *factorization* (5.6) of the group $\mathbf{S}_3$, just as the quotient equation (5.3) entailed a factorization (5.4) of the group $\mathbb{Z}_{/6}$.

Examples **(B)** and **(C)** show an important difference between factoring groups and factoring numbers. Any number with prime factorization "2×3" must be equal to 6. But two *different* groups can both 'factor' into the same elementary components, without being the same. The groups $\mathbb{Z}_{/6}$ and $\mathbf{S}_3$ *both* factor into elementary components $\mathbb{Z}_{/3}$ and $\mathbb{Z}_{/3}$, but $\mathbb{Z}_{/6} \not\cong \mathbf{S}_3$.

5.2 Simple Groups

Prerequisites: §1.4, §1.2 **Recommended:** §5.1

A group $\mathcal{G}$ is **simple** if it has no nontrivial proper normal subgroups. For example, $\mathbb{Z}_{/3}$ is simple (because the only proper subgroup of $\mathbb{Z}_{/3}$ is $\{0\}$, which is trivial). Simple groups are the 'elementary components' of group theory, and are analogous to prime numbers. Indeed, we have the following:

Proposition 52 *For any $n \in \mathbb{N}$,* $\Big(\ \mathbb{Z}_{/n}$ *is simple* $\Big) \iff \Big(\ n$ *is prime* $\Big)$.

[1] In fact, $\mathbf{S}_3$ is actually a *semidirect product* of $\mathbb{Z}_{/3}$ and $\mathbb{Z}_{/2}$.

Proof: '$\Longleftarrow$' Let $\mathcal{A}$ be any subgroup of $\mathbb{Z}_{/n}$. Then Lagrange's Theorem (Theorem 7 on page 6) says $|\mathcal{A}|$ divides $|\mathbb{Z}_{/n}| = n$. But n is prime, so either $|\mathcal{A}| = 1$ (in which case $\mathcal{A} = \{\bar{0}\}$) or $|\mathcal{A}| = n$ (in which case $\mathcal{A} = \mathbb{Z}_{/n}$).

Thus, $\mathcal{A}$ has no nontrivial proper subgroups —in particular, $\mathcal{A}$ has no nontrivial proper *normal* subgroups. Hence, $\mathcal{A}$ is simple.

'$\Longrightarrow$' Suppose n was *not* prime; then $n = k{\cdot}m$ for some numbers k and m, with $1 < k, m < n$. Thus, the subgroup $\langle \overline{m} \rangle = \left\{ \bar{0},\ \overline{m},\ \overline{2m},\ \ldots,\ \overline{(k-1)m} \right\}$ is a subgroup with k elements, so it is a nontrivial subgroup of $\mathbb{Z}_{/n}$. But $\mathbb{Z}_{/n}$ is abelian, so $\langle \overline{m} \rangle$ is automatically a *normal* subgroup. Hence, $\mathbb{Z}_{/n}$ cannot be simple. □

In the proof of Proposition 52, we implicitly used the following result:

Lemma 53 *Let $\mathcal{G}$ be a group. Then:*

(a) $\Big($ *$\mathcal{G}$ has no nontrivial subgroups* $\Big) \Longrightarrow \Big($ *$\mathcal{G}$ is simple* $\Big)$.

(b) *If $\mathcal{G}$ is* abelian, *then* $\Big($ *$\mathcal{G}$ has no nontrivial subgroups* $\Big) \Longleftrightarrow \Big($ *$\mathcal{G}$ is simple* $\Big)$.

Proof: **Exercise 23** □

Observe that the 'if and only if' is *not* true if $\mathcal{G}$ is not abelian. It isn't obvious, because we haven't yet seen any examples of *nonabelian* simple groups. We will show in §6.3 that the alternating group $\mathbf{A}_N$ is simple for all $N \geq 5$. Thus, $\mathbf{A}_5$ is an example of a simple group with many nontrivial subgroups.

Lemma 54 *Let $\mathcal{G}$ be a simple group. If $\phi : \mathcal{G} \longrightarrow \mathcal{H}$ is any group homomorphism, then either ϕ is trivial or ϕ is a monomorphism.*

Proof: **Exercise 24** □

5.3 Composition Series; The Hölder program

Prerequisites: §5.2, §2.1

Let $\mathcal{G}$ be a group. A **normal series** is a sequence of subgroups:

$$\mathcal{N}_1 \ \lhd\ \mathcal{N}_2 \ \lhd\ \mathcal{N}_3 \ \lhd \ldots \ \lhd\ \mathcal{N}_J \ \lhd\ \mathcal{G}.$$

Note that $\mathcal{N}_1$ is normal in $\mathcal{N}_2$, but it is not necessarily true that $\mathcal{N}_1$ is normal in $\mathcal{G}$. Likewise, for every $j \in [1..J)$, $\mathcal{N}_j$ is normal in $\mathcal{N}_{j+1}$, but not necessarily normal in $\mathcal{G}$.

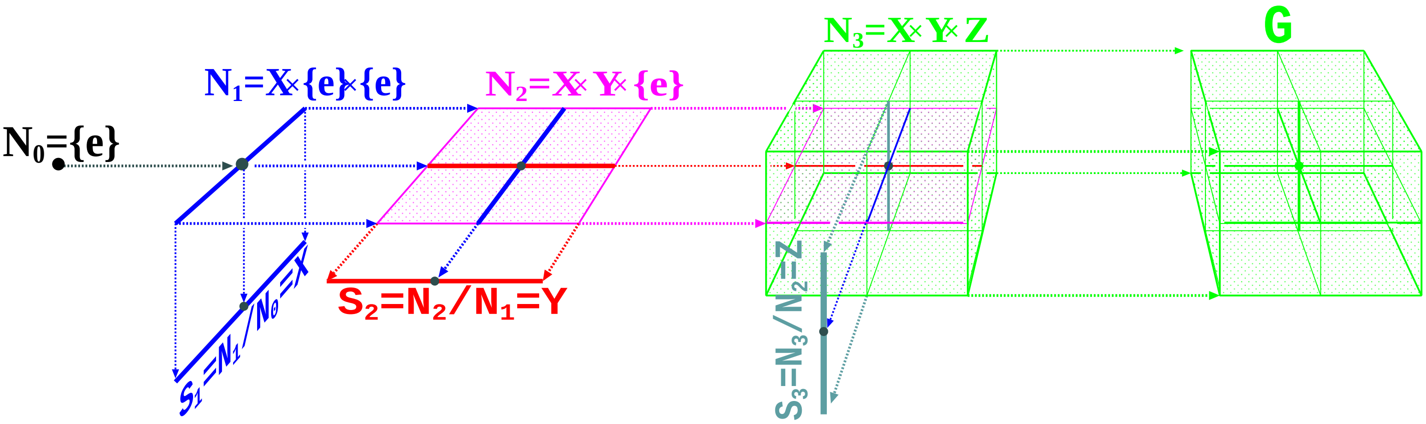


Figure 5.1: Example ⟨56a⟩

Example 55: Let $\mathcal{G} = \mathbb{Z}_{/120}$, and let

$$\begin{array}{rclcl}
\mathcal{N}_3 &=& \langle\bar{2}\rangle &=& \{\bar{0}, \bar{2},\ \bar{4},\ \bar{6},\ \bar{8}, \ldots, \overline{116},\ \overline{118}\}; \\
\mathcal{N}_2 &=& \langle\overline{12}\rangle &=& \{\bar{0}, \overline{12}, \overline{24}, \overline{36}, \overline{48}, \overline{60}, \overline{72}, \overline{84}, \overline{96}, \overline{108}\}; \\
\mathcal{N}_1 &=& \langle\overline{24}\rangle &=& \{\bar{0},\quad \overline{24},\quad \overline{48},\quad \overline{72},\quad \overline{96}\}.
\end{array}$$

Thus, $\mathcal{N}_1 \lhd \mathcal{N}_2 \lhd \mathcal{N}_3 \lhd \mathbb{Z}_{/120}$.

A **composition series** is a normal series

$$\{e_{\mathcal{G}}\} \;=\; \mathcal{N}_0 \;\lhd\; \mathcal{N}_1 \;\lhd\; \mathcal{N}_2 \;\lhd \ldots \;\lhd\; \mathcal{N}_J \;=\; \mathcal{G}, \tag{5.7}$$

such that the quotient groups $\left\{\begin{array}{rcl} \mathcal{S}_1 &=& \mathcal{N}_1/\mathcal{N}_0 \;=\; \mathcal{N}_1 \\ \mathcal{S}_2 &=& \mathcal{N}_2/\mathcal{N}_1 \\ \mathcal{S}_3 &=& \mathcal{N}_3/\mathcal{N}_2 \\ & \vdots & \\ \mathcal{S}_J &=& \mathcal{N}_J/\mathcal{N}_{J-1} \;=\; \mathcal{G}/\mathcal{N}_{J-1} \end{array}\right\}$ are all *simple* groups.

We say this series has **rank** J. The groups $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_J$ are called **composition factors**, and the set $\{\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_J\}$ is called a **composition factor set**[2].

The idea is that $\mathcal{G}$ is in some way a 'product' of its composition factors:

$$\mathcal{G} \quad = \quad \mathcal{S}_1 \boxtimes \mathcal{S}_2 \leftthreetimes \mathcal{S}_3 \rightthreetimes \mathcal{S}_4 \odot \ldots \circledast \mathcal{S}_J$$

where '$\boxtimes$', '$\leftthreetimes$', '$\rightthreetimes$', etc. represent some ways of combining groups together (eg. direct product, semidirect product, etc.). Thus, a composition series is for a group what a *prime factorization* is for an integer.

Example 56:

[2] Technically, this is a *multiset*, because we allow the same element to appear more than once.

(a) Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be simple groups, and let $\mathcal{G} = \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$. As shown in Figure 5.1, define:

$$\begin{array}{rcccccccl} \mathcal{N}_3 & = & \mathcal{X} & \times & \mathcal{Y} & \times & \mathcal{Z} & = & \big\{(x, y, c, e)\ ;\ x \in \mathcal{X},\ y \in \mathcal{Y} \text{ and } \ z \in \mathcal{Z}\big\} \\ \mathcal{N}_2 & = & \mathcal{X} & \times & \mathcal{Y} & \times & \{e_{\mathcal{Z}}\} & = & \big\{(x, y, e_{\mathcal{Z}})\ ;\ x \in \mathcal{X} \text{ and } \ y \in \mathcal{Y}\big\} \\ \mathcal{N}_1 & = & \mathcal{X} & \times & \{e_{\mathcal{Y}}\} & \times & \{e_{\mathcal{Z}}\} & = & \big\{(x, e_{\mathcal{Y}}, e_{\mathcal{Z}})\ ;\ x \in \mathcal{X}\big\} \\ \mathcal{N}_0 & = & \{e_{\mathcal{X}}\} & \times & \{e_{\mathcal{Y}}\} & \times & \{e_{\mathcal{Z}}\} & = & \{e_{\mathcal{G}}\} \end{array}$$

Then this is a composition series for $\mathcal{G}$, with composition factors:

$$\begin{array}{rclcl} \mathcal{S}_3 & = & \mathcal{N}_3/\mathcal{N}_2 & \cong & \mathcal{Z}; \\ \mathcal{S}_2 & = & \mathcal{N}_2/\mathcal{N}_1 & \cong & \mathcal{Y}; \\ \mathcal{S}_1 & = & \mathcal{N}_1/\mathcal{N}_0 & \cong & \mathcal{X}. \end{array}$$

(b) Let $\mathcal{G} = \mathbb{Z}_{/12}$, and define:

$$\begin{array}{rclcl} \mathcal{N}_3 & = & \mathbb{Z}_{/12} & = & \{\bar{0},\ \bar{1},\ \ \bar{2},\ \bar{3},\ \ \bar{4},\ \bar{5}\ \ \bar{6},\ \bar{7},\ \ \bar{8},\ \bar{9},\ \ \overline{10},\ \overline{11}\}; \\ \mathcal{N}_2 & = & \langle\bar{2}\rangle & = & \{\bar{0},\ \ \bar{2},\ \ \bar{4},\ \ \bar{6},\ \ \bar{8},\ \ \overline{10}\}; \\ \mathcal{N}_1 & = & \langle\bar{4}\rangle & = & \{\bar{0},\ \ \bar{4},\ \ \bar{8}\}; \\ \mathcal{N}_0 & = & \{\bar{0}\}. & & \end{array}$$

Then this is a composition series for $\mathbb{Z}_{/12}$, with composition factors:

$$\begin{array}{rclcl} \mathcal{S}_3 & = & \mathcal{N}_3/\mathcal{N}_2 & \cong & \mathbb{Z}_{/2}; \\ \mathcal{S}_2 & = & \mathcal{N}_2/\mathcal{N}_1 & \cong & \mathbb{Z}_{/2}; \\ \mathcal{S}_1 & = & \mathcal{N}_1/\mathcal{N}_0 & \cong & \mathbb{Z}_{/3}. \end{array}$$

(c) Again, let $\mathcal{G} = \mathbb{Z}_{/12}$, but now define:

$$\begin{array}{rclcl} \mathcal{N}_3 & = & \mathbb{Z}_{/12} & = & \{\bar{0},\ \bar{1},\ \bar{2},\ \ \bar{3},\ \bar{4},\ \bar{5}\ \ \bar{6},\ \bar{7},\ \bar{8},\ \ \bar{9}, \overline{10},\ \overline{11}\}; \\ \mathcal{N}_2 & = & \langle\bar{3}\rangle & = & \{\bar{0},\ \ \bar{3},\ \ \bar{6},\ \ \bar{9}, \}; \\ \mathcal{N}_1 & = & \langle\bar{6}\rangle & = & \{\bar{0},\ \ \bar{6}\}; \\ \mathcal{N}_0 & = & \{\bar{0}\}. & & \end{array}$$

Then this is *another* composition series for $\mathbb{Z}_{/12}$, with composition factors:

$$\begin{array}{rclcl} \mathcal{S}_3 & = & \mathcal{N}_3/\mathcal{N}_2 & \cong & \mathbb{Z}_{/3}; \\ \mathcal{S}_2 & = & \mathcal{N}_2/\mathcal{N}_1 & \cong & \mathbb{Z}_{/2}; \\ \mathcal{S}_1 & = & \mathcal{N}_1/\mathcal{N}_0 & \cong & \mathbb{Z}_{/2}. \end{array}$$

(d) The normal series in Example ⟨55⟩ is *not* a composition series (even if we add terms $\mathcal{N}_4 = \mathbb{Z}_{/120}$ and $\mathcal{N}_0 = \{\bar{0}\}$). The reason: $\mathcal{S}_3 = \mathcal{N}_3/\mathcal{N}_2 \cong \mathbb{Z}_{/6}$ is *not* simple. However, we can 'refine' this normal chain into a composition series by inserting the normal subgroup

$$\mathcal{N}_{2\frac{1}{2}} \quad = \quad \langle\bar{6}\rangle \quad = \quad \{\bar{6}, \overline{12}, \overline{18}, \ldots, \overline{124}\}.$$

Now we have a composition series: $\{\bar{0}\} = \mathcal{N}_0 \lhd \mathcal{N}_1 \lhd \mathcal{N}_2 \lhd \mathcal{N}_{2\frac{1}{2}} \lhd \mathcal{N}_3 \lhd \mathcal{N}_4 = \mathbb{Z}_{/120}$,

with composition factors: $$\left\{\begin{array}{rcccl} \mathcal{S}_4 &=& \mathcal{N}_4/\mathcal{N}_3 &\cong& \mathbb{Z}_{/2}; \\ \mathcal{S}_3 &=& \mathcal{N}_3/\mathcal{N}_{2\frac{1}{2}} &\cong& \mathbb{Z}_{/3}; \\ \mathcal{S}_{2\frac{1}{2}} &=& \mathcal{N}_{2\frac{1}{2}}/\mathcal{N}_2 &\cong& \mathbb{Z}_{/2}; \\ \mathcal{S}_2 &=& \mathcal{N}_2/\mathcal{N}_1 &\cong& \mathbb{Z}_{/2}; \\ \mathcal{S}_1 &=& \mathcal{N}_1/\mathcal{N}_0 &\cong& \mathbb{Z}_{/5}. \end{array}\right. .$$

(e) If $\mathcal{G}$ is a *simple* group, then $\mathcal{G}$ has a very short composition series: $\{e_{\mathcal{G}}\} = \mathcal{N}_0 \lhd \mathcal{N}_1 = \mathcal{G}$, with *one* composition factor: $\mathcal{S}_1 = \mathcal{N}_1/\mathcal{N}_0 \cong \mathcal{G}$.

In Examples (56b) and (56c), we constructed two different composition series for $\mathbb{Z}_{/12}$, but they had the same rank, and yielded the *same* composition factor set: $\{\mathbb{Z}_{/2}, \mathbb{Z}_{/2}, \mathbb{Z}_{/3}\}$ (albeit in a different order). This is reassuring; if a composition series is supposed to be a 'prime factorization' of a group, it would be unfortunate if different series yielded different factors for the same group. The following theorem says this is true in general.

Theorem 57 (Jordan-Hölder)

Let $\mathcal{G}$ be a finite group. Then:

(a) *$\mathcal{G}$ has a composition series.*

(b) *Any two composition series for $\mathcal{G}$ have the same rank, and yield the same composition factor set. In other words, if*

$$\begin{array}{rcccccccccccc} & \{e_{\mathcal{G}}\} &=& \mathcal{N}_0 &\lhd& \mathcal{N}_1 &\lhd& \mathcal{N}_2 &\lhd \ldots &\lhd& \mathcal{N}_J &=& \mathcal{G} \\ \text{and} & \{e_{\mathcal{G}}\} &=& \mathcal{M}_0 &\lhd& \mathcal{M}_1 &\lhd& \mathcal{M}_2 &\lhd \ldots &\lhd& \mathcal{M}_K &=& \mathcal{G} \end{array}$$

are two composition series for $\mathcal{G}$, with composition factors

$$\begin{array}{rclcrcl} \mathcal{S}_1 &=& \mathcal{N}_1/\mathcal{N}_0 & & \mathcal{T}_1 &=& \mathcal{M}_1/\mathcal{M}_0 \\ \mathcal{S}_2 &=& \mathcal{N}_2/\mathcal{N}_1 & & \mathcal{T}_2 &=& \mathcal{M}_2/\mathcal{M}_1 \\ \mathcal{S}_3 &=& \mathcal{N}_3/\mathcal{N}_2 & \text{and} & \mathcal{T}_3 &=& \mathcal{M}_3/\mathcal{M}_2 \\ &\vdots& & & &\vdots& \\ \mathcal{S}_J &=& \mathcal{N}_J/\mathcal{N}_{J-1} & & \mathcal{T}_K &=& \mathcal{M}_K/\mathcal{M}_{K-1} \end{array}$$

then actually, $J = K$, and (up to reordering), $\{\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_J\} = \{\mathcal{T}_1, \mathcal{T}_2, \ldots, \mathcal{T}_K\}$.

Proof: (a) (*by complete induction on $|\mathcal{G}|$*).

Base Case: ($|\mathcal{G}| = 2$) The only group of order 2 is $\mathbb{Z}_{/2}$, which is simple and therefore has composition series $\{\bar{0}\} = \mathcal{N}_0 \lhd \mathcal{N}_1 = \mathbb{Z}_{/2}$.

Induction: Suppose that *every group* of order less than n has a composition series, and let $|\mathcal{G}| = n$. We claim that $\mathcal{G}$ must also have a composition series. This is **Exercise 25**

Hint: there are two cases: $\begin{cases} 1. & \mathcal{G} \text{ is simple.} \\ 2. & \mathcal{G} \text{ is not simple.} \end{cases}$

If $\mathcal{G}$ is simple, it has a very short composition series. If $\mathcal{G}$ is not simple, let $\mathcal{N}$ be a maximal normal subgroup of $\mathcal{G}$. Then $|\mathcal{N}| < n$, so apply the induction hypothesis to $\mathcal{N}$ and proceed.

(b) *(by induction on J)*

Base Case: ($J = 2$) In this case, we have two composition series

$$\begin{array}{rcl} \{e_{\mathcal{G}}\} & = & \mathcal{N}_0 \vartriangleleft \mathcal{N}_1 \vartriangleleft \mathcal{N}_2 \;=\; \mathcal{G}; \\ \text{and} \quad \{e_{\mathcal{G}}\} & = & \mathcal{M}_0 \vartriangleleft \mathcal{M}_1 \vartriangleleft \mathcal{M}_2 \vartriangleleft \ldots \vartriangleleft \mathcal{M}_K \;=\; \mathcal{G}. \end{array}$$

Our goal is to prove that $K = 2$, and that $\{\mathcal{S}_1, \mathcal{S}_2\} = \{\mathcal{T}_1, \mathcal{T}_2\}$. Consider the following diamond:

$$\begin{tikzcd}
& \mathcal{M}_{K-1}\mathcal{N}_1 & \\
\mathcal{M}_{K-1} \arrow[ur] & & \mathcal{N}1 \arrow[ul] \\
& \mathcal{M}_{K-1} \cap \mathcal{N}_1 \arrow[ul] \arrow[ur] & .
\end{tikzcd}$$

Claim 1: $\mathcal{M}_{K-1}\mathcal{N}_1 = \mathcal{G}$.

Claim 2: $\mathcal{M}_{K-1} \cap \mathcal{N}_1 = \{e_{\mathcal{G}}\}$.

Claim 3: *$\mathcal{M}_{K-1} \cong \mathcal{S}_1$. Thus, $\mathcal{M}_{K-1}$ is simple.*

Claim 4: $K = 2$, $\mathcal{T}_0 \cong \mathcal{S}_1$, *and* $\mathcal{T}_1 \cong \mathcal{S}_0$.

Exercise 26 Prove Claim 1 using Claim 14; prove Claim 2 using Claim 15. Prove Claims 3 and 4 using the Diamond Isomorphism Theorem (Theorem 16 on page 14).

Induction: Suppose that part **(b)** is true for *any* group possessing a composition series of rank $J - 1$.

Define $\widetilde{\mathcal{G}} = \mathcal{M}_{K-1} \cap \mathcal{N}_{J-1}$.

For all $k \in [1..K)$, define $\widetilde{\mathcal{M}}_k = \mathcal{M}_k \cap \widetilde{\mathcal{G}}$.

For all $j \in [1..J)$, define $\widetilde{\mathcal{N}}_j = \widetilde{\mathcal{G}} \cap \mathcal{N}_j$.

Claim 5: $\widetilde{\mathcal{M}}_1 \;=\; \{e_{\mathcal{G}}\} \;=\; \widetilde{\mathcal{N}}_1$.

Claim 6:
$\{e_{\mathcal{G}}\} = \widetilde{\mathcal{N}}_1 \lhd \widetilde{\mathcal{N}}_2 \lhd \ldots \lhd \widetilde{\mathcal{N}}_{J-1} = \widetilde{\mathcal{G}}$
and

$\{e_{\mathcal{G}}\} = \widetilde{\mathcal{M}}_1 \lhd \widetilde{\mathcal{M}}_2 \lhd \ldots \lhd \widetilde{\mathcal{M}}_{K-1} = \widetilde{\mathcal{G}}$

are composition series for the group $\widetilde{\mathcal{G}}$.

Let the corresponding composition factors be:

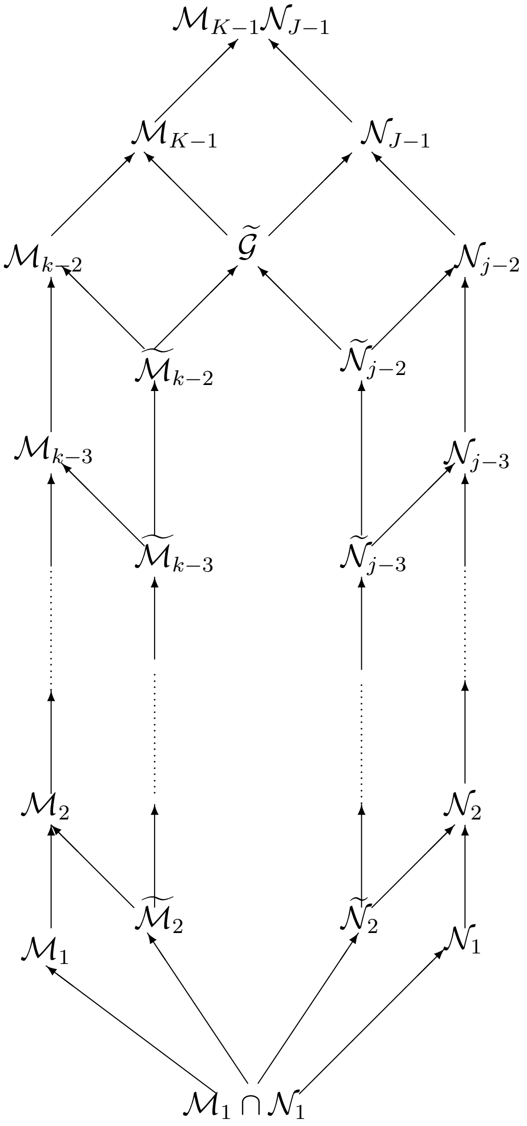
$$\begin{array}{rcl}
\widetilde{\mathcal{S}}_{J-1} & = & \widetilde{\mathcal{N}}_{J-1}/\widetilde{\mathcal{N}}_{J-2} \\
\widetilde{\mathcal{S}}_{J-2} & = & \widetilde{\mathcal{N}}_{J-2}/\widetilde{\mathcal{N}}_{J-3} \\
\vdots & & \\
\widetilde{\mathcal{S}}_{3} & = & \widetilde{\mathcal{N}}_{3}/\widetilde{\mathcal{N}}_{2} \\
\widetilde{\mathcal{S}}_{2} & = & \widetilde{\mathcal{N}}_{2}/\widetilde{\mathcal{N}}_{1}
\end{array}
\quad \text{and} \quad
\begin{array}{rcl}
\widetilde{\mathcal{T}}_{K-1} & = & \widetilde{\mathcal{M}}_{K-1}/\widetilde{\mathcal{M}}_{K-2} \\
\widetilde{\mathcal{T}}_{K-2} & = & \widetilde{\mathcal{M}}_{K-2}/\widetilde{\mathcal{M}}_{K-3} \\
\vdots & & \\
\widetilde{\mathcal{T}}_{3} & = & \widetilde{\mathcal{M}}_{3}/\widetilde{\mathcal{M}}_{2} \\
\widetilde{\mathcal{T}}_{2} & = & \widetilde{\mathcal{M}}_{2}/\widetilde{\mathcal{M}}_{1}
\end{array}$$

Claim 7: *$J = K$, and*

$\{\widetilde{\mathcal{S}}_2, \widetilde{\mathcal{S}}_3, \ldots, \widetilde{\mathcal{S}}_{J-1}\} \;=\; \{\widetilde{\mathcal{T}}_2, \widetilde{\mathcal{T}}_3, \ldots, \widetilde{\mathcal{T}}_{K-1}\}$.

Claim 8:
$\{e_{\mathcal{G}}\} = \widetilde{\mathcal{N}}_1 \lhd \widetilde{\mathcal{N}}_2 \lhd \widetilde{\mathcal{N}}_3 \lhd \ldots \lhd \widetilde{\mathcal{N}}_{J-1} = \widetilde{\mathcal{G}} \lhd \mathcal{N}_J$

is a composition series for the group $\mathcal{N}_J$.


$\mathcal{M}_{K-1}\mathcal{N}_{J-1}$
$\mathcal{M}_{K-1}$
$\mathcal{N}_{J-1}$
$\mathcal{M}_{k-2}$
$\widetilde{\mathcal{G}}$
$\mathcal{N}_{j-2}$
$\widetilde{\mathcal{M}}_{k-2}$
$\widetilde{\mathcal{N}}_{j-2}$
$\mathcal{M}_{k-3}$
$\mathcal{N}_{j-3}$
$\widetilde{\mathcal{M}}_{k-3}$
$\widetilde{\mathcal{N}}_{j-3}$
$\mathcal{M}_2$
$\mathcal{N}_2$
$\widetilde{\mathcal{M}}_2$
$\widetilde{\mathcal{N}}_2$
$\mathcal{M}_1$
$\mathcal{N}_1$
$\mathcal{M}_1 \cap \mathcal{N}_1$

Define $\widetilde{\mathcal{S}}_J \;=\; \mathcal{N}_J/\widetilde{\mathcal{G}}$

Claim 9: $\{\widetilde{\mathcal{S}}_2, \widetilde{\mathcal{S}}_3, \ldots, \widetilde{\mathcal{S}}_{J-1}, \widetilde{\mathcal{S}}_J\} \;=\; \{\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_{J-2}, \mathcal{S}_{J-1}\}$.

Claim 10: $\{e_{\mathcal{G}}\} = \widetilde{\mathcal{M}}_1 \lhd \widetilde{\mathcal{M}}_2 \lhd \widetilde{\mathcal{M}}_3 \lhd \ldots \lhd \widetilde{\mathcal{M}}_{K-1} = \widetilde{\mathcal{G}} \lhd \mathcal{M}_K$

is a composition series for the group $\mathcal{M}_K$.

Define $\widetilde{\mathcal{T}}_K \;=\; \mathcal{M}_{K-1}/\widetilde{\mathcal{G}}$.

Claim 11: $\{\widetilde{\mathcal{T}}_2, \widetilde{\mathcal{T}}_3, \ldots, \widetilde{\mathcal{T}}_{K-1}, \widetilde{\mathcal{T}}_K\} \;=\; \{\mathcal{T}_1, \mathcal{T}_2, \ldots, \mathcal{T}_{K-2}, \mathcal{T}_{K-1}\}$.

Claim 12: $\mathcal{M}_{K-1}\mathcal{N}_{J-1} \;=\; \mathcal{G}$.

Claim 13: $\widetilde{\mathcal{S}}_J \;=\; \mathcal{T}_K$, *and* $\widetilde{\mathcal{T}}_K \;=\; \mathcal{S}_J$.

The proofs of several of these claims depend upon certain basic facts about normal subgroups. A **maximal normal subgroup** of $\mathcal{G}$ is a normal subgroup $\mathcal{A} \lhd \mathcal{G}$ so that there exists no normal subgroup $\mathcal{B} \lhd \mathcal{G}$ with $\mathcal{A} < \mathcal{B}$.

Claim 14: **(a)** *$\mathcal{N}_{j-1}$ and $\mathcal{M}_{k-1}$ are maximal normal subgroups of $\mathcal{G}$.*

(b) *If $\mathcal{N}$ and $\mathcal{M}$ are distinct maximal normal subgroups of $\mathcal{G}$, then $\mathcal{N}\mathcal{M} = \mathcal{G}$.*

A **minimal normal subgroup** of $\mathcal{G}$ is a normal subgroup $\mathcal{A} \lhd \mathcal{G}$ which is also *simple* —ie. there exists no nontrivial normal subgroup $\mathcal{B} \lhd \mathcal{A}$.

Claim 15: **(a)** *$\mathcal{N}_1$ and $\mathcal{N}_1$ are minimal normal subgroups of $\mathcal{G}$.*

(b) *If $\mathcal{N}$ and $\mathcal{M}$ are distinct minimal normal subgroups of $\mathcal{G}$, then $\mathcal{N} \cap \mathcal{M} = \{e_{\mathcal{G}}\}$.*

Finally, we're using the following fact about composition series:

Claim 16: *Let $\{e_{\mathcal{G}}\} = \mathcal{A}_1 \lhd \mathcal{A}_2 \lhd \ldots \lhd \mathcal{A}_R = \mathcal{G}$ be any composition series for $\mathcal{G}$, and let $\mathcal{H} < \mathcal{G}$. For all $r \in [1..R]$, let $\mathcal{B}_r = \mathcal{A}_r \cap \mathcal{H}$. Then $\{e_{\mathcal{G}}\} = \mathcal{B}_1 \lhd \mathcal{B}_2 \lhd \ldots \lhd \mathcal{B}_R = \mathcal{H}$ is a composition series for $\mathcal{B}$.*

Exercise 27 First prove Claims 14, 15, and 16. Then prove Claim 5 using Claim 14. Prove Claim 12 using Claim 15. Prove Claims 6, 8, and 10 using Claim 16. Prove Claims 7, 9, and 11 using the induction hypothesis. Prove Claim 13 using the Diamond Isomorphism Theorem (Theorem 16 on page 14). Now bring all this together to prove the theorem. ______________ □

The Jordan-Hölder theorem means that we can talk about *the* composition factors of the group $\mathcal{G}$, without making reference to any specific composition series. For instance, as we saw in Examples (56b) and (56c), the composition factor set for $\mathbb{Z}_{/12}$ is $\{\mathbb{Z}_{/2}, \mathbb{Z}_{/2}, \mathbb{Z}_{/3}\}$. Is this because the integer 12 has prime factor set $\{2, 2, 3\}$? Yes; the same pattern holds for any *abelian* group...

Proposition 58 *Let $m \in \mathbb{N}$ have prime factorization: $m = p_1^{\mu_1} p_2^{\mu_2} \cdots p_J^{\mu_J}$ (where $p_1, \ldots, p_J$ are prime, and $\mu_1, \ldots, \mu_J \in \mathbb{N}$). Then any abelian group of order m has composition factor set*

$$\{\overbrace{\mathbb{Z}_{/p_1}, \ldots, \mathbb{Z}_{/p_1}}^{\mu_1}, \overbrace{\mathbb{Z}_{/p_2}, \ldots, \mathbb{Z}_{/p_2}}^{\mu_2}, \ldots, \overbrace{\mathbb{Z}_{/p_J}, \ldots, \mathbb{Z}_{/p_J}}^{\mu_J}\}.$$

Proof: We will prove this theorem by complete induction on m.

Claim 1: *If $n = p_1^{\nu_1} p_2^{\nu_2} \cdots p_J^{\nu_J}$, then the cyclic group $\mathbb{Z}_{/n}$ has composition factor set*

$$\{\overbrace{\mathbb{Z}_{/p_1}, \ldots, \mathbb{Z}_{/p_1}}^{\nu_1}, \overbrace{\mathbb{Z}_{/p_2}, \ldots, \mathbb{Z}_{/p_2}}^{\nu_2}, \ldots, \overbrace{\mathbb{Z}_{/p_J}, \ldots, \mathbb{Z}_{/p_J}}^{\nu_J}\}.$$

Proof: **Exercise 28** .... □ [Claim 1]

Claim 2: *Let $\mathcal{G}$ be a group; let $\mathcal{N} \lhd \mathcal{G}$ be a normal subgroup, and let $\mathcal{Q}$ be a quotient group. If $\mathcal{N}$ has composition factor set $\{\mathcal{N}_1, \mathcal{N}_2, \ldots, \mathcal{N}_L\}$, and $\mathcal{Q}$ has composition factor set $\{\mathcal{Q}_1, \mathcal{Q}_2, \ldots, \mathcal{Q}_K\}$, then $\mathcal{G}$ has composition factor set $\{\mathcal{N}_1, \mathcal{N}_2, \ldots, \mathcal{N}_L, \mathcal{Q}_1, \mathcal{Q}_2, \ldots, \mathcal{Q}_K\}$.*

Proof: **Exercise 29** Hint: Suppose $\mathcal{Q}$ had composition series: $\{e\} = \mathcal{R}_0 \lhd \mathcal{R}_1 \lhd \ldots \lhd \mathcal{R}_K = \mathcal{Q}$, with $\mathcal{R}_k/\mathcal{R}_{k-1} = \mathcal{Q}_k$, for all $k \in [1..K]$. Use the Chain Isomorphism Theorem (Theorem 18 on page 15) to transform this into a normal series

$$\mathcal{N} = \widetilde{\mathcal{R}}_0 \lhd \widetilde{\mathcal{R}}_1 \lhd \widetilde{\mathcal{R}}_2 \lhd \ldots \lhd \widetilde{\mathcal{R}}_K = \mathcal{G},$$

such that $\widetilde{\mathcal{R}}_k/\widetilde{\mathcal{R}}_{k-1} = \mathcal{R}_k/\mathcal{R}_{k-1} = \mathcal{Q}_k$, for all $k \in [1..K]$.

Now suppose $\mathcal{N}$ had composition series: $\{e\} = \mathcal{M}_0 \lhd \mathcal{M}_1 \lhd \ldots \lhd \mathcal{M}_L = \mathcal{N}$, such that $\mathcal{M}_\ell/\mathcal{M}_{\ell-1} = \mathcal{N}_\ell$, for all $\ell \in [1..L]$. Conclude that $\mathcal{G}$ has composition series:

$\{e\} = \mathcal{M}_0 \lhd \mathcal{M}_1 \lhd \ldots \lhd \mathcal{M}_L = \mathcal{N} = \widetilde{\mathcal{R}}_0 \lhd \widetilde{\mathcal{R}}_1 \lhd \ldots \lhd \widetilde{\mathcal{R}}_K = \mathcal{G}$. . . . . □ `[Claim 2]`

Base Case: $(m = 2)$ The only group of order 2 is $\mathbb{Z}_{/2}$; its only composition factor is $\mathbb{Z}_{/2}$.

Induction: Suppose the theorem is true for *all* abelian groups of order less than m, and let $\mathcal{A}$ be an abelian group of order m. Let $a \in \mathcal{A}$ be any nontrivial element, and let $\mathcal{N} = \langle a \rangle$. Then $\mathcal{N}$ is a cyclic subgroup of order n, for some $n < m$, and Lagrange's Theorem (Theorem 7 on page 6) says that n divides m. Thus, $n = p_1^{\nu_1} p_2^{\nu_2} \cdots p_J^{\nu_J}$ for some $\nu_1 \leq \mu_1$, $\nu_2 \leq \mu_1$, ..., $\nu_J \leq \mu_J$. Thus, Claim 1 says that $\mathcal{N}$ has composition factor set $\{\underbrace{\mathbb{Z}_{/p_1}, \ldots, \mathbb{Z}_{/p_1}}_{\nu_1}, \ \ldots, \underbrace{\mathbb{Z}_{/p_J}, \ldots, \mathbb{Z}_{/p_J}}_{\nu_J}\}$.

Now, $\mathcal{A}$ is abelian, so $\mathcal{N}$ is automatically normal in $\mathcal{A}$. Let $\mathcal{Q} = \mathcal{A}/\mathcal{N}$. Then $\mathcal{Q}$ is an abelian group of order $q = m/n = p_1^{\chi_1} p_2^{\chi_2} \cdots p_J^{\chi_J}$, where $\chi_1 = \mu_1 - \nu_1$, ..., $\chi_J = \mu_J - \nu_J$.

Thus, $\mathcal{Q}$ is an abelian group of order less than m, so the induction hypothesis says that $\mathcal{Q}$ has composition factor set $\{\underbrace{\mathbb{Z}_{/p_1}, \ldots, \mathbb{Z}_{/p_1}}_{\chi_1}, \ \ldots, \underbrace{\mathbb{Z}_{/p_J}, \ldots, \mathbb{Z}_{/p_J}}_{\chi_J}\}$. Now apply Claim 2 to conclude that $\mathcal{G}$ has composition factor set $\{\underbrace{\mathbb{Z}_{/p_1}, \ldots, \mathbb{Z}_{/p_1}}_{\nu_1+\chi_1}, \ \ldots, \underbrace{\mathbb{Z}_{/p_J}, \ldots, \mathbb{Z}_{/p_J}}_{\nu_J+\chi_J}\}$, as desired. □

The Hölder Program: The Jordan-Hölder theorem is the starting point of the *Hölder program*, a master strategy to build an 'atlas' of all finite groups, in two stages:

1. Construct a complete list of all finite simple groups.
2. Characterize all possible ways that two groups $\mathcal{N}$ and $\mathcal{Q}$ can be 'combined' to create a larger group, $\mathcal{G}$, such that $\mathcal{N} \lhd \mathcal{G}$, and $\mathcal{G}/\mathcal{N} = \mathcal{Q}$.

Stage 1 of the Hölder program is called the *Classification of Finite Simple Groups*, and was a massive, century-long effort which was finally completed in 1980 (see §5.4). If we think of groups as 'molecules', then simple groups are 'atoms', and the *Classification* is analogous to Mendeleev's construction of a *Periodic Table of Elements*.

Stage 2 is called the *Extension Problem*, and is still not completely solved (see §5.5). If we think of groups as 'molecules', and simple groups as 'atoms', then the *Extension Problem* is analogous to the entire subject of chemistry.

Refinement: In Example ⟨56d⟩ on page 45, we saw that a normal series of $\mathbb{Z}_{/120}$ could be 'refined' into a composition series, by judiciously inserting additional terms in the series. Another example of this sort of 'refinement' is Claim 2 of Proposition 58, which we generalize as follows:

Proposition 59 Refinement Lemma

Let $\mathcal{G}$ be a group, with a normal series: $\{e\} = \mathcal{N}_0 \lhd \mathcal{N}_1 \lhd \mathcal{N}_2 \lhd \ldots \lhd \mathcal{N}_J = \mathcal{G}$. Suppose that :

$$\begin{array}{rcll}
\mathcal{G}^1 & = & \mathcal{N}_1/\mathcal{N}_0 & \textit{has composition series: } \{e\} = \mathcal{M}^1_0 \lhd \mathcal{M}^1_1 \lhd \mathcal{M}^1_2 \lhd \ldots \lhd \mathcal{M}^1_{K_1} = \mathcal{G}^1. \\
\mathcal{G}^2 & = & \mathcal{N}_2/\mathcal{N}_1 & \textit{has composition series: } \{e\} = \mathcal{M}^2_0 \lhd \mathcal{M}^2_1 \lhd \mathcal{M}^2_2 \lhd \ldots \lhd \mathcal{M}^2_{K_2} = \mathcal{G}^2. \\
\vdots & & & \qquad\qquad\vdots \qquad\qquad\qquad \vdots \\
\mathcal{G}^J & = & \mathcal{N}_J/\mathcal{N}_{J-1} & \textit{has composition series: } \{e\} = \mathcal{M}^J_0 \lhd \mathcal{M}^J_1 \lhd \mathcal{M}^J_2 \lhd \ldots \lhd \mathcal{M}^J_{K_J} = \mathcal{G}^J.
\end{array}$$

Then $\mathcal{G}$ has a composition series:

$$\begin{aligned}
\{e\} = \widetilde{\mathcal{M}}^1_0 \lhd \widetilde{\mathcal{M}}^1_1 \lhd \widetilde{\mathcal{M}}^1_2 \lhd \ldots \lhd \widetilde{\mathcal{M}}^1_{K_1} = \mathcal{N}_1 = \widetilde{\mathcal{M}}^2_0 \lhd \widetilde{\mathcal{M}}^2_1 \lhd \ldots \lhd \widetilde{\mathcal{M}}^2_{K_2} = \mathcal{N}_2 \\
= \widetilde{\mathcal{M}}^3_0 \lhd \widetilde{\mathcal{M}}^3_1 \lhd \ldots \lhd \widetilde{\mathcal{M}}^3_{K_3} = \mathcal{N}_3 = \widetilde{\mathcal{M}}^4_0 \lhd \widetilde{\mathcal{M}}^4_1 \lhd \ldots \\
\ldots \lhd \mathcal{N}_{J-1} = \widetilde{\mathcal{M}}^J_0 \lhd \widetilde{\mathcal{M}}^J_1 \lhd \widetilde{\mathcal{M}}^J_2 \lhd \ldots \lhd \widetilde{\mathcal{M}}^J_{K_J} = \mathcal{N}_J = \mathcal{G}
\end{aligned}$$

such that, for all $j \in [1..J]$ and $k \in [1..K_j]$, $\left(\widetilde{\mathcal{M}}^k_j / \widetilde{\mathcal{M}}^k_{j-1}\right) \cong \left(\mathcal{M}^k_j / \mathcal{M}^k_{j-1}\right)$. *Hence,*

$$\left(\begin{array}{l}
\mathcal{G}^1 \textit{ has composition factor set } \{\mathcal{S}^1_1, \mathcal{S}^1_2, \ldots, \mathcal{S}^1_{K_1}\} \\
\mathcal{G}^2 \textit{ has composition factor set } \{\mathcal{S}^2_1, \mathcal{S}^2_2, \ldots, \mathcal{S}^2_{K_2}\} \\
\vdots \qquad\quad \vdots \qquad\quad \vdots \\
\mathcal{G}^J \textit{ has composition factor set } \{\mathcal{S}^J_1, \mathcal{S}^J_2, \ldots, \mathcal{S}^J_{K_J}\}
\end{array}\right) \Longrightarrow \left(\begin{array}{l}
\mathcal{G} \textit{ has composition factor set} \\
\\
\{\mathcal{S}^1_1, \mathcal{S}^1_2, \ldots, \mathcal{S}^1_{K_1}, \mathcal{S}^2_1, \mathcal{S}^2_2, \ldots, \mathcal{S}^2_{K_2}, \ldots, \\
\qquad\qquad \ldots, \mathcal{S}^J_1, \mathcal{S}^J_2, \ldots, \mathcal{S}^J_{K_J}\}
\end{array}\right)$$

Proof: **Exercise 30** Hint: Generalize the proof of Claim 2 of Proposition 58. For each $j \in [1..J]$, use the Chain Isomorphism Theorem (Theorem 18 on page 15) to transform the normal series

$$\{e\} = \mathcal{M}^j_0 \lhd \mathcal{M}^j_1 \lhd \mathcal{M}^j_2 \lhd \ldots \lhd \mathcal{M}^j_{K_j} = \mathcal{G}^j$$

into a normal series $\mathcal{N}_{j-1} = \widetilde{\mathcal{M}}^j_0 \lhd \widetilde{\mathcal{M}}^j_1 \lhd \widetilde{\mathcal{M}}^j_2 \lhd \ldots \lhd \widetilde{\mathcal{M}}^j_{K_j} = \mathcal{N}_j$. □

5.4 (∗) The Classification of Finite Simple Groups

Prerequisites: §5.2 **Recommended:** §5.3

Part 1 of the Hölder program called for a complete list of all finite simple groups, a project which ultimately spanned one hundred years and generated almost ten thousand pages of published research. At this point, the *Classification* is complete, but the details are beyond the scope of these notes. Instead, we will simply summarize the main results:

1. There are 18 infinite *families* of finite simple groups. Each of these families is a collection of simple groups built from some common 'template', indexed by one or more natural numbers. For example, here are three families of simple groups:

   (a) The **prime cyclic groups**: $\{\mathbb{Z}_{/p}\ ;\ \text{where } p \text{ is any prime number}\}$. (Proposition 52 on page 42).

   (b) The **alternating groups**: $\{\mathbf{A}_N\ ;\ \text{where } N \geq 5\}$. (see §6.3).

   (c) The **finite projective groups** [1, Thm 8.8.3]:
   $\{\mathbb{PSL}^n(\mathbb{F})\ ;\ \text{where } \mathbb{F} \text{ is any finite field, and } n \geq 2\ \ (\text{with } n \geq 3 \text{ if } \mathbb{F} = \mathbb{Z}_{/2} \text{ or } \mathbb{Z}_{/3})\}$,
   Here, $\mathbb{PSL}^n(\mathbb{F}) = \mathbb{SL}^n[\mathbb{F}]/\mathcal{Z}$, where $\mathbb{SL}^n[\mathbb{F}]$ is the group of $\mathbb{F}$-valued matrices with determinant 1, and $\mathcal{Z} = \mathcal{Z}(\mathbb{SL}^n[\mathbb{F}]) = \{\pm\mathbf{Id}\}$ is its center.

2. There are also 26 *sporadic* simple groups, which do not fit into any of these families.

3. The smallest of these sporadic simple groups is the *Matthieu group*, $\mathbf{M}_{11}$, which has cardinality 7920.

4. The largest sporadic simple group is called *the Monster*, and has cardinality approximately 10^{53}. The mathematical 'footsteps' the Monster have been discovered in surprisingly diverse areas of mathematics. The significance of the Monster is not well understood; the study of its footsteps is sometimes called *(Monstrous) Moonshine.* [?]

5. **Feit-Thompson Theorem:** *The only simple groups of* odd *order are the prime cyclic groups $\mathbb{Z}_{/p}$ (where p is prime).*

   (the proof of the Feit-Thompson is 255 pages long, and occupies an entire volume of *Pacific Journal of Mathematics*).

6. Likewise, the only *abelian* simple groups are the prime cyclic groups.

   (This follows easily from Fundamental Theorem of Finitely Generated Abelian Groups; see §4.3).

7. The smallest *nonabelian* simple group is the *icosahedral* group[3] $\mathbf{A}_5$, (which is also $\mathbb{PSL}^2(\mathbb{Z}_{/5})$), which has order 60 (see §6.3). The next largest nonabelian simple groups have orders 168, 360, 504, 660, 1092, and 2448.

5.5 (∗) The Extension Problem; Short Exact Sequences

Prerequisites: §1.4 **Recommended:** §5.1

If $\mathcal{G}$ is a group, and $\mathcal{N} \lhd \mathcal{G}$, and $\mathcal{Q} = \mathcal{G}/\mathcal{N}$ is the quotient group, then $\mathcal{G}$ is somehow a 'combination' of $\mathcal{N}$ and $\mathcal{Q}$. We call $\mathcal{G}$ an **extension** of $\mathcal{Q}$ by $\mathcal{N}$.

[3] So called because it is the group of symmetries of the regular icosahedron and regular dodecahedron.

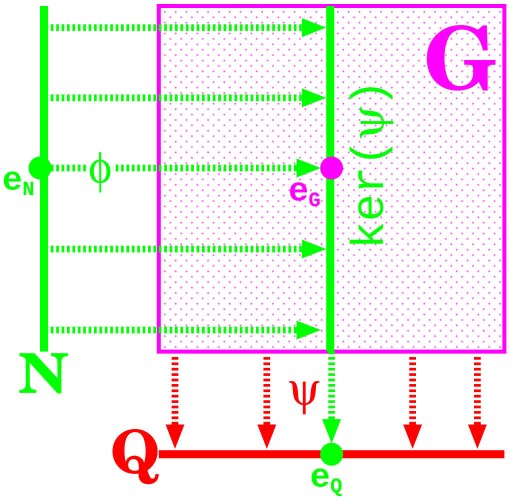
Now consider the 'inverse' problem: Given two groups $\mathcal{N}$ and $\mathcal{Q}$, what are all the ways in which $\mathcal{N}$ and $\mathcal{Q}$ can be 'combined' to yield a group $\mathcal{G}$, so that $\mathcal{N} \lhd \mathcal{G}$ and $\mathcal{Q} \cong \mathcal{G}/\mathcal{N}$? That is, what are all the *extensions* of $\mathcal{Q}$ by $\mathcal{N}$? The most obvious extension of $\mathcal{Q}$ by $\mathcal{N}$ is just the *direct product*, $\mathcal{G} = \mathcal{N} \times \mathcal{Q}$. However, this is rarely the *only* extension. The problem of characterizing these extensions is the *Extension Problem*.

An extension is sometimes described using a **short exact sequence**, a sequence of two homomorphisms and three groups:

$$\mathcal{N} \overset{\phi}{\rightarrowtail} \mathcal{G} \overset{\psi}{\twoheadrightarrow} \mathcal{Q},$$

such that:

1. $\phi : \mathcal{N} \rightarrowtail \mathcal{G}$ is a *monomorphism.*
2. $\psi : \mathcal{G} \twoheadrightarrow \mathcal{Q}$ is an *epimorphism.*
3. $\phi(\mathcal{N}) = \ker(\psi)$.



In other words, if $\widetilde{\mathcal{N}} = \phi(\mathcal{N})$, then $\widetilde{\mathcal{N}}$ is a normal subgroup of $\mathcal{G}$, there is an isomorphism $\mathcal{N} \cong \widetilde{\mathcal{N}}$ (via ϕ), and $\mathcal{Q} \cong \mathcal{G}/\widetilde{\mathcal{N}}$. The advantage of this notation is that we can apply the following lemma.

Proposition 60 The Five Lemma

Let $(\mathcal{N} \overset{\phi}{\rightarrowtail} \mathcal{G} \overset{\psi}{\twoheadrightarrow} \mathcal{Q})$ *and* $(\widetilde{\mathcal{N}} \overset{\widetilde{\phi}}{\rightarrowtail} \widetilde{\mathcal{G}} \overset{\widetilde{\psi}}{\twoheadrightarrow} \widetilde{\mathcal{Q}})$ *be two short exact sequences of groups, linked together in the following commuting diagram:*

$$\begin{tikzcd}
\mathcal{N} \arrow[r, tail, "\phi"] \arrow[d, "\nu"'] & \mathcal{G} \arrow[r, two heads, "\psi"] \arrow[d, "\gamma"'] & \mathcal{Q} \arrow[d, "\chi"] \\
\widetilde{\mathcal{N}} \arrow[r, tail, "\widetilde{\phi}"] & \widetilde{\mathcal{G}} \arrow[r, two heads, "\widetilde{\psi}"] & \widetilde{\mathcal{Q}}
\end{tikzcd}$$

If any two of the morphisms ν, γ, and ξ are isomorphisms, then the third one is, as well.

In particular, suppose that $\mathcal{G}$ and $\widetilde{\mathcal{G}}$ are two extensions of $\mathcal{Q}$ by $\mathcal{N}$, and that there is a homomorphism $\gamma : \mathcal{G} \longrightarrow \widetilde{\mathcal{G}}$ so that the following diagram commutes:

$$\begin{tikzcd}
& \mathcal{G} \arrow[dr, two heads, "\psi"] \arrow[dd, "\gamma"'] & \\
\mathcal{N} \arrow[ur, tail, "\phi"] \arrow[dr, tail, "\widetilde{\phi}"'] & & \mathcal{Q} \\
& \widetilde{\mathcal{G}} \arrow[ur, two heads, "\widetilde{\psi}"'] &
\end{tikzcd}$$

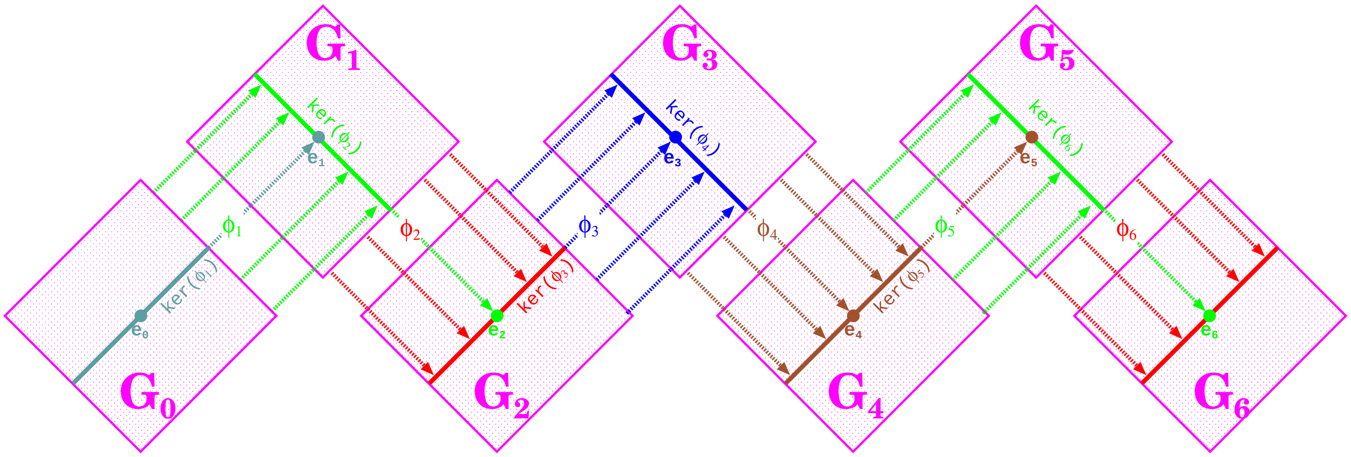


Figure 5.2: An exact sequence

Then γ *is an* *isomorphism, so that* $\mathcal{G} \cong \widetilde{\mathcal{G}}$.

Proof: **Exercise 31** □

The Extension Problem is an important part of the Hölder program (§5.3), and also arises frequently in *algebraic topology*, *differential geometry*, and *algebraic geometry*, where we often characterize the 'shape' of a space $\mathcal{X}$ using group $\mathcal{G}$ (eg. a *homotopy* group, *(co)homology* group, *holonomy* group, etc.). Our analysis of $\mathcal{X}$ often provides only 'indirect' information about $\mathcal{G}$, in the form of two groups $\mathcal{N}$ and $\mathcal{Q}$ so that $\mathcal{G}$ is an extension of $\mathcal{Q}$ by $\mathcal{N}$. Thus, to compute $\mathcal{G}$, we must solve the extension problem.

5.6 (∗) Exact Sequences

Prerequisites: §5.5

An **morphism sequence** is a sequence of groups and group homomorphisms:

$$\mathcal{G}_0 \xrightarrow{\phi_1} \mathcal{G}_1 \xrightarrow{\phi_2} \mathcal{G}_2 \xrightarrow{\phi_3} \ldots \xrightarrow{\phi_N} \mathcal{G}_N.$$

We say this sequence is **exact** if (as shown in Figure 5.2),

$$\mathsf{image}\,[\phi_1] = \ker(\phi_2), \qquad \mathsf{image}\,[\phi_2] = \ker(\phi_3), \qquad \ldots\ldots, \qquad \mathsf{image}\,[\phi_{N-1}] = \ker(\phi_N).$$

For example, a *short exact sequence* is an exact sequence containing exactly 3 nontrivial elements:

$$0 \overset{\xi}{\hookrightarrow} \mathcal{N} \overset{\phi}{\rightarrowtail} \mathcal{G} \overset{\psi}{\twoheadrightarrow} \mathcal{Q} \xrightarrow{\zeta} 0.$$

In other words:

1. $\ker(\phi) = \xi(0) = \{e\}$ —hence, ϕ is a monomorphism.
2. $\zeta(\mathcal{Q}) = 0$ —hence, $\ker(\zeta) = \mathcal{Q}$.
3. $\psi(\mathcal{G}) = \ker(\zeta) = \mathcal{Q}$ —hence, ψ is an epimorphism.
4. $\ker(\psi) = \phi(\mathcal{N})$.

5.7 Solvability

Prerequisites: §5.3

A group $\mathcal{G}$ is called **solvable** if it has a normal series:

$$\{e_{\mathcal{G}}\} = \mathcal{N}_0 \lhd \mathcal{N}_1 \lhd \mathcal{N}_2 \lhd \ldots \lhd \mathcal{N}_J = \mathcal{G}. \tag{5.8}$$

such that the quotient groups $\left\{\begin{array}{lcl} \mathcal{A}_1 &=& \mathcal{N}_1/\mathcal{N}_0 = \mathcal{N}_1 \\ \mathcal{A}_2 &=& \mathcal{N}_2/\mathcal{N}_1 \\ \mathcal{A}_3 &=& \mathcal{N}_3/\mathcal{N}_2 \\ & \vdots & \\ \mathcal{A}_J &=& \mathcal{N}_J/\mathcal{N}_{J-1} = \mathcal{G}/\mathcal{N}_{J-1} \end{array}\right\}$ are all *abelian* groups.

The series (5.8) is called a **solution series** for $\mathcal{G}$. Note that the solution series (5.8) is not necessarily a *composition* series, because the groups $\mathcal{A}_1, \ldots, \mathcal{A}_J$ are not necessarily *simple*.

Example 61:

(a) Any *abelian* group is solvable. If $\mathcal{G}$ is an abelian group, then the normal series

$$\{0\} = \mathcal{N}_0 \lhd \mathcal{N}_1 = \mathcal{G}$$

is a solution series, because $\mathcal{A}_1 = \mathcal{N}_1/\mathcal{N}_0 = \mathcal{G}$ is abelian.

(b) Let $\mathcal{G} = \mathbf{S}_3$. Then $\mathbf{S}_3$ is nonabelian, but it is still solvable. To see this, let $\mathcal{N}_1 = \mathbf{A}_3$. Then the normal series $\{0\} = \mathcal{N}_0 \lhd \mathcal{N}_1 \lhd \mathcal{N}_2 = \mathcal{G}$ is a solution series, because

$$\begin{array}{rcccccl} \mathcal{A}_1 &=& \mathcal{N}_1/\mathcal{N}_0 &=& \mathbf{A}_3 &\cong& \mathbb{Z}_{/3} \text{ is abelian.} \\ \text{and } \mathcal{A}_2 &=& \mathcal{N}_2/\mathcal{N}_1 &=& \mathbf{S}_3/\mathbf{A}_3 &\cong& \mathbb{Z}_{/2} \text{ is abelian.} \end{array}$$

(c) For any $N \in \mathbb{N}$, the dihedral group $\mathbf{D}_{2\times N}$ is nonabelian but solvable. To see this, let ρ be a rotation and κ be a reflection, so that $\mathbf{D}_{2\times N}$ is generated by $\{\rho, \kappa\}$. Let $\mathcal{N}_1 = \langle\rho\rangle \cong \mathbb{Z}_{/N}$. Then $\mathcal{N}_1 \lhd \mathbf{D}_{2\times N}$, and $\mathbf{D}_{2\times N}/\mathcal{N}_1$ is a two-element group, isomorphic to $\mathbb{Z}_{/2}$. Thus, $\mathcal{A}_1 = \mathcal{N}_1/\{e\} = \mathbb{Z}_{/N}$ is abelian, and $\mathcal{A}_2 = \mathbf{D}_{2\times N}/\mathcal{N}_1 = \mathbb{Z}_{/2}$ is abelian.

Theorem 62 *Let $\mathcal{G}$ be a group. The following are equivalent:*

(a) *$\mathcal{G}$ is solvable.*

(b) *The composition factors of $\mathcal{G}$ are all cyclic groups of prime order.*

(c) *$\mathcal{G}$ has a normal series $\{e_{\mathcal{G}}\} = \mathcal{N}_0 \lhd \mathcal{N}_1 \lhd \mathcal{N}_2 \lhd \ldots \lhd \mathcal{N}_J = \mathcal{G}$. where each of the factors $\mathcal{A}_j = \mathcal{N}_j/\mathcal{N}_{j-1}$ is a cyclic group.*

(d) *$\mathcal{G}$ has a solution series $\{e_{\mathcal{G}}\} = \mathcal{N}_0 \lhd \mathcal{N}_1 \lhd \mathcal{N}_2 \lhd \ldots \lhd \mathcal{N}_J = \mathcal{G}$ so that $\mathcal{N}_j$ is not only normal in $\mathcal{N}_{j+1}$, but is actually normal in $\mathcal{G}$.*

Proof: '$(a)\Longrightarrow(b)$' **Exercise 32** Hint: combine Proposition 58 on page 49 with Proposition 59 on page 51.

'$(b)\Longrightarrow(c)$' follows immediately from the definition of composition factors.

'$(c)\Longrightarrow(d)$' **Exercise 33** .

'$(d)\Longrightarrow(a)$' is immediate. □

Chapter 6

Group Actions

6.1 Introduction

Let $\mathcal{A}$ be a set, and let $\mathbf{S}_{\mathcal{A}} = \{\phi : \mathcal{A} \longrightarrow \mathcal{A}\ ;\ \phi \text{ any bijection}\}$ be the group of all permutations of $\mathcal{A}$. Thus, if $\mathcal{A} = [1..N]$, then $\mathbf{S}_{\mathcal{A}} = \mathbf{S}_N$ is the Nth *symmetry group.*

If $\mathbf{G}$ is a group, then an **action** of $\mathbf{G}$ **upon** $\mathcal{A}$ is a group homomorphism $\sigma : \mathbf{G} \longrightarrow \mathbf{S}_{\mathcal{A}}$. In other words, to every element $g \in \mathbf{G}$, σ associates a permutation $\sigma_g : \mathcal{A} \longrightarrow \mathcal{A}$. This is a way of realizing $\mathbf{G}$ as a group of transformations of $\mathcal{A}$.

Notation: If $a \in \mathcal{A}$ and $g \in \mathbf{G}$, we usually write '$g.a$' to indicate the $\sigma_g(a)$. This notation has a convenient 'associativity' property: If $g, h \in \mathbf{G}$, then

$$\begin{aligned} (g \cdot h).a &= \sigma_{g \cdot h}(a) \quad = \quad \sigma_g \circ \sigma_h(a) \quad = \quad \sigma_g\Big(\sigma_h(a)\Big) \quad = \quad \sigma_g\,(h.a) \\ &= g.(h.a). \end{aligned}$$

Because of this 'associativity', we can be careless about bracketting when describing group actions: there is no ambiguity in just writing "$g \cdot h.a$".

Note: The 'product' $g.a$ is just a notational convention, and does *not* really represent 'multiplication', since $\mathcal{A}$ is not, in general, a group.

Example 63:

(a) Let $\mathbf{G} = \mathbf{S}_N$ and let $\mathcal{A} = [1..N]$. Then $\mathbf{S}_N$ acts on $[1..N]$ by permutations. For example, if $g = (123)$ and $a = 2$, then $g.a = 3$.

(b) **Permutation groups:** Let $\mathbf{H} < \mathbf{S}_N$ be any subgroup of $\mathbf{S}_N$. Then $\mathbf{H}$ also acts on $[1..N]$. A subgroup of $\mathbf{S}_N$ is called a *permutation group.*

(c) Let $\mathbf{G} = \mathbf{D}_{2\times5}$ be the symmetries of a pentagon, and let $\mathcal{A} = \{1, 2, 3, 4, 5\}$ represent the five vertices of the pentagon. Then $\mathbf{G}$ acts by permuting these vertices. For example, if ρ is counter-clockwise rotation, then $\rho.1 = 2$, $\rho.2 = 3$, etc.

(d) **Rotation:** Let $\mathbf{G} = \mathbb{S}^1$ be the unit circle group, and let $\mathcal{A} = \mathbb{R}^2$. Then $\mathbb{S}^1$ acts on $\mathbb{R}^2$ by rotations about zero. To be precise, if $\theta \in \mathbb{S}^1$ is any angle (between 0 and 2π), and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{R}^2$, then $\theta.\mathbf{v} \ = \ \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$, where $s = \sin(\theta)$ and $c = \cos(\theta)$.

(e) **Conjugation:** Any group acts upon *itself* by conjugation. Here, $\mathcal{A} = \mathbf{G}$, and for any $g \in \mathbf{G}$, we define $\sigma_g : \mathbf{G} \longrightarrow \mathbf{G}$ to be the map: $\sigma_g(a) \ = \ gag^{-1}$.

(f) **Left-multiplication:** Any group acts upon *itself* by left-multiplication. Here, $\mathcal{A} = \mathbf{G}$, and for any $g \in \mathbf{G}$, we define $\sigma_g : \mathbf{G} \longrightarrow \mathbf{G}$ to be the map: $\sigma_g(a) \ = \ g \cdot a$.

(g) Let $\mathbf{A}$ be a group, and let $\mathbf{G} \subset \mathbf{A}$ be a subgroup. Then $\mathbf{G}$ acts on $\mathbf{A}$ by left-multiplication. For any $g \in \mathbf{G}$, we define $\sigma_g : \mathbf{A} \longrightarrow \mathbf{A}$ to be the map: $\sigma_g(a) \ = \ g \cdot a$.

6.2 Orbits and Stabilizers

Prerequisites: §6.1

If $a \in \mathcal{A}$, then the **orbit** of a is the set

$$\mathsf{Orbit}_{\mathbf{G}}(a) \quad = \quad \{g.a \ ; \ g \in \mathbf{G}\}$$

Sometimes the orbit of a is denoted "$\mathbf{G}.a$".

Example 64:

(a) Let $\mathbf{G} = \mathbb{S}^1$ be the unit circle group, acting on $\mathcal{A} = \mathbb{R}^2$ by rotations about zero, as in Example ⟨63d⟩. Let $\mathbf{v} = (v_1, v_2)$, and let $r \ = \ \|\mathbf{v}\| \ = \ \sqrt{v_1^2 + v_2^2}$ be the norm of $\mathbf{v}$. Then

$$\mathsf{Orbit}(\mathbf{v}) \quad = \quad \left\{\mathbf{w} \in \mathbb{R}^2 \ ; \ \|\mathbf{w}\| = r\right\}$$

is the circle of radius r about zero. In particular, note that $\mathsf{Orbit}((0,0)) \ = \ \{(0,0)\}$.

(b) **Conjugation:** Let $\mathbf{G}$ act upon itself by conjugation, as in Example ⟨63e⟩. If $a \in \mathbf{G}$, then the orbit of a is the set:

$$\mathcal{K}(a) \quad = \quad \left\{gag^{-1} \ ; \ g \in \mathbf{G}\right\}$$

This is called the **conjugacy class** of a. Observe that, if $a \in \mathcal{Z}(\mathbf{G})$, then $\mathcal{K}(a) \ = \ \{a\}$.

(c) Let $\mathbf{A}$ be a group, and let $\mathbf{G} \subset \mathbf{A}$ be a subgroup acting on $\mathbf{A}$ by left-multiplication. as in Example ⟨63g⟩. For any $a \in \mathbf{G}$, $\mathsf{Orbit}(a)$ is just the *right coset* $\mathbf{G}.a$.

Exercise 34 Verify these examples.

Lemma 65 *Let* $\mathbf{G}$ *act on* $\mathcal{A}$.

(a) *For any* $a \in \mathcal{A}$*,* $\quad a \in \mathsf{Orbit}\,(a)$.

(b) *For any* $a, b, c \in \mathcal{A}$*, if* $a \in \mathsf{Orbit}\,(b)$ *and* $b \in \mathsf{Orbit}\,(c)$*, then* $a \in \mathsf{Orbit}\,(c)$.

(c) *For any* $a, b \in \mathcal{A}$*, the following are equivalent:*

- $a \in \mathsf{Orbit}\,(b)$.
- $b \in \mathsf{Orbit}\,(a)$.
- $\mathsf{Orbit}\,(a) = \mathsf{Orbit}\,(b)$.

(d) $\mathcal{A}$ *is a disjoint union of orbits. For example, if* $\mathcal{A}$ *is finite, then* $\mathcal{A} = \mathcal{O}_1 \sqcup \mathcal{O}_2 \sqcup \ldots \sqcup \mathcal{O}_M$*, where* $\mathcal{O}_m = \mathsf{Orbit}\,(a_m)$ *for some* $a_m \in \mathcal{A}$.

Proof: (**Exercise 35**) □

We say that the action of $\mathbf{G}$ on $\mathcal{A}$ is **transitive** if $\mathcal{A} = \mathsf{Orbit}\,(a)$ for some $a \in \mathcal{A}$.

Lemma 66 *The following are equivalent:*

1. $\mathbf{G}$ *acts transitively on* $\mathcal{A}$.
2. *For any* $a \in \mathcal{A}$*,* $\quad \mathcal{A} = \mathsf{Orbit}\,(a)$.
3. *For any* $a, b \in \mathcal{A}$*, there is some* $g \in \mathbf{G}$ *so that* $g.a = b$.

Proof: (**Exercise 36**) □

Example 67:

(a) Let $\mathbf{G} = \mathbf{S}_N$ act on $\mathcal{A} = [1..N]$, as in Example ⟨63a⟩. This action is transitive. To see this, let $m, n \in [1..N]$. If $\sigma = (nm)$, then $\sigma.n = m$.

(b) Let $\mathbf{G} = \mathbf{D}_{2\times 5}$ act on $\mathcal{A} = \{1, 2, 3, 4, 5\}$, as in Example (63c). Then this action is transitive.

(c) **Left-multiplication:** Let $\mathbf{G}$ act upon itself by left-multiplication, as in Example ⟨63f⟩. This action is transitive. To see this, let $a, b \in \mathbf{G}$. Let $g = b \cdot a^{-1}$. Then $g.a = b$.

If $a \in \mathcal{A}$, then the **stabilizer** of a is the set of elements which act trivially on a:

$$\mathsf{Stab}\,(a) \quad = \quad \{g \in \mathbf{G}\ ;\ g.a \ = \ a\}$$

Sometimes the stabilizer of a is denoted "$\mathbf{G}_a$".

Example 68:

(a) Let $\mathbf{G} = \mathbf{S}_4$, acting by permutations on $\mathcal{A} = [1..4]$, as in Example ⟨63a⟩. Then $\mathsf{Stab}(4)$ is the set of all permutations which fix 4. Explicitly,

$$\begin{aligned}\mathsf{Stab}(4) \quad = \quad \Big\{ & e,\ (12),\ (13),\ (15),\ (23),\ (25),\ (35),\ (12)(35),\ (13)(25),\ (15)(23),\ (123),\\ & (125),\ (135),\ (132),\ (152),\ (153),\ (235),\ (253),\ (1235),\ (1532),\\ & (1325),\ (1523),\ (1352),\ (1253)\Big\}.\end{aligned}$$

(b) Let $\mathbf{G} = \mathbb{S}^1$ be the unit circle group, acting on $\mathcal{A} = \mathbb{R}^2$ by rotations about zero, as in Example ⟨63d⟩. If $\mathbf{v}$ is any nonzero vector, then $\mathsf{Stab}(\mathbf{v}) = \{e\}$. However, $\mathsf{Stab}((0,0)) = \mathbb{S}^1$ —ie. all of $\mathbb{S}^1$ acts trivially upon $(0,0)$.

(c) **Conjugation:** Let $\mathbf{G}$ act upon itself by conjugation, as in Example ⟨63e⟩. If $a \in \mathbf{G}$, then the stabilizer of a is the set:

$$\mathcal{C}_{\mathbf{G}}(a) \quad = \quad \{g \in \mathbf{G}\ ;\ ga \ = \ ag\}$$

This is called the **centralizer** of a in $\mathbf{G}$. Observe that, if $a \in \mathcal{Z}(\mathbf{G})$, then $\mathcal{C}_{\mathbf{G}}(a) \ = \ \mathbf{G}$.

Lemma 69 *If $b = g.a$, then $\mathsf{Stab}(b) \ = \ g \cdot \mathsf{Stab}(a) \cdot g^{-1}$.*

Proof: (**Exercise 37**) □

We say that the action of $\mathbf{G}$ upon $\mathcal{A}$ is **free** if $\mathsf{Stab}(a) = \{e\}$ for all $a \in \mathcal{A}$. In other words, *every* element of $\mathbf{G}$ acts nontrivially on *every* element of $\mathcal{A}$.

Example 70:

(a) Let $\mathbf{G}$ act on itself by left-multiplication, as in Example (63f). This action is free: for any $g \in \mathbf{G}$ and $a \in \mathbf{G}$, if $g.a = a$, then $g = e$ by left-cancellation.

If $\mathbf{G}$ acts on the set $\mathcal{A}$ via a homomorphism $\sigma : \mathbf{G} \longrightarrow \mathbf{S}_{\mathcal{A}}$, then the **kernel** of the action is the kernel of σ. In other words, the kernel is the set of all elements in $\mathbf{G}$ which act trivially upon *all* elements of $\mathcal{A}$.

Example 71: Let $\mathbf{G}$ act upon itself by left-multiplication, as in Example ⟨63e⟩. The kernel of this action is just the **center** of $\mathbf{G}$: $\mathcal{Z}(\mathbf{G}) \ = \ \{g \in \mathbf{G}\ ;\ ga \ = \ ag \text{ for all } a \in \mathbf{G}\}$.

Lemma 72 *Let $\mathbf{G}$ act on $\mathcal{A}$, and let $\mathbf{K}$ be the kernel of the action. Then $\mathbf{K} \ = \ \bigcap_{a \in \mathcal{A}} \mathsf{Stab}(a)$.*

In particular, if $\mathbf{G}$ acts freely, then $\mathbf{K} = \{e\}$, so the map $\sigma : \mathbf{G} \longrightarrow \mathbf{S}_{\mathcal{A}}$ is a monomorphism.

Proof: (**Exercise 38**) □

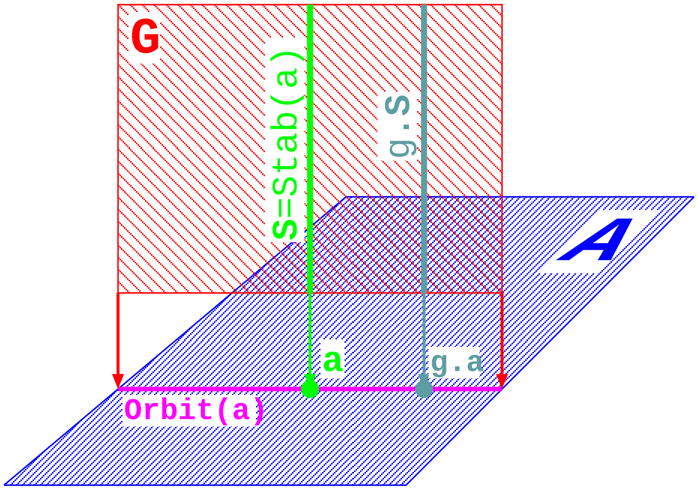


Figure 6.1: The Orbit-Stabilizer Theorem

Corollary 73 (Cayley's Theorem)

Every group is isomorphic to a permutation group. In particular, if $|\mathbf{G}| = N$, then $\mathbf{G}$ is isomorphic to a subgroup of $\mathbf{S}_N$.

Proof: Let $\mathbf{G}$ act upon itself by left-multiplication. Let $\sigma : \mathbf{G} \longrightarrow \mathbf{S}_{\mathbf{G}}$ be the associated homomorphism, and let $\mathbf{I} = \sigma(\mathbf{G}) \subset \mathbf{S}_{\mathbf{G}}$.

As we saw in Example ⟨71⟩, this action is free. Thus, Lemma 72 says that $\ker(\sigma) = \{e\}$; hence, σ is an isomorphism between $\mathbf{G}$ and $\mathbf{I}$. □

Cayley's theorem is nice because it establishes a connection between 'abstract' (algebraic) groups and 'concrete' (transformational) groups. It also says that the permutation groups have a sort of *Universal Property*: in a sense, *every* group is a subgroup of some permutation group. Thus, if one knew everything about permutation groups, then one would know everything about *all* groups. Although philosophically interesting, this result has surprisingly little application. Partly this is because it suggests that we study a group of order N by embedding it in a (much larger) group of order $N!$.

You may have noticed, in the previous examples, that the larger the stabilizer of a gets, the smaller the orbit of a becomes. The precise formulation of this is as follows:

Theorem 74 (Orbit-Stabilizer Theorem)

Let $\mathbf{G}$ act on $\mathcal{A}$, and let $a \in \mathcal{A}$. Let $\mathbf{S} = \mathsf{Stab}\,(a)$.

1. *There is a bijection between the elements of $\mathsf{Orbit}\,(a)$ and the cosets of $\mathbf{S}$ in $\mathbf{G}$, given by:*

$$\Phi : \mathbf{G}/\mathbf{S} \ni g\mathbf{S} \ \mapsto\ g.a \in \mathsf{Orbit}\,(a) \qquad \textit{(see Figure 6.1)}$$

2. *Thus,* $|\mathsf{Orbit}\,(a)| = |\mathbf{G}/\mathbf{S}|$. *If* $\mathbf{G}$ *is finite, then* $|\mathsf{Orbit}\,(a)| = \frac{|\mathbf{G}|}{|\mathbf{S}|}$.

Proof:

Claim 1: *For any* $g, h \in \mathbf{G}$, $\Big(g.a = h.a \Big) \iff \Big(h \in g.\mathbf{S} \Big)$.

Proof: $\Big(g.a = h.a \Big) \iff \Big(a = g^{-1}.h.a \Big) \iff \Big(g^{-1}h \in \mathsf{Stab}\,(a) = \mathbf{S} \Big) \iff \Big(h \in g\mathbf{S} \Big)$. .... □ `[Claim 1]`

(1) Φ *is well-defined and injective:* Let $g, h \in \mathbf{G}$. We want to show $\Big(g\mathbf{S} = h\mathbf{S} \Big) \iff \Big(\Phi(g\mathbf{S}) = \Phi(h\mathbf{S}) \Big)$. But,

$$\Big(g\mathbf{S} = h\mathbf{S} \Big) \iff \Big(h \in g\mathbf{S} \Big) \overset{}{\Leftarrow_{\text{Claim1}}\Rightarrow} \Big(g.a = h.a \Big) \iff \Big(\Phi(g\mathbf{S}) = \Phi(h\mathbf{S}) \Big).$$

Φ *is surjective:* Let $b \in \mathsf{Orbit}\,(a)$. Then $b = g.a$ for some g. Thus, $b = \Phi(g\mathbf{S})$.

(2) The first equality follows from **(1)**. The second is just Lagrange's Theorem. □

Example 75: Conjugacy vs. Centralizers

Let $\mathbf{G}$ act upon itself by conjugation, as in Example (63e). If $a \in \mathbf{G}$, then Example (68c) says that the stabilizer of a is its *centralizer*:

$$\mathsf{Stab}\,(a) = \mathcal{C}_{\mathbf{G}}\,(a) = \{g \in \mathbf{G}\,;\ ga = ag\}$$

Let $\mathcal{K}\,(a) = \mathsf{Orbit}\,(a)$ be the conjugacy class of a, as in Example (64b). Then the Orbit-Stabilizer Theorem tells us: $|\mathcal{K}\,(a)| = \dfrac{|\mathbf{G}|}{|\mathcal{C}_{\mathbf{G}}\,(a)|}$.

Corollary 76 (The Class Equation)

Let $\mathbf{G}$ *be a finite group. Then*

$$|\mathbf{G}| = |\mathcal{Z}\,(\mathbf{G})| + \frac{|\mathbf{G}|}{|\mathcal{C}_{\mathbf{G}}\,(a_1)|} + \frac{|\mathbf{G}|}{|\mathcal{C}_{\mathbf{G}}\,(a_2)|} + \ldots + \frac{|\mathbf{G}|}{|\mathcal{C}_{\mathbf{G}}\,(a_N)|}$$

where $a_1, \ldots, a_N$ *represent the distinct nontrivial conjugacy classes of* $\mathbf{G}$.

Proof: Lemma 65(d) says that $\mathbf{G}$ is a disjoint union of conjugacy classes:

$$\mathbf{G} = \mathcal{K}_1 \sqcup \mathcal{K}_2 \sqcup \ldots \sqcup \mathcal{K}_M \tag{6.1}$$

Recall that, if $z \in \mathcal{Z}(\mathbf{G})$, then $\mathcal{K}(z) = \{z\}$. Suppose $\mathcal{Z}(\mathbf{G}) = \{z_1, z_2, \ldots, z_L\}$. Let $N = M - L$, assume without loss of generality that $\mathcal{K}_{N+1} = \{z_1\}, \mathcal{K}_{N+2} = \{z_2\}, , \ldots, \mathcal{K}_M = \{z_L\}$. Then we can rewrite eqn. (6.3) as:

$$\mathbf{G} \quad = \quad \mathcal{K}_1 \sqcup \mathcal{K}_2 \sqcup \ldots \sqcup \mathcal{K}_N \sqcup \{z_1\} \sqcup \{z_2\} \sqcup \ldots \sqcup \{z_L\} \quad = \quad \mathcal{K}_1 \sqcup \mathcal{K}_2 \sqcup \ldots \sqcup \mathcal{K}_N \sqcup \mathcal{Z}(\mathbf{G}).$$

Thus,

$$|\mathbf{G}| \quad = \quad |\mathcal{K}_1| + |\mathcal{K}_2| + \ldots + |\mathcal{K}_N| + |\mathcal{Z}(\mathbf{G})|. \tag{6.2}$$

For all $n \in [1..N]$, suppose that $\mathcal{K}_n = \mathcal{K}(a_n)$ for some $a_n \in \mathbf{G}$. Then Example 75 says that

$$|\mathcal{K}_n| \quad = \quad \frac{|\mathbf{G}|}{|\mathcal{C}_{\mathbf{G}}(a_n)|}. \tag{6.3}$$

Substituting eqn. (6.3) into eqn. (6.2) yields the Class Equation. □

Corollary 77 *If $|\mathbf{G}| = p^k$ for some prime p, then $\mathcal{Z}(\mathbf{G})$ is nontrivial.*

Proof: Let $Z = |\mathcal{Z}(\mathbf{G})|$. Since $e \in \mathcal{Z}(\mathbf{G})$, we know that $Z \geq 1$. Thus, the Class Equation reads:

$$p^k \quad = \quad Z + \frac{p^k}{C_1} + \frac{p^k}{C_2} + \ldots + \frac{p^k}{C_N}$$

where $C_n = |\mathcal{C}_{\mathbf{G}}(a_n)|$. Observe that all the terms $\frac{p^k}{C_n}$ must be divisible by p, and of course, p^k is divisible by p. Thus, Z must also be divisible by p. Since $Z \geq 1$, we conclude that $Z = p^m$ for some $m \geq 1$. □

6.3 The Simplicity of $\mathbf{A}_N$

Prerequisites: §5.2, §6.2 **Recommended:** §5.3

To realise the Hölder program, we must identify all finite simple groups. So far, the only simple groups we know are the *prime-power cyclic groups*: $\mathbb{Z}_{/2}$, $\mathbb{Z}_{/3}$, $\mathbb{Z}_{/5}$, $\mathbb{Z}_{/7}$, $\mathbb{Z}_{/11}$, $\mathbb{Z}_{/13}$, $\mathbb{Z}_{/17}, \ldots$. All of these are abelian. Are there any nonabelian simple groups?

The answer is, 'Yes'. The smallest nonabelian simple group is the alternating group $\mathbf{A}_5$, which has order 60 (also called the *icosahedral* group, because it is the group of symmetries of the regular icosahedron and regular dodecahedron). We will show that, for any $N \geq 5$, the group $\mathbf{A}_N$ is simple.

Let $\mathbf{G}$ be a group. If $g \in \mathbf{G}$, recall that the **conjugacy class** of g is the set $\mathcal{K}(g) = \{hgh^{-1} \; ; \; h \in \mathbf{G}\}$. If $g \in \mathcal{Z}(\mathbf{G})$, then $\mathcal{K}(g) = \{g\}$. In particular, $\mathcal{K}(e) = \{e\}$.

Recall from Lemma 65(d) (page 58) that $\mathbf{G}$ is a *disjoint union* of its conjugacy classes[1]. In other words,

$$\mathbf{G} = \mathcal{Z}(\mathbf{G}) \sqcup \mathcal{K}_1 \sqcup \mathcal{K}_2 \sqcup \ldots \sqcup \mathcal{K}_J, \tag{6.4}$$

where, for $j \in [1..J]$, $\quad \mathcal{K}_j = \mathcal{K}(g_j)$ for some $g_j \in \mathbf{G}$.

Lemma 78 *Let $\mathbf{G}$ be a group with conjugacy class decomposition (6.4), and let $\mathbf{N} \lhd \mathbf{G}$ be a normal subgroup.*

(a) *If $g \in \mathbf{N}$, then $\mathcal{K}(g) \subset \mathbf{N}$.*

(b) *Thus, $\mathbf{N}$ is a disjoint union of conjugacy classes. In other words, $\mathbf{N} = \mathbf{Y} \sqcup \mathcal{K}_{j_1} \sqcup \ldots \sqcup \mathcal{K}_{j_I}$, where $\mathbf{Y} \subset \mathcal{Z}(\mathcal{G})$, and where $\{j_1, j_2, \ldots, j_I\} \subset [1..J]$ is some subcollection.*

(c) *Thus, $|\mathbf{N}| = Y + K_{j_1} + K_{j_2} + \ldots + K_{j_I}$, where $Y = |\mathbf{Y}|$, and $K_j = |\mathcal{K}_j|$ for $j \in [1..J]$.*

Proof: Exercise □

Corollary 79 (Combinatorial Simplicity Criterion)

Let $\mathbf{G}$ be a group with conjugacy class decomposition (6.4), and let $K_j = |\mathcal{K}_j|$ for $j \in [1..J]$. Let $Z = |\mathcal{Z}(\mathbf{G})|$.

Suppose that there exists no number $Y \in [1..Z]$, and no proper subcollection $\{j_1, j_2, \ldots, j_I\} \subset [1..J]$ so that the sum $(Y + K_{j_1} + K_{j_2} + \ldots + K_{j_I})$ divides $|\mathbf{G}|$. Then $\mathbf{G}$ is simple.

Proof: If $\mathbf{N}$ was any proper normal subgroup of $\mathbf{G}$, then $|\mathbf{N}| = Y + K_{j_1} + K_{j_2} + \ldots + K_{j_I}$, as in Lemma 78(c). Here, $Y = |\mathbf{Y}|$, where $\mathbf{Y} \subset \mathcal{Z}(\mathbf{G})$ and $e \in \mathbf{Y}$. Thus, $1 \leq |\mathbf{Y}| \leq Z$.

On the other hand, $\mathbf{N}$ is a subgroup of $\mathbf{G}$, so by Lagrange's theorem, $|\mathbf{N}|$ must divide $|\mathbf{G}|$. If no sum $(Y + K_{j_1} + K_{j_2} + \ldots + K_{j_I})$ divides $|\mathbf{G}|$, then no such $\mathbf{N}$ can exist. □

Example 80: Suppose $\mathbf{G}$ was a group of order 60, with trivial center. Thus, $Z = 1$. Suppose the conjugacy classes of $\mathbf{G}$ are $\mathcal{K}_1$, $\mathcal{K}_2$, $\mathcal{K}_3$, and $\mathcal{K}_4$, where $K_1 = |\mathcal{K}_1| = 20$, $K_2 = |\mathcal{K}_2| = 12$, $K_3 = |\mathcal{K}_3| = 12$, and $K_4 = |\mathcal{K}_4| = 15$. Observe that none of the numbers

$$\begin{array}{ccccc} 1+20 & 1+12 & 1+20+12 & 1+15 & 1+20+15 \\ 1+20+12 & 1+20+12+15 & 1+12+12 & 1+15+12+12 & 1+20+12+12 \end{array}$$

divides 60. Thus, $\mathbf{G}$ is simple by Corollary 79.

[1] See also the proof of the Class Equation, Corollary 76 on page 62.

We will prove that $\mathbf{A}_5$ is simple by applying Corollary 79. Indeed, we will show that $\mathbf{A}_5$ satisfies exactly the description of Example ⟨80⟩. First we need a way of measuring the conjugacy classes of $\mathbf{A}_5$. To start, we'll look at conjugacy classes in $\mathbf{S}_5$ (which are not the same, but are related).

Lemma 81 (Conjugacy in $\mathbf{S}_N$)

Let $\sigma, \tau \in \mathbf{S}_N$. If $\sigma = (a_1 a_2 \ldots a_n)(b_1 b_2 \ldots b_m) \ldots (c_1 c_2 \ldots c_\ell)$, then $\tau\sigma\tau^{-1} = \Big(\tau(a_1)\tau(a_2)\ldots\tau(a_n)\Big)\Big(\tau(b_1)\tau(b_2)\ldots\tau(b_m)\Big)\ldots\Big(\tau(c_1)\tau(c_2)\ldots\tau(c_\ell)\Big)$.

Proof: Exercise □

Example 82: Suppose $\sigma = (1\ 2\ 3)\ (4\ 5)\ (6\ 7)$, and $\tau = (1\ 5\ 2\ 3)\ (7\ 4\ 6)$. Then $\tau\sigma\tau^{-1} = (5\ 3\ 1)\ (6\ 2)\ (7\ 4)$.

Let $\sigma \in \mathbf{S}_N$, and write σ as a product of disjoint cycles:

$$\sigma = (a_1 a_2 \ldots a_n)(b_1 b_2 \ldots b_m) \ldots (c_1 c_2 \ldots c_\ell) \tag{6.5}$$

Assume without loss of generality that $n \leq m \leq \ldots \leq \ell$. The **cycle type** of σ is the list of numbers $(n, m, \ldots, \ell)$. For example, if $\sigma = (1\ 2\ 3)\ (4\ 5\ 6\ 7)\ (8\ 9\ 10\ 11)\ (12\ 13\ 14\ 15\ 16)$, then its cycle type is $(3, 4, 4, 5)$

Proposition 83 *Let $\sigma, \sigma_1 \in \mathbf{S}_N$. Then*

$$\Big(\ \sigma \text{ and } \sigma' \text{ are conjugate in } \mathbf{S}_N\ \Big) \iff \Big(\ \sigma \text{ and } \sigma' \text{ have the same cycle type}\ \Big)$$

Proof: Suppose $\sigma = (a_1 a_2 \ldots a_n)(b_1 b_2 \ldots b_m) \ldots (c_1 c_2 \ldots c_\ell)$.

'$\Longrightarrow$': Suppose $\sigma' = \tau\sigma_1\tau^{-1}$. Let $a'_1 = \tau(a_1)$, $a'_2 = \tau(a_2)$, etc. Then Lemma 81 says that $\sigma' = (a'_1 a'_2 \ldots a'_n)(b'_1 b'_2 \ldots b'_m) \ldots (c'_1 c'_2 \ldots c'_\ell)$, which clearly has the same cycle type as σ.

'$\Longleftarrow$': Suppose σ' has the same cycle type as σ. Thus, $\sigma' = (a'_1 a'_2 \ldots a'_n)(b'_1 b'_2 \ldots b'_m) \ldots (c'_1 c'_2 \ldots c'_\ell)$, for some elements $a'_1, a'_2, \ldots, a'_n, b'_1, \ldots, b'_m, \ldots \in [1..N]$. Let τ be the permutation defined by the property: $\tau(a_1) = a'_1$, $\tau(a_2) = a'_2$, etc. Then Lemma 81 says that $\tau\sigma\tau^{-1} = \sigma'$. □

Corollary 84 *Let $\sigma \in \mathbf{S}_N$ have cycle type $(m_1, m_2, \ldots, m_k)$, and let $M = m_1 + m_2 + \ldots + m_K$. Then* $|\mathcal{K}(\sigma)| = \dfrac{N!}{(m_1 m_2 \ldots m_k) \cdot (N-M)!}$.

Proof: To keep notation simple, we'll give the proof when σ has cycle type (n, m, ℓ) —the general case is much the same. By Proposition 83, $\mathcal{K}(\sigma)$ consists of all permutations with this same cycle type. How many of these are there? I want to build a permutation with cycle decomposition

$$(a_1 a_2 \ldots a_n)(b_1 b_2 \ldots b_m)(c_1 c_2 \ldots c_\ell)$$

where $a_1 a_2 \ldots a_n, b_1 b_2 \ldots b_m, c_1 c_2 \ldots c_\ell$ are distinct elements of $[1..N]$. I have N choices for a_1. Once a_1 is chosen, I have $(N-1)$ choices left for a_2. Then I have $(N-2)$ choices left for a_3, and so on. At the end, I have $(N-M+1)$ choices for c_ℓ, where $M = n + m + \ell$. The total number of choices is

$$N \cdot (N-1) \cdot (N-2) \cdot \ldots \cdot (N-M+1) \quad = \quad \frac{N \cdot (N-1) \cdot \ldots \cdot 2 \cdot 1}{(N-M) \cdot (N-M-1) \cdot \ldots \cdot 2 \cdot 1} \quad = \quad \frac{N!}{(N-M)!}$$

But we're not done yet, because the permutation $(a_1 a_2 a_3 \ldots a_n)$ is really the same as the permutation $(a_2 a_3 a_4 \ldots a_{n-1} a_n a_1)$. Thus, in the previous argument, we have counted the *same* permutation many times. To be precise, we have n times overcounted each cycle $(a_1 a_2 \ldots a_n)$, because there are n distinct *cyclically permuted* ways of writing this cycle which really represent the same permutation. Likewise we have m times overcounted each cycle $(b_1 b_2 \ldots b_m)$ and we have ℓ times overcounted each cycle $(c_1 c_2 \ldots c_\ell)$. In total we have overcounted by a factor of $n \cdot m \cdot \ell$. Thus, we must divide by $n \cdot m \cdot \ell$ to conclude: $|\mathcal{K}(\sigma)| = \dfrac{N!}{n \cdot m \cdot \ell \cdot (N-M)!}$.
□

Example 85: If $\sigma = (1\ 2\ 3\ 4\ 5)$ in $\mathbf{S}_8$, then $M = m_1 = 5$, so that $\mathcal{K}(\sigma)$ has cardinality $\dfrac{8!}{5 \cdot (8-5)!} = \dfrac{8 \cdot 7 \cdot 6 \cdot 5 \cdot 4}{5} = 1344$.

Recall that the **centralizer** of an element g in group $\mathbf{G}$ is the set $\mathcal{C}_{\mathbf{G}}(g) = \{c \in \mathbf{G} \ ; \ cg = gc\}$.

Lemma 86 *If* $\mathbf{G}$ *is any group, and* $g \in \mathbf{G}$*, then* $|\mathcal{K}(g)| = \dfrac{|\mathbf{G}|}{|\mathcal{C}_{\mathbf{G}}(g)|}$.

Proof: Let $\mathbf{G}$ act on itself by conjugation. Then $\mathcal{K}(g)$ is just the *orbit* of g under this group action, and $\mathcal{C}_{\mathbf{G}}(g)$ is just the stabilizer of g. The result now follows from the Orbit-Stabilizer theorem. □

Corollary 87 *If* $\sigma \in \mathbf{S}_N$ *has cycle type* $(m_1, m_2, \ldots, m_K)$ *and* $M = m_1 + \ldots + m_K$*, then* $|\mathcal{C}_{\mathbf{S}_N}(\sigma)| = m_1 \cdot m_2 \cdot \ldots \cdot m_K \cdot (N-M)!$.

Proof: Combine Corollary 84 and Lemma 86. □

Example 88: If $\sigma = (1\ 2\ 3\ 4\ 5)$ in $\mathbf{S}_8$ then $\mathcal{C}_{\mathbf{S}_8}(\sigma)$ has cardinality $5\cdot(8-5)! \ = \ 5\cdot 3! \ = \ 30$.

Corollary 87 provides information only about the *size* of $\mathcal{C}_{\mathbf{S}_N}(\sigma)$ but in some cases this is enough to exactly determine $\mathcal{C}_{\mathbf{S}_N}(\sigma)$....

Corollary 89 *Suppose $\sigma = (a_1 a_2 \ldots a_M)$ is a single M-cycle, for some $M \leq N$. Let $\mathcal{A} = \{a_1, a_2, \ldots, a_n\}$, and let $\mathcal{B} = [1..N] \setminus \mathcal{A}$. Then*

$$\mathcal{C}_{\mathbf{S}_N}(\sigma) \quad = \quad \{\sigma^m \tau \ ; \ m \in [0..M), \ \tau \text{ a permutation of } \mathcal{B}\}$$

Proof: Let $\mathcal{S} = \{\sigma^m \tau \ ; \ m \in [0..M), \ \tau \text{ a permutation of } \mathcal{B}\}$.

Claim 1: $\mathcal{S} \ \subset \ \mathcal{C}_{\mathbf{S}_N}(\sigma)$.

Proof: Clearly, any power of σ commutes with σ, and if τ is a permutation of $\mathcal{B}$, then τ and σ do not touch any of the same elements of $[1..N]$, so τ also commutes with σ. Thus, any element $\sigma^m\tau$ in $\mathcal{S}$ commutes with σ. .... □ [Claim 1]

Now, Corollary 87 says that $|\mathcal{C}_{\mathbf{S}_N}(\sigma)| \ = \ M \cdot (N-M)!$.

Claim 2: $|\mathcal{S}| = M \cdot (N-M)!$ *also.*

Proof: Observe that $|\mathcal{B}| \ = \ N-M$, so the number of permutations τ of $\mathcal{B}$ is $(N-M)!$. The number of distinct powers of σ is M. Thus, the number of products $\sigma^m \cdot \tau$ is $M \cdot (N-M)!$. □ [Claim 2]

Combining Claims 1 and 2, we conclude that $\mathcal{S} \ = \ \mathcal{C}_{\mathbf{S}_N}(\sigma)$. □

Example 90: Consider $\mathbf{S}_5$. If $\sigma = (1\ 2\ 3)$ then

$$\begin{aligned}\mathcal{C}_{\mathbf{S}_5}(\sigma) &= \{\sigma^n\tau \ ; \ n = 0,1,2, \ \tau \text{ permutes } \{4,5\}\} \ = \ \{(1\ 2\ 3)^n (4\ 5)^m \ ; \ n = 0,1,2, \ m = 0,1\} \\ &= \Big\{e, \ (4\,5), \ (1\,2\,3), \ (1\,2\,3)(4\,5), \ (1\,3\,2), \ (1\,3\,2)(4\,5)\Big\}.\end{aligned}$$

If $\sigma = (12345)$ then $\mathcal{C}_{\mathbf{S}_5}(\sigma) = \{\sigma^n \ ; \ n = 0,1,2,3,4\} = \{e, \ (12345), \ (13524), \ (14253), \ (15432)\}$.

Proposition 91 $\mathbf{A}_5$ *is simple.*

Proof: We will apply Corollary 79. We need to measure the conjugacy classes of $\mathbf{A}_5$.

Claim 1: *Let $\sigma, \tau \in \mathbf{A}_N$. If σ and τ are conjugate in $\mathbf{A}_N$, then they have the same cycle type.*

Proof: If σ and τ are conjugate in $\mathbf{A}_N$, then $\sigma = \alpha\tau\alpha^{-1}$ for some $\alpha \in \mathbf{A}_N$. Since $\mathbf{A}_N \subset \mathbf{S}_N$, this means that If σ and τ are also conjugate in $\mathbf{S}_N$. Now apply Proposition 83. □

(Note that, unlike Proposition 83, the statement of Claim 1 is *not* 'if and only if'.)

Claim 2: *The distinct cycle types of $\mathbf{A}_5$ are:*

$$(a\ b\ c), \qquad (a\ b\ c\ d\ e), \qquad (a\ b)(c\ d)$$

and the identity element e.

Proof: Exercise. .... □ `[Claim 2]`

Claim 3: *There are exactly 20 elements in $\mathbf{A}_5$ of cycle type $(a\ b\ c)$, and all are conjugate.*

Proof: As in Corollary 87, the number of elements of type $(a\ b\ c)$ is $\frac{5\cdot4\cdot3}{3} = 20$. To see that all are conjugate, consider the element $\sigma = (123)$. We want to show that $|\mathcal{K}(\sigma)| = 20$. Lemma 86 says

$$|\mathcal{K}(\sigma)| \quad = \quad \frac{|\mathbf{A}_5|}{|\mathcal{C}_{\mathbf{A}_5}(\sigma)} \quad = \quad \frac{60}{|\mathcal{C}_{\mathbf{A}_5}(\sigma)|},$$

so it suffices to show that $|\mathcal{C}_{\mathbf{A}_5}(\sigma)| = 3$.

Now, $\mathcal{C}_{\mathbf{A}_5}(\sigma) = \mathbf{A}_5 \cap \mathcal{C}_{\mathbf{S}_5}(\sigma)$, and in Example 90 we found that

$$\mathcal{C}_{\mathbf{S}_5}(\sigma) \ = \ \Big\{e,\ (45),\ (123),\ (123)(45),\ (132),\ (132)(45)\Big\}.$$

Of the six elements in this list, only three are in $\mathbf{A}_5$, namely, e, (123), and (132). Thus, $\mathcal{C}_{\mathbf{A}_5}(\sigma) = \{e,\ (123),\ (132)\}$ has cardinality 3, so that $|\mathcal{K}(\sigma)|$ has cardinality $60/3 = 20$, as desired. .... □ `[Claim 3]`

Claim 4: *There are exactly 24 elements in $\mathbf{A}_5$ of cycle type $(a\ b\ c\ d\ e)$, and they fall into two conjugacy classes, having 12 elements each.*

Proof: As in Corollary 87, the number of elements of type $(a\ b\ c\ d\ e)$ is $\frac{5\cdot4\cdot3\cdot2}{5} = 24$. Consider the element $\sigma_1 = (1\ 2\ 3\ 4\ 5)$; we'll show that $|\mathcal{K}(\sigma_1)| = 12$. As in Claim 3, we observe that

$$|\mathcal{K}(\sigma_1)| \quad = \quad \frac{|\mathbf{A}_5|}{|\mathcal{C}_{\mathbf{A}_5}(\sigma_1)} \quad = \quad \frac{60}{|\mathcal{C}_{\mathbf{A}_5}(\sigma)|},$$

so it suffices to show that $|\mathcal{C}_{\mathbf{A}_5}(\sigma_1)| = 5$. Again, $\mathcal{C}_{\mathbf{A}_5}(\sigma_1) = \mathbf{A}_5 \cap \mathcal{C}_{\mathbf{S}_5}(\sigma)$, and in Example 90 we found that $\mathcal{C}_{\mathbf{S}_5}(\sigma_1) = \{e,\ (12345),\ (13524),\ (14253),\ (15432)\}$. All five of these elements are in $\mathbf{A}_5$, so we're done.

Next, let $\sigma_2 = (13524)$; by similar reasoning, $|\mathcal{K}(\sigma_2)| = 12$. Thus, between them, $\mathcal{K}(\sigma_1)$ and $\mathcal{K}(\sigma_2)$ cover all 24 elements of type $(a\ b\ c\ d\ e)$.

It remains to show that σ_1 and σ_2 are not conjugate in $\mathbf{A}_5$. This is **Exercise 39** (Hint: Show that, if $\tau\sigma_2\tau^{-1} = \sigma_1$, then $\tau = \sigma_1^n \cdot \beta$, where $\beta = (2453)$. Note that β is odd —thus, τ must be odd. Thus, σ_1 and σ_2 cannot be conjugated with an *even* permutation, so they are not conjugate in $\mathbf{A}_5$.) .... □ `[Claim 4]`

Claim 5: *There are exactly 15 elements in $\mathbf{A}_5$ of cycle type $(a\ b)(c\ d)$, and all are conjugate.*

Proof: Exercise. Hint: let $\sigma = (12)(34)$, and show that $|\mathcal{C}_{\mathbf{A}_5}(\sigma)| = 4$. .. □ `[Claim 5]`

Combining Claims 2, 3, 4, and 5, we see that $\mathbf{A}_5$ fits the description of Example ⟨80⟩. Thus, it must be simple. ________ □

Proposition 92 *For any $N \geq 5$, $\mathbf{A}_N$ is simple.*

Proof: (by induction on N)

Base Case: $(N = 5)$ This is Proposition 91.

Induction: Let $N \geq 6$, and suppose $\mathbf{A}_{(N-1)}$ is simple. Let $\mathbf{H} \lhd \mathbf{A}_N$ be some normal subgroup. To show that $\mathbf{A}_N$ is simple, we must show that either $\mathbf{H} = \{e\}$ or $\mathbf{H} = \mathbf{A}_N$. Suppose, by contradiction, that $\{e\} \neq \mathbf{H} \neq \mathbf{A}_N$.

Recall that $\mathbf{A}_N \subset \mathbf{S}_N$, and that $\mathbf{S}_N$ acts on $[1..N]$ by permutations.

Claim 1: *If $\tau \in \mathbf{H}$, and $\tau \neq e$, then τ fixes no element of $[1..N]$. That is, $\tau(n) \neq n$ for any $n \in [1..N]$.*

Proof: Let $\mathbf{G}_n = \mathsf{Stab}(n) = \{\alpha \in \mathbf{A}_N \ ; \ \alpha(n) = n\}$.

Claim 1.1: $\mathbf{G}_n \cong \mathbf{A}_{N-1}$.

Proof: Let $\mathbf{F}_n = \{\sigma \in \mathbf{S}_N \ ; \ \sigma(n) = n\}$. Then $\mathbf{F}_n$ is really the set of all permutations of $\{1, 2, \ldots, n-1, \cancel{n}, n+1, \ldots, N\}$, a set having $N-1$ elements, so $\mathbf{F}_n \cong \mathbf{S}_{N-1}$. But $\mathbf{G}_n = \mathbf{A}_N \cap \mathbf{F}_n$ is just the set of *even* elements in $\mathbf{F}_n$ —hence, $\mathbf{G}_n \cong \mathbf{A}_{N-1}$. □ `[Claim 1.1]`

Claim 1.2: *If $e \neq \tau \in \mathbf{H}$, and $\tau(n) = n$, then $\mathbf{G}_n \subset \mathbf{H}$.*

Proof: If $\tau(n) = n$, then $\tau \in \mathbf{H} \cap \mathbf{G}_n$. Thus, $\mathbf{H} \cap \mathbf{G}_n \neq \{e\}$. But $\mathbf{H} \lhd \mathbf{A}_N$, so that $\mathbf{H} \cap \mathbf{G}_n \lhd \mathbf{G}_n$. But by Claim 1, $\mathbf{G}_n$ is isomorphic to $\mathbf{A}_{N-1}$, and by induction hypothesis, $\mathbf{A}_{N-1}$ is simple. Thus, $\mathbf{G}_n$ has no nontrivial proper normal subgroups. Hence, we must have $\mathbf{H} \cap \mathbf{G}_n = \mathbf{G}_n$, which means $\mathbf{G}_n \subset \mathbf{H}$. .... □ `[Claim 1.2]`

Claim 1.3: *If $\mathbf{G}_n \subset \mathbf{H}$ for some $n \in [1..N]$, then $\mathbf{G}_m \subset \mathbf{H}$ for all $m \in [1..N]$.*

Proof: Let $\sigma \in \mathbf{A}_N$ be such that $\sigma(n) = m$. Thus, $\sigma \mathbf{G}_n \sigma^{-1} = \sigma \left(\mathsf{Stab}(n)\right) \sigma^{-1} \underset{\text{Lemma 69}}{=} \mathsf{Stab}(\sigma(n)) = \mathsf{Stab}(m) = \mathbf{G}_m$. Since $\mathbf{G}_n \subset \mathbf{H}$, it follows that $\mathbf{G}_m \subset \sigma \mathbf{H} \sigma^{-1}$. However, $\sigma \mathbf{H} \sigma^{-1} = \mathbf{H}$, because $\mathbf{H}$ is normal. .... □ `[Claim 1.3]`

Thus, $\mathbf{G}_1, \mathbf{G}_2, \ldots, \mathbf{G}_N \subset \mathbf{H}$, so that $\langle \mathbf{G}_1, \mathbf{G}_2, \ldots, \mathbf{G}_N \rangle \subset \mathbf{H}$.

Claim 1.4: $\langle \mathbf{G}_1, \mathbf{G}_2, \ldots, \mathbf{G}_N \rangle = \mathbf{A}_N$.

Proof: We want to show that any element of $\mathbf{A}_N$ can be written as a product of elements from $\mathbf{G}_1, \mathbf{G}_2, \ldots, \mathbf{G}_N$. So, let $\sigma \in \mathbf{A}_N$. Since σ is even, we can write it as a product of an even number of 2-cycles:

$$\begin{array}{rcl} \sigma & = & \underbrace{\alpha_1\beta_1}_{} \ \underbrace{\alpha_2\beta_2}_{} \ \ldots \ \underbrace{\alpha_K\beta_K}_{} \\ & = & \lambda_1 \quad \lambda_2 \quad \ldots \quad \lambda_K \end{array}$$

Here, $\lambda_k = \alpha_k\beta_k$ is a product of two 2-cycles.

Claim 1.4.1: *For all k, $\lambda_k \in \mathbf{G}_{n_k}$ for some n_k*

Proof: Suppose $\alpha_k = (a_1a_2)$ and $\beta_k = (b_1b_2)$, for some $a_1, a_2, b_1, b_2 \in [1..N]$. Thus, $\lambda_k = (a_1a_2)(b_1b_2)$ permutes the elements a_1, a_2, b_1, b_2, but fixes all other elements in $[1..N]$. But $N \geq 5$ by hypothesis, so λ_k must fix at least one $n \in [1..N]$. Thus, $\lambda_k \in \mathbf{G}_n$. .... □ `[Claim 1.4.1]`

Thus, $\sigma = \lambda_1\lambda_2\ldots\lambda_K$ is a product of elements from $\mathbf{G}_1, \mathbf{G}_2, \ldots, \mathbf{G}_N$. □ `[Claim 1.4]`

So, if $\tau \in \mathbf{H}$, and $\tau(n) = n$, then Claim 1.2 says $\mathbf{G}_n \subset \mathbf{H}$. Then Claim 1.3 says that $\mathbf{G}_m \subset \mathbf{H}$ for all m. But then Claim 1.4 says $\mathbf{A}_N \subset \mathbf{H}$ —in other words, $\mathbf{H} = \mathbf{A}_N$. □ `[Claim 1]`

Claim 2: *Let $\tau_1, \tau_2 \in \mathbf{H}$. If there is any $n \in [1..N]$ so that $\tau_1(n) = \tau_2(n)$, then $\tau_1 = \tau_2$.*

Proof: If τ_1 and τ_2 are elements of $\mathbf{H}$, then $\tau_1^{-1}\tau_2 \in \mathbf{H}$ also. But if $\tau_1(n) = \tau_2(n)$, then $\tau_1^{-1}\tau_2(n) = n$. By Claim 1, this means that $\tau_1^{-1}\tau_2 = e$. Thus, $\tau_1 = \tau_2$. ... □ `[Claim 2]`

Claim 3: *All elements of $\mathbf{H}$ are products of disjoint 2-cycles. In other words, if $\tau \in \mathbf{H}$, then $\tau = (ab)(cd)(ef)\ldots$, where $a, b, c, d, e, f, \ldots$ are all distinct.*

Proof: Suppose τ contained a cycle of length 3 or longer. Thus,

$$\tau \quad = \quad (a_1a_2a_3\ldots)(b_1b_2\ldots)\ldots(c_1c_2\ldots).$$

for some $a_1, a_2, a_3, \ldots, b_1, b_2, \ldots$ in $[1..N]$. Let $x, y \in [1..N]$ be any elements, such that $x \neq a_3$, and let $\sigma = (a_3xy)$. Then $\sigma \in \mathbf{A}_N$.

Let $\tau_1 = \sigma\tau\sigma^{-1}$. Since $\mathbf{H}$ is normal, and $\tau \in \mathbf{H}$, we know that $\tau_1 \in \mathbf{H}$ also. By Lemma 81,

$$\begin{array}{rcl} \tau_1 & = & \Big(\sigma(a_1)\sigma(a_2)\sigma(a_3)\ldots\Big)\Big(\sigma(b_1)\sigma(b_2)\ldots\Big)\ldots\Big(\sigma(c_1)\sigma(c_2)\ldots\Big) \\ & = & \Big(a_1a_2\sigma(a_3)\ldots\Big)\Big(\sigma(b_1)\sigma(b_2)\ldots\Big)\ldots\Big(\sigma(c_1)\sigma(c_2)\ldots\Big). \end{array}$$

Thus, $\tau_1(a_1) = a_2 = \tau(a_1)$. On the other hand, $\tau_1(a_2) = x \neq a_3 = \tau(a_2)$, so τ_1 and τ_2 cannot be equal. This contradicts Claim 2. .... □ `[Claim 3]`

Let $a, b, c, d, e, f \in [1..N]$ be distinct elements (this is possible because $N \geq 6$ by hypothesis), and let $\tau = (ab)(cd)(ef)\ldots$ be some element of $\mathbf{H}$ (by Claim 3, any element of $\mathbf{H}$ must have this form). Let $\sigma = (ab)(de)$. Then $\sigma \in \mathbf{A}_N$, so $\tau_1 = \sigma\tau\sigma^{-1} \in \mathbf{H}$ also. But Lemma 81 says that $\tau_1 = (ab)(ce)(df)$. Thus, $\tau_1 \neq \tau$, but $\tau(a) = b = \tau_1(a)$. Again, we have a contradiction of Claim 2.

By contradiction, we must conclude that $\mathbf{H} = \mathbf{A}_N$. Hence, $\mathbf{A}_N$ is simple. ______________ □

Part II

Rings

Chapter 7

Introduction

7.1 Basic Definitions and Examples

Recall the basic properties of integer arithmetic. The set of integers $\mathbb{Z}$ comes with two operations, *addition* and *multiplication*, which have the following properties:

1. $(\mathbb{Z}, +)$ is an *abelian group*. That is:

   (a) '+' is *associative*: For all $n, m, \ell \in \mathbb{Z}$, $\quad n + (m + \ell) \ = \ (n + m) + \ell$.

   (b) '+' is *commutative*: For all $n, m \in \mathbb{Z}$, $\quad n + m \ = \ m + n$.

   (c) There is an *additive identity* element (namely, 0), such that for any $n \in \mathbb{Z}$, $\quad 0+n \ = \ n \ = \ n + 0$.

   (d) Every element $n \in \mathbb{Z}$ has an *additive inverse* (namely $-n$), so that $n+(-n) \ = \ 0 \ = \ (-n) + n$.

2. $(\mathbb{Z}, \cdot)$ is not a group. For example, the element 2 has no multiplicative inverse in $\mathbb{Z}$ (since $\frac{1}{2}$ is not an integer).

   However, '$\cdot$' is still *associative*: For all $n, m, \ell \in \mathbb{Z}$, $\quad n \cdot (m \cdot \ell) \ = \ (n \cdot m) \cdot \ell$.

3. $\mathbb{Z}$ satisfies the *distributive law*: For all $n, m, \ell \in \mathbb{Z}$, $\quad n \cdot (m + \ell) \ = \ (n \cdot m) + (n \cdot \ell)$.

A ring is an abstract algebraic object which mimics these properties. To be precise, a **ring** is a set $\mathcal{R}$, equipped with two binary operations, '+' and '$\cdot$', so that:

1. $(\mathcal{R}, +)$ is an *abelian group*. That is:

   (a) '+' is *associative*: For all $r, s, t \in \mathcal{R}$, $\quad r + (s + t) \ = \ (r + s) + t$.

   (b) '+' is *commutative*: For all $r, s \in \mathcal{R}$, $\quad r + s \ = \ s + r$.

   (c) There is an *additive identity* element $0 \in \mathcal{R}$ so that for any $r \in \mathcal{R}$, $\quad 0 + r \ = \ r \ = \ r + 0$.

(d) Every element $r \in \mathcal{R}$ has an *additive inverse*, denoted $-r$, so that $r + (-r) = 0 = (-r) + r$

2. $(\mathcal{R}, \cdot)$ is not a group. However, '$\cdot$' is still *associative*: For all $r, s, t \in \mathcal{R}$, $r\cdot(s\cdot t) = (r\cdot s)\cdot t$.

3. $\mathcal{R}$ satisfies the *distributive law*: For all $r, s, t \in \mathcal{R}$, $r \cdot (s + t) = (r \cdot s) + (r \cdot t)$, and also $(s + t) \cdot r = (s \cdot r) + (t \cdot r)$.

Notice:

1. The operation '$\cdot$' need not be *commutative* (although it is for integers). If '$\cdot$' is commutative, then we call $\mathcal{R}$ a **commutative ring**, or a **domain**.

2. A **multiplicative identity** is an element $e \in \mathcal{R}$ so that, for any $r \in \mathcal{R}$, $e\cdot r = r = r\cdot e$. If such an element exists, it is unique, and is denoted by '1' (as in the integers).

   The set $\mathcal{R}$ may not have a multiplicative identity (although it almost always does). A ring *without* a multiplicative identity is sometimes called a **rng**. A ring *having* a multiplicative identity is sometimes called a **ring with identity**.

Example 93:

(a) Let $2\mathbb{Z} = \{\ldots, -4, -2, 0, 2, 4, 6, 8, \ldots\}$ be the set of all even numbers. Then $2\mathbb{Z}$ is a commutative ring under the normal arithmetic operations of addition and multiplication. However, $2\mathbb{Z}$ has no identity element.

(b) More generally, for any $n \in \mathbb{N}$, let $n\mathbb{Z} = \{nz\ ;\ z \in \mathbb{Z}\}$ be the set of all multiples of n. Then $n\mathbb{Z}$ is a commutative ring under the normal arithmetic operations of addition and multiplication. However, unless $n = 1$, the ring $n\mathbb{Z}$ has no identity element.

(c) **Rational numbers:** Let $\mathbb{Q} = \left\{\frac{a}{b}\ ;\ a, b \in \mathbb{Z}\right\}$ be the set of rational numbers. We define addition and multiplication on $\mathbb{Q}$ in the obvious way:

$$\frac{a_1}{b_1} + \frac{a_2}{b_2} = \frac{a_1 b_2 + a_2 b_1}{b_1 b_2}, \qquad \text{and} \qquad \frac{a_1}{b_1} \cdot \frac{a_2}{b_2} = \frac{a_1 \cdot a_2}{b_1 \cdot b_2}. \tag{7.1}$$

Then $\mathbb{Q}$ is a commutative ring, with additive identity 0 and multiplicative identity 1.

(d) **Real numbers:** Let $\mathbb{R}$ be the set of real numbers. Then $\mathbb{R}$ is a commutative ring under the normal arithmetic operations of addition and multiplication. It has a multiplicative identity, namely 1.

(e) **Complex numbers:** Let $\mathbb{C}$ be the set of complex numbers. That is, $\mathbb{C} = \{x + y\mathbf{i}\ ;\ x, y \in \mathbb{R}\}$, where $\mathbf{i}$ is the square root of negative one, satisfying the equation: $\mathbf{i}^2 = -1$. The arithmetic operations on $\mathbb{C}$ are defined:

$$\left.\begin{array}{rcl} (x_1 + y_1\mathbf{i}) + (x_2 + y_2\mathbf{i}) & = & (x_1 + x_2) + (y_1 + y_2)\mathbf{i} \\ (x_1 + y_1\mathbf{i}) \cdot (x_2 + y_2\mathbf{i}) & = & (x_1 y_1 - x_2 y_2) + (x_1 y_2 + x_2 y_1)\mathbf{i} \end{array}\right\} \tag{7.2}$$

Then $\mathbb{C}$ is a commutative ring, with multiplicative identity $1 = 1 + 0\mathbf{i}$.

(f) **Gaussian Integers:** Let $\mathbb{Z}[\mathbf{i}] = \{z + y\mathbf{i}\ ;\ z, y \in \mathbb{Z}\}$, where $\mathbf{i}$ is the square root of negative one, satisfying the equation: $\mathbf{i}^2 = -1$. Thus, $\mathbb{Z}[\mathbf{i}]$ is a subset of $\mathbb{C}$, and the arithmetic operations on $\mathbb{Z}[\mathbf{i}]$ are again defined by the equations (7.2). It is left as **Exercise 40** to show that $\mathbb{Z}[\mathbf{i}]$ is closed under these operations and forms a commutative ring. Num.Thr.

(g) **Hamiltonians:** The ring of Hamiltonions is defined as follows. First, we introduce three formal elements, $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$, all of which are square roots of -1. That is:

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 \;=\; -1.$$

Furthermore, $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$, are related by the following expressions:

$$\left.\begin{array}{rclcrclcrcl} \mathbf{ij} & = & \mathbf{k}; & \qquad & \mathbf{jk} & = & \mathbf{i}; & \qquad & \mathbf{ki} & = & \mathbf{j}. \\ \mathbf{ji} & = & -\mathbf{k}; & & \mathbf{kj} & = & -\mathbf{i}; & & \mathbf{ik} & = & -\mathbf{j}. \end{array}\right\} \tag{7.3}$$

(Note that this multiplication is *not* commutative)

The **Hamiltonions** are the set $\mathbb{H}$ of all formal linear combinations of the form

$$w + x\mathbf{i} + y\mathbf{j} + z\mathbf{k}, \qquad \text{where } w, x, y, z \text{ are any real numbers.}$$

We define addition in the obvious fashion:

$$(w_1{+}x_1\mathbf{i}{+}y_1\mathbf{j}{+}z_1\mathbf{k}) + (w_2{+}x_2\mathbf{i}{+}y_2\mathbf{j}{+}z_2\mathbf{k}) \;=\; (w_1{+}w_2) + (x_1{+}x_2)\mathbf{i} + (y_1{+}y_2)\mathbf{j} + (z_1{+}z_2)\mathbf{k}.$$

We define addition by applying the formulae (7.3), and legislating that multiplication is distributive. For example:

$$\begin{aligned} (3 + 4\mathbf{i} - 5\mathbf{j}) \cdot (2 + 6\mathbf{i}) &= 3(2 + 6\mathbf{i}) \;-\; 4\mathbf{i}(2 + 6\mathbf{i}) \;-\; 5\mathbf{j}(2 + 6\mathbf{i}) \\ &= (6 + 12\mathbf{i}) \;-\; (8\mathbf{i} + 24\mathbf{i}^2) \;-\; (10\mathbf{j} + 30\mathbf{ji}) \\ &= 6 + 12\mathbf{i} \;-\; 8\mathbf{i} - 24 \;-\; 10\mathbf{j} - 30\mathbf{k} \\ &= (6 - 24) \;+\; (12 - 8)\mathbf{i} \;-\; 10\mathbf{j} \;-\; 30\mathbf{k}. \end{aligned}$$

$\mathbb{H}$ is a *noncommutative* ring, with additive identity $0 = 0 + 0\mathbf{i} + 0\mathbf{j} + 0\mathbf{k}$ and multiplicative identity element $1 = 1 + 0\mathbf{i} + 0\mathbf{j} + 0\mathbf{k}$. Note that $\mathbb{H}$ and is an 'extension' of $\mathbb{C}$ in the same way that $\mathbb{C}$ is an extension of $\mathbb{R}$.

(h) Let $\mathbb{Z}_{/5} = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}, \bar{4}\}$ be the set of congruence classes of integers, mod 5, and let '+' and '$\cdot$' be addition and multiplication, mod 5. For example:

$$\bar{2} + \bar{4} \;=\; \bar{1}, \qquad \text{and} \qquad \bar{2} \cdot \bar{4} \;=\; \bar{3}.$$

Then $\mathbb{Z}_{/5}$ forms a commutative ring, with additive identity element $\bar{0}$, and multiplicative identity $\bar{1}$.

(i) More generally, for any $n \in \mathbb{N}$, let $\mathbb{Z}_{/n}$ be the set of congruence classes of integers, mod n, and let '+' and '·' be addition and multiplication, mod n. Then $\mathbb{Z}_{/n}$ forms a commutative ring, with additive identity element $\bar{0}$, and multiplicative identity $\bar{1}$.

(j) **Trivial Ring:** The set $\{0\}$ is a ring, where $0+0=0=0\cdot 0$. It is called the *trivial* ring.

The axioms of ring arithmetic have some basic consequences:

Proposition 94 *Let $\mathcal{R}$ be a ring.*

(a) *If $\mathcal{R}$ has additive identity 0, then for any $r \in \mathcal{R}$,* $\quad r\cdot 0 = 0 = 0\cdot r.$

(b) *Each element in $\mathcal{R}$ has a* unique *additive inverse.*

(c) *If $s \in \mathcal{R}$ has additive inverse $-s$, then for any $r \in \mathcal{R}$,* $\quad (-s)\cdot r \;=\; -(s\cdot r) \;=\; s\cdot(-r).$

In particular, for any $r \in \mathcal{R}$, $\quad (-1)\cdot r \;=\; -r \;=\; r\cdot(-1).$

(d) *If $\mathcal{R}$ has a multiplicative identity, it is* unique. *In other words, if e_1 and e_2 are two elements of $\mathcal{R}$ such that $e_1 \cdot r = r$ and $e_2 \cdot r = r$ for any $r \in \mathcal{R}$, then $e_1 = e_2$.*

The unique additive inverse is denoted 1, or $1_{\mathcal{R}}$.

Note: *If $e \in \mathcal{R}$, then the fact that $e\cdot r = r$ for* some *$r \in \mathcal{R}$, is* not *sufficient to conclude that e is the multiplicative identity. See* **Remark B** *of Example ⟨95a⟩ below.*

(e) *If $r \in \mathcal{R}$ has a multiplicative inverse, then it is* unique. *In other words, if i_1 and i_2 are two elements of $\mathcal{R}$ such that $i_1 \cdot r = 1$ and $i_2 \cdot r = 1$, then $i_1 = i_2$.*

Proof: **Exercise 41** □

7.2 (∗) Many Examples of Rings

Prerequisites: §7.1

[This section contains many examples. It is not necessary to read and understand all these examples right now; just read a few in order to get your intuititions working. These examples will be referred to throughout the text.]

Example 95: (Product Rings)

⟨a⟩ Let $\mathcal{R} = \mathbb{R}^2$, and define addition and multiplication *componentwise.* That is, for any $(x_1, y_1) \in \mathbb{R}^2$ and $(x_2, y_2) \in \mathbb{R}^2$,

$$\begin{aligned}(x_1,y_1)+(x_2,y_2) &= (x_1+y_1,\ x_2+y_2)\\ (x_1,y_1)\cdot(x_2,y_2) &= (x_1\cdot y_1,\ x_2\cdot y_2)\end{aligned}$$

Then $\mathbb{R}^2$ is a commutative ring, with additive identity $\mathbf{0} = (0,0)$ and multiplicative identity $\mathbf{1} = (1,1)$.

Remark A: The *abelian group* $(\mathbb{R}^2,+)$ is isomorphic to the group $(\mathbb{C},+)$ (see Example ⟨93e⟩). However, *rings* $(\mathbb{R}^2,+,\cdot)$ is *not* isomorphic to the ring $(\mathbb{C},+,\cdot)$. Thus, we have two rings which have identical additive structures, but which are not the same as rings.

Remark B: Let $\mathbf{e} = (1,0)$. Observe that, for any element $\mathbf{x} = (x,0)$ in $\mathbb{R}^2$, we have $\mathbf{e}\cdot\mathbf{x} = \mathbf{x}$. However, $\mathbf{e}$ is *not* the multiplicative identity of $\mathbb{R}^2$. This is an example of the warning in Proposition (94d).

⟨b⟩ More generally, if $\mathcal{R}$ and $\mathcal{S}$ are any two rings, we define the **product ring** $\mathcal{R}\times\mathcal{S}$ to have componentwise addition and multiplication. That is, for any $r_1,r_2\in\mathcal{R}$ and $s_1,s_2\in\mathcal{S}$,

$$\begin{aligned}(r_1,s_1)+(r_2,s_2) &= (r_1+s_1,\ r_2+s_2)\\ (r_1,s_1)\cdot(r_2,s_2) &= (r_1\cdot s_1,\ r_2\cdot s_2)\end{aligned}$$

Then $\mathcal{R}\times\mathcal{S}$ is a ring. Furthermore,

- If $\mathcal{R}$ has additive identity $0_{\mathcal{R}}$ and $\mathcal{S}$ has additive identity $0_{\mathcal{S}}$, then $\mathcal{R}\times\mathcal{S}$ has additive identity $\mathbf{0} = (0_{\mathcal{R}},\ 0_{\mathcal{S}})$.
- If $\mathcal{R}$ has multiplicative identity $1_{\mathcal{R}}$ and $\mathcal{S}$ has multiplicative identity $1_{\mathcal{S}}$, then $\mathcal{R}\times\mathcal{S}$ has multiplicative identity $\mathbf{1} = (1_{\mathcal{R}},\ 1_{\mathcal{S}})$.
- $\mathcal{R}\times\mathcal{S}$ is commutative if and only if $\mathcal{R}$ and $\mathcal{S}$ are both commutative.

Exercise 42 Verify these claims.

⟨c⟩ Let $\mathcal{R}_1,\mathcal{R}_2,\ldots,\mathcal{R}_n$ be rings; by iterating this construction, we can define the product ring $\mathcal{R}_1\times\mathcal{R}_2\times\ldots\times\mathcal{R}_n$.

⟨d⟩ In particular, if $\mathcal{R}$ is any ring, then we define $\mathcal{R}^n = \underbrace{\mathcal{R}\times\mathcal{R}\times\ldots\times\mathcal{R}}_{n}$.

Example 96: (Matrix Rings)

Grp.Repr.

⟨a⟩ Let $\mathcal{M}_2(\mathbb{R})$ be the set of all 2×2 matrices with real coefficients. We define matrix addition and multiplication in the familiar manner:

$$\begin{aligned}\begin{bmatrix}a_1 & b_1\\ c_1 & d_1\end{bmatrix}+\begin{bmatrix}a_2 & b_2\\ c_2 & d_2\end{bmatrix} &= \begin{bmatrix}(a_1+a_2) & (b_1+b_2)\\ (c_1+c_2) & (d_1+d_2)\end{bmatrix};\\ \begin{bmatrix}a_1 & b_1\\ c_1 & d_1\end{bmatrix}\cdot\begin{bmatrix}a_2 & b_2\\ c_2 & d_2\end{bmatrix} &= \begin{bmatrix}(a_1a_2+b_1c_2) & (a_1b_2+b_1d_2)\\ (c_1a_2+d_1c_2) & (c_1b_2+d_1d_2)\end{bmatrix}.\end{aligned}$$

Then $\mathcal{M}_2(\mathbb{R})$ is a *noncommutative* ring, with additive identity element $\mathbf{0} = \left[\begin{smallmatrix}0&0\\0&0\end{smallmatrix}\right]$, and multiplicative identity element $\mathbf{Id} = \left[\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right]$ (**Exercise 43**).

Remark: The *abelian groups* $(\mathbb{R}^4, +)$, $(\mathbb{H}, +)$ and $(\mathcal{M}_2(\mathbb{R}), +)$ are all isomorphic as groups (see Examples (93g) and (95)). However, the *rings* $(\mathbb{R}^4, +, \cdot)$, $(\mathbb{H}, +, \cdot)$ and $(\mathcal{M}_2(\mathbb{R}), +, \cdot)$ are all different.

⟨b⟩ More generally, let $\mathcal{M}_n(\mathbb{R})$ be the set of all $n \times n$ matrices with real coefficients. We define matrix addition and multiplication in the familiar manner. Then $\mathcal{M}_n(\mathbb{R})$ is a *noncommutative* ring (**Exercise 44**), with

$$\text{Additive identity} \quad \mathbf{0} = \begin{bmatrix} 0 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & 0 \end{bmatrix} \quad \text{and multiplicative identity} \quad \mathbf{Id} = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix}.$$

⟨c⟩ Let $\mathcal{M}_n(\mathbb{Z})$ be the set of all $n \times n$ matrices with integer coefficients. For example, $\left[\begin{smallmatrix} 1 & -9 \\ 3 & 7 \end{smallmatrix}\right]$ is an element of $\mathcal{M}_2(\mathbb{Z})$, but $\left[\begin{smallmatrix} \frac{1}{2} & \frac{9}{5} \\ 3 & \frac{7}{2} \end{smallmatrix}\right]$ is not. Then $\mathcal{M}_n(\mathbb{Z})$ is closed under matrix multiplication, addition, and negation (**Exercise 45**), so it is also a (noncommutative) ring.

⟨d⟩ Let $\mathcal{M}_n(\mathbb{Q})$ be the set of all $n \times n$ matrices with rational coefficients. For example, $\left[\begin{smallmatrix} \frac{1}{2} & \frac{9}{5} \\ 3 & \frac{7}{2} \end{smallmatrix}\right]$ is an element of $\mathcal{M}_2(\mathbb{Q})$, but $\left[\begin{smallmatrix} \pi & \sqrt{3} \\ 3 & 7 \end{smallmatrix}\right]$ is not. Then $\mathcal{M}_n(\mathbb{Q})$ is closed under matrix multiplication, addition, and negation (**Exercise 46**), so it is also a (noncommutative) ring.

⟨e⟩ Let $\mathcal{M}_n(\mathbb{C})$ be the set of all $n \times n$ matrices with complex coefficients, and let $\mathcal{M}_n(\mathbb{H})$ be the set of all $n \times n$ matrices with Hamiltonian coefficients. Then $\mathcal{M}_n(\mathbb{C})$ and $\mathcal{M}_n(\mathbb{H})$ are (noncommutative) rings.

⟨f⟩ In general, if $\mathcal{R}$ is any ring, let $\mathcal{M}_n(\mathcal{R})$ be the set of all $n \times n$ matrices with coefficients in $\mathcal{R}$. We define the multiplication and addition of these matrices in a manner exactly analogous to that for real-valued matrices. Then $\mathcal{M}_n(\mathcal{R})$ is a ring. Observe:

1. Set $n = 1$; then $\mathcal{M}_1(\mathcal{R})$ is identical to $\mathcal{R}$.
2. If $n \geq 2$, then $\mathcal{M}_n(\mathcal{R})$ is never commutative (even if $\mathcal{R}$ is commutative).
3. If $\mathcal{R}$ has multiplicative identity 1, then $\mathcal{M}_n(\mathcal{R})$ has multiplicative identity $\mathbf{Id} = \left[\begin{smallmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{smallmatrix}\right]$.

Exercise 47 Verify these claims.

Example 97: (Continuous Function Rings)

⟨a⟩ Let $\mathcal{C}(\mathbb{R})$ be the set of all *continuous* real-valued functions $f : \mathbb{R}\longrightarrow\mathbb{R}$. We define addition and multiplication of functions *pointwise*. That is, for any $f, g : \mathbb{R}\longrightarrow\mathbb{R}$ and any $r \in \mathbb{R}$,

$$(f+g)(r) \;=\; f(r)+g(r) \quad\text{and}\quad (f\cdot g)(r) \;=\; f(r)\cdot g(r).$$

It is proved in first-year calculus courses that the sum of two continous functions is continuous, and the product of two continuous functions is continuous. Thus, $\mathcal{C}(\mathbb{R})$ is a commutative ring. Observe:

1. The additive identity element is the constant zero function $\mathbf{0}$, defined by the property: $\mathbf{0}(r) = 0$ for any $r \in \mathbb{R}$.
2. The multiplicative identity element is the constant one function $\mathbb{1}$, defined by the property: $\mathbb{1}(r) = 1$ for any $r \in \mathbb{R}$.
3. The additive inverse of $f : \mathbb{R}\longrightarrow\mathbb{R}$ is the function $-f$ (also continuous).

(**Exercise 48**)[Verify these claims.]

The algebraic structure of $\mathcal{C}(\mathbb{R})$ somehow 'encodes' the topological structure of $\mathbb{R}$. For example, there exists no element $f \in \mathcal{C}(\mathbb{R})$ such that f is zero everywhere in $(-\infty, 1)$, but $f(1) \neq 0$. This reflects the fact that the point 1 lies on the 'boundary' of the domain $(\infty, 1)$. This relationship between topological and algebraic structure is the starting point of *algebraic geometry*.

⟨b⟩ Let $-\infty \leq a < b \leq \infty$, and let $\mathcal{C}(a, b)$ be the set of all *continuous* real-valued functions $f : (a, b)\longrightarrow\mathbb{R}$, with pointwise addition and multiplication. Then $\mathcal{C}(a, b)$ is a commutative ring.

⟨c⟩ If $\mathbf{U} \subset \mathbb{R}^n$ is any open subset, then let $\mathcal{C}(\mathbf{U})$ be the set of continuous real-valued functions $f : \mathbf{U}\longrightarrow\mathbb{R}$, with pointwise addition and multiplication. Then $\mathcal{C}(\mathbf{U})$ is a commutative ring.

⟨d⟩ If $\mathbf{X}$ is any topological space, then let $\mathcal{C}(\mathbf{X})$ is the set of continuous real-valued functions $f : \mathbf{X}\longrightarrow\mathbb{R}$, with pointwise addition and multiplication. Then $\mathcal{C}(\mathbf{X})$ is a commutative ring.

⟨e⟩ Similarly, if $\mathbf{X}$ is any topological space, let $\mathcal{C}(\mathbf{X}; \mathbb{C})$ be the set of all continuous, *complex*-valued functions $f : \mathbf{X}\longrightarrow\mathbb{C}$, with pointwise addition and multiplication. Then $\mathcal{C}(\mathbf{X}; \mathbb{C})$ is a commutative ring.

⟨f⟩ Let $-\infty \leq a < b \leq \infty$, and let $\mathcal{C}^1(a, b)$ be the set of all *continuously differentiable* real-valued functions $f : (a, b)\longrightarrow\mathbb{R}$, with pointwise addition and multiplication. It is proved in first-year calculus that the sum of two differentiable functions is differentiable, and the product of two differentiable functions is differentiable. Thus, $\mathcal{C}^1(a, b)$ is a commutative ring, with additive identity $\mathbf{0}$ and multiplicative identity $\mathbb{1}$.

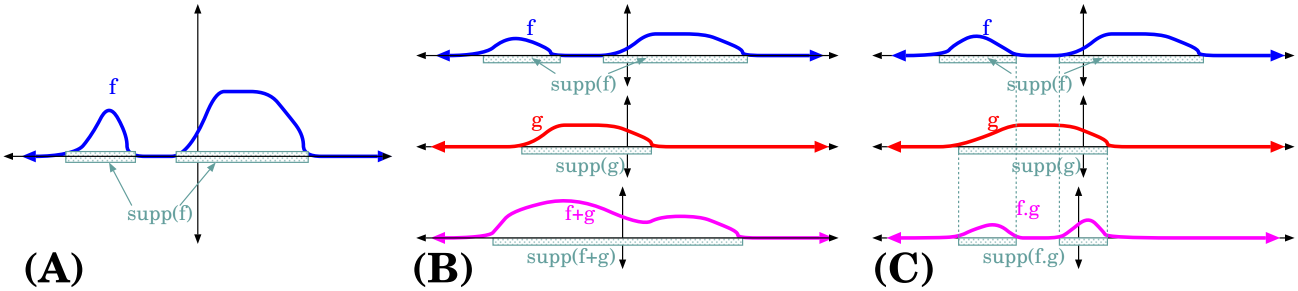


Figure 7.1: Functions with compact support.

⟨g⟩ Let $-\infty \le a < b \le \infty$. For any $k \in \mathbb{N}$, let $\mathcal{C}^k(a,b)$ be the set of all k *times continuously differentiable* real-valued functions $f:(a,b)\longrightarrow\mathbb{R}$, with pointwise addition and multiplication. Then $\mathcal{C}^k(a,b)$ is a commutative ring, with additive identity $\mathbf{0}$ and multiplicative identity $\mathbb{1}$ (**Exercise 49**).

⟨h⟩ Let $-\infty \le a < b \le \infty$, and let $\mathcal{C}^\infty(a,b)$ be the set of all *infinitely differentiable* real-valued functions $f:(a,b)\longrightarrow\mathbb{R}$, with pointwise addition and multiplication. Then $\mathcal{C}^\infty(a,b)$ is a commutative ring, with additive identity $\mathbf{0}$ and multiplicative identity $\mathbb{1}$ (**Exercise 50**).

⟨i⟩ If $\mathbf{U} \subset \mathbb{R}^n$ is an open subset, then

- $\mathcal{C}(\mathbf{U})$ is the set of continuous real-valued functions $f:\mathbf{U}\longrightarrow\mathbb{R}$.
- $\mathcal{C}^1(\mathbf{U})$ is the set of continuously differentiable real-valued functions $f:\mathbf{U}\longrightarrow\mathbb{R}$.
- $\mathcal{C}^k(\mathbf{U})$ is the set of k times continuously differentiable real-valued functions $f:\mathbf{U}\longrightarrow\mathbb{R}$.
- $\mathcal{C}^\infty(\mathbf{U})$ is the set of infinitely differentiable real-valued functions $f:\mathbf{U}\longrightarrow\mathbb{R}$.
- $\mathcal{C}^\omega(\mathbf{U})$ is the set of analytic real-valued functions $f:\mathbf{U}\longrightarrow\mathbb{R}$.

Then all of these are commutative rings.

⟨j⟩ If $\mathcal{M}$ is any differentiable manifold (eg. a sphere, torus, etc), then let $\mathcal{C}^\infty(\mathbf{U})$ be the set of infinitely differentiable real-valued functions $f:\mathbf{U}\longrightarrow\mathbb{R}$, with pointwise addition and multiplication. Then $\mathcal{C}^\infty(\mathbf{U})$ is a commutative ring.

⟨k⟩ **Functions of Compact Support:** If $f:\mathbb{R}\longrightarrow\mathbb{R}$, then the **support** of f is the set of all points where f is nonzero:

$$\mathsf{supp}[f] \quad = \quad \{r \in \mathbb{R}\,;\; f(r) \neq 0\} \qquad\qquad \text{(see Figure 7.1A)}$$

We say that f has **compact support** if $\mathsf{supp}[f]$ is a bounded subset of $\mathbb{R}$. Let $\mathcal{C}_0(\mathbb{R})$ be the set of all continuous functions $f:\mathbb{R}\longrightarrow\mathbb{R}$ with compact support.

If $f,g:\mathbb{R}\longrightarrow\mathbb{R}$, it is not hard to show (**Exercise 51**)that:

- $\mathsf{supp}\,[f+g] \subset \mathsf{supp}\,[f] \cup \mathsf{supp}\,[g]$ (see Figure 7.1B).
- $\mathsf{supp}\,[f\cdot g] \;= \mathsf{supp}\,[f] \cap \mathsf{supp}\,[g]$ (see Figure 7.1C).

Thus, if f and g both have compact support, then so do the functions $f+g$ and $f\cdot g$. Thus, $\mathcal{C}_0(\mathbb{R})$ is a commutative ring. Note, however, that $\mathcal{C}_0(\mathbb{R})$ has no multiplicative identity, because the function $1\!\!1$ does *not* have compact support: $\mathsf{supp}\,[1\!\!1] = \mathbb{R}$.

Example 98: (Polynomial Rings)

Alg.Geo. ,

Galois

⟨a⟩ **Polynomials of one variable:** Let $\mathbb{R}[x]$ denote the set of all *polynomial functions* on $\mathbb{R}$ —that is, functions of the form

$$P(x) \quad = \quad p_n x^n + p_{n-1}x^{n-1} + \ldots + p_2 x^2 + p_1 x + p_0,$$

where $p_n, p_{n-1}, \ldots, p_1, p_0$ are real numbers.

The sum or product of two polynomials is also a polynomial, so $\mathbb{R}[x]$ is a ring. Polynomial arithmetic can be described by the following formulae:

$$\begin{aligned}\Big(p_N x^N + \ldots + p_2x^2 + p_1x + p_0\Big) \;&+\; \Big(q_N x^N + \ldots + q_2x^2 + q_1x + q_0\Big)\\ &= \; (p_N+q_N)x^N \;+\; \ldots \;+\; (p_2+q_2)x^2 \;+\; (p_1+q_1)x \;+\; (p_0+q_0). \qquad (7.4)\\ \Big(p_N x^N + \ldots + p_2x^2 + p_1x + p_0\Big) \;&\cdot\; \Big(q_M x^M + \ldots + q_2x^2 + q_1x + q_0\Big)\\ &= \; p_n q_M x^{N+M} \;+\; (p_N q_{M-1} + p_{N-1}q^M)x^{N+M-1} \;+\; \ldots\\ &\qquad \ldots \;+\; (p_2q_0 + p_1q_1 + p_0q_2)x^2 \;+\; (p_1q_0 + p_0q_1)x \;+\; p_0q_0. \qquad (7.5)\end{aligned}$$

We can write formulae (7.4) and (7.5) more abstractly:

$$\left(\sum_{n=1}^{N} p_n x^n\right) + \left(\sum_{n=1}^{N} q_n x^n\right) = \sum_{n=1}^{N}(p_n+q_n)x^n \qquad (7.6)$$

$$\left(\sum_{n=1}^{N} p_n x^n\right) \cdot \left(\sum_{m=1}^{M} q_m x^m\right) = \sum_{k=1}^{N+M}\left(\sum_{n+m=k} p_n q_m\right) x^k \qquad (7.7)$$

The **degree** of a polynomial is the highest exponent. For example, if $p(x) = 4x^3 - 7x^2 + 2$, then $\mathsf{degree}\,(p) = 3$. If $p(x)$ and $q(x)$ are polynomials, then it is easy to verify:

$$\mathsf{degree}\,(p\cdot q) \quad = \quad \mathsf{degree}\,(p) + \mathsf{degree}\,(q) \qquad (\underline{\textbf{Exercise 52}}) \qquad (7.8)$$

⟨b⟩ **Complex Polynomials:** Let $\mathbb{C}[x]$ be the set of all functions $f : \mathbb{C}{\longrightarrow}\mathbb{C}$ which are polynomials with complex coefficients. Then $\mathbb{C}[x]$ is a commutative ring with identity.

⟨c⟩ Let $\mathbb{Z}[x]$ be the set of all functions $f : \mathbb{Z}{\longrightarrow}\mathbb{Z}$ which are polynomials with integer coefficients. Then $\mathbb{Z}[x]$ is a commutative ring with identity.

⟨d⟩ Let $\mathbb{Q}[x]$ be the set of all functions $f : \mathbb{Q} \longrightarrow \mathbb{Q}$ which are polynomials with rational coefficients. Then $\mathbb{Q}[x]$ is a commutative ring with identity.

⟨e⟩ **Polynomials of Many Variables:** Let $\mathbb{R}[x, y]$ be the set of all functions $f : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$ which are polynomials in *both* variables. We define addition and multiplication in the obvious way. For example,

$$(3x + 4y) \cdot (x^2 - 2y) \quad = \quad 3x^3 - 6xy + 4x^2y - 2y^2.$$

More generally, for any $n \in \mathbb{N}$, let $\mathbb{R}[x_1, x_2, \ldots, x_n]$ be the set of all functions $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ which are polynomials in n variables. Then $\mathbb{R}[x_1, x_2, \ldots, x_n]$ is a commutative ring with identity.

⟨f⟩ **Abstract Polynomial Rings:** If $\mathcal{R}$ is any ring, let $\mathcal{R}[x]$ be the set of all *formal polynomials*

$$P(x) \quad = \quad p_n x^n + p_{n-1} x^{n-1} + \ldots + p_2 x^2 + p_1 x + p_0,$$

where $p_n, p_{n-1}, \ldots, p_1, p_0$ are elements of $\mathcal{R}$. Such an expression defines a function $\mathcal{R} \longrightarrow \mathcal{R}$. However, we do *not* regard $P(x)$ as a function, but instead as an abstract algebraic expression. We refer to the symbol x as an *indeterminant*, which means it is just an abstract algebraic 'placeholder'.

Addition and multiplication are defined using formulae formulae (7.4) and (7.5). Then:

- $\mathcal{R}[x]$ is a ring.
- $\mathcal{R}[x]$ is commutative if and only if $\mathcal{R}$ is commutative.
- $\mathcal{R}[x]$ has a multiplicative identity if and only if $\mathcal{R}$ has one

Exercise 53 Verify these claims.

⟨g⟩ Let $\mathcal{S} = \mathcal{R}[x]$ is a ring, and consider the polynomial ring $\mathcal{S}[y]$ (where y is another indeterminant). Elements of this ring are formal polynomials in the variables x and y, with coefficients in $\mathcal{R}$. Thus, we denote this ring by $\mathcal{R}[x, y]$.

More generally, for any $n \in \mathbb{N}$, let $\mathcal{R}[x_1, x_2, \ldots, x_n]$ be the ring of all formal polynomials in n variables, and coefficients in $\mathcal{R}$.

⟨h⟩ **Rational functions:** Let $\mathbb{R}(x)$ denote the set of *rational functions* –that is, all functions of the form $f(x) = \frac{p(x)}{q(x)}$, where $p, q \in \mathbb{R}[x]$ are polynomials. We define addition and multiplication on $\mathbb{R}(x)$ in the obvious way:

$$\frac{p_1(x)}{q_1(x)} + \frac{p_2(x)}{q_2(x)} = \frac{p_1(x)q_2(x) + p_2(x)q_1(x)}{q_1(x)q_2(x)}, \quad \text{and} \quad \frac{p_1(x)}{q_1(x)} \cdot \frac{p_2(x)}{q_2(x)} = \frac{p_1(x) \cdot p_2(x)}{q_1(x) \cdot q_2(x)}. \quad (7.9)$$

Then $\mathbb{R}(x)$ is a commutative ring.

Observe that $\mathbb{R}(x)$ is related to $\mathbb{R}[x]$ in the same way that $\mathbb{Q}$ is related to $\mathbb{Z}$; indeed, formula (7.9) is exactly analogous to formula (7.1) on page 76.

Example 99: (Analytic Functions and Power Series)

⟨a⟩ **Real Analytic Functions of one variable:** Let $0 < r \leq \infty$, and let $\mathcal{C}^\omega(-r,r)$ be the set of all *analytic* real-valued functions $F : (-r,r)\longrightarrow\mathbb{R}$. A function F is *analytic* if F is infinitely differentiable, and F has a *Taylor series*

$$F(x) \quad = \quad \sum_{n=0}^\infty f_n x^n \qquad \text{(where } f_0, f_1, f_2, \ldots \text{ are real numbers)},$$

which converges everywhere on the interval $(-r,r)$. For example, the exponential function $\exp(x)$ is an element of $\mathcal{C}^\omega(-\infty,\infty) = \mathcal{C}^\omega(\mathbb{R})$, because, for any $x \in \mathbb{R}$,

$$\exp(x) \quad = \quad \sum_{n=0}^\infty \frac{1}{n!} x^n$$

and this series converges uniformly on $(-\infty,\infty)$.

Elementary calculus can be used to prove that the sum or product of two analytic functions is also analytic. To be precise, if

$$F(x) \quad = \quad \sum_{n=0}^\infty f_n x^n, \qquad \text{and} \qquad G(x) \quad = \quad \sum_{n=0}^\infty g_n x^n,$$

then

$$F(x)+G(x) \; = \; \sum_{n=0}^\infty (f_n+g_n)x^n, \qquad \text{and} \qquad F(x)\cdot G(x) \; = \; \sum_{n=0}^\infty \left(\sum_{k=0}^n f_k g_{n-k} \right) x^n. \tag{7.10}$$

$\mathcal{C}^\omega(-r,r)$ is a commutative ring, with additive identity $\mathbf{0}$ and multiplicative identity $\mathbb{1}$.

⟨b⟩ **Holomorphic functions:** If $\mathbf{U} \subset \mathbb{C}$ is an open subset, then let $\mathcal{C}^\omega(\mathbf{U};\mathbb{C})$ be the set of holomorphic functions $F : \mathbf{U}\longrightarrow\mathbb{C}$. A function F is *holomorphic* if it is complex-differentiable; it then has a Taylor series

$$F(x) \quad = \quad \sum_{n=0}^\infty f_n x^n \qquad \text{(where } f_0, f_1, f_2, \ldots \text{ are complex numbers)},$$

Then $\mathcal{C}^\omega(\mathbf{U};\mathbb{C})$ is a commutative ring.

⟨c⟩ **Formal Power Series:** We define $\mathbb{R}[[x]]$ to be the ring of all *formal power series*. In other words,

$$\mathbb{R}[[x]] \quad = \quad \left\{ \sum_{n=0}^\infty f_n x^n \ ; \ \text{where} \quad f_0, f_1, f_2, \ldots \quad \text{are any real numbers} \right\}$$

Observe that we place no constraints on the coefficients $f_0, f_1, f_2, \ldots$. Thus, in general the power series $\displaystyle\sum_{n=0}^{\infty} f_n$ may not converge *anywhere*. For example, the series

$$1 - x + x^2 - x^3 - x^4 + x^5 - \ldots$$

is an element of $\mathbb{R}[[x]]$, despite the fact that this series has zero radius of convergence.

We define addition and multiplication in $\mathbb{R}[[x]]$ using formulae (7.10). Then $\mathbb{R}[[x]]$ is a commutative ring, with additive identity $0 = 0 + 0x + 0x^2 + \ldots$ and multiplicative identity $1 = 1 + 0x + 0x^2 + \ldots$.

⟨d⟩ **Abstract Power Series:** If $\mathcal{R}$ is any ring, then let $\mathcal{R}[[x]]$ be the set of *formal power series* with coefficients in $\mathcal{R}$. In other words,

$$\mathcal{R}[[x]] \quad = \quad \left\{ \sum_{n=0}^{\infty} f_n x^n \text{ ; where } \quad f_0, f_1, f_2, \ldots \quad \text{are any elements of } \mathcal{R} \right\}$$

Note: in general, there is no topological structure on $\mathcal{R}$. In other words, there is no notion of 'convergence' on $\mathcal{R}$, so it is meaningless to talk about the 'convergence' of this power series. It is a purely formal construction.

We define addition and multiplication in $\mathcal{R}[[x]]$ using formulae (7.10). Then:

- $\mathcal{R}[[x]]$ is a ring.
- $\mathcal{R}[[x]]$ is commutative if and only if $\mathcal{R}$ is commutative.
- $\mathcal{R}[[x]]$ has a multiplicative identity if and only if $\mathcal{R}$ has one.

Exercise 54 Verify these claims.

⟨e⟩ **Meromorphic functions:** If $\mathbf{U} \subset \mathbb{C}$ is an open subset, then let $\mathcal{C}^{\omega}(\mathbf{U}; \bar{\mathbb{C}})$ be the set of meromorphic functions $F : \mathbf{U} \longrightarrow \bar{\mathbb{C}}$.

Here, $\bar{\mathbb{C}} = \mathbb{C} \sqcup \{\infty\}$ is the *Riemann Sphere*, consisting of the complex plane and a 'point at infinity'. A function F is *meromorphic* if F is complex-differentiable everywhere except at some discrete collection of *poles*, where F is infinite. For example, the function $F(x) = \frac{1}{x - \mathbf{i}}$ is meromorphic (and $F(\mathbf{i}) = \infty$).

If $\zeta \in \mathbf{U}$, then F has a *Laurent series*

$$F(x) \quad = \quad \sum_{n=-N}^{\infty} f_n (x - \zeta)^n \qquad \left(\begin{array}{l} \text{Here, } N \in \mathbb{N}\text{, and } f_{-N}, f_{1-N}, \ldots, f_{-1}, f_0, f_1, f_2, \ldots \text{ are} \\ \text{any complex numbers.} \end{array} \right)$$

which converges in a neighbourhood of ζ. If ζ is a pole, then $N > 0$. If ζ is *not* a pole (ie. $F(\zeta)$ is finite) then $N = 0$.

$\mathcal{C}^{\omega}(\mathbf{U};\bar{\mathbb{C}})$ is a ring. Furthermore, if $\zeta \in \mathbf{U}$, and

$$F(x) \quad = \quad \sum_{n=-N}^{\infty} f_n(x-\zeta)^n, \qquad \text{and} \qquad G(x) \quad = \quad \sum_{n=-M}^{\infty} g_m(x-\zeta)^m,$$

then

$$F(x)+G(x) \quad = \quad \sum_{n=-N}^{\infty} (f_n+g_n)(x-\zeta)^n. \qquad \left(\begin{array}{l}\text{Assume WOLOG that } M \leq N\\ \text{and define } g_n = 0 \text{ if } n < -M\end{array}\right) \tag{7.11}$$

$$\text{and} \quad F(x)\cdot G(x) \quad = \quad \sum_{n=-(N+M)}^{\infty} \left(\sum_{k=-N}^{n} f_k g_{n-k}\right)(x-\zeta)^n. \tag{7.12}$$

(**Exercise 55**).

⟨f⟩ **Formal Laurent Series:** We define $\mathbb{C}((x))$ to be the ring of all *formal Laurent series*. In other words,

$$\mathbb{C}((x)) \quad = \quad \left\{\sum_{n=-N}^{\infty} f_n x^n \text{ ; where } N \in \mathbb{N}, \text{ and } f_{-K}, f_{1-K}, \ldots \quad \text{are complex numbers}\right\}$$

Observe that we place no constraints on the coefficients $f_0, f_1, f_2, \ldots$. Thus, in general, the Laurent series may not converge anywhere.

We define addition and multiplication in $\mathbb{C}((x))$ using formulae (7.11) and (7.12). Then $\mathbb{C}((x))$ is a ring, with additive identity $0 = 0 + 0x + 0x^2 + \ldots$ and multiplicative identity $1 = 1 + 0x + 0x^2 + \ldots$. (**Exercise 56**)

⟨g⟩ **Abstract Laurent Series:** If $\mathcal{R}$ is any ring, then let $\mathcal{R}((x))$ be the set of *formal Laurent series* with coefficients in $\mathcal{R}$. In other words,

$$\mathcal{R}((x)) \quad = \quad \left\{\sum_{n=-N}^{\infty} f_n x^n \text{ ; where } N \in \mathbb{N}, \text{ and } f_0, f_1, f_2, \ldots \quad \text{are any elements of } \mathcal{R}\right\}$$

Note: in general, there is no topological structure on $\mathcal{R}$. In other words, there is no notion of 'convergence' on $\mathcal{R}$, so it is meaningless to talk about the 'convergence' of the Laurent series. It is a purely formal construction.

We define addition and multiplication in $\mathcal{R}((x))$ using formulae (7.11) and (7.12). Then:

- $\mathcal{R}((x))$ is a ring.
- $\mathcal{R}((x))$ is commutative if and only if $\mathcal{R}$ is commutative.
- $\mathcal{R}((x))$ has a multiplicative identity if and only if $\mathcal{R}$ has one.

Exercise 57 Verify these claims.

Example 100: (Abstract Function Rings)

⟨a⟩ Consider a set $\{s_1, s_2\}$ with two elements. Let $\mathbb{R}^{\{s_1,s_2\}}$ be the set of all functions from $\{s_1, s_2\}$ into $\mathbb{R}$. Thus, an element $f \in \mathbb{R}^{\{s_1,s_2\}}$ is a function with two values: $f(s_1)$ and $f(s_2)$. We define addition and multiplication pointwise: $(f+g)(s_1) = f(s_1) + g(s_1)$, etc. Then $\mathbb{R}^{\{s_1,s_2\}}$ is a commutative ring (**Exercise 58**).

Observe that any element f of $\mathbb{R}^{\{s_1,s_2\}}$ defines an element (x_1, x_2) of $\mathbb{R}^2$ (Example ⟨95a⟩), where

$$x_1 \quad = \quad f(s_1) \quad \text{and} \quad x_2 \quad = \quad f(s_2). \tag{7.13}$$

Conversely, any element (x_1, x_2) of $\mathbb{R}^2$ defines an element f of $\mathbb{R}^{\{s_1,s_2\}}$ via equation (7.13). Thus, there is a natural bijection between $\mathbb{R}^{\{s_1,s_2\}}$ and $\mathbb{R}^2$.

⟨b⟩ More generally, let $\mathcal{N} = \{s_1, s_2, \ldots, s_N\}$ be a set with N elements, and let $\mathbb{R}^{\mathcal{N}}$ be the set of all functions from $\mathcal{N}$ into $\mathbb{R}$. We define addition and multiplication pointwise. Then $\mathbb{R}^{\mathcal{N}}$ is a commutative ring (**Exercise 59**), and again, there is a natural correspondence between $\mathbb{R}^{\mathcal{N}}$ and $\mathbb{R}^N$ (Example ⟨95d⟩).

⟨c⟩ Let $\mathbb{R}^{\mathbb{R}}$ be the set of all functions from $\mathbb{R}$ to $\mathbb{R}$. Thus, $\mathbb{R}^{\mathbb{R}}$ includes all elements of $\mathcal{C}(\mathbb{R})$ (Example ⟨97a⟩), but also includes all *discontinuous* functions. Then $\mathbb{R}^{\mathbb{R}}$ is a commutative ring under componentwise addition and multiplication.

⟨d⟩ Let $\mathbf{X}$ be any set, and let $\mathcal{R}$ be any ring. Let $\mathcal{R}^{\mathbf{X}}$ be the set of all functions from $\mathbf{X}$ into $\mathcal{R}$. Then:

- $\mathcal{R}^{\mathbf{X}}$ is a ring under componentwise addition and multiplication.
- $\mathcal{R}^{\mathbf{X}}$ is commutative if and only if $\mathcal{R}$ is commutative.
- $\mathcal{R}^{\mathbf{X}}$ is has a multiplicative identity if and only if $\mathcal{R}$ has one.

Exercise 60 Verify these claims.

Example 101: (Endomorphism Rings)

⟨a⟩ Consider the abelian group $\mathbb{Z}^2$, and let $\mathbf{End}\,[\mathbb{Z}^2]$ be the set of all *endomorphisms* of $\mathbb{Z}^2$; that is, functions $\phi : \mathbb{Z}^2 \longrightarrow \mathbb{Z}^2$ such that $\phi(\mathbf{y}+\mathbf{z}) = \phi(\mathbf{y}) + \phi(\mathbf{z})$ for any $\mathbf{y}, \mathbf{z} \in \mathbb{Z}^2$.

We define addition on $\mathbf{End}\,[\mathbb{Z}^2]$ pointwise: if $\phi, \gamma \in \mathbf{End}\,[\mathbb{Z}^2]$, and $\mathbf{z} \in \mathbb{Z}^2$, then $(\phi + \gamma)(\mathbf{z}) \;=\; \phi(\mathbf{z}) + \gamma(\mathbf{z})$.

However, we do *not* define multiplication componentwise; instead, we define it via *function composition*: if $\phi, \gamma \in \mathbf{End}\,[\mathbb{Z}^2]$, and $\mathbf{z} \in \mathbb{Z}^2$, then $(\phi \circ \gamma)(\mathbf{z}) \;=\; \phi\Big(\gamma(\mathbf{z})\Big)$.

It is left as **Exercise 61** to verify:

- If ϕ and γ are endomorphisms of $\mathbb{Z}^2$, then $\phi+\gamma$ and $\phi\circ\gamma$ are also endomorphisms of $\mathbb{Z}^2$.
- The set $\mathbf{End}\,[\mathbb{Z}^2]$ forms a *noncommutative* ring under these operations.
- The additive identity is the endomorphism $\mathbf{O}$ defined: $\mathbf{O}(\mathbf{z}) = 0$ for all $\mathbf{z} \in \mathbb{Z}^2$.
- The multiplicative identity is the identity endomorphism $\mathbf{Id}$ defined: $\mathbf{Id}(\mathbf{z}) = \mathbf{z}$ for all $\mathbf{z} \in \mathbb{Z}^2$.

Endomorphism can be understood via matrices. If $\phi \in \mathbf{End}\,[\mathbb{Z}^2]$, then there is a 2×2 integer matrix $\mathbf{F} = \left[\begin{smallmatrix} f_1 & f_2 \\ f_3 & f_4 \end{smallmatrix}\right]$ so that, for any element $\mathbf{z} = (z_1, z_2)$ in $\mathbb{Z}^2$,

$$\phi(\mathbf{z}) \quad = \quad \left[\begin{array}{cc} f_1 & f_2 \\ f_3 & f_4 \end{array}\right]\left[\begin{array}{c} z_1 \\ z_2 \end{array}\right] \quad = \quad \left[\begin{array}{c} f_1z_1 + f_2z_2 \\ f_3z_1 + f_4z_2 \end{array}\right]$$

If γ is another endomorphism, with matrix $\mathbf{G}$, then the endomorphism $(\phi+\gamma)$ has matrix $\mathbf{F}+\mathbf{G}$, and the endomorphism $(\phi\circ\gamma)$ has matrix $\mathbf{F}\cdot\mathbf{G}$ (**Exercise 62**). Thus, the algebra of endomorphisms closely resembles the algebra of matrices.

⟨b⟩ More generally, let $(\mathcal{A}, +)$ be any *additive abelian group*, and let $\mathbf{End}\,[\mathcal{A}]$ be the set of all *endomorphisms* of $\mathcal{A}$; that is, functions $\phi : \mathcal{A}\longrightarrow\mathcal{A}$ such that $\phi(a + b) = \phi(a) + \phi(b)$ for any $a, b \in \mathcal{A}$.

Again, we define addition on $\mathbf{End}\,[\mathcal{A}]$ pointwise, and we define multiplication via function composition. In other words, if $\phi, \gamma \in \mathbf{End}\,[\mathcal{A}]$, and $a \in \mathcal{A}$, then

$$(\phi+\gamma)(a) \;=\; \phi(a) + \gamma(a) \quad \text{and} \quad (\phi\circ\gamma)(a) \;=\; \phi\Big(\gamma(a)\Big).$$

It is left as **Exercise 63** to verify:

- If ϕ and γ are endomorphisms of $\mathcal{A}$, then $\phi+\gamma$ and $\phi\circ\gamma$ are also endomorphisms of $\mathcal{A}$.
- The set $\mathbf{End}\,[\mathcal{A}]$ forms a *noncommutative* ring under these operations.
- The additive identity is the endomorphism $\mathbf{O}$ defined: $\mathbf{O}(a) = 0$ for all $a \in \mathcal{A}$.
- The multiplicative identity is the identity endomorphism $\mathbf{Id}$ defined: $\mathbf{Id}(a) = a$ for all $a \in \mathcal{A}$.

⟨c⟩ A subring of $\mathbf{End}\,[\mathcal{A}]$ is called an *endomorphism ring*.

⟨d⟩ If $\mathcal{G}$ is a *nonabelian* group, then $\mathbf{End}\,[\mathcal{G}]$ does *not* form a ring (**Exercise 64**). ______

Example 102: Rings of Subsets

Let $\mathbf{X}$ be some set, and let $\mathcal{P}(\mathbf{X})$ be the set of all subsets of $\mathbf{X}$:

$$\mathcal{P}(\mathbf{X}) \quad = \quad \{\mathbf{P}\ ;\ \mathbf{P} \subset \mathbf{X}\}.$$

If $\mathbf{P}_1, \mathbf{P}_2 \in \mathcal{P}(\mathbf{X})$, then we define their *sum* as follows:

$$\mathbf{P}_1 + \mathbf{P}_2 \quad = \quad \{x \in \mathbf{X}\ ;\ \text{either } x \in \mathbf{P}_1 \text{ or } x \in \mathbf{P}_2, \text{ but not in both}\}.$$

(Often, this is called the *symmetric difference* of $\mathbf{P}_1$ and $\mathbf{P}_2$, and written as "$\mathbf{P}_1 \triangle \mathbf{P}_2$") We define the *product* of $\mathbf{P}_1$ and $\mathbf{P}_2$ to be their intersection:

$$\mathbf{P}_1 \cdot \mathbf{P}_2 \quad = \quad \mathbf{P}_1 \cap \mathbf{P}_2.$$

Then:

- The set $\mathcal{P}(\mathbf{X})$ is a *commutative ring* under these operations,
- The *additive* identity is the empty set $\emptyset$.
- The *multiplicative* identity is the set $\mathbf{X}$.

$\mathcal{P}(\mathbf{X})$ is called the **Ring of Subsets** of $\mathbf{X}$.

Exercise 65 Show that $\mathbf{P}_1 \cup \mathbf{P}_2 \ = \ \mathbf{P}_1 + \mathbf{P}_2 + (\mathbf{P}_1 \cdot \mathbf{P}_2)$.

Example 103: (Group Rings)

⟨a⟩ Let $\mathbb{R}[x, x^{-1}]$ to be the set of all polynomials in x and x^{-1} with real coefficients. An example element of $\mathbb{R}[x, x^{-1}]$ is

$$3x^2 - 6x + \pi + \frac{5}{2}x^{-1} - 7x^{-2} \tag{7.14}$$

We define multiplication and addition of these polynomials in the obvious way, similar to in Example ⟨98a⟩.

⟨b⟩ Another way to think about $\mathbb{R}[x, x^{-1}]$ is as follows. Let $\mathcal{X} = \{\ldots, x^{-2}, x^{-1}, x^0, x^1, x^2, \ldots\}$ be the multiplicative group generated by a single element x. (Thus, $\mathcal{X}$ is isomorphic to $\mathbb{Z}$). We define $\mathbb{R}\mathcal{X}$ to be the set of all *formal linear combinations* of elements in $\mathcal{X}$. In other words, an element of $\mathbb{R}\mathcal{X}$ has the form

$$r_1 x^{n_1} + r_2 x^{n_2} + \ldots + r_k x^{n_k}$$

where $r_1, \ldots, r_k \in \mathbb{R}$, and $n_1, \ldots, n_k \in \mathbb{Z}$. For example the polynomial (7.14) could be an element of $\mathbb{R}\mathcal{X}$.

We define addition and multiplication of elements in $\mathbb{R}\mathcal{X}$ exactly as for polynomials. That is:

$$\Big(r_1 x^{n_1} + r_2 x^{n_2} + \ldots + r_k x^{n_k}\Big) \ + \ \Big(s_1 x^{n_1} + s_2 x^{n_2} + \ldots + s_k x^{n_k}\Big)$$
$$= \ (r_1 + s_1)x^{n_1} \ + \ (r_2 + s_2)x^{n_2} + \ldots + (r_k + s_k)x^{n_k},$$

and

$$\begin{aligned}
&\Big(r_1x^{n_1} + r_2x^{n_2} + \ldots + r_kx^{n_k}\Big) \cdot \Big(s_1x^{m_1} + s_2x^{m_2} + \ldots + s_jx^{m_j}\Big) \\
= \quad & r_1s_1x^{n_1+m_1} + r_2s_1x^{n_2+m_1} + \ldots + r_ks_1x^{n_k+m_1} \\
& \quad + r_1s_2x^{n_1+m_2} + r_2s_2x^{n_2+m_2} + \ldots + r_ks_2x^{n_k+m_2} + \ldots \\
& \quad \ldots + r_1s_jx^{n_1+m_j} + r_2s_jx^{n_2+m_j} + \ldots + r_ks_jx^{n_k+m_j}
\end{aligned}$$

In other words, the algebra of $\mathbb{R}\mathcal{X}$ is identical to that of $\mathbb{R}[x, x^{-1}]$.

⟨c⟩ Let $(\mathcal{G}, \cdot)$ be any (multiplicative) group, and let $\mathcal{R}$ be any commutative ring. We define $\mathcal{R}\mathcal{X}$ to be the set of all *formal linear combinations* of elements in $\mathcal{X}$ with coefficients in $\mathcal{R}$. In other words, an element of $\mathbb{R}\mathcal{X}$ has the form

$$r_1\mathbf{g}_1 + r_2\mathbf{g}_2 + \ldots + r_k\mathbf{g}_k$$

where $r_1, \ldots, r_k \in \mathcal{R}$ and $\mathbf{g}_1, \ldots, \mathbf{g}_k \in \mathcal{G}$.

We define addition and multiplication of elements in $\mathbb{R}\mathcal{X}$ exactly as for polynomials. That is,

$$\begin{aligned}
&\Big(r_1\mathbf{g}_1 + r_2\mathbf{g}_2 + \ldots + r_k\mathbf{g}_k\Big) \; + \; \Big(s_1\mathbf{g}_1 + s_2\mathbf{g}_2 + \ldots + s_k\mathbf{g}_k\Big) \\
&\qquad\qquad = \quad (r_1 + s_1)\mathbf{g}_1 \; + \; (r_2 + s_2)\mathbf{g}_2 + \ldots + (r_k + s_k)\mathbf{g}_k,
\end{aligned}$$

and

$$\begin{aligned}
&\Big(r_1\mathbf{g}_1 + r_2\mathbf{g}_2 + \ldots + r_k\mathbf{g}_k\Big) \cdot \Big(s_1\mathbf{h}_1 + s_2\mathbf{h}_2 + \ldots + s_j\mathbf{h}_j\Big) \\
= \quad & r_1s_1\mathbf{g}_1\mathbf{h}_1 + r_2s_1\mathbf{g}_2\mathbf{h}_1 + \ldots + r_ks_1\mathbf{g}_k\mathbf{h}_1 \; + \; r_1s_2\mathbf{g}_1\mathbf{h}_2 + r_2s_2\mathbf{g}_2\mathbf{h}_2 + \ldots + r_ks_2\mathbf{g}_k\mathbf{h}_2 + \ldots \\
& \quad \ldots + r_1s_j\mathbf{g}_1\mathbf{h}_j + r_2s_j\mathbf{g}_2\mathbf{h}_j + \ldots + r_ks_j\mathbf{g}_k\mathbf{h}_j
\end{aligned}$$

Observe that this is the natural generalization of Example ⟨103b⟩. Then:

- $\mathcal{R}\mathcal{G}$ is a ring.
- $\mathcal{R}\mathcal{G}$ is commutative if and only if $\mathcal{G}$ is abelian.
- $\mathcal{R}\mathcal{G}$ has multiplicative identity $0 = 0\mathbf{g}$ for any $\mathbf{g} \in \mathcal{G}$.
- If $\mathcal{G}$ has identity $e_{\mathcal{G}}$, then $\mathcal{R}\mathcal{G}$ has multiplicative identity $1 \cdot e_{\mathcal{G}}$.

(**Exercise 66** Verify these claims). $\mathcal{R}\mathcal{G}$ is called the **Group ring** of $\mathcal{G}$ with coefficients in $\mathcal{R}$. It is very important in the *representation theory* of groups (see § 7.3.5 on page 96) ⎯

7.3 (*) Applications of Ring Theory

Prerequisites: §7.1

Here we briefly survey the major mathematical applications of ring theory.

7.3.1 Algebraic Number Theory

Number theory studies the mathematical properties of the ring of integers. Abstract ring theory sheds light on classical number theory in two ways:

1. We can contextualize the problems and results of number theory within a larger framework. Given any result in number theory, we can ask, 'Is this theorem true only for integers, or does it generalize to some other rings as well?'

2. Problems about integers can often be solved by relating the integers to some other ring, such as the Gaussian Integers $\mathbb{Z}[\mathbf{i}]$ of Example ⟨93f⟩, or the ring $\mathbb{Z}_{/n}$ of Example ⟨93i⟩.

For example, a **Diophantine Equation** is a polynomial equation of the form

$$p(x_1, x_2, \ldots, x_n) \quad = \quad 0,$$

where $p \in \mathbb{Z}[x_1, x_2, \ldots, x_n]$ is a polynomial in n variables with integer coefficients, and we impose the constraint that $x_1, \ldots, x_n$ must be integers. For example, the equation

$$x^2 - y \quad = \quad 0, \tag{7.15}$$

says that $y = x^2$, ie. that y is a *perfect square*. Equation (7.15) has solutions $(x, y) = (1, 1)$, $(2, 4)$, $(3, 9)$, etc. The *Pythagorean equation*

$$x^2 + y^2 - z^2 \quad = \quad 0, \tag{7.16}$$

says that $x^2 + y^2 = z^2$ —in other words, (x, y, z) form a *Pythagorean triple*. One solution is $(x, y, z) = (3, 4, 5)$.

The obvious generalization of the Pythagorean Equation is the *Fermat equation*

$$x^n + y^n - z^n \quad = \quad 0, \tag{7.17}$$

(where $n \in \mathbb{N}$ is fixed), which says that $x^n + y^n = z^n$. *Fermat's Last Theorem* says there are no nontrivial solutions for any $n \geq 3$. This theorem was finally proved by Andrew Wiles in 1998, after almost 4 centuries of mathematical effort.

It is (relatively) easy to find real or complex solutions to a polynomial equations like (7.15), (7.16) or (7.17), but it is often very difficult to isolate the *integer* solutions —or indeed, even to determine whether any exist. The study of Diophantine equations is thus a major area of number theory, and ring theory provides important tools. The simplest tool is just *reduction mod p*. For example, the equation (7.17) is true only if the congruence equation

$$x^n + y^n - z^n \quad \equiv \quad 0 \pmod{p} \tag{7.18}$$

is true, for any fixed $p \in \mathbb{N}$. Equation (7.18) is an equation in the finite ring $\mathbb{Z}_{/p}$ (Example ⟨93i⟩), and is probably easier to solve than (7.17).

We will indicate when a particular topic is relevant to algebraic number theory by placing the flag $\boxed{\text{Num.Thr.}}$ in the margin.

7.3.2 Field Theory & Galois Theory

Consider the following three problems:

Solution formulae for polynomial equations: If $a, b, c \in \mathbb{R}$ are constants, then the quadratic equation

$$ax^2 + bx + c \quad = \quad 0 \tag{7.19}$$

can always be solved by recourse to the *quadratic formula:*

$$x \quad = \quad \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{7.20}$$

Geronimo Cardano (1501-1576) developed a similar (but much more complicated) formula for solving a *cubic* equation:

$$ax^3 + bx^2 + cx + d \quad = \quad 0.$$

Later, an even more complicated formula was developed solve the *quartic* equation:

$$ax^4 + bx^3 + cx^2 + dx + e \quad = \quad 0.$$

This raised the question: does there exist a formula to solve an arbitrary *quintic* equation

$$ax^5 + bx^4 + cx^3 + dx^2 + ex + f \quad = \quad 0? \tag{7.21}$$

What about higher-order polynomials?

Note that the question here is *not* whether a solution *exists*, but whether that solution can always be found using an algebraic formula like (7.20). For example, the transcendental equation

$$e^x \quad = \quad 17 \tag{7.22}$$

has a solution —namely, $x = \log(17)$. However, there is no *finite algebraic expression* for x in terms of 17 and 'radicals' of 17, such as $\sqrt{17}$, $\sqrt[3]{17}$ etc. So we're asking,

Is equation (7.21) solvable via a finite expression involving radicals of a, b, c, etc? Or is it like (7.22), possessing a solution, but not one we can easily describe?

The nature of the complex numbers: If $4ac > b^2$, then the quadratic equation (7.19) has no real solution. In this case, the quadratic formula (7.20) produces a 'nonsensical' result, involving the square root of a negative number. The complex numbers were introduced to provide an interpretation for these 'nonsense' solutions; they 'extend' the real numbers to a larger field, where the quadratic formula (7.20) always produces a meaningful answer. Thus, in the complex numbers, *every quadratic polynomial has a solution.*

Does *every* polynomial equation have a solution in the complex numbers? Or must we pass to an even *larger* field in order to solve cubics, quartics, quintics, etc.? C.F. Gauss answered this:

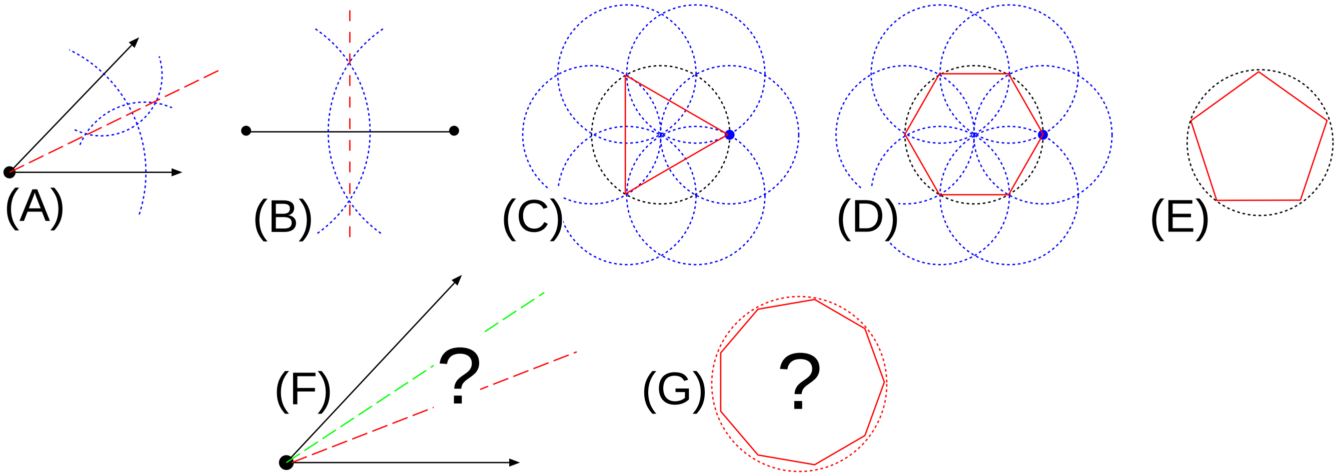


Figure 7.2: Classical compass and straightedge constructions.

Fundamental Theorem of Algebra: *Let* $P(x) = p_n x^n + p_{n-1}x^{n-1} + \ldots + p_2x^2 + p_1x + p_0$ *be any polynomial with complex coefficients* $p_n, \ldots, p_1, p_0$*. Then:*

1. *There is a* root $z \in \mathbb{C}$ *so that* $P(z) = 0$*.*
2. P *can be completely* factored *over the complex numbers. In other words, there exist* $z_1, z_2, \ldots, z_n \in \mathbb{C}$ *so that* $P(x) = (x - z_1) \cdot (x - z_2) \cdots (x - z_n)$*.*

This raises the question: Given an arbitrary field $\mathbb{F}$ and some polynomial equation "$p(x) = 0$" over $\mathbb{F}$, can we solve this equation by 'extending' $\mathbb{F}$ to a larger field $\mathbb{E}$, in the same way that we extended $\mathbb{R}$ to $\mathbb{C}$? Is there something analogous to the Fundamental Theorem of Algebra for these extensions?

Classical Geometric Constructions: Since the time of Euclid, certain geometric figures have been constructed using a compass and straightedge. A *compass* is a string with a pin at one end and a pencil at the other, used to inscribe circles of any radius. A *straightedge* is just a peice of wood with a perfectly straight edge, used to inscribe lines. Using these two simple tools, one can bisect angles (Figure 7.2A) , bisect lines and construct their perpendiculars (Figure 7.2B) , and inscribe an equilateral triangle (Figure 7.2C), regular hexagon (Figure 7.2D), or regular pentagon (Figure 7.2E), inside a given circle.

However, many things *cannot* be constructed. For example, after two thousand years of effort, there is no known method to *trisect* an angle (Figure 7.2F), or to construct a regular nonagon (a nine-sided polygon; see Figure 7.2G). Is this merely a failure of ingenuity? Or is there a fundamental obstruction which makes these things impossible?

In 1832, while in prison, a twenty-one year old mathematician named Evariste Galois found the key to solving these three problems, by studying the symmetry groups of the roots of polynomials. Shortly afterwards, Galois was killed in a duel, but his ideas became the foundation of the vast and beautiful theory which bears his name.

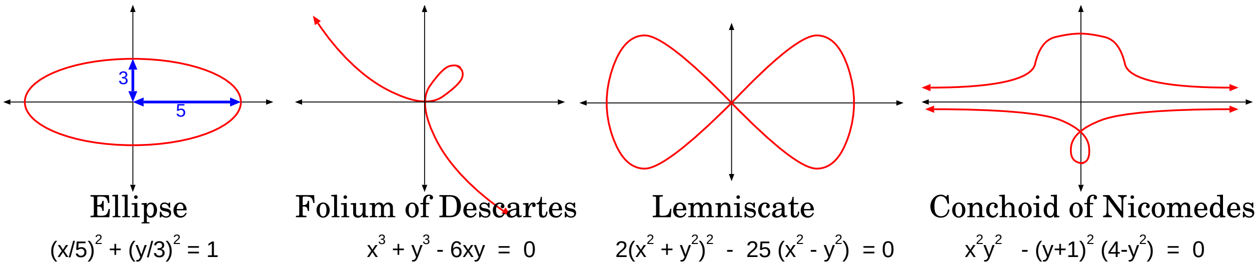


Figure 7.3: Algebraic varieties.

We will indicate when a particular topic is relevant to galois theory by placing the flag $\boxed{\text{Galois}}$ in the margin.

7.3.3 Algebraic Geometry & Commutative Algebra

An **algebraic variety** is the geometric figure determined by the solutions to a polynomial equation. For example, the *unit sphere* is the set of solutions $(x, y, z) \in \mathbb{R}^3$ to the equation

$$x^2 + y^2 + z^2 - 1 \quad = \quad 0.$$

while an *ellipse* is the set of solutions $(x, y) \in \mathbb{R}^2$ to the equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1^2 \quad = \quad 0.$$

where $a, b \in \mathbb{R}$ are constants. Figure 7.3 depicts several algebraic varieties and their defining equations.

The geometric properties of algebraic varieties are closely related to the algebra the polynomial ring $\mathbb{R}[x_1, x_2, \ldots, x_n]$ (Example ⟨98e⟩), and classical *algebraic geometry* explored this relationship. The scope of modern algebraic geometry is far greater. For example, a **differentiable variety** is a solution to an equation

$$f(x_1, x_2, \ldots, x_n) \quad = \quad 0, \tag{7.23}$$

where $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is any differentiable function. The geometry of such varieties is related to ring $\mathcal{C}^\infty(\mathbb{R}^n)$ of differentiable functions (Example ⟨3i⟩). Any differentiable manifold in $\mathbb{R}^n$ can be represented as a differentiable variety; thus, modern algebraic geometry subsumes a large part of differential geometry.

At a topological level, any closed subset of $\mathbb{R}^n$ can be represented as the set of solutions to an equation like (7.23), where $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is some *continuous* function. Thus, the *topology* of $\mathbb{R}^n$ is 'encoded' by the structure of the ring $\mathcal{C}(\mathbb{R}^n)$ of continuous functions (Example ⟨3c⟩). More generally, if $\mathbf{X}$ is any 'regular' topological space (eg. any manifold, metric space, etc.), then the entire topological structure of $\mathbf{X}$ is encoded in the structure of the ring $\mathcal{C}(\mathbf{X})$ of continuous real-valued functions on $\mathbf{X}$ (Example (3d)).

We can also reverse the direction of this correspondence. Given an arbitrary commutative ring $\mathcal{R}$, we ask, 'Is there some space $\mathbf{X}$ so that $\mathcal{R}$ is the ring of continuous functions on $\mathbf{X}$?' Under suitable conditions, the answer is 'yes'; this space is called the *spectrum* or *Zariski topology* of $\mathcal{R}$, and we can learn much about $\mathcal{R}$ by studying this space. Thus, *there is a geometric interpretation for any commutative ring.*

The study of this relationship is sometimes called **commutative algebra**, since it concerns commutative rings. However, recently, there has been an effort to generalize these methods to *noncommutative* rings. Certain geometric structures (for example, foliations of manifolds) can be represented with a noncommutative ring, just as the topology of a space $\mathbf{X}$ can be represented using the commutative ring $\mathcal{C}(\mathbf{X})$. This burgeoning field is called *noncommutative geometry.*

We will indicate when a particular topic is relevant to algebraic geometry by placing the flag $\boxed{\text{Alg.Geo.}}$ in the margin.

7.3.4 Algebraic Topology & Homological Algebra

Algebraic topology studies the global topological properties of spaces by 'measuring' their structure using algebraic objects such as *homotopy groups* and *(co)homology groups.* These objects are called *homotopy invariants*, because for a given space $\mathbf{X}$, you will always 'measure' the same homotopy or homology groups for $\mathbf{X}$, no matter how $\mathbf{X}$ is deformed.

Given two spaces $\mathbf{X}$ and $\mathbf{Y}$, we might ask: *Is* $\mathbf{Y}$ *really just a deformed version of* $\mathbf{X}$*, or are they fundamentally different?* We can answer this question by computing homotopy invariants. If $\mathbf{X}$ and $\mathbf{Y}$ yield *different* invariants, then they *cannot* be the same[1].

This endeavour requires a lot of algebraic machinery, for two reasons:

- First, to *compute* the homotopy invariants of topological spaces.

- Second, to *compare* the invariants of two spaces, to see if they are the same or not.

This algebraic machinery involves rings, modules, and homomorphisms between modules, and is called *homological algebra.*

We will indicate when a particular topic is relevant to algebraic topology or homological algebra by placing the flag $\boxed{\text{Alg.Top.}}$ in the margin.

7.3.5 Group Representation Theory & Ring Structure Theory

A *representation* of a group $\mathcal{G}$ is a homomorphism $\rho : \mathcal{G} \longrightarrow \mathbb{GL}^n[\mathbb{R}]$. This allows us to study the structure of $\mathcal{G}$ using the well-understood algebra of matrices. Given a group $\mathcal{G}$, we ask, 'What representations of $\mathcal{G}$ exist, and what can they tell us about $\mathcal{G}$?'

Representation theory has many important applications. For example:

[1] On the other hand, if they yield the *same* invariants, $\mathbf{X}$ and $\mathbf{Y}$ might be the same, but they also might not be.

Physics: Groups represent the *symmetries* of physical laws. Thus, group representation theory plays central role in the 'Standard model' of subatomic physics.

Nonabelian Harmonic Analysis: Fourier theory (aka harmonic analysis) provides a powerful tool for studying functions and probability distributions on the real line $\mathbb{R}$, Euclidean space $\mathbb{R}^n$, the circle $\mathbb{T}^1$, or the torus $\mathbb{T}^n$. Fourier theory is essential for solving differential equations, and ubiquitous in probability theory.

Fourier theory is actually the representation theory of certain *abelian* groups. To generalize the methods of Fourier Theory to *nonabelian* groups (eg. linear groups, Lie groups), we must study the representation theory of these groups.

Given a representation $\rho : \mathcal{G} \longrightarrow \mathbb{GL}^n[\mathbb{R}]$, we can immediately extend ρ to a *ring homomorphism* $\rho^* : \mathbb{R}\mathcal{G} \longrightarrow \mathcal{M}_n(\mathbb{R})$ from the group ring over $\mathcal{G}$ to the ring of $n \times n$ matrices (see Example ⟨96b⟩ on 80, and Example ⟨103⟩ on page 90). Conversely, any ring homomorphism $\rho^* : \mathbb{R}\mathcal{G} \longrightarrow \mathcal{M}_n(\mathbb{R})$ determines a representation of the group $\mathcal{G}$. Thus, the representation theory of the group $\mathcal{G}$ is closely related to the group ring $\mathbb{R}\mathcal{G}$. Indeed, we can completely classify the representations of $\mathcal{G}$ using the *structure theory* of $\mathbb{R}\mathcal{G}$.

We will indicate when a particular topic is relevant to group representation theory or ring structure theory by placing the flag $\boxed{\text{Grp.Repr.}}$ in the margin.

7.4 Subrings

Prerequisites: §7.1

If $\mathcal{R}$ is a ring with operations '+' and '·', then a **subring** of $\mathcal{R}$ is any subset $\mathcal{S} \subset \mathcal{R}$ which is also a ring, under the same operations. We indicate this by writing: "$\mathcal{S} < \mathcal{R}$".

Example 104:

(a) $2\mathbb{Z} < \mathbb{Z} < \mathbb{Q} < \mathbb{R} < \mathbb{C} < \mathbb{H}$ (see Examples (93a-93g)).

(b) $\mathbb{Z} < \mathbb{Z}[\mathbf{i}]$ (see Example ⟨93f⟩).

(c) If $n \in \mathbb{N}$, then $n\mathbb{Z}$ is a subring of $\mathbb{Z}$.

(d) For any $n \in \mathbb{N}$, $\mathcal{M}_n(\mathbb{Z}) < \mathcal{M}_n(\mathbb{Q}) < \mathcal{M}_n(\mathbb{R}) < \mathcal{M}_n(\mathbb{C}) < \mathcal{M}_n(\mathbb{H})$ (see Examples (96b-96e)).

In general, if $\mathcal{S}$ is a subring of $\mathcal{R}$, then $\mathcal{M}_n(\mathcal{S})$ is a subring of $\mathcal{M}_n(\mathcal{R})$ (see Example ⟨96f⟩.

(e) $\mathbb{R}[x] < \mathcal{C}^\omega(\mathbb{R}) < \mathcal{C}^\infty(\mathbb{R}) < \ldots < \mathcal{C}^3(\mathbb{R}) < \mathcal{C}^2(\mathbb{R}) < \mathcal{C}^1(\mathbb{R}) < \mathcal{C}(\mathbb{R}) < \mathbb{R}^\mathbb{R}$ (see Examples (97a), (3f- 3h), (98a), (99a), and (100c).

(f) $\mathbb{Z}[x] \;<\; \mathbb{Q}[x] \;<\; \mathbb{R}[x] \;<\; \mathbb{C}[x]$ (see Examples (98a-98d)).

In general, if $\mathcal{S}$ is a subring of $\mathcal{R}$, then $\mathcal{S}[x]$ is a subring of $\mathcal{R}[x]$ (see Example ⟨98f⟩). (**Exercise 67**).

(g) $\mathbb{R}[x] \;<\; \mathbb{R}[[x]]$. (Examples (98a) and (99c)).

More generally, for any ring, $\mathcal{R}[x] \;<\; \mathcal{R}[[x]]$ (Examples (98f) and (99d)) (**Exercise 68**).

(h) $\mathbb{C}[[x]] \;<\; \mathbb{C}((x))$. (Examples (99c) and (99f)).

More generally, for any ring, $\mathcal{R}[[x]] \;<\; \mathcal{R}((x))$. (Examples (99d) and (99g)) (**Exercise 69**).

7.5 Units and Fields

Prerequisites: §7.1

We have seen that, in a general ring, not all elements have multiplicative inverses. A **unit** of $\mathcal{R}$ is an element with a multiplicative inverse. The set of all units of $\mathcal{R}$ is denoted $\mathcal{R}^{\times}$.

A **field** is a *commutative* ring $\mathcal{R}$ such that *all* nonzero elements of $\mathcal{R}$ are units. In other words: $\mathcal{R}^{\times} \;=\; \mathcal{R} \setminus \{0\}$.

Example 105:

(a) $\mathbb{Z}$ is *not* a field: the only units of $\mathbb{Z}$ are 1 and -1.

(b) $\mathbb{Q}$ and $\mathbb{R}$ are fields. If $r \neq 0$, then it is a unit, with multiplicative inverse $1/r$.

(c) $\mathbb{C}$ is a field; if $x+y\mathbf{i} \neq 0$, then it is a unit, with multiplicative inverse $\dfrac{x - y\mathbf{i}}{x^2 + y^2}$ (**Exercise 70**).

(d) $\mathbb{H}$ is not commutative, so it can't be a field. However, $\mathbb{H}$ 'wants' to be a field, because every nonzero element is a unit (**Exercise 71**). We say that $\mathbb{H}$ is a **division ring**.

(e) Let $\mathcal{R} = \mathcal{M}_n(\mathbb{R})$ from Example ⟨96b⟩. The set of units of $\mathcal{M}_n(\mathbb{R})$ is the set of all *invertible* $n \times n$ matrices. In other words, $\mathcal{M}_n^{\times}(\mathbb{R}) \;=\; \mathbb{GL}^n[\mathbb{R}]$.

(f) Let $\mathcal{R} \;=\; \mathbb{R}[x]$ from Example ⟨98a⟩. The units of $\mathbb{R}[x]$ are just the nonzero *constant* polynomials —that is, polynomials of the form: $p(x) = p_0 + 0x + 0x^2 + \ldots$, where $p_0 \neq 0$. Thus, $\mathbb{R}[x]^{\times}$ is identical to $\mathbb{R}^{\times}$. Thus $\mathbb{R}[x]$ is *not* a field.

(g) Let $\mathcal{R}$ be any ring, and consider $\mathcal{R}[x]$ from Example ⟨98f⟩. The units of $\mathcal{R}[x]$ are just the *unit constant* polynomials —that is, polynomials of the form: $p(x) = p_0 + 0x + 0x^2 + \ldots$, where $p_0 \in \mathcal{R}^{\times}$. Thus, $\mathcal{R}[x]^{\times}$ is identical to $\mathcal{R}^{\times}$, and thus, $\mathcal{R}[x]$ is never itself a field.

(h) Let $\mathcal{R} = \mathcal{C}(\mathbb{R})$ from Example $\langle$97a$\rangle$. An element of $\mathcal{C}(\mathbb{R})$ has a multiplicative inverse if and only if f is nonzero everywhere; in this case, the inverse of f is the function $g(x) = \frac{1}{f(x)}$.

The same holds for $\mathcal{R} = \mathcal{C}^k(\mathbb{R})$ for any $1 \leq k \leq \infty$ (see Examples (3f-3h)).

(i) Let $\mathcal{R} = \mathcal{C}^\omega(\mathbf{U}; \bar{\mathbb{C}})$, as in Example (99e). Then $\mathcal{R}$ is a field: if f is a nonzero meromorphic function, and $g(x) = \frac{1}{f(x)}$, then g is also meromorphic. The zeros of f become the poles of g and vice versa (**Exercise 72** Verify these claims.)

(j) Let $\mathcal{F}$ be a field, and let $\mathcal{R} = \mathcal{F}((x))$ be the ring of formal Laurent series over $\mathcal{F}$, as in Example $\langle$99g$\rangle$. Then $\mathcal{R}$ is *also* a field (**Exercise 73**).

(k) Let $\mathbb{R}(x)$ denote the set of rational functions, as in Example $\langle$98h$\rangle$. Then $\mathbb{R}(x)$ is a field (**Exercise 74**).

Lemma 106 *Let $\mathcal{R}$ be any ring. Then $\mathcal{R}^\times$ forms a* group *under multiplication. This is called the* **group of units** *of $\mathcal{R}$.*

Proof: **Exercise 75** □

Example 107: $\mathbb{Z}_{/n}{}^\times$

Let $n \in \mathbb{N}$. If $\overline{k} \in \mathbb{Z}_{/n}$ then we saw earlier that

$$\Big(\ \overline{k} \text{ has a multiplicative inverse, mod } n\ \Big) \iff \Big(\ k \text{ is relatively prime to } n \text{ —ie. } \gcd(n,k) = 1\ \Big)$$

Thus,

$$\mathbb{Z}_{/n}{}^\times \quad = \quad \big\{\overline{k} \in \mathbb{Z}_{/n}\ ;\ k \text{ is relatively prime to } n\big\}.$$

In particular, it follows that

$$\Big(\ \mathbb{Z}_{/n} \text{ is a field}\ \Big) \iff \Big(\ \mathbb{Z}_{/n}{}^\times \ = \ \{\overline{1}, \overline{2}, \ldots, \overline{n-1}\}\ \Big) \iff \Big(\ n \text{ is prime}\ \Big).$$

7.6 Zero Divisors and Integral Domains

Num.Thr.

Alg.Geo.

Alg.Top.

Prerequisites: §7.5

Let $r \in \mathcal{R}$ be nonzero. We say r is a **zero divisor** if there is some element $s \in \mathcal{R}$ so that $s \neq 0$, but $r \cdot s = 0$.

Example 108:

(a) Let $\mathcal{R} = \mathbb{Z}_{/10}$, as in Example $\langle$93i$\rangle$. Let $r = \bar{5}$, and let $s = \bar{2}$. Then $r \neq 0 \neq s$, but $\overline{r} \cdot \overline{s} = \overline{10} = \bar{0}$. Thus, $\bar{5}$ and $\bar{2}$ are both zero divisors in $\mathbb{Z}_{/10}$.

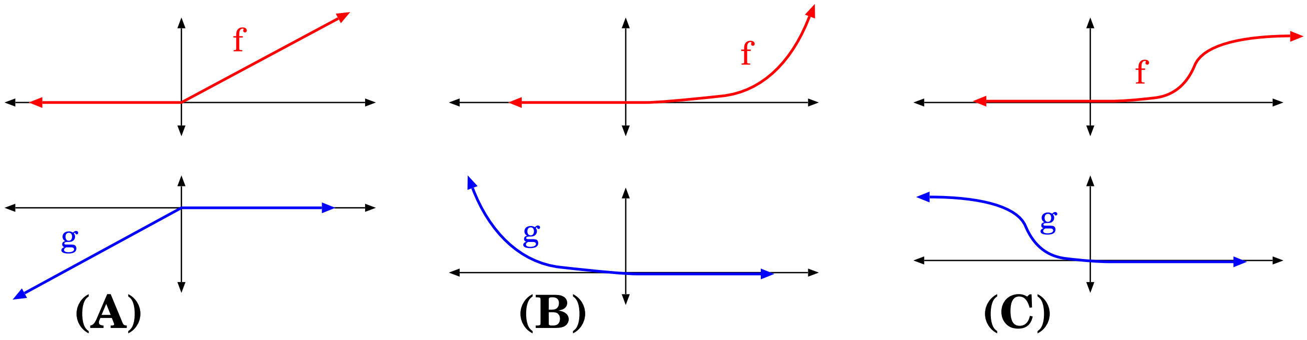


Figure 7.4: Zero divisors in the rings $\mathcal{C}(\mathbb{R})$, $\mathcal{C}^K(\mathbb{R})$, and $\mathcal{C}^\infty(\mathbb{R})$.

(b) Let $\mathcal{R} = \mathbb{Z}_{/n}$, as in Example ⟨93i⟩. If $n = p \cdot q$ for any integers $p, q > 1$, then $\overline{p}$ and $\overline{q}$ are zero divisors, because $\overline{p} \cdot \overline{q} = \overline{0}$.

(c) Let $\mathcal{R} = \mathbb{R}^2$ with the product ring structure, as in Example ⟨95a⟩. Let $r = (1, 0)$ and $s = (0, 1)$. Then $r \cdot s = (0, 0)$, so r and s are zero divisors.

(d) Let $\mathcal{R} = \mathcal{M}_2(\mathbb{R})$, as in Example ⟨96a⟩. Let $\mathbf{A} = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right]$ and let $\mathbf{B} = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right]$. Then $\mathbf{A} \cdot \mathbf{B} = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}\right]$, so $\mathbf{A}$ and $\mathbf{B}$ are zero-divisors.

(e) Let $\mathcal{R} = \mathcal{C}(\mathbb{R})$, as in Example ⟨97a⟩. As shown in Figure 7.4(A), let

$$f(x) \;=\; \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } x > 0 \end{cases} \quad \text{and} \quad g(x) \;=\; \begin{cases} x & \text{if } x \leq 0 \\ 0 & \text{if } x > 0 \end{cases}$$

Then f and g are both nonzero continuous functions, but $f \cdot g = 0$. Thus, f and g are zero-divisors.

(f) Let $\mathcal{R} = \mathcal{C}^k(\mathbb{R})$ for some $k \in \mathbb{N}$, as in Example (3g). As shown in Figure 7.4(B), let

$$f(x) \;=\; \begin{cases} 0 & \text{if } x \leq 0 \\ x^{k+1} & \text{if } x > 0 \end{cases} \quad \text{and} \quad g(x) \;=\; \begin{cases} x^{k+1} & \text{if } x \leq 0 \\ 0 & \text{if } x > 0 \end{cases}$$

Then $f \in \mathcal{C}^{k+1}(\mathbb{R})$ and $g \in \mathcal{C}^{k+1}(\mathbb{R})$ are both nonzero, but $f \cdot g = 0$. Thus, f and g are zero-divisors.

(g) Let $\mathcal{R} = \mathcal{C}^\infty(\mathbb{R})$, as in Example (3h). As shown in Figure 7.4(C), let

$$f(x) \;=\; \begin{cases} 0 & \text{if } x \leq 0 \\ \exp\left(\frac{-1}{x^2}\right) & \text{if } x > 0 \end{cases} \quad \text{and} \quad g(x) \;=\; \begin{cases} \exp\left(\frac{-1}{x^2}\right) & \text{if } x \leq 0 \\ 0 & \text{if } x > 0 \end{cases}$$

Then $f \in \mathcal{C}^\infty(\mathbb{R})$ and $g \in \mathcal{C}^\infty(\mathbb{R})$ are both nonzero, but $f \cdot g = 0$. Thus, f and g are zero-divisors.

(h) Let $\mathcal{R} = \mathbb{R}[x]$, as in Example ⟨98a⟩. Then $\mathcal{R}$ has *no* zero divisors, because the product of any nonzero polynomials is nonzero.

(**Exercise 76** Verify this. Hint: use formula (7.8) on page 83.)

(i) Let $\mathcal{R} = \mathbb{Z}$. Then $\mathcal{R}$ has *no* zero divisors, because the product of any nonzero integers is nonzero.

An **integral domain** is a commutative ring with no zero divisors.

Example 109:

(a) $\mathbb{Z}$ is an integral domain, by Example ⟨108i⟩.

(b) Likewise, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are all integral domains.

(c) Any *field* is an integral domain (**Exercise 77**).

(d) $\mathbb{Z}_{/10}$ is not an integral domain, because $\bar{2}$ and $\bar{5}$ are zero divisors, from Example ⟨108a⟩.

(e) In general, if $n = p \cdot q$ for any integers $p, q > 1$, then $\mathbb{Z}_{/n}$ is *not* an integral domain, because $\overline{p}$ and $\overline{q}$ are zero divisors, from Example ⟨108b⟩.

(f) Thus, combining Examples (107), (109c) and (109e), we conclude:

$$\Big(\mathbb{Z}_{/n} \text{ is an integral domain} \Big) \iff \Big(n \text{ is prime} \Big) \iff \Big(\mathbb{Z}_{/n} \text{ is a field} \Big).$$

(g) The polynomial ring $\mathbb{R}[x]$ is an integral domain, by Example (108h).

(h) Likewise, $\mathbb{Z}[x]$, $\mathbb{Q}[x]$, and $\mathbb{C}[x]$ are all integral domains.

(i) If $\mathcal{R}$ is a commutative ring, then the following are equivalent (**Exercise 78**):

- $\mathcal{R}$ is an integral domain.
- $\mathcal{R}[x]$ is an integral domain.
- $\mathcal{R}[x_1, x_2, \ldots, x_n]$ is an integral domain, for all $n \in \mathbb{N}$.

(j) The ring $\mathcal{C}^\infty(\mathbb{R})$ is *not* an integral domain, by example (108g).

(k) However, $\mathcal{C}^\omega(\mathbb{R})$ *is* an integral domain. To see this, we combine the following two assertions:

(i) If $f, g : \mathbb{R} \longrightarrow \mathbb{R}$ are continuous functions, and $f \cdot g = 0$, then there exist nonempty open sets $\mathbf{U} \subset \mathbb{R}$ and $\mathbf{V} \subset \mathbb{R}$ so that $f(u) = 0$ for all $u \in \mathbf{U}$, and $g(v) = 0$ for all $v \in \mathbf{V}$

(ii) (The Identity Theorem) If $f : \mathbb{R} \longrightarrow \mathbb{R}$ is an analytic function, and there is a nonempty open set $\mathbf{U} \subset \mathbb{R}$ so that $f(u) = 0$ for all $u \in \mathbf{U}$, then $f = 0$ everywhere.

Exercise 79 Use (i) and (ii) to show that $\mathcal{C}^\omega(\mathbb{R})$ has no zero divisors.

(The proofs of assertions (i) and (ii) are beyond the scope of these notes.)

(l) Let $\mathcal{D}$ be any integral domain, and let $\mathcal{R} = \mathcal{D}[[x]]$ be the ring of formal power series over $\mathcal{D}$, as in Example ⟨99d⟩. Then $\mathcal{R}$ is *also* an integral domain (**Exercise 80**).

Integral domains are useful because they allow us to 'cancel' common factors out of equations. When confronted with an algebraic expression of the form

$$r \cdot s \quad = \quad r \cdot t,$$

we are tempted to 'cancel' the r, and conclude that $s = t$. This 'cancellation law' is valid in a group; we simply multiply both sides by r^{-1} to get $s \;=\; r^{-1}rs \;=\; r^{-1}rt \;=\; t$. Cancellation also works if $\mathcal{R}$ is a field, and r is a nonzero element (therefore invertible).

However, the 'cancellation law' is *not* valid in an arbitrary ring. For example, if $\mathcal{R} = \mathbb{Z}_{/10}$, then

$$\bar{2} \cdot \bar{3} \quad = \quad \bar{6} \quad = \quad \bar{2} \cdot \bar{8} \qquad \text{even though} \qquad \bar{2} \quad \neq \quad \bar{8}.$$

The problem here is that $\bar{2}$ is a zero divisor in $\mathbb{Z}_{/10}$. To obtain a 'cancellation law', we must restrict ourselves to integral domains.

Proposition 110 (Cancellation Law for Integral Domains)

Let $\mathcal{R}$ be an integral domain, and let $r \in \mathcal{R}$ be nonzero. Then for any $s, t \in \mathcal{R}$,

$$\Big(rs = rt \Big) \iff \Big(s = t \Big).$$

Proof: $\Big(rs = rt \Big) \iff \Big(rs - rt \;=\; 0 \Big) \iff \Big(r(s-t) \;=\; 0 \Big).$

Since r is nonzero, and cannot be a zero divisor, we conclude that $(s - t) \;=\; 0$. In other words, $s = t$. □

Chapter 8

Homomorphisms, Quotients, and Ideals

8.1 Homomorphisms

Prerequisites: §7.1

Consider the ring $\mathbb{Z}_{/5}$ from Example ⟨93h⟩. For any $z \in \mathbb{Z}$, let $\overline{z} \in \mathbb{Z}_{/5}$ be the congruence class of z, mod 5. The arithmetic of $\mathbb{Z}_{/5}$ has the following convenient property: for any $z_1, z_2 \in \mathbb{Z}$,

$$\overline{z_1 + z_2} \quad = \quad \overline{z}_1 + \overline{z}_2; \qquad \text{and} \qquad \overline{z_1 \cdot z_2} \quad = \quad \overline{z}_1 \cdot \overline{z}_2. \tag{8.1}$$

Define the function $\phi : \mathbb{Z} \longrightarrow \mathbb{Z}_{/5}$ by $\phi(z) = \overline{z}$. Then the equations (8.1) can be rewritten:

$$\begin{array}{rcccccc} \phi(z_1 + z_2) & = & \overline{z_1 + z_2} & = & \overline{z}_1 + \overline{z}_2 & = & \phi(z_1) + \phi(z_2); \\ \text{and } \phi(z_1 \cdot z_2) & = & \overline{z_1 \cdot z_2} & = & \overline{z}_1 \cdot \overline{z}_2 & = & \phi(z_1) \cdot \phi(z_2); \end{array}$$

We say that ϕ is a *homomorphism* from the ring $\mathbb{Z}$ to the ring $\mathbb{Z}_{/5}$.

In general, let $\mathcal{R}$ and $\mathcal{S}$ be rings. A **ring homomorphism** is a map $\phi : \mathcal{R} \longrightarrow \mathcal{S}$ so that, for any $r_1, r_2 \in \mathcal{R}$

$$\phi(r_1 + r_2) \quad = \quad \phi(r_1) + \phi(r_2) \qquad \text{and} \qquad \phi(r_1 \cdot r_2) \quad = \quad \phi(r_1) \cdot \phi(r_2).$$

Example 111:

(a) Let $\mathcal{R} = \mathbb{Z}$ and let $\mathcal{S} = \mathbb{Z}_{/5}$ (Example ⟨93h⟩), and define $\phi : \mathbb{Z} \longrightarrow \mathbb{Z}_{/5}$ by $\phi(z) = \overline{z}$. Then ϕ is a ring homomorphism.

(b) More generally, let $n \in \mathbb{N}$, and define $\phi : \mathbb{Z} \longrightarrow \mathbb{Z}_{/n}$ by $\phi(z) = \overline{z}$, where $\overline{z}$ is the congruence class of z, modulo n. Then ϕ is a ring homomorphism.

(c) Let $\mathcal{R} = \mathbb{R}^2$ be the product ring from Example ⟨95a⟩, and let $\mathcal{S} = \mathbb{R}$. Let $\pi_1 : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be the projection into the first coordinate. Then π_1 is a homomorphism, because for any

(x_1, y_1) and (x_2, y_2) in $\mathbb{R}^2$,

$$\begin{array}{rcccl} \pi_1\Big((x_1,y_1)+(x_2,y_2)\Big) & = & \pi_1\Big((x_1+x_2),\ (y_1+y_2)\Big) & = & x_1+x_2 \\ & = & \pi_1(x_1,y_1)+\pi_1(x_2,y_2). & & \\ \text{and}\ \ \pi_1\Big((x_1,y_1)\cdot(x_2,y_2)\Big) & = & \pi_1\Big((x_1\cdot x_2),\ (y_1\cdot y_2)\Big) & = & x_1\cdot x_2 \\ & = & \pi_1(x_1,y_1)\cdot\pi_1(x_2,y_2). & & \end{array}$$

(d) Let $\mathcal{R} = \mathcal{C}(\mathbb{R})$, as in Example ⟨97a⟩, and let $\mathcal{S} = \mathbb{R}$. Define the **evaluation map** $\epsilon_0 : \mathcal{C}(\mathbb{R}) \longrightarrow \mathbb{R}$ by $\epsilon_0(f) = f(0)$. Then ϵ_0 is a ring homomorphism, because

$$\begin{array}{rcccccl} \epsilon_0(f_1+f_2) & = & (f_1+f_2)(0) & = & f_1(0)+f_2(0) & = & \epsilon_0(f_1)+\epsilon_0(f_2); \\ \text{and}\ \ \epsilon_0(f_1\cdot f_2) & = & (f_1\cdot f_2)(0) & = & f_1(0)\cdot f_2(0) & = & \epsilon_0(f_1)\cdot\epsilon_0(f_2); \end{array}$$

(e) More generally, let $r \in \mathbb{R}$ be any fixed real number, and define the **evaluation map** $\epsilon_r : \mathcal{C}(\mathbb{R}) \longrightarrow \mathbb{R}$ by $\epsilon_r(f) = f(r)$. Then ϵ_r is a ring homomorphism.

(f) Let $\mathcal{R} = \mathbb{R}[x]$, as in Example ⟨98a⟩, and let $\epsilon_0 : \mathbb{R}[x] \longrightarrow \mathbb{R}$ be the evaluation map, as before. Again, ϵ_0 is a homomorphism.

Observe that, if $P(x) = p_n x^n + \ldots + p_2 x^2 + p_1 x + p_0$, then $\epsilon_0(P) = p_0$.

(g) Let $\mathcal{S}$ be any ring, and let $\mathcal{R} = \mathcal{S}[x]$, as in Example (98f). Now define $\epsilon_0 : \mathcal{S}[x] \longrightarrow \mathcal{S}$ as follows: If $P(x) = p_n x^n + \ldots + p_2 x^2 + p_1 x + p_0$, then $\epsilon_0(P) = p_0$. Then ϵ_0 is a homomorphism (**Exercise 81**).

(h) Let $-\infty \leq a < b \leq \infty$, and define the **restriction** map $\rho_{(a,b)} : \mathcal{C}(\mathbb{R}) \longrightarrow \mathcal{C}(a,b)$ where $\rho_{(a,b)}(f)$ is just the restriction of f to a function $f\big|_{(a,b)} : (a,b) \longrightarrow \mathbb{R}$. That is, for any $r \in (a,b)$, $f\big|_{(a,b)}(r) = f(r)$, but $f\big|_{(a,b)}$ is not defined outside of the domain (a,b). Then $\rho_{(a,b)}$ is a homomorphism (**Exercise 82**).

(i) More generally, let $\mathbf{X}$ be any topological space and let $\mathbf{U} \subset \mathbf{X}$ be any subspace. Define the **restriction** map $\rho_{\mathbf{U}} : \mathcal{C}(\mathbf{X}) \longrightarrow \mathcal{C}(\mathbf{U})$ by $\rho_{\mathbf{U}}(f) = f\big|_{\mathbf{U}}$. Then $\rho_{\mathbf{U}}$ is a homomorphism.

(j) Let $\mathcal{R} = \mathcal{M}_2(\mathbb{Z})$ be the ring of 2×2 integer matrices, as in Example ⟨96c⟩, and let $\mathcal{S} = \mathcal{M}_2(\mathbb{Z}_{/5})$ be the ring of 2×2 matrices with coefficients in $\mathbb{Z}_{/5}$. Define the homomorphism $\phi : \mathcal{R} \longrightarrow \mathcal{S}$ by reducing each coefficient mod 5. That is,

$$\phi\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \overline{a} & \overline{b} \\ \overline{c} & \overline{d} \end{bmatrix},$$

where $\overline{a}$ is the congruence class of a, mod 5, etc. Then ϕ is a ring homomorphism.

(k) Let $\mathcal{R}$ be any ring with multiplicative identity $1_\mathcal{R}$. There is a natural homomomorphism $\phi:\mathbb{Z}\longrightarrow\mathcal{R}$ given by $\phi(n) = \underbrace{1_\mathcal{R}+\ldots+1_\mathcal{R}}_{n}$ for any $n\in\mathbb{N}$, and $\phi(-n) = -\phi(n)$.

Indeed, this is the *only* homomorphism from $\mathbb{Z}$ into $\mathcal{R}$. (**Exercise 83**).

Lemma 112 *Let $\phi:\mathcal{R}\longrightarrow\mathcal{S}$ be a ring homomorphism. Then:*

1. $\phi(0_\mathcal{R}) = 0_\mathcal{S}$.
2. *For any $r\in\mathcal{R}$, $\phi(-r) = -\phi(r)$. If r is a unit, then $\phi(r^{-1}) = \phi(r)^{-1}$.*
3. *$\phi(\mathcal{R})$ is a subring of $\mathcal{S}$.*

Proof: **Exercise 84** □

Let $0_\mathcal{S}$ be the zero of the ring $\mathcal{S}$. The **kernel** of ϕ is the preimage of $0_\mathcal{S}$:

$$\ker[\phi] \quad = \quad \phi^{-1}\{0\} \quad = \quad \{r\in\mathcal{R}\ ;\ \phi(r) = 0_\mathcal{S}\}.$$

Example 113:

(a) Define $\phi:\mathbb{Z}\longrightarrow\mathbb{Z}_{/5}$ by $\phi(z) = \overline{z}$, as in Example ⟨111a⟩. Then

$$\ker(\phi) \quad = \quad \{z\in\mathbb{Z}\ ;\ \overline{z} = 0 \pmod 5\} \quad = \quad \{\ldots,-5,0,5,10,\ldots\} \quad = \quad 5\mathbb{Z}.$$

(b) More generally, let $n\in\mathbb{N}$, and define $\phi:\mathbb{Z}\longrightarrow\mathbb{Z}_{/n}$ by $\phi(z) = \overline{z}$, as in Example ⟨111b⟩. Then $\ker(\phi) \ = \ n\mathbb{Z}$.

(c) Let $\pi_1:\mathbb{R}^2\longrightarrow\mathbb{R}$ be the projection into the first coordinate, as in Example ⟨111c⟩. Then

$$\ker(\phi) \quad = \quad \left\{(x,y)\in\mathbb{R}^2\ ;\ x = 0\right\} \quad = \quad \{(0,y)\ ;\ y\in\mathbb{R}\}.$$

(d) Let $r\in\mathbb{R}$, and Let $\epsilon_r:\mathcal{C}(\mathbb{R})\longrightarrow\mathbb{R}$ be as in Example ⟨111e⟩. Then $\ker(\epsilon_r) \ = \ \{f\in\mathcal{C}(\mathbb{R})\ ;\ f(r) = 0\}$.

(e) Let $\epsilon_0:\mathcal{R}[x]\longrightarrow\mathcal{R}$ be the evaluation map, as in Example ⟨111g⟩. If $P(x) = p_nx^n+\ldots+p_2x^2+p_1x+p_0$, then

$$\begin{aligned}\Big(\ P\in\ker(\epsilon_0)\ \Big) &\iff \Big(\ p_0 = 0\ \Big)\\ &\iff \Big(\ P(x) \ = \ p_nx^n+\ldots+p_2x^2+p_1x \ = \ x\cdot(p_nx^{n-1}+\ldots+p_2x^1+p_1)\ \Big)\end{aligned}$$

Thus, $\ker(\epsilon_0) \quad = \quad \{x\cdot q(x)\ ;\ q\in\mathcal{R}[x]$ any polynomial$\}$.

(f) Let $-\infty\leq a<b\leq\infty$, let $\rho_{(a,b)}:\mathcal{C}(\mathbb{R})\longrightarrow\mathcal{C}(a,b)$ be the **restriction map**, as in Example ⟨111h⟩. Then $\ker(\rho_{(a,b)}) \ = \ \{f\in\mathcal{C}(\mathbb{R})\ ;\ f(x) = 0$ for all $x\in(a,b)\}$.

(g) Define the homomorphism $\phi:\mathcal{M}_2(\mathbb{Z})\longrightarrow\mathcal{M}_2(\mathbb{Z}_{/5})$ by reducing each coefficient mod 5, as in Example ⟨111j⟩. Then $\ker(\phi)$ consists of matrices whose entries are all multiples of 5. That is, $\ker(\phi) = \mathcal{M}_2(5\mathbb{Z})$. (**Exercise 85**)

8.2 (∗) Many Examples of Ring Homomorphisms

Prerequisites: §8.1

[This section contains many examples. It is not necessary to read and understand all these examples right now; just read a few in order to get your intuititions working. These examples will be referred to throughout the text.]

A **ring monomorphism** is a ring homomorphism that is *injective* (ie. one-to-one).

Example 114: (Ring Monomorphisms)

⟨a⟩ Define the map $\iota : \mathbb{Z} \longrightarrow \mathbb{R}$ by $\iota(z) = z$ for any $z \in \mathbb{Z}$. (In other words, ι transforms z considered as an element of $\mathbb{Z}$ to z considered as an element of $\mathbb{R}$). Then ι is a ring monomorphism.

⟨b⟩ More generally, if $\mathcal{S}$ is a subring of $\mathcal{R}$, we define the **inclusion map** $\iota : \mathcal{S} \longrightarrow \mathcal{R}$ by $\iota(s) = s$ for any $s \in \mathcal{S}$. Then ι is a monomorphism.

⟨c⟩ Define $\phi : \mathbb{R} \longrightarrow \mathbb{R}[x]$ as follows: for any $r \in \mathbb{R}$, $\phi(r)$ is the constant polynomial $r + 0x + 0x^2 + \ldots$. Then ϕ is a monomorphism.

⟨d⟩ Define $\phi : \mathbb{R} \longrightarrow \mathcal{C}(\mathbb{R})$ as follows: for any $r \in \mathbb{R}$, $\phi(r)$ is the constant function with value r. Then ϕ is a monomorphism.

A **ring epimorphism** is a homomorphism that is *surjective* (ie. onto).

Example 115: (Ring Epimorphisms)

⟨a⟩ Let $\epsilon_0 : \mathbb{R}[x] \longrightarrow \mathbb{R}$ be the *evaluation map* of Example ⟨111g⟩. Then ϵ_0 is an epimorphism (**Exercise 86**).

⟨b⟩ Define $\epsilon_{\mathbf{i}} : \mathbb{R}[x] \longrightarrow \mathbb{C}$ by mapping any polynomial $p(x) = p_n x^n + \ldots + p_2 x^2 + p_1 x + p_0$ to the complex number $p(\mathbf{i}) = p_n \mathbf{i}^n + \ldots + p_2 \mathbf{i}^2 + p_1 \mathbf{i} + p_0$ For example, if $p(x) = 3x^5 - 7x^4 + 2x^3 - 3x^2 + 4x - 5$, then

$$\epsilon_{\mathbf{i}}(p) \;=\; p(\mathbf{i}) \;=\; 3(\mathbf{i}) - 7(1) + 2(-\mathbf{i}) - 3(-1) + 4\mathbf{i} - 4 \;=\; -8 + 5\mathbf{i}$$

Then $\epsilon_{\mathbf{i}}$ is an epimorphism (**Exercise 87**).

⟨c⟩ Let $\mathbb{Z}[x]$ be as in Example ⟨98c⟩ and let $\mathbb{Z}[\mathbf{i}]$ be as in Example ⟨93f⟩. Define $\epsilon_{\mathbf{i}} : \mathbb{Z}[x] \longrightarrow \mathbb{Z}[\mathbf{i}]$ by mapping any polynomial $p(x) = p_n x^n + \ldots + p_1 x + p_0$ to the complex number $p(\mathbf{i}) = p_n \mathbf{i}^n + \ldots + p_1 \mathbf{i} + p_0$. Then $\epsilon_{\mathbf{i}}$ is an epimorphism (**Exercise 88**).

A **ring isomorphism** is a homomorphism that is *bijective* (ie. one-to-one and onto).

Example 116: (Ring Isomorphisms)

⟨a⟩ Let $\mathcal{M}_2(\mathbb{Z})$ be the ring of 2×2 integer matrices (Example ⟨96c⟩ on page 80) and let $\mathbf{End}\left[\mathbb{Z}^2\right]$ be the ring of endomorphisms of $\mathbb{Z}^2$ (Example ⟨101a⟩ on page 88).

If $\mathbf{M} = \left[\begin{smallmatrix} m_1 & m_2 \\ m_3 & m_4 \end{smallmatrix}\right]$ is a matrix in $\mathcal{M}_2(\mathbb{Z})$, then define endomorphism $\phi_{\mathbf{M}} \in \mathbf{End}\left[\mathbb{Z}^2\right]$ as follows: for any element $\mathbf{z} = (z_1, z_2)$ in $\mathbb{Z}^2$,

$$\phi_{\mathbf{M}}(\mathbf{z}) \quad = \quad \begin{bmatrix} m_1 & m_2 \\ m_3 & m_4 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \quad = \quad \begin{bmatrix} m_1 z_1 + m_2 z_2 \\ m_3 z_1 + m_4 z_2 \end{bmatrix}$$

The map $\mathbf{M} \mapsto \phi_{\mathbf{M}}$ is an isomorphism $\mathcal{M}_2(\mathbb{Z})$ into $\mathbf{End}\left[\mathbb{Z}^2\right]$ (**Exercise 89**).

⟨b⟩ More generally, let $\mathcal{M}_n(\mathbb{Z})$ be the ring of $n \times n$ integer matrices and $\mathbf{End}\left[\mathbb{Z}^n\right]$ be the endomorphism ring of $\mathbb{Z}^n$, and for any $\mathbf{M} \in \mathcal{M}_n(\mathbb{Z})$, let $\phi_{\mathbf{M}} \in \mathbf{End}\left[\mathbb{Z}^n\right]$ by multiplication-by-$\mathbf{M}$. Then the map $\mathbf{M} \mapsto \phi_{\mathbf{M}}$ is an isomorphism $\mathcal{M}_n(\mathbb{Z})$ into $\mathbf{End}\left[\mathbb{Z}^n\right]$ (**Exercise 90**).

⟨c⟩ Fix a number $N \in \mathbb{N}$. Let $\mathcal{R}$ be any ring, and let $\mathcal{R}^N$ be the product ring from Example ⟨95d⟩ on page 79. Let $\mathcal{N} = \{1, 2, \ldots, N\}$, and let $\mathbb{R}^{\mathcal{N}}$ be the set of all functions from $\mathcal{N}$ into $\mathbb{R}$ (Example ⟨100b⟩ on page 88). Define $\Phi : \mathcal{R}^{\mathcal{N}} \longrightarrow \mathcal{R}^N$ as follows:

Given any function $F \in \mathcal{R}^{\mathcal{N}}$, define the N-tuple $\mathbf{f} = (f_1, f_2, \ldots, f_N)$ by $f_n = F(n)$ for all $n \in \mathcal{N}$. Then the map $F \mapsto \mathbf{f}$ is an isomorphism of $\mathcal{R}^{\mathcal{N}}$ into $\mathcal{R}^N$ (**Exercise 91**).

⟨d⟩ Let $\mathbf{X}$ be a set, and let $\mathcal{P}(\mathbf{X})$ be the ring of subsets of $\mathbf{X}$ (Example ⟨102⟩ on page 89). Let ${\mathbb{Z}_{/2}}^{\mathbf{X}}$ be the ring of all functions from $\mathbf{X}$ into $\mathbb{Z}_{/2}$ (Example ⟨100d⟩ on page 88). Define $\Phi : {\mathbb{Z}_{/2}}^{\mathbf{X}} \longrightarrow \mathcal{P}(\mathbf{X})$ as follows: for any function $f \in {\mathbb{Z}_{/2}}^{\mathbf{X}}$, let $\Phi(f)$ be the set $\{x \in \mathbf{X} \ ; \ f(x) = 1\}$. Then Φ is a ring isomorphism (**Exercise 92**).

⟨e⟩ Let $\mathcal{E} = \left\{ \left[\begin{smallmatrix} x & y \\ -y & x \end{smallmatrix}\right] \ ; \ x, y \in \mathbb{R} \right\}$ be the set of all 2×2 *conformal* matrices. Then $\mathcal{E}$ is a subring of $\mathcal{M}_2(\mathbb{R})$ (**Exercise 93**).

Define the map $\phi : \mathbb{C} \longrightarrow \mathcal{E}$ by $\phi(x + y\mathbf{i}) = \left[\begin{smallmatrix} x & y \\ -y & x \end{smallmatrix}\right]$. Then ϕ is a ring isomorphism (**Exercise 94**).

⟨f⟩ **Cayley's Theorem for Rings:** *Any ring with identity is isomorphic to the endomorphism ring of some abelian group.*

Proof: Let $\mathcal{R}$ be any ring with a multiplicative identity. Let $\widetilde{\mathcal{R}}$ be the *additive abelian group* obtained by forgetting the multiplicative structure on $\mathcal{R}$, and just treating $\mathcal{R}$ as a group under addition. If r is any element of $\mathcal{R}$, let $\widetilde{r}$ denote the corresponding element of $\widetilde{\mathcal{R}}$.

(For example, if $\mathcal{R} = \mathbb{C}$, then $\widetilde{\mathcal{R}} = \mathbb{R}^2$. If $z = x + y\mathbf{i}$ is a complex number, then $\widetilde{z} = (x, y) \in \mathbb{R}^2$).

Let $\mathbf{End}\left[\widetilde{\mathcal{R}}\right]$ be the ring of endomorphisms of $\widetilde{\mathcal{R}}$ from Example ⟨101⟩ on page 88. There is a natural homomorphism $\phi : \mathcal{R} \longrightarrow \mathbf{End}\left[\widetilde{\mathcal{R}}\right]$, defined as follows: For any $r \in \mathcal{R}$, let

$\phi(r) = \phi_r$ be the map $\phi_r : \widetilde{\mathcal{R}} \longrightarrow \widetilde{\mathcal{R}}$ defined: $\phi_r(\widetilde{s}) = \widetilde{r \cdot s}$. (**Exercise 95** Verify that this is a homomorphism.)

(For example, if $\mathcal{R} = \mathbb{C}$ and $\widetilde{\mathcal{R}} = \mathbb{R}^2$, then for any complex number $z = x + y\mathbf{i}$, $\phi_z : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ is just multiplication by the conformal matrix $\begin{bmatrix} x & y \\ -y & x \end{bmatrix}$ from Example ⟨116e⟩.)

If $\mathcal{E} = \phi(\mathcal{R})$, then $\mathcal{E}$ is a subring of $\mathbf{End}\left[\widetilde{\mathcal{R}}\right]$, and ϕ is an isomorphism from $\mathcal{R}$ to $\mathcal{E}$. (**Exercise 96** Hint: it is important that $\mathcal{R}$ has an identity element.)

(For example, if $\mathcal{R} = \mathbb{C}$, and $\widetilde{\mathcal{R}} = \mathbb{R}^2$, then $\mathcal{E}$ is just the ring of 2×2 real-valued *conformal* matrices, as in Example ⟨116e⟩.) □

Cayley's theorem for *groups* says that any group $\mathcal{G}$ can be seen as a *permutation group*; this is done by letting $\mathcal{G}$ act upon itself. Thus, the class of permutation groups are 'universal' for groups.

Similarly, Cayley's theorem for *rings* says that any ring $\mathcal{R}$ can be seen as an *endomorphism ring*; again this is done by letting $\mathcal{R}$ act upon itself. Thus, the class of endomorphism rings are 'universal' for rings.

8.3 Ideals

Prerequisites: §7.4, §8.1

Consider the set $2\mathbb{Z}$ of all *even numbers*, and recall:

(a) The sum of any two even numbers is an even number.

(b) The product of any two even numbers is an even number.

(c) If e is even, and n is any other number, then $e \cdot n$ is also even.

Facts (a) and (b) together imply that $2\mathbb{Z}$ is a subring. Fact (c) says that $2\mathbb{Z}$ has an interesting 'contagious' property; multiplying any number by an even number makes it even. We say that $2\mathbb{Z}$ is an *ideal* of $\mathbb{Z}$.

Let $\mathcal{R}$ be any ring. An **ideal** is a subring $\mathcal{I} < \mathcal{R}$ such that:

$$\text{For any } r \in \mathcal{R}, \text{ and any } i \in \mathcal{I}, \qquad r \cdot i \in \mathcal{I}, \text{ and } \ i \cdot r \in \mathcal{I}.$$

We then write: "$\mathcal{I} \triangleleft \mathcal{R}$".

Example 117:

(a) $2\mathbb{Z}$ is an ideal of $\mathbb{Z}$.

(b) More generally, let $n \in \mathbb{N}$, and let $\mathcal{I} = n\mathbb{Z}$ be the set of all multiples of n. Then $n\mathbb{Z}$ is an ideal. To see this, first recall from Example ⟨104c⟩ that $n\mathbb{Z}$ is a subring of $\mathbb{Z}$. Next, observe that, if $i = n \cdot z$ is any element of $n\mathbb{Z}$, and $r \in \mathbb{Z}$ is any other integer, then $i \cdot r = n \cdot (z \cdot r)$ is also in $n\mathbb{Z}$.

(c) Let $\mathcal{R} = \mathbb{R}^2$ be the product ring from Example ⟨95a⟩, and let $\mathcal{I} = \{(0, y) \; ; \; y \in \mathbb{R}\}$. Then $\mathcal{I}$ is an ideal. To see this, first observe that $\mathcal{I}$ is a subring: for any $(0, y_1)$ and $(0, y_2)$ in $\mathcal{I}$,

$$(0, y_1) + (0, y_2) \;=\; (0,\; y_1 + y_2) \quad \text{and} \quad (0, y_1) \cdot (0, y_2) \;=\; (0,\; y_1 \cdot y_2)$$

are also in $\mathcal{I}$. If (x, y) is any element of $\mathcal{R}$, and $(0, i) \in \mathcal{I}$, then $(0, i) \cdot (x, y) \;=\; (0, i \cdot y)$ is also in $\mathcal{I}$.

(d) Let $\mathcal{R} = \mathcal{C}(\mathbb{R})$, as in Example ⟨97a⟩. Fix $r \in \mathcal{R}$, and define

$$\mathcal{M}_r \;=\; \{f \in \mathcal{C}(\mathbb{R}) \; ; \; f(r) = 0\}.$$

Then $\mathcal{M}_r$ is an ideal. To see this, suppose m_1, m_2 are elements of $\mathcal{M}_r$, and let $f \in \mathcal{C}(\mathbb{R})$ be any other function. Then

$$\begin{array}{rclclcll}
(m_1 + m_2)(r) & = & m_1(r) + m_2(r) & = & 0 + 0 & = & 0, & \text{so } m_1 + m_2 \text{ is in } \mathcal{M}_r. \\
(m_1 \cdot m_2)(r) & = & m_1(r) \cdot m_2(r) & = & 0 \cdot 0 & = & 0, & \text{so } m_1 \cdot m_2 \text{ is in } \mathcal{M}_r. \\
(m_1 \cdot f)(r) & = & m_1(r) \cdot f(r) & = & 0 \cdot f(r) & = & 0, & \text{so } m_1 \cdot f \text{ is in } \mathcal{M}_r.
\end{array}$$

(e) Let $\mathcal{S}$ be any ring, and let $\mathcal{R} \;=\; \mathcal{S}[x]$, as in Example (98f). Let

$$\mathcal{I} \;=\; \{x \cdot q(x) \; ; \; q \in \mathcal{R}[x] \text{ any polynomial}\}.$$

Then $\mathcal{I}$ is an ideal (**Exercise 97**).

(f) Let $\mathcal{R} \;=\; \mathcal{C}(\mathbb{R})$, as in Example ⟨97a⟩. Let $-\infty \leq a < b \leq \infty$, and define $\mathcal{I} \;=\; \{f \in \mathcal{C}(\mathbb{R}) \; ; \; f(x) = 0 \text{ for all } x \in (a, b)\}$. Then $\mathcal{I}$ is an ideal (**Exercise 98**).

(g) Let $\mathcal{R} \;=\; \mathcal{M}_2(\mathbb{Z})$ be the ring of 2×2 integer matrices, as in Example ⟨96c⟩, and let $\mathcal{I} = \mathcal{M}_2(5\mathbb{Z})$ be the ring of 2×2 matrices with coefficients in $5\mathbb{Z}$ —for example $\left[\begin{smallmatrix} 5 & 10 \\ -20 & 15 \end{smallmatrix}\right]$. Then $\mathcal{M}_2(5\mathbb{Z})$ is an ideal of $\mathcal{M}_2(\mathbb{Z})$ (**Exercise 99**).

(h) More generally, let $\mathcal{R}$ be any ring and let $\mathcal{I} \lhd \mathcal{R}$. Let $\mathcal{M}_n(\mathcal{R})$ be the set of $n \times n$ matrices with coefficients in $\mathcal{R}$, as in Example ⟨96f⟩. Then $\mathcal{M}_n(\mathcal{I})$ is an ideal of $\mathcal{M}_n(\mathcal{R})$ (**Exercise 100**).

(i) Let $\mathcal{R} = \mathbb{Z}[x]$ be the ring of polynomials with integer coefficients, as in Example ⟨98c⟩. Let $5\mathbb{Z}[x]$ be the ring of polynomials with coefficients in $5\mathbb{Z}$ (for example, $p(x) = 15x^3 - 5x = 10$). Then $5\mathbb{Z}[x]$ is an ideal of $\mathbb{Z}[x]$ (**Exercise 101**).

(j) More generally, let $\mathcal{R}$ be any ring and let $\mathcal{I} \lhd \mathcal{R}$. Let $\mathcal{R}[x]$ be the set of all formal polynomials with coefficients in $\mathcal{R}$, as in Example ⟨98f⟩. Then $\mathcal{I}[x]$ is an ideal of $\mathcal{R}[x]$ (**Exercise 102**).

(k) **Principal ideals in commutative rings:** In Example ⟨117a⟩, we saw that $2\mathbb{Z} = \{2z \; ; \; z \in \mathbb{Z}\}$ is an ideal of $\mathbb{Z}$. In Example ⟨117e⟩, we saw that $\mathcal{I} = \{x \cdot q(x) \; ; \; q \in \mathcal{R}[x]\}$ is an ideal of $\mathcal{R}[x]$; we could have written this ideal as $\mathcal{I} = x \cdot \mathbb{R}[x]$.

In general, if $\mathcal{R}$ is any commutative ring, and $p \in \mathcal{R}$, then the *principal ideal* generated by p is the set

$$p\mathcal{R} \quad = \quad \{pr \; ; \; r \in \mathcal{R}\}.$$

Exercise 103 Show that this is an ideal.

We normally denote this ideal by '(p)'. For example, if $\mathcal{R} = \mathbb{Z}$, then $(3) = 3\mathbb{Z}$.

(l) Let $\mathcal{R} = \mathcal{C}(\mathbb{R})$, as in Example ⟨97a⟩ on page 81, and let $\mathcal{I} = \mathcal{C}_0(\mathbb{R})$, as in Example ⟨3k⟩. Then $\mathcal{I} \lhd \mathcal{R}$ (**Exercise 104**).

Ideals are to rings as *normal subgroups* are to groups...

Proposition 118 *Let $\phi : \mathcal{R} \longrightarrow \mathcal{I}$ be a ring homomorphism. Then* $\ker(\phi)$ *is an ideal of* $\mathcal{R}$.

Proof: Let $k_1, k_2 \in \ker(\phi)$, and let $r \in \mathcal{R}$. Then

$$\begin{array}{rcccccl} \phi(k_1 + k_2) & = & \phi(k_1) + \phi(k_2) & = & 0 + 0 & = & 0, \text{ so } k_1 + k_2 \text{ is in } \ker(\phi). \\ \phi(k_1 \cdot k_2) & = & \phi(k_1) \cdot \phi(k_2) & = & 0 \cdot 0 & = & 0, \text{ so } k_1 \cdot k_2 \text{ is in } \ker(\phi). \\ \phi(k_1 \cdot f) & = & \phi(k_1) \cdot \phi(f) & = & 0 \cdot \phi(f) & = & 0, \text{ so } k_1 \cdot f \text{ is in } \ker(\phi). \\ \text{and } \phi(f \cdot k_1) & = & \phi(f) \cdot \phi(k_1) & = & \phi(f) \cdot 0 & = & 0, \text{ so } f \cdot k_1 \text{ is in } \ker(\phi). \end{array}$$

□

Example 119:

(a) Example ⟨117a⟩ is just the kernel of the homomorphism $\phi : \mathbb{Z} \longrightarrow \mathbb{Z}_{/2}$.

(b) Example ⟨117b⟩ is just the kernel of the homomorphism $\phi : \mathbb{Z} \longrightarrow \mathbb{Z}_{/n}$, from Example (113b).

(c) Example ⟨117c⟩ is just the kernel of the homomorphism $\pi_1 : \mathbb{R}^2 \longrightarrow \mathbb{R}$, from Example ⟨113c⟩

(d) Example ⟨117d⟩ is the kernel of the evaluation map $\epsilon_r : \mathcal{C}(\mathbb{R}) \longrightarrow \mathbb{R}$, from Example ⟨113d⟩.

(e) Example ⟨117e⟩ is the kernel of the evaluation map $\epsilon_0 : \mathcal{R}[x] \longrightarrow \mathcal{R}$, from Example (113e).

(f) Example ⟨117f⟩ is the kernel of the restriction map $\rho_{(a,b)} : \mathcal{C}(\mathbb{R}) \longrightarrow \mathcal{C}(a, b)$ from Example ⟨113f⟩.

(g) Example ⟨117g⟩ is the kernel of the homomorphism $\phi : \mathcal{M}_2(\mathbb{Z}) \longrightarrow \mathcal{M}_2(\mathbb{Z}_{/5})$ from Example ⟨113g⟩.

Ideals are 'allergic' to multiplicative inverses, in the following sense:

Lemma 120 *Let $\mathcal{R}$ be a ring with multiplicative identity* 1, *and let $\mathcal{I}$ be an ideal. The following are equivalent:*

(a) *$\mathcal{I}$ contains some unit element u.*

(b) $1 \in \mathcal{I}$.

(c) $\mathcal{I} = \mathcal{R}$.

Proof: **Exercise 105** □

Corollary 121 *Let $\mathbb{F}$ be a field*[1]*. Then $\mathbb{F}$ contains no proper ideals.*

Proof: Let $\mathcal{I}$ be any nontrivial ideal of $\mathbb{F}$. Thus, $\mathcal{I}$ contains some nonzero element $u \in \mathbb{F}$. But $\mathbb{F}$ is a field, so u is a unit, so $\mathcal{I} = \mathbb{F}$ by the previous theorem. □

We say that a ring $\mathcal{R}$ is **simple** if it contains no proper ideals. Thus, Corollary 121 can be reformulated: *Every field is a simple ring.*

8.4 Quotient Rings

Prerequisites: §8.3 **Recommended:** §1.4

Let $\mathcal{R}$ be a ring. If $\mathcal{I} < \mathcal{R}$ is any subring, and $r \in \mathcal{R}$, then the corresponding **coset** of $\mathcal{I}$ is the set:

$$r + \mathcal{I} \quad = \quad \{r + i \; ; \; i \in \mathcal{I}\}.$$

Example 122:

(a) Let $\mathcal{R} = \mathbb{Z}$, and let $\mathcal{I} = 5\mathbb{Z} = \{5z \; ; \; z \in \mathbb{Z}\} = \{\ldots, -5, 0, 5, 10, 15, \ldots\}$ as in Example ⟨117b⟩. Then $3 + 5\mathbb{Z} = \{3 + 5z \; ; \; z \in \mathbb{Z}\} = \{\ldots, -2, 3, 8, 13, 18, \ldots\}$.

(b) Let $\mathcal{S}$ and $\mathcal{T}$ be rings, and consider the product ring

$$\mathcal{R} \; = \; \mathcal{S} \times \mathcal{T} \; = \; \{(s,t) \; ; \; s \in \mathcal{S} \text{ and } t \in \mathcal{T}\}. \qquad \text{(Example ⟨95⟩ on page 78)}$$

Let $\mathcal{I} = \{(0_{\mathcal{S}}, t) \; ; \; t \in \mathcal{T}\} = \{0_{\mathcal{S}}\} \times \mathcal{T}$. Then for any $(s, t_1) \in \mathcal{R}$,

$(s,t_1)+\mathcal{I} = \{(s,t_1) + (0_{\mathcal{S}}, t_2) \; ; \; t_2 \in \mathcal{T}\} = \{(s, \; t_1 + t_2) \; ; \; t \in \mathcal{T}\} = \{(s,t) \; ; \; t \in \mathcal{T}\}$
$= \{s\} \times \mathcal{T}$.

[1] ...or a division ring.

Lemma 123 *Let $\mathcal{R}$ be a ring and let $\mathcal{I} < \mathcal{R}$ be a subring. For any $r \in \mathcal{R}$:*

(a) $\left(r + \mathcal{I} \ = \ \mathcal{I} \right) \iff \left(r \in \mathcal{I} \right)$.

(b) $\left(r \in \mathcal{I} \right) \Longrightarrow \left(r \cdot \mathcal{I} \ \subset \ \mathcal{I} \right)$. *However, the converse is not generally true.*

(For example, if $\mathcal{I}$ is an ideal, then $r\mathcal{I} \subset \mathcal{I}$ for any $r \in \mathcal{R}$.)

Proof: **Exercise 106** ______________ □

Example 124:

(a) Let $\mathcal{R} = \mathbb{Z}$, and let $\mathcal{I} \ = \ 5\mathbb{Z}$ as in Example ⟨122a⟩. Then $10+5\mathbb{Z} \ = \ \{10 + 5z \ ; \ z \in \mathbb{Z}\} \ = \ \{\ldots, 5, 10, 15, 20, 25, \ldots\} \ = \ 5\mathbb{Z}$.

(b) Let $\mathcal{S}$ and $\mathcal{T}$ be groups; let $\mathcal{R} \ = \ \mathcal{S} \times \mathcal{T}$ and let $\mathcal{I} \ = \ \{0_{\mathcal{S}}\} \times \mathcal{T}$, as in Example ⟨122b⟩. Then for any $t \in \mathcal{T}$, $(0_{\mathcal{S}}, t) + \mathcal{I} \ = \ \{0_{\mathcal{S}}\} \times \mathcal{T} \ = \ \mathcal{I}$. ______________

The **coset space** of $\mathcal{I}$ is the set of all its left cosets:

$$\mathcal{R}/\mathcal{I} \quad = \quad \{(r + \mathcal{S}) \ ; \ r \in \mathcal{R}\}.$$

Example 125:

(a) Let $\mathcal{R} = \mathbb{Z}$, and let $\mathcal{I} = 5\mathbb{Z}$ as in Example ⟨122a⟩. Then

$$\frac{\mathbb{Z}}{5\mathbb{Z}} \quad = \quad \{5\mathbb{Z}, \quad 1 + 5\mathbb{Z}, \quad 2 + 5\mathbb{Z}, \quad 3 + 5\mathbb{Z}, \quad 4 + 5\mathbb{Z}\}.$$

(b) Let $\mathcal{S}$ and $\mathcal{T}$ be rings; let $\mathcal{R} = \mathcal{S} \times \mathcal{T}$, and let $\mathcal{I} \ = \ \{0_{\mathcal{S}}\} \times \mathcal{T}$, as in Example ⟨122b⟩. Then $\dfrac{\mathcal{R}}{\mathcal{I}} \ = \ \{\mathcal{I}_s \ ; \ s \in \mathcal{S}\}$, where, for any fixed $s \in \mathcal{S}$, $\mathcal{I}_s = \{(s, t) \ ; \ t \in \mathcal{T}\}$. ______________

If $\mathcal{A}, \mathcal{B} \subset \mathcal{R}$ are subsets, then their **sum** is the set

$$\mathcal{A} + \mathcal{B} \ = \ \{a + b \ ; \ a \in \mathcal{A} \text{ and } \ b \in \mathcal{B}\}, \tag{8.2}$$

and their **product** is the set

$$\mathcal{A} \cdot \mathcal{B} \ = \ \{a \cdot b \ ; \ a \in \mathcal{A} \text{ and } \ b \in \mathcal{B}\}. \tag{8.3}$$

Lemma 126 *Let $\mathcal{R}$ be a ring.*

(a) *Subset addition in $\mathcal{R}$ is* <u>associative</u> *and* <u>commutative</u>. *That is:*

$$\begin{array}{rrcl} \text{for any subsets } \mathcal{A},\mathcal{B},\mathcal{C} \subset \mathcal{R}, & \mathcal{A}+(\mathcal{B}+\mathcal{C}) & = & (\mathcal{A}+\mathcal{B})+\mathcal{C}, \\ \text{and, for any subsets } \mathcal{A},\mathcal{B} \subset \mathcal{R}, & \mathcal{A}+\mathcal{B} & = & \mathcal{B}+\mathcal{A}. \end{array}$$

(b) *Subset multiplication in $\mathcal{R}$ is* <u>associative</u> *and* <u>distributive</u>. *That is, for any subsets $\mathcal{A},\mathcal{B},\mathcal{C} \subset \mathcal{R}$,*

$$\begin{array}{rcl} \mathcal{A}\cdot(\mathcal{B}\cdot\mathcal{C}) & = & (\mathcal{A}\cdot\mathcal{B})\cdot\mathcal{C}, \\ \mathcal{A}\cdot(\mathcal{B}+\mathcal{C}) & = & (\mathcal{A}\cdot\mathcal{B})+(\mathcal{A}\cdot\mathcal{C}), \\ \text{and } (\mathcal{B}+\mathcal{C})\cdot\mathcal{A} & = & (\mathcal{B}\cdot\mathcal{A})+(\mathcal{C}\cdot\mathcal{A}), \end{array}$$

(c) *If $\mathcal{A} < \mathcal{R}$ is a subring of $\mathcal{R}$, then $\mathcal{A}+\mathcal{A}=\mathcal{A}$ and $\mathcal{A}\cdot\mathcal{A}\subset\mathcal{A}$. Furthermore,*

$$\Big(\mathcal{A} \text{ is an ideal of } \mathcal{R}\Big) \iff \Big(\mathcal{R}\cdot\mathcal{A} \;=\; \mathcal{A} \;=\; \mathcal{A}\cdot\mathcal{R}\Big)$$

Proof: **<u>Exercise 107</u>** □

Proposition 127 *Let $\mathcal{R}$ be a ring, and let $\mathcal{I} < \mathcal{R}$ be a subring. The following are equivalent:*

(a) *$\mathcal{I} = \ker(\Phi)$ for some ring homomorphism $\Phi : \mathcal{R}\longrightarrow\mathcal{S}$ (where $\mathcal{S}$ is some ring).*

(b) *$\mathcal{I} \lhd \mathcal{R}$.*

(c) *The coset space $\mathcal{R}/\mathcal{I}$ is a* <u>ring</u> *under the addition operation (8.2) and multiplication operation (8.3). Furthermore:*

1. *If $(a+\mathcal{I})$ and $(b+\mathcal{I})$ are cosets of $\mathcal{I}$, then*

$$(a+\mathcal{I})+(b+\mathcal{I}) \;=\; (a+b)+\mathcal{I}, \quad \text{and} \quad (a+\mathcal{I})\cdot(b+\mathcal{I}) \;=\; (a\cdot b)+\mathcal{I}. \tag{8.4}$$

2. *Define $\pi : \mathcal{R}\longrightarrow\mathcal{R}/\mathcal{I}$ by: $\pi(r) = r+\mathcal{I}$. Then π is a* <u>ring epimorphism</u>, *and $\ker(\pi)=\mathcal{I}$.*

Proof: ‘**(a)**$\Longrightarrow$**(b)**’ This is just Proposition 118 on page 110.

‘**(b)**$\Longrightarrow$**(c)**’ Lemma 126 says these operations are associative etc.

<u>Exercise 108</u> Verify equations (8.4). Check:

$$\text{The additive identity of } \mathcal{R}/\mathcal{I} \text{ is the coset } \mathcal{I} = (0+\mathcal{I}). \tag{8.5}$$

Show that the additive inverse of the coset $(r+\mathcal{I})$ is the coset $(-r+\mathcal{I})$. If $\mathcal{R}$ has a multiplicative identity 1, show that $\mathcal{R}/\mathcal{I}$ has multiplicative identity $(1+\mathcal{I})$.

The fact that π is a homomorphism follows immediately from equations (8.4):

$$\begin{array}{rcccccc} \pi(a+b) & = & (a+b)+\mathcal{I} & = & (a+\mathcal{I})+(b+\mathcal{I}) & = & \pi(a)+\pi(b), \\ \text{and} \;\; \pi(a\cdot b) & = & (a\cdot b)+\mathcal{I} & = & (a+\mathcal{I})\cdot(b+\mathcal{I}) & = & \pi(a)\cdot\pi(b). \end{array}$$

Also, for any $r \in \mathcal{R}$,

$$\Big(r \in \ker(\pi) \Big) \iff \Big(\pi(r) = 0 \Big) \Leftarrow_{\text{by (8.5)}}\Rightarrow \Big((r+\mathcal{I}) \ = \ \mathcal{I} \Big) \Leftarrow_{\text{Lem.123(a)}}\Rightarrow \Big(r \in \mathcal{I} \Big).$$

Hence, $\ker(\pi) = \mathcal{I}$.

'**(c)**$\Longrightarrow$**(a)**' Let $\mathcal{S} = \mathcal{R}/\mathcal{I}$, and let $\Phi = \pi$. □

The ring $\mathcal{R}/\mathcal{I}$ is called the **quotient ring**, and the epimorphism $\pi : \mathcal{R} \longrightarrow \mathcal{R}/\mathcal{I}$ is called the **projection map** or **quotient map**.

Example 128: Let $\mathcal{R} = \mathbb{Z}$ and let $\mathcal{I} = 5\mathbb{Z}$. Then $\mathcal{R}/\mathcal{I}$ is the ring $\mathbb{Z}_{/5}$ of congruence classes, mod 5.

8.5 The Fundamental Isomorphism Theorems

Prerequisites: §8.4 **Recommended:** §2.1, §2.2, §2.3

Theorem 129 Fundamental Isomorphism Theorem

Let $\mathcal{R}$ and $\mathcal{S}$ be rings, and let $\phi : \mathcal{R} \longrightarrow \mathcal{S}$ be a ring homomorphism, with image $\mathcal{T} = \phi(\mathcal{R}) \subset \mathcal{S}$, and kernel $\mathcal{K}$. Then:

(a) $\mathcal{T} \ \cong \ \mathcal{R}/\mathcal{K}$.

(b) *For any $r \in \mathcal{R}$ with $t = \phi(r)$, the ϕ-preimage of t is the r-coset of $\mathcal{K}$. That is:* $\phi^{-1}\{t\} \ = \ (r+\mathcal{K})$.

Proof: <u>**Exercise 109**</u> □

Corollary 130 *Let $\mathcal{R}$ and $\mathcal{S}$ be rings, and let $\phi : \mathcal{R} \longrightarrow \mathcal{S}$ be a ring homomorphism. Then*

$$\Big(\phi \textit{ is injective} \Big) \iff \Big(\ker(\phi) = \{0\} \Big)$$

Proof: <u>**Exercise 110**</u> □

Corollary 131 *Let $\mathbb{F}$ be a field*[2]*, and let $\mathcal{R}$ be a ring. Let $\Phi : \mathbb{F} \longrightarrow \mathcal{R}$ be a group homomorphism. Then:*

either *Φ is trivial (ie. $\Phi(\mathbb{F}) = 0$)*

or *Φ is injective, in which case $\Phi(\mathbb{F})$ is isomorphic to $\mathbb{F}$.*

Proof: <u>**Exercise 111**</u> Hint: Combine Corollaries 130 and 121. □

8.6 The Ring Isomorphism Theorems

Prerequisites: §8.4 **Recommended:** §2.1, §2.2, §2.3

The three isomorphism theorems for groups have analogies for rings.

Theorem 132 Diamond Isomorphism Theorem

Let $\mathcal{R}$ be a ring. Let $\mathcal{S} < \mathcal{R}$ be a subring, and let $\mathcal{I} \lhd \mathcal{R}$ be an ideal. Then:

(a) *$\mathcal{S} + \mathcal{I}$ is a subring of $\mathcal{R}$.*

(b) *$\mathcal{I} \lhd (\mathcal{S} + \mathcal{I})$.*

(c) *$(\mathcal{S} \cap \mathcal{I}) \lhd \mathcal{S}$.*

(d) *There is an isomorphism:* $\dfrac{\mathcal{S}+\mathcal{I}}{\mathcal{I}} \cong \dfrac{\mathcal{S}}{\mathcal{S}\cap\mathcal{I}}$ *given by the map*

$$\begin{array}{rcl} \Phi : \dfrac{\mathcal{S}+\mathcal{I}}{\mathcal{I}} & \longrightarrow & \dfrac{\mathcal{S}}{\mathcal{S}\cap\mathcal{I}} \\ (s+i)+\mathcal{I} & \mapsto & s+(\mathcal{S}\cap\mathcal{I}) \end{array}$$

Proof: <u>**Exercise 112**</u> □

Let $\mathcal{R}$ be a ring, with ideal $\mathcal{I} \lhd \mathcal{R}$ and subring $\mathcal{J} < \mathcal{R}$. Suppose $\mathcal{I} < \mathcal{J} < \mathcal{R}$. Then $\mathcal{I}$ is also a ideal of $\mathcal{J}$, and the quotient ring

$$\frac{\mathcal{J}}{\mathcal{I}} = \{j + \mathcal{I} \; ; \; j \in \mathcal{J}\}$$

is a subset of the quotient ring $\dfrac{\mathcal{R}}{\mathcal{I}} = \{r + \mathcal{I} \; ; \; r \in \mathcal{R}\}$.

[2] ...or any simple ring, for that matter.

Theorem 133 Chain Isomorphism Theorem

Let $\mathcal{R}$ be a ring, with ideals $\mathcal{I} \lhd \mathcal{R}$ and $\mathcal{J} \lhd \mathcal{R}$. Suppose $\mathcal{I} < \mathcal{J}$. Then:

(a) $\dfrac{\mathcal{J}}{\mathcal{I}}$ *is a ideal of* $\dfrac{\mathcal{R}}{\mathcal{I}}$.

(b) *There is an isomorphism* $\dfrac{(\mathcal{R}/\mathcal{I})}{(\mathcal{J}/\mathcal{I})} \quad \cong \quad \dfrac{\mathcal{R}}{\mathcal{J}}$.

(c) *Use 'bar' notation to denote elements of $\mathcal{R}/\mathcal{I}$. Thus, $(r+\mathcal{I}) = \overline{r}$, $\quad \mathcal{J}/\mathcal{I} = \overline{\mathcal{J}}$, $\mathcal{R}/\mathcal{I} = \overline{\mathcal{R}}$, and $\dfrac{\mathcal{R}/\mathcal{I}}{\mathcal{J}/\mathcal{I}} = \overline{\mathcal{R}}/\overline{\mathcal{J}}$. Then the isomorphism $\Phi : \overline{\mathcal{R}}/\overline{\mathcal{J}} \longrightarrow \mathcal{R}/\mathcal{J}$ is defined: $\Phi\left(\overline{r}+\overline{\mathcal{J}}\right) = r+\mathcal{J}$.*

Proof: **Exercise 113** □

Theorem 134 Lattice Isomorphism Theorem

Let $\mathcal{R}$ be a ring and let $\mathcal{I} \lhd \mathcal{R}$ be an ideal. Let $\overline{\mathcal{R}} = \mathcal{R}/\mathcal{I}$. Let $\mathfrak{L}(\overline{\mathcal{R}})$ be the subring lattice of $\overline{\mathcal{R}}$, and let $\mathfrak{L}_{\mathcal{I}}(\mathcal{R})$ be the 'fragment' of $\mathfrak{L}(\mathcal{R})$ consisting of all subrings which contain $\mathcal{I}$. That is:

$$\mathfrak{L}_{\mathcal{I}}(\mathcal{R}) \quad = \quad \{\mathcal{A} < \mathcal{R} \,;\, \mathcal{I} < \mathcal{A}\}.$$

Then there is an order-preserving bijection from $\mathfrak{L}_{\mathcal{I}}(\mathcal{R})$ into $\mathfrak{L}(\overline{\mathcal{R}})$, given:

$$\mathfrak{L}_{\mathcal{I}}(\mathcal{R}) \ \ni \mathcal{A} \mapsto \overline{\mathcal{A}} \in \ \mathfrak{L}(\overline{\mathcal{R}}).$$

Furthermore, for any $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D} \in \mathfrak{L}_{\mathcal{I}}(\mathcal{R})$,

(a) $\Big(\ \mathcal{A} < \mathcal{B}\ \Big) \iff \Big(\ \overline{\mathcal{A}} < \overline{\mathcal{B}}\ \Big)$.

(b) $\overline{\mathcal{C} \cap \mathcal{D}} = \overline{\mathcal{C}} \cap \overline{\mathcal{D}}$.

(c) $\Big(\ \mathcal{A} \lhd \mathcal{R}\ \Big) \iff \Big(\ \overline{\mathcal{A}} \lhd \overline{\mathcal{R}}\ \Big)$, *and in this case,* $\mathcal{R}/\mathcal{A} \ \cong\ \overline{\mathcal{R}}/\overline{\mathcal{A}}$.

Proof: **Exercise 114** □

Chapter 9

Algebraic Geometry

9.1 Algebraic Varieties

Prerequisites: §7.5 **Recommended:** §7.3.3

Let $\mathbb{F}$ be a field –for example, $\mathbb{F} = \mathbb{Q}$, $\mathbb{R}$, or $\mathbb{C}$ (usually $\mathbb{F} = \mathbb{C}$). We refer to the set $\mathbb{F}^n$ as the **affine** n**-space** over $\mathbb{F}$ (we consider $\mathbb{F}^n$ merely as a set of points, *not* as a vector space or a ring). If $p(x_1, \ldots, x_n)$ is a polynomial in $\mathbb{F}[x_1,..., x_n]$, then p determines a function $p : \mathbb{F}^n \longrightarrow \mathbb{F}$ in the obvious way.

An **(affine) algebraic variety** in $\mathbb{F}^n$ is a subset $\mathbf{V} \subset \mathbb{F}^n$ which is the set of common zeros of some collection $\mathcal{P} \subset \mathbb{F}[x_1,..., x_n]$ of polynomials. That is:

$$\mathbf{V} \quad = \quad \mathbb{V}(\mathcal{P}) \quad = \quad \{\mathbf{f} \in \mathbb{F}^n \; ; \; p(\mathbf{f}) = 0 \text{ for all } p \in \mathcal{P}\}.$$

Example 135: Real Algebraic Varieties

In each of the following examples, let $\mathbb{F} = \mathbb{R}$.

(a) Circles: Consider the singleton set $\mathcal{P} = \{p(x, y)\}$, where $p(x, y) \; = \; x^2 + y^2 - 1$. Then

$$\mathbb{V}(\mathcal{P}) \quad = \quad \left\{(x, y) \in \mathbb{R}^2 \; ; \; p(x, y) = 0\right\} \quad = \quad \left\{(x, y) \in \mathbb{R}^2 \; ; \; x^2 + y^2 = 1\right\}$$

is the circle about the origin of radius 1, which we denote $\mathbb{S}^1$ (Figure 9.1A).

(b) Ellipses: Let $\mathcal{P} = \{p(x, y)\}$, where now, $p(x, y) \; = \; (x/5)^2 + (y/2)^2 - 1$. Then

$$\mathbb{V}(\mathcal{P}) \quad = \quad \left\{(x, y) \in \mathbb{R}^2 \; ; \; p(x, y) = 0\right\} \quad = \quad \left\{(x, y) \in \mathbb{R}^2 \; ; \; (x/5)^2 + (y/2)^2 = 1\right\}$$

is an ellipse about the origin with minor axis 2 and major axis 5 (Figure 9.1B).

(c) Spheres: Let $\mathcal{P} = \{p(x_1, \ldots, x_n)\}$, where $p(x_1, \ldots, x_n) \; = \; x_1^2 + \ldots + x_n^2 - 1$. Then

$$\mathbb{V}(\mathcal{P}) \quad = \quad \{\mathbf{x} \in \mathbb{R}^n \; ; \; p(\mathbf{x}) = 0\} \quad = \quad \left\{\mathbf{x} \in \mathbb{R}^2 \; ; \; x_1^2 + \ldots + x_n = 1\right\}$$

is the $(n-1)$-dimensional sphere about the origin of radius 1, which we denote $\mathbb{S}^{n-1}$ (Figure 9.1C).

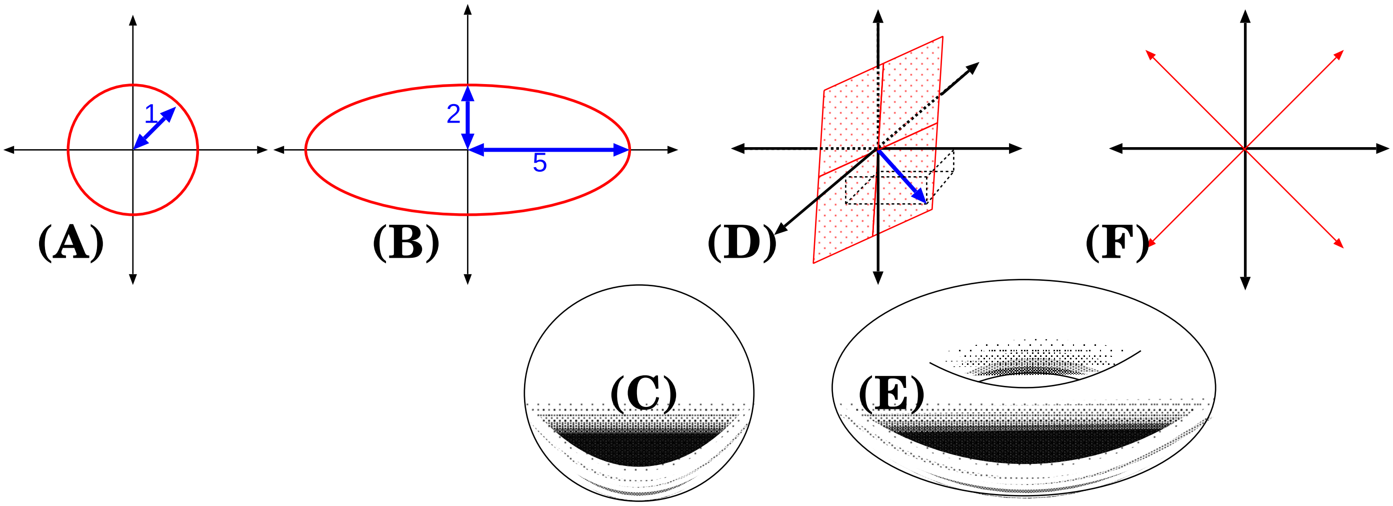


Figure 9.1: Some algebraic varieties

(d) Planes: Let $\mathcal{P} = \{\ell(x,y,z)\}$, where $\ell(x,y,z) = 3x - 2y - z$. Then

$$\mathbb{V}(\mathcal{P}) \quad = \quad \left\{(x,y,z) \in \mathbb{R}^3 \; ; \; \ell(x,y,z) = 0\right\} \quad = \quad \left\{(x,y,z) \in \mathbb{R}^2 \; ; \; 3x - 2y - z \;=\; 0\right\}$$

is the plane through the origin, orthogonal to the vector $(3,-2,-1)$ (Figure 9.1D).

(e) Torii: Let $\mathcal{P} = \{p(w,x,y,z),\ q(w,x,y,z)\}$, where $p(w,x,y,z) = w^2 + x^2 - 1$ and $q(w,x,y,z) = y^2 + z^2 - 1$. Then

$$\begin{aligned} \mathbb{V}(\mathcal{P}) &= \left\{(w,x,y,z) \in \mathbb{R}^4 \; ; \; p(w,x,y,z) \;=\; 0 \;=\; q(w,x,y,z)\right\} \\ &= \left\{(w,x,y,z) \in \mathbb{R}^2 \; ; \; w^2 + x^2 \;=\; 1 \;=\; y^2 + z^2\right\} \end{aligned}$$

is the 2-dimensional **torus**, which we denote $\mathbb{T}^2$ (Figure 9.1E).

(f) X: Let $\mathcal{P} = \{p(x,y)$, where $p(x,y) \;=\; x^2 - y^2 \;=\; (x+y)\cdot(x-y)$. Then

$$\mathbb{V}(\mathcal{P}) \quad = \quad \left\{(x,y) \in \mathbb{R}^2 \; ; \; x^2 - y^2 = 0\right\} \quad = \quad \left\{(x,y) \in \mathbb{R}^2 \; ; \; x = \pm y\right\}$$

is two diagonal lines which cross to make an ‘X’ shape (Figure 9.1F).

Example 136: Algebraic Groups

An **algebraic group** is an algebraic variety with a natural group structure. We will not give a formal definition now, but instead provide some simple examples.

(a) The special linear group $\mathbb{SL}^2[\mathbb{R}]$: Let $\mathcal{M}_2(\mathbb{R})$ be the set of 2×2 real matrices. Identify $\mathcal{M}_2(\mathbb{R})$ with $\mathbb{R}^4$ in the obvious way, so that a 2×2 matrix is written $\begin{bmatrix} w & x \\ y & z \end{bmatrix}$. and let

$\mathcal{P} = \{d(w,x,y,z)\}$, where $d(w,x,y,z) \;=\; wz - yx - 1 \;=\; \det\left[\begin{smallmatrix} w & x \\ y & z\end{smallmatrix}\right] - 1$. Then

$$\begin{aligned}\mathbb{V}(\{d\}) &= \left\{\begin{bmatrix} w & x \\ y & z\end{bmatrix} \in \mathcal{M}_2(\mathbb{R}) \;;\; d(w,x,y,z) = 0\right\} \\ &= \left\{\begin{bmatrix} w & x \\ y & z\end{bmatrix} \in \mathcal{M}_2(\mathbb{R}) \;;\; \det\begin{bmatrix} w & x \\ y & z\end{bmatrix} \;=\; 1\right\}\end{aligned}$$

is the **special linear group** of 2×2 matrices, which we denote $\mathbb{SL}^2[\mathbb{R}]$.

(b) The special linear group $\mathbb{SL}^n[\mathbb{R}]$: Now consider $\mathcal{M}_n(\mathbb{R})$, which we identify with $\mathbb{R}^{n\times n}$ in the obvious way. It is left as an exercise to check that the determinant function

$\det : \mathcal{M}_n(\mathbb{R}) \longrightarrow \mathbb{R}$ is a polynomial. Thus, the function $d(\mathbf{M}) = \det(\mathbf{M}) - 1$ is also a polynomial, and thus, the $n \times n$ special linear group

$$\mathbb{SL}^n[\mathbb{R}] \quad = \quad \{\mathbf{M} \in \mathcal{M}_n(\mathbb{R}) \;;\; d(\mathbf{M}) = 0\}$$

is an algebraic variety.

(c) The orthogonal group $\mathbb{O}^2(\mathbb{R})$: Again, consider $\mathcal{M}_2(\mathbb{R})$, and now let $\mathcal{P} = \{p,q,r,s\}$, where

$$\begin{aligned} p(w,x,y,z) &= w^2 + x^2 - 1; & \qquad q(w,x,y,z) &= wy + xz; \\ r(w,x,y,z) &= yw + zx; & \text{and} \qquad s(w,x,y,z) &= y^2 + z^2 - 1. \end{aligned}$$

Then

$$\begin{aligned}\mathbb{V}(\{p,q,r,s\}) &= \left\{\begin{bmatrix} w & x \\ y & z\end{bmatrix} \in \mathcal{M}_2(\mathbb{R}) \;;\; \begin{matrix} p(w,x,y,z) &=& q(w,x,y,z) &=& 0 \\ r(w,x,y,z) &=& s(w,x,y,z) &=& 0 \end{matrix}\right\} \\ &= \left\{\begin{bmatrix} w & x \\ y & z\end{bmatrix} \in \mathcal{M}_2(\mathbb{R}) \;;\; \begin{bmatrix} w & x \\ y & z\end{bmatrix} \cdot \begin{bmatrix} w & y \\ x & z\end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1\end{bmatrix}\right\} \\ &= \left\{\begin{bmatrix} w & x \\ y & z\end{bmatrix} \in \mathcal{M}_2(\mathbb{R}) \;;\; \begin{bmatrix} w & x \\ y & z\end{bmatrix}^{-1} = \begin{bmatrix} w & y \\ x & z\end{bmatrix}\right\}\end{aligned}$$

is the **orthogonal group** of 2×2 matrices, which we denote $\mathbb{O}^2(\mathbb{R})$.

(d) The orthogonal group $\mathbb{O}^n(\mathbb{R})$: In a similar fashion, we can formulate a set of n^2 polynomials $\{p_{11}, \ldots, p_{nn}\}$ on $\mathcal{M}_n(\mathbb{R})$ so that

$$\mathbb{V}(\mathcal{P}) \quad = \quad \{\mathbf{M} \in \mathcal{M}_n(\mathbb{R}) \;;\; p_{ij}(\mathbf{M}) = 0 \text{ for all } i,j\} \quad = \quad \left\{\mathbf{M} \in \mathcal{M}_n(\mathbb{R}) \;;\; \mathbf{M}^{-1} \;=\; \mathbf{M}^t\right\}$$

is the **orthogonal group** of $n \times n$ matrices, which we denote $\mathbb{O}^n(\mathbb{R})$. The details are **Exercise 115** .

(e) The special orthogonal group $\mathbb{SO}^2[\mathbb{R}]$ Let $\mathcal{P} = \{p,q,r,s,d\}$, where p,q,r and s are as in Example **(c)**, and d is as in Example **(a)**. Then

$$\mathbb{V}(\{p,q,r,s,d\}) \quad = \quad \mathbb{SL}^2[\mathbb{R}] \cap \mathbb{O}^2(\mathbb{R}) \quad = \quad \mathbb{SO}^2[\mathbb{R}]$$

is the **special orthogonal group** of 2×2 matrices.

Example 137: The Fermat Varieties

Let $\mathbb{F} = \mathbb{Q}$, and let $p_n(x,y,z) \ = \ x^n + y^n - z^n$. Clearly, $p_n(0,0,0) = 0$. Fermat's famous *Last Theorem* states:

For any $n \geq 3$, there exist no nonzero $x,y,z \in \mathbb{Z}$ so that $p_n(x,y,z) \ = \ 0$.

This is actually equivalent to the (apparently stronger) statement:

For any $n \geq 3$, there exist no nonzero $x,y,z \in \mathbb{Q}$ so that $p_n(x,y,z) \ = \ 0$.

To see this, suppose that $x = \frac{x_1}{x_2}$, $y = \frac{y_1}{y_2}$, and $z = \frac{z_1}{z_2}$ are rational numbers, and let L be the lowest common multiple of the denominators x_2, y_2 and z_2. Then $X = Lx$, $Y = Ly$, and $Z = Lz$ are integers, and

$$\Big(x^n + y^n \ = \ z^n \Big) \quad \Longleftrightarrow \quad \Big(L^n x^n + L^n y^n \ = \ L^n z^n \Big) \quad \Longleftrightarrow \quad \Big(X^n + Y^n \ = \ Z^n \Big).$$

Hence, *any rational solution to Fermat's equation yields an integer solution.* Conversely, any integer solution to Fermat's equation yields a rational solution, because integers *are* rational numbers.

Hence, if we define the rational algebraic variety

$$\mathbf{V}_n \quad = \quad \big\{(x,y,z) \in \mathbb{Q}^3 \ ; \ x^n + y^n = z^n\big\},$$

then Fermat's *Last Theorem* can be reformulated: "$\mathbf{V}_n = \{(0,0,0)\}$" which is a statement of algebraic geometry.

Many other Diophantine problems can be translated into algebraic geometry in this way. Thus, there is a close relationship between number theory and algebraic geometry.

Example 138: Riemann Surfaces

Let $\mathbb{F} = \mathbb{C}$. A **Riemann surface** is an algebraic variety in $\mathbb{C}^2$ determined by the solutions to a single polynomial equation $p(x,y) = 0$. For example:

(a) The graph of a function: Let $q : \mathbb{C} \longrightarrow \mathbb{C}$ be a polynomial function (say, $q(x) = x^2$), and define $p(x,y) = y - q(x)$. Then

$$\mathbb{V}(\{p\}) \quad = \quad \big\{(x,y) \in \mathbb{C}^2 \ ; \ p(x,y) \ = \ 0\big\} \quad = \quad \big\{(x,y) \in \mathbb{C}^2 \ ; \ y \ = \ q(x)\big\}$$

is the **graph** of the function q. For example, if $q(x) = x^2$, then

$$\mathbf{V}_2 \quad = \quad \left\{(x,\ x^2)\ ;\ x \in \mathbb{C}\right\}.$$

Observe that there is a natural bijection $\pi : \mathbf{V} \longrightarrow \mathbb{C}$ given by $\pi(x, y) = x$. Thus, $\mathbf{V}$ can be seen as a sort of 'deformation' of the complex plane.

(b) The graph of a branched function: The complex square root function $\sqrt{\bullet}$ has two 'branches', each of which constitutes a function defined almost everywhere on the complex plane. To simultaneously visualize both branches, we imagine graphing them both, to obtain a surface

$$\mathbf{V}_{1/2} \quad = \quad \{(y,\ \sqrt{y})\ ;\ y \in \mathbb{C}\},$$

where, for any $y \in \mathbb{C}$, we understand $\sqrt{y}$ to represent two distinct values.

The surface $\mathbf{V}_{1/2}$ forms a *branched double covering* of the complex plane, in the following sense. Consider the projection map $\pi : \mathbf{V}_{1/2} \longrightarrow \mathbb{C}$ defined $\pi(y, \sqrt{y}) \ = \ y$, and observe that π is two-to-one everywhere except at $(0, 0)$, where π is one-to-one.

Observe that the Riemann surface $\mathbf{V}_{1/2}$ is actually the same as the Riemann surface $\mathbf{V}_2$ from Example **(a)**, just with the coordinates reversed. In other words, we can define a natural bijection

$$\mathbf{V}_2 \ni \ (x, x^2) \ \mapsto \ (y, \sqrt{y}) \ \in \mathbf{V}_{1/2}$$

by $y = x^2$.

These examples illustrate an important fact about Riemann surfaces, which we state without proof:

Let $p(x, y) \in \mathbb{C}[x, y]$ be a polynomial of degree n. Then the Riemann surface

$$\mathbf{V} \quad = \quad \left\{(x, y) \in \mathbb{C}^2\ ;\ p(x, y) = 0\right\}$$

is always nonempty, and is a branched n-fold covering of the complex plane.

9.2 The Coordinate Ring

Prerequisites: §9.1, §8.6

Let $\mathbb{F}$ be a field and let $\mathbf{V} \subset \mathbb{F}^n$ be an algebraic variety. A **coordinate function** on $\mathbf{V}$ is the restriction to $\mathbf{V}$ of some polynomial function $p : \mathbb{F}^n \longrightarrow \mathbb{F}$. The set of all coordinate functions on $\mathbf{V}$ is called the **coordinate ring** of $\mathbf{V}$, and denoted $\mathcal{C}oord\,(\mathbf{V})$. That is:

$$\mathcal{C}oord\,(\mathbf{V}) \quad = \quad \left\{p_{|\mathbf{V}}\ ;\ p \in \mathbb{F}[x_1, ..., x_n]\right\}.$$

If $p, q \in \mathbb{F}[x_1, ..., x_n]$ are two polynomials such that $p(v) = q(v)$ for all $v \in \mathbf{V}$, then $p_{|\mathbf{V}}$ and $q_{|\mathbf{V}}$ are really the same function. Hence, elements of $\mathcal{C}oord\,(\mathbf{V})$ can be seen as *equivalence classes* of polynomials in $\mathbb{F}[x_1, ..., x_n]$, where $\Big(\ p\ \sim\ q\ \Big) \iff \Big(\ p(v) = q(v)$ for all $v \in \mathbf{V}\ \Big)$.

Example 139:

(a) Let $\mathbf{V} = \mathbb{F}^n$. Then $\mathcal{C}oord(\mathbb{F}^n) \;=\; \mathbb{F}[x_1,\ldots,x_n]$.

(b) Consider $\mathbb{F}^2$, and let $\mathcal{P} = \{p(x,y)\}$, where $p(x,y) = y$. Then

$$\mathbb{V}(p) \;=\; \left\{(x,y) \in \mathbb{F}^2 \;;\; y = 0\right\} \;=\; \{(x,0)\;;\; x \in \mathbb{F}\}$$

is just the natural embedding of $\mathbb{F}$ into $\mathbb{F}^2$. If $q(x,y)$ and $r(x,y)$ are any polynomials in $\mathbb{F}[x,y]$, then

$$\begin{aligned}\left(\, q_{|\mathbf{V}} \;=\; r_{|\mathbf{V}} \,\right) \quad &\Longleftrightarrow \quad \left(\, q(v,w) = r(v,w) \text{ for all } (v,w) \in \mathbf{V} \,\right) \\ &\Longleftrightarrow \quad \left(\, q(x,0) = r(x,0) \text{ for all } x \in \mathbb{F} \,\right).\end{aligned}$$

(c) Let $\mathbb{F} = \mathbb{C}$. Let $\mathbf{c} = (c_1, c_2, \ldots, c_n)$ be a single point in $\mathbb{C}^n$, and consider the singleton set $\mathbf{V} = \{\mathbf{c}\}$. Then $\mathbf{V} = \mathbb{V}(p_1, p_2, \ldots, p_n)$, where, for any $\mathbf{x} = (x_1, \ldots, x_n)$ in $\mathbb{C}^n$,

$$p_1(\mathbf{x}) \;=\; (x_1 - c_1); \quad p_2(\mathbf{x}) \;=\; (x_2 - c_2); \quad \ldots \quad p_n(\mathbf{x}) \;=\; (x_n - c_n).$$

If $q(x,y)$ and $r(x,y)$ are any polynomials in $\mathbb{C}[x,y]$, then

$$\begin{aligned}\left(\, q_{|\mathbf{V}} \;=\; r_{|\mathbf{V}} \,\right) \quad \Longleftrightarrow \quad \left(\, q(\mathbf{c}) = r(\mathbf{c}) \,\right) \quad &\Longleftrightarrow \quad \left(\, q(\mathbf{c}) - r(\mathbf{c}) \;=\; 0 \,\right) \\ \Longleftrightarrow \quad \left(\, (q - r) \in \mathcal{M}_{\mathbf{c}} \,\right)&,\end{aligned}$$

where $\mathcal{M}_{\mathbf{c}} \;=\; \{f \in \mathbb{C}[x_1,\ldots,x_n]\;;\; f(\mathbf{c}) \;=\; 0\}$ is the maximal ideal from `Hilbert's Nullstellensatz` (p. 149). In other words,

$$\left(\, q_{|\mathbf{V}} \;=\; r_{|\mathbf{V}} \,\right) \;\Longleftrightarrow\; \left(\, q \text{ and } r \text{ belong to the same coset of } \mathcal{M}_{\mathbf{c}} \,\right).$$

The **annihilator** of $\mathbf{V}$ is the ideal

$$\mathcal{A}nn(\mathbf{V}) \;=\; \{f \in \mathbb{F}[x_1,\ldots,x_n]\;;\; f(v) \;=\; 0 \text{ for all } v \in \mathbf{V}\}. \qquad \text{(see §10.7.4)}$$

For example, if $\mathbf{V} = \{\mathbf{c}\}$ as in Example ⟨139c⟩, then $\mathcal{A}nn(\mathbf{V}) = \mathcal{M}_{\mathbf{c}}$. Example ⟨139c⟩ then generalizes as follows:

Proposition 140 *Let $\mathbf{V} \subset \mathbb{F}^n$ be an algebraic variety, and let $\mathcal{A}nn(\mathbf{V})$ be its annihilator. There is a natural ring isomorphism $\mathcal{C}oord(\mathbf{V}) \;\cong\; \dfrac{\mathbb{F}[x_1,\ldots,x_n]}{\mathcal{A}nn(\mathbf{V})}$, given by the map:*

$$\mathcal{C}oord(\mathbf{V}) \ni p_{|\mathbf{V}} \quad \mapsto \quad \left(p + \mathcal{A}nn(\mathbf{V})\right) \;\in\; \frac{\mathbb{F}[x_1,\ldots,x_n]}{\mathcal{A}nn(\mathbf{V})}.$$

Proof: Define $\Phi : \mathbb{F}[x_1,\ldots,x_n]\longrightarrow\mathcal{C}oord(\mathbf{V})$ by $\Phi(p) \;=\; p_{|\mathbf{V}}$. Clearly, this is a surjective ring homomorphism. Thus, the `Fundamental Isomorphism Theorem` (Theorem **??** on page ??) says that $\mathcal{C}oord(\mathbf{V}) \;\cong\; \dfrac{\mathbb{F}[x_1,\ldots,x_n]}{\ker(\Phi)}$. It therefore suffices to show that $\ker(\Phi) = \mathcal{A}nn(\mathbf{V})$. To see this, observe that:

$$\Big(\, p \in \ker(\Phi) \,\Big) \iff \Big(\, \Phi(p) \;=\; 0 \,\Big) \iff \Big(\, p_{|\mathbf{V}} \;=\; 0 \,\Big)$$
$$\iff \Big(\, \text{For all } v \in \mathbf{V},\quad p(v) = 0 \,\Big) \iff \Big(\, p \in \mathcal{A}nn(\mathbf{V}) \,\Big).$$

□

Recall that a ring $\mathcal{R}$ is *perfect* if $\mathcal{R}$ has no nilpotent elements —ie. $\sqrt[*]{0_{\mathcal{R}}} \;=\; \{0_{\mathcal{R}}\}$ (see §10.7.3).

Corollary 141 *Let $\mathcal{R}$ be a quotient ring of $\mathbb{C}[x_1,\ldots,x_n]$. Then*

$$\Big(\, \mathcal{R} \cong \mathcal{C}oord(\mathbf{V}) \textit{ for some algebraic variety } \mathbf{V} \subset \mathbb{C}^n \,\Big) \iff \Big(\, \mathcal{R} \textit{ is perfect.} \,\Big)$$

Proof: Suppose $\mathcal{R} = \mathbb{C}[x_1,\ldots,x_n]/\mathcal{I}$ for some ideal $\mathcal{I} \lhd \mathbb{C}[x_1,\ldots,x_n]$. Let $\mathbf{V} = \mathbb{V}(\mathcal{I})$ be the variety induced by $\mathcal{I}$. Then

$$\Big(\, \mathcal{R} \text{ is a perfect ring} \,\Big) \overset{\text{Lem.219}}{\Longleftarrow\!=\!\Longrightarrow} \Big(\, \mathcal{I} \text{ is a radical ideal} \,\Big) \overset{\text{(Nlstz)}}{\Longleftarrow\!=\!\Longrightarrow} \Big(\, \mathcal{I} \;=\; \mathcal{A}nn(\mathbf{V}) \,\Big)$$
$$\iff \Big(\, \mathcal{R} \;\cong\; \frac{\mathbb{C}[x_1,\ldots,x_n]}{\mathcal{A}nn(\mathbf{V})} \,\Big) \overset{\text{Prop.140}}{\Longleftarrow\!=\!\Longrightarrow} \Big(\, \mathcal{R} \;\cong\; \mathcal{C}oord(\mathbf{V}) \,\Big).$$

Here, **(Nlstz)** is the ‘Radical’ `Nullstellensatz` (Theorem 17 on page 174). □

Corollary 142 *There is a natural bijective correspondence:*

$$\Big\{ \textit{Algebraic varieties in } \mathbb{C}^n \Big\} \longleftrightarrow \Big\{ \textit{Perfect quotient rings of } \mathbb{C}[x_1,\ldots,x_n] \Big\}.$$

9.3 Morphisms

Prerequisites: §9.2 **Recommended:** §10.4.1,§10.7.4

If $\mathbf{V} \subset \mathbb{F}^n$ is an algebraic variety, then a **(geometric) morphism** from $\mathbf{V}$ to $\mathbb{F}^m$ is a function $\Phi : \mathbf{V}\longrightarrow\mathbb{F}^m$ given by

$$\Phi(\mathbf{v}) \quad=\quad \Big(\phi_1(\mathbf{v}), \phi_2(\mathbf{v}), \ldots, \phi_m(\mathbf{v})\Big),$$

where $\phi_1, \phi_2, \ldots, \phi_m \in \mathcal{C}oord(\mathbf{V})$.

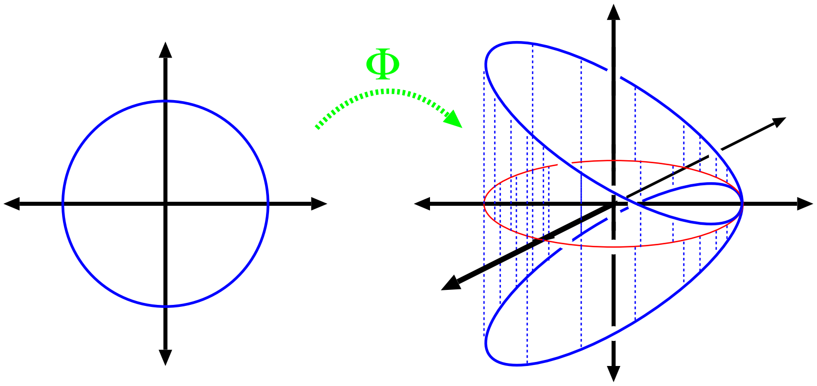


Figure 9.2: The map $\Phi(x,y) \;=\; (x^2-y^2,\ 2xy,\ x)$ transforms a circle into a double-loop.

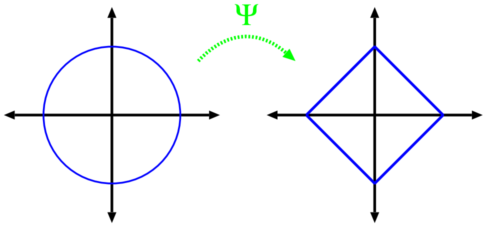


Figure 9.3: The map $\Psi(x,y) \;=\; (x^2,y^2)$ transforms a circle into a diamond.

Example 143: Define $\Phi:\mathbb{S}^1\longrightarrow\mathbb{R}^3$ by $\Phi(x,y) \;=\; (x^2-y^2,\ 2xy,\ x)$. This map transforms the circle into a double-loop in three-dimensional space; see Figure 9.2.

If $\mathbf{W}\subset\mathbb{F}^m$ is another algebraic variety, then a **morphism** from $\mathbf{V}$ to $\mathbf{W}$ is a morphism $\Phi:\mathbf{V}\longrightarrow\mathbb{F}^m$ such that $\Phi(\mathbf{V})\subset\mathbf{W}$.

Example 144: Let $\mathbf{V}=\mathbb{S}^1=\mathbf{W}$, and define $\Phi:\mathbb{S}^1\longrightarrow\mathbb{S}^1$ by $\Phi(x,y) \;=\; (x^2-y^2,\ 2xy)$. Then Φ is a morphism which wraps the circle twice around itself (**Exercise 116** Hint: write $x=\cos(\theta)$ and $y=\sin(\theta)$.). This is the second *Legendre polynomial.*

Note that the image of variety under a morphism is not necessarily a variety:

Example 145: Let $\mathbb{S}^1 \;=\; \{(x,y)\in\mathbb{R}^2\ ;\ x^2+y^2=1\}$ be the unit circle, and define $\Psi:\mathbb{S}^1\longrightarrow\mathbb{R}^2$ by $\Psi(x,y)=(x^2,y^2)$. Then the image $\Psi(\mathbb{S}^1)$ is the 'diamond' $\{(x,y)\in\mathbb{R}^2\ ;\ |x|+|y|=1\}$, which is not an algebraic variety. See Figure 9.3. (**Exercise 117** Check this.)

If $\Phi:\mathbf{V}\longrightarrow\mathbb{F}^m$ is a morphism, let $\overline{\mathsf{image}}\,(\Phi)$ be the *smallest* algebraic variety in $\mathbb{F}^m$ containing the image $\Phi(\mathbf{V})$. In other words:

$$\overline{\mathsf{image}}\,(\Phi) \quad=\quad \bigcap_{\substack{\mathbf{W}\subset\mathbb{F}^m\\ \mathbf{W}\text{ a variety}\\ \Phi(\mathbf{V})\subset\mathbf{W}}} \mathbf{W}.$$

(this is called the *Zariski closure* of $\Phi(\mathbf{V})$). It follows from Lemma 222 on page 174 that

$$\overline{\mathsf{image}}\,(\Phi) \quad = \quad \mathbb{V}\left(\mathcal{A}^{m}\,(\Phi(\mathbf{V}))\right).$$

We say that $\Phi : \mathbf{V} \longrightarrow \mathbf{W}$ is a **(geometric) epimorphism** if $\overline{\mathsf{image}}\,(\mathbf{V}) = \mathbf{W}$. Note that this does *not* necessarily mean Φ is surjective onto $\mathbf{W}$.

Example 146: Let $\mathbf{V} = \mathbb{R} = \mathbf{W}$, and let $\Phi(x) = x^2$. Then $\Phi(\mathbb{R}) = \{r \in \mathbb{R}\ ;\ r \geq 0\}$ is the positive real line. Thus, the only algebraic variety in $\mathbb{R}$ which contains $\Phi(\mathbb{R})$ is $\mathbb{R}$ itself. Hence, $\overline{\mathsf{image}}\,(\Phi) = \mathbb{R}$, so Φ is a geometric epimorphism from $\mathbb{R}$ to itself, even though Φ is not surjective.

Note that, even if the morphism Φ is *invertible*, the inverse function Φ^{-1} is not necessarily a morphism:

Example 147: Let $\mathbf{V} = \mathbb{R} = \mathbf{W}$, and define $\Phi : \mathbb{R} \longrightarrow \mathbb{R}$ by $\Phi(x) = x^3$. Thus, Φ is bijective, therefore invertible. However, $\Phi^{-1}(x) = x^{1/3}$ is *not* a geometric morphism (because it is not a polynomial).

We say that $\Phi : \mathbf{V} \longrightarrow \mathbf{W}$ is an **(geometric) monomorphism** if:

1. $\Phi : \mathbf{V} \longrightarrow \mathbf{W}$ is injective, and
2. If $\mathbf{U} = \Phi(\mathbf{V}) \subset \mathbf{W}$, then the inverse map $\Phi^{-1} : \mathbf{U} \longrightarrow \mathbf{V}$ is the restriction of some polynomial function. —ie. $\Phi^{-1}(\mathbf{u}) = \Big(\psi_1(\mathbf{u}), \psi_2(\mathbf{u}), \ldots, \psi_m(\mathbf{u})\Big)$ for all $\mathbf{u} \in \mathbf{U}$, where $\psi_1, \psi_2, \ldots, \psi_n \in \mathbb{F}[x_1, ..., x_m]$ are polynomial functions on $\mathbb{F}^m$.

We say that Φ is a **(geometric) isomorphism** if:

1. $\Phi : \mathbf{V} \longrightarrow \mathbf{W}$ is bijective, and
2. The inverse map $\Phi^{-1} : \mathbf{W} \longrightarrow \mathbf{V}$ is also a morphism —ie. $\Phi^{-1}(\mathbf{w}) = \Big(\psi_1(\mathbf{w}), \psi_2(\mathbf{w}), \ldots, \psi_m(\mathbf{w})\Big)$, where $\psi_1, \psi_2, \ldots, \psi_n \in \mathcal{C}oord\,(\mathbf{W})$.

We then say that $\mathbf{V}$ and $\mathbf{W}$ are **(geometrically) isomorphic**.

Example 148:

(a) Let $\mathbb{S}^1 \subset \mathbb{R}^2$ be the unit circle, as in Example $\langle 135 \rangle$**(a)**. Let $\mathbf{E} \subset \mathbb{R}^2$ be the ellipse with minor axis 2 and major axis 5, as in Example $\langle 135 \rangle$**(b)**. Define $\Phi : \mathbb{S}^1 \longrightarrow \mathbf{E}$ by $\Phi(x, y) = (5x, 2y)$ (in other words, $\phi_1(x, y) = 5x$ and $\phi_2(x, y) = 2y$). Then Φ is a geometric isomorphism, with inverse $\Phi^{-1}(x, y) = (\frac{x}{5}, \frac{y}{2})$.

(b) Let $\mathbb{SO}^2[\mathbb{R}] \subset \mathcal{M}_2(\mathbb{R})$ be the **special orthogonal group** from Example ⟨136⟩**(e)**, and let $\mathbb{S}^1 \subset \mathbb{R}^2$ be the circle. Define $\Phi : \mathbb{S}^1 \longrightarrow \mathbb{SO}^2[\mathbb{R}]$ by

$$\Phi(x, y) \quad = \quad \begin{bmatrix} x & y \\ -y & x \end{bmatrix}.$$

Then Φ is a geometric isomorphism (**Exercise 118**).

Proposition 149 *Let* $\mathbf{V}$ *and* $\mathbf{W}$ *be algebraic varieties, and let* $\phi : \mathbf{V} \longrightarrow \mathbf{W}$ *be a morphism.*

(a) *If* $f : \mathbf{W} \longrightarrow \mathbb{F}$ *is a coordinate function, then* $f \circ \phi : \mathbf{V} \longrightarrow \mathbb{F}$ *is also a coordinate function.*

(b) *Define* $\Phi : \mathcal{C}oord(\mathbf{W}) \longrightarrow \mathcal{C}oord(\mathbf{V})$ *by* $\Phi(f) = f \circ \phi$. *Then:*

1. Φ *is a ring homomorphism, and* $\Phi(\mathbb{1}_{\mathbf{Y}}) = \mathbb{1}_{\mathbf{X}}$.
2. $\Big($ ϕ *is a geometric monomorphism* $\Big) \iff \Big($ Φ *is a ring epimorphism* $\Big)$.
3. $\Big($ ϕ *is a geometric epimorphism* $\Big) \iff \Big($ Φ *is a ring monomorphism* $\Big)$.
4. $\Big($ ϕ *is a geometric isomorphism* $\Big) \iff \Big($ Φ *is a ring isomorphism* $\Big)$.

(c) **(Functorial Property)** *Suppose* $\mathbf{U}$, $\mathbf{V}$ *and* $\mathbf{W}$ *are algebraic varieties, and that* $\phi : \mathbf{U} \longrightarrow \mathbf{V}$ *and* $\psi : \mathbf{V} \longrightarrow \mathbf{W}$ *are morphisms. Let* $\gamma = \psi \circ \phi : \mathbf{U} \longrightarrow \mathbf{W}$, *so that diagram* **(A)** *below commutes.*

Define homomorphisms $\Phi : \mathcal{C}oord(\mathbf{V}) \longrightarrow \mathcal{C}oord(\mathbf{U})$, $\quad \Psi : \mathcal{C}oord(\mathbf{W}) \longrightarrow \mathcal{C}oord(\mathbf{V})$, $\quad$ *and* $\Gamma : \mathcal{C}oord(\mathbf{W}) \longrightarrow \mathcal{C}oord(\mathbf{U})$ *as in part* **(b)**. *Then* $\Gamma = \Phi \circ \Psi$. *In other words, diagram* **(B)** *commutes:*

$$\begin{tikzcd} \mathbf{U} \arrow[r, "\phi"] \arrow[dr, "\gamma"'] & \mathbf{V} \arrow[d, "\psi"] \\ \text{(A)} & \mathbf{W} \end{tikzcd} \quad \overset{}{=\!=\!=\!=\!=\!\Longrightarrow} \quad \begin{tikzcd} \mathcal{C}oord(\mathbf{U}) & \mathcal{C}oord(\mathbf{V}) \arrow[l, "\Phi"'] \\ \text{(B)} & \mathcal{C}oord(\mathbf{W}) \arrow[u, "\Psi"'] \arrow[ul, "\Gamma"] \end{tikzcd}$$

Proof: **(a)** Suppose $\mathbf{V} \subset \mathbb{F}^n$ and $\mathbf{W} \subset \mathbb{F}^m$. The construction is illustrated by the following diagram:

$$\begin{tikzcd} \mathbf{V} \arrow[r, "\phi"] \arrow[dr, "\Phi(f) = f \circ \phi"'] & \mathbf{W} \arrow[d, "f"] \\ & \mathbb{F} \end{tikzcd}$$

By hypothesis, f is a coordinate function —that is, $f = p_{|\mathbf{W}}$, where $p \in \mathbb{F}[x_1, ..., x_m]$ is some polynomial. Likewise, $\phi(\mathbf{v}) = \Big(\varphi_1(\mathbf{v}), \ldots, \varphi_m(\mathbf{v})\Big)$, where $\varphi_1, \varphi_2, \ldots, \varphi_m \in \mathbb{F}[x_1,..., x_n]$ are polynomials. Thus, $f \circ \phi(\mathbf{v}) \;=\; p\Big(\varphi_1(\mathbf{v}), \ldots, \varphi_m(\mathbf{v})\Big)$. It is **<u>Exercise 119</u>** to check that $p\Big(\varphi_1(\mathbf{v}), \ldots, \varphi_m(\mathbf{v})\Big)$ is a polynomial.

(b1) **<u>Exercise 120</u>** .

(b2) "$\Longrightarrow$" Suppose ϕ is a monomorphism. Let $\mathbf{U} = \phi(\mathbf{V}) \subset \mathbf{W}$. Thus, the map $\phi^{-1} : \mathbf{U} \longrightarrow \mathbf{V}$ is well-defined, and $\phi^{-1} = q_{|\mathbf{U}}$ for some polynomial function $q : \mathbb{F}^m \longrightarrow \mathbb{F}^n$.

We want to show that Φ is surjective; hence, given any $f \in \mathcal{C}\mathrm{oord}(\mathbf{V})$, we want some $g \in \mathcal{C}\mathrm{oord}(\mathbf{W})$ so that $\Phi(g) = f$.

Suppose $f = p_{|\mathbf{W}}$, where $p \in \mathbb{F}[x_1, ..., x_m]$. Let $g = (p \circ q)_{|\mathbf{W}}$. Then g is a coordinate function on $\mathbf{W}$ (because $p \circ q$ is a polynomial), and, for any $\mathbf{v} \in \mathbf{V}$, we have:

$$\Phi(g)(\mathbf{v}) \;=\; g \circ \phi(\mathbf{v}) \;=\; p \circ q\,(\phi(\mathbf{v})) \;\underset{(*)}{=}\; p \circ \phi^{-1}(\phi(\mathbf{v})) \;=\; p(\mathbf{v}) \;\underset{(\dagger)}{=}\; f(\mathbf{v}).$$

Here, $(*)$ is because $\phi(\mathbf{v}) \in \mathbf{U}$ and $q_{|\mathbf{U}} = \phi^{-1}$; $(\dagger)$ is because $p_{|\mathbf{V}} = f$.

(b2) "$\Longleftarrow$" Suppose that $\Phi : \mathcal{C}\mathrm{oord}(\mathbf{W}) \longrightarrow \mathcal{C}\mathrm{oord}(\mathbf{V})$ is surjective.

Claim 1: *ϕ is injective.*

Proof: Suppose not; then there exist points $\mathbf{v}, \mathbf{v}' \in \mathbf{V}$ so that $\phi(\mathbf{v}) = \mathbf{w} = \phi(\mathbf{v}')$. Now, let $f \in \mathcal{C}\mathrm{oord}(\mathbf{V})$ be any polynomial such that $f(\mathbf{v}) \neq f(\mathbf{v}')$. For example, if $\mathbf{v} = (v_1, \ldots, v_n)$ and $\mathbf{v}' = (v_1', \ldots, v_n')$, then $\mathbf{v}$ and $\mathbf{v}'$ must differ in some coordinate —say $v_1 \neq v_1'$. Then let $f(x_1, \ldots, x_n) = x_1 - v_1$. Then $f(\mathbf{v}) = 0 \neq f(\mathbf{v}')$.

I claim there is no function $g \in \mathcal{C}\mathrm{oord}(\mathbf{W})$ such that $\Phi(g) = f$. To see this, observe that, for any $g \in \mathcal{C}\mathrm{oord}(\mathbf{W})$,

$$\Phi(g)(\mathbf{v}) \;=\; g\Big(\phi(\mathbf{v})\Big) \;=\; g(\mathbf{w}) \;=\; g\Big(\phi(\mathbf{v}')\Big) \;=\; \Phi(g)(\mathbf{v}').$$

Hence, $\Phi(g)(\mathbf{v}) = \Phi(g)(\mathbf{v}')$, whereas $f(\mathbf{v}) \neq f(\mathbf{v}')$. Thus, $\Phi(g)$ cannot equal f.

Thus, f is *not* in the image of Φ, contradicting surjectivity. By contradiction, ϕ must be injective. .... □ `[Claim 1]`

Now, let $\mathbf{U} = \phi(\mathbf{V})$, and let $\phi^{-1} : \mathbf{U} \longrightarrow \mathbf{V}$ be the inverse function.

Claim 2: *$\phi^{-1} = q_{|\mathbf{U}}$ for some polynomial function $q : \mathbb{F}^m \longrightarrow \mathbb{F}^n$.*

Proof: Let $\pi_1 : \mathbb{F}^n \longrightarrow \mathbb{F}$ be projection into the first coordinate; ie. $\pi_1(x_1, \ldots, x_n) = x_1$. This is clearly a polynomial, so $(\pi_1)_{|\mathbf{V}}$ is a coordinate function. Thus, since Φ is surjective, there is some element $g_1 \in \mathcal{C}\!\mathit{ord}(\mathbf{W})$ so that $\Phi(g_1) = (\pi_1)_{|\mathbf{V}}$. But g_1 is the restriction of some polynomial $q_1 \in \mathbb{F}[x_1, ..., x_m]$; hence, we have:

$$q_1 \circ \phi \quad = \quad g_1 \circ \phi \quad = \quad (\pi_1)_{|\mathbf{V}}.$$

Likewise, if $\pi_k : \mathbb{F}^n \longrightarrow \mathbb{F}$ is projection into the kth coordinate, then there is some polynomial $q_k \in \mathbb{F}[x_1, ..., x_m]$ so that

$$q_k \circ \phi \quad = \quad (\pi_k)_{|\mathbf{V}}. \tag{9.1}$$

I claim $\phi^{-1} = (q_1, \ldots, q_n)$. To see this, observe that, for any $\mathbf{v} = (v_1, \ldots, v_n) \in \mathbf{V}$,

$$\begin{aligned}(q_1, \ldots, q_n) \circ \phi(\mathbf{v}) \quad &= \quad \Big(q_1 \circ \phi(\mathbf{v}), \ldots, q_n \circ \phi(\mathbf{v})\Big) \quad \overline{\overline{\scriptstyle \text{by}(9.1)}} \quad \Big(\pi_1(\mathbf{v}), \ldots, \pi_n(\mathbf{v})\Big) \\ &= \quad (v_1, \ldots, v_n) \quad = \quad \mathbf{v}.\end{aligned}$$

In other words, $(q_1, \ldots, q_n) \circ \phi = \mathbf{Id}$ on $\mathbf{V}$. .... □ `[Claim 2]`

(b3) "$\Longrightarrow$" Suppose ϕ is an geometric epimorphism (ie. $\overline{\mathsf{image}}(\phi) = \mathbf{W}$); we claim Φ is a ring monomorphism. To prove this, it suffices to show that $\ker(\Phi) = \{0\}$. So, suppose $g \in \mathcal{C}\!\mathit{ord}(\mathbf{W})$ and $\Phi(g) = 0$; we want to show that $g \equiv 0$ —in other words, that $g(\mathbf{w}) = 0$ for all $\mathbf{w} \in \mathbf{W}$.

Let $\mathbb{V}(g) = \{\mathbf{w} \in \mathbf{W} \ ; \ g(\mathbf{w}) = 0\}$ be the algebraic variety induced by g.

Claim 3: $\phi(\mathbf{V}) \subset \mathbb{V}(g)$.

Proof: If $\mathbf{w} \in \phi(\mathbf{V})$, then there is some $\mathbf{v} \in \mathbf{V}$ so that $\phi(\mathbf{v}) = \mathbf{w}$. But then $g(\mathbf{w}) = g(\phi(\mathbf{v})) = \Phi(g)(\mathbf{v}) = 0$ (because $\Phi(g) = 0$). This holds for all $\mathbf{w} \in \phi(\mathbf{V})$, so $\phi(\mathbf{V}) \subset \mathbb{V}(g)$. .... □ `[Claim 3]`

Thus, $\overline{\mathsf{image}}(\phi) \subset \mathbb{V}(g)$. But by hypothesis, $\overline{\mathsf{image}}(\phi) = \mathbf{W}$; hence $\mathbb{V}(g) = \mathbf{W}$, which means $g \equiv 0$.

(b3) "$\Longleftarrow$" Suppose ϕ was *not* a geometric epimorphism. We claim Φ is *not* a monomorphism —in other words, $\ker(\Phi) \neq \{0\}$.

Since ϕ is not an epimorphism, it follows that $\mathbf{U} = \overline{\mathsf{image}}(\phi)$ is a proper subset of $\mathbf{W}$.

Let $\mathcal{A}^m(\mathbf{W}) = \Big\{p \in \mathbb{F}[x_1, ..., x_m] \ ; \ p_{|\mathbf{W}} \equiv 0\Big\}$ be the annihilator of $\mathbf{W}$. It follows:

Claim 4: $\mathcal{A}^m(\mathbf{W}) \subsetneq \mathcal{A}^m(\mathbf{U})$.

Proof: <u>**Exercise 121**</u> Hint: Apply Lemmas 220 and 222 on page 174 ..... □ `[Claim 4]`

So, let $p \in \mathbb{F}[x_1, ..., x_m]$ be a polynomial so that $p \in \mathcal{A}^m(\mathbf{U})$ but $p \notin \mathcal{A}^m(\mathbf{W})$. Hence, if $g = p_{|\mathbf{W}}$, then g is a coordinate function, and $g \neq 0$.

Claim 5: $\Phi(g) = 0$.

Proof: Let $\mathbf{v} \in \mathbf{V}$, and let $\mathbf{u} = \phi(\mathbf{v}) \in \mathbf{U}$. Then $\Phi(g)(\mathbf{v}) = g \circ \phi(\mathbf{v}) = g(\mathbf{u}) = 0$, because $g \in \mathcal{A}\mathit{nn}(\mathbf{U})$. .... □ `[Claim 5]`

Hence, $g \in \ker(\Phi)$ is nonzero, so Φ cannot be a monomorphism.

(b4) Combine **(b2)** and **(b3)**.

(c) **Exercise 122** . ______________________________ □

Corollary 150 *Let* $\mathbf{V}$ *and* $\mathbf{W}$ *be algebraic varieties. Then*

$$\Big(\mathbf{V} \textit{ and } \mathbf{W} \textit{ are geometrically isomorphic} \Big) \iff \Big(\mathcal{C}\mathit{oord}(\mathbf{V}) \textit{ and } \mathcal{C}\mathit{oord}(\mathbf{W}) \textit{ are isomorphic as rings} \Big).$$

Thus, *all geometric information about a variety is encoded in its coordinate ring.*

Recall that a ring $\mathcal{R}$ is *perfect* if $\mathcal{R}$ has no nilpotent elements —ie. $\sqrt[*]{0_{\mathcal{R}}} = \{0\}$ (see §10.7.3). Recall from Corollary 142 on page 123 that every perfect quotient of $\mathbb{C}[x_1,..., x_n]$ is the coordinate ring of some algebraic variety $\mathbf{V} \subset \mathbb{C}^n$. It follows:

Corollary 151 *There is a natural bijective correspondence:*

$$\left\{ \begin{array}{c} \textit{(Geometric) Isomorphism classes} \\ \textit{of algebraic varieties in } \mathbb{C}^n \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{l} \textit{(Algebraic) Isomorphism classes of} \\ \textit{perfect quotient rings of } \mathbb{C}[x_1,..., x_n] \end{array} \right\}.$$

So, instead of defining complex algebraic geometry as 'the study of complex algebraic varieties', we could define it as the study of 'the study of perfect quotients of complex polynomial rings, with a geometric interpretation.'

9.4 Irreducible Varieties

Prerequisites: §9.1, §10.5

Let $\mathbb{F}$ be a field, and let $\mathbf{V} \subset \mathbb{F}^n$ be an algebraic variety. We say $\mathbf{V}$ is **irreducible** if $\mathbf{V}$ cannot be written as a union of two smaller varieties. In other words, there are no varieties $\mathbf{U}$ and $\mathbf{W}$ so that

$$\mathbf{V} = \mathbf{U} \cup \mathbf{W}.$$

Example 152:

(a) Let $\mathbf{c} \in \mathbb{C}^n$ be a point. Then the singleton variety $\mathbf{V} = \{\mathbf{c}\}$ is irreducible.

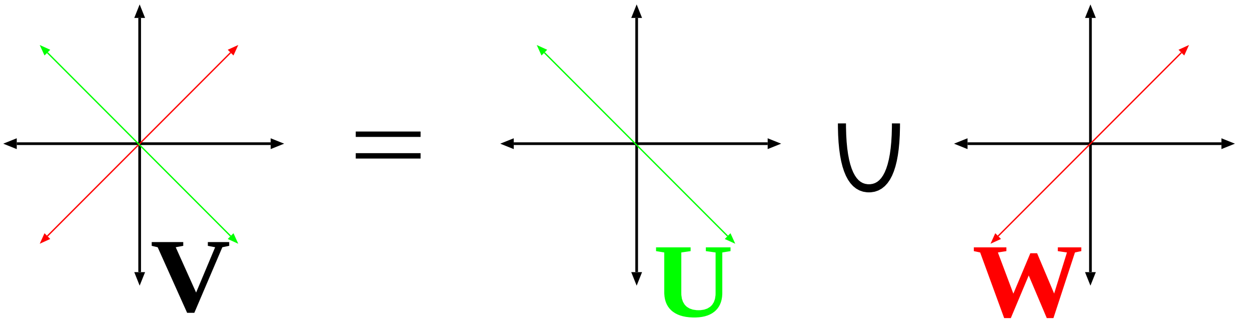


Figure 9.4: $\mathbf{V} = \mathbf{U} \cup \mathbf{W}$

(b) As in Example ⟨135⟩**(f)**, let $\mathbf{V} = \mathbb{V}(p)$, where $p(x,y) = x^2 - y^2 = (x+y)\cdot(x-y)$. Then $\mathbf{V} = \mathbf{U} \cup \mathbf{W}$, where

$$\begin{array}{rrcccl} & \mathbf{U} & = & \mathbb{V}(x+y) & = & \{(x,y) \in \mathbb{R}^2 \; ; \; x = -y\}, \\ \text{and} & \mathbf{W} & = & \mathbb{V}(x-y) & = & \{(x,y) \in \mathbb{R}^2 \; ; \; x = y\}. \qquad \text{(See Figure 9.4)} \end{array}$$

Hence, $\mathbf{V}$ is *not* irreducible.

(c) More generally, suppose that $p(x_1,\ldots,x_n) \in \mathbb{F}[x_1,...,x_n]$ is a polynomial which factors:

$$p(x_1,\ldots,x_n) \quad = \quad q(x_1,\ldots,x_n)\cdot r(x_1,\ldots,x_n)$$

Then $\mathbb{V}(p) \;=\; \mathbb{V}(q) \cup \mathbb{V}(r)$, so if $\mathbb{V}(q)$ and $\mathbb{V}(r)$ are both nonempty, then $\mathbb{V}(p)$ is *not* irreducible.

It appears from Examples ⟨152b⟩ and ⟨152c⟩ that irreducibility of a *variety* is related to irreducibility of the *polynomials* which generate it. The precise formulation of this is as follows:

Proposition 153 *The following are equivalent:*

(a) $\mathbf{V}$ *is irreducible.*

(b) *If* $\mathbf{U}$ *and* $\mathbf{W}$ *are varieties such that* $\mathbf{V} \;\subset\; \mathbf{U} \cup \mathbf{W}$, *then either* $\mathbf{V} \subset \mathbf{U}$ *or* $\mathbf{V} \subset \mathbf{W}$.

(c) $\mathcal{A}nn(\mathbf{V})$ *is a prime ideal in* $\mathbb{F}[x_1,...,x_n]$.

(d) $\mathcal{C}oord(\mathbf{V})$ *is an integral domain.*

In particular, if $\mathbf{V} = \mathbb{V}(p)$ *for some* $p \in \mathbb{F}[x_1,...,x_n]$, *then*

$$\Big(\ \mathbf{V} \text{ is irreducible}\ \Big) \iff \Big(\ p \text{ is irreducible}\ \Big).$$

Proof: "(**a**)$\Longrightarrow$(**b**)" Suppose $\mathbf{V} \subset \mathbf{U} \cup \mathbf{W}$. Recall that the intersection of two varieties is a variety. Thus, $\mathbf{U}' = \mathbf{V} \cap \mathbf{U}$ is a variety, and so is $\mathbf{W}' = \mathbf{V} \cap \mathbf{W}$. Not that $\mathbf{V} = \mathbf{U}' \cup \mathbf{W}'$. But $\mathbf{V}$ is irreducible, so either $\mathbf{U}' = \mathbf{V}$ or $\mathbf{W}' = \mathbf{V}$; in other words, either $\mathbf{V} \subset \mathbf{U}$ or $\mathbf{V} \subset \mathbf{W}$.

"(**b**)$\Longrightarrow$(**a**)" Suppose $\mathbf{V} = \mathbf{U} \cup \mathbf{W}$; then $\mathbf{V} \subset \mathbf{U} \cup \mathbf{W}$, so **(b)** implies that either $\mathbf{V} \subset \mathbf{U}$ or $\mathbf{V} \subset \mathbf{W}$. Suppose that $\mathbf{V} \subset \mathbf{U}$. But $\mathbf{V} = \mathbf{U} \cup \mathbf{W}$, so clearly $\mathbf{U} \subset \mathbf{V}$. Hence, $\mathbf{U} = \mathbf{V}$. (Likewise, if $\mathbf{V} \subset \mathbf{W}$, then $\mathbf{V} = \mathbf{W}$.)

We conclude that $\mathbf{V}$ cannot be written as a union of two subvarieties; hence it is irreducible.

"(**b**)$\Longrightarrow$(**c**)" Suppose $f, g \in \mathbb{F}[x_1, \ldots, x_n]$ are polynomials such that $f \cdot g \in \mathcal{A}nn(\mathbf{V})$. I claim that either $f \in \mathcal{A}nn(\mathbf{V})$ or $g \in \mathcal{A}nn(\mathbf{V})$. To see this observe that

$$\begin{aligned} \Big((f \cdot g) \in \mathcal{A}nn(\mathbf{V}) \Big) \quad &\Longleftrightarrow \quad \Big(\mathbf{V} \subset \mathbb{V}(f \cdot g) = \mathbb{V}(f) \cup \mathbb{V}(g). \Big) \\ &=_{(\mathrm{b})}\Rightarrow \quad \Big(\text{Either } \mathbf{V} \subset \mathbb{V}(f) \text{ or } \mathbf{V} \subset \mathbb{V}(g). \Big) \\ &\Longleftrightarrow \quad \Big(\text{Either } f \in \mathcal{A}nn(\mathbf{V}) \text{ or } g \in \mathcal{A}nn(\mathbf{V}) \Big). \end{aligned}$$

"(**c**)$\Longrightarrow$(**a**)" Suppose $\mathbf{V}$ is *not* irreducible; we'll show that $\mathcal{A}nn(\mathbf{V})$ cannot be prime.

Suppose $\mathbf{V} = \mathbf{U} \cup \mathbf{W}$, but $\mathbf{V} \neq \mathbf{U}$ and $\mathbf{V} \neq \mathbf{W}$. Thus, $\mathbf{U} \subsetneq \mathbf{V}$, so Lemma 220 on page 173 implies that $\mathcal{A}nn(\mathbf{V}) \subsetneq \mathcal{A}nn(\mathbf{U})$. So, find $f \in \mathcal{A}nn(\mathbf{U})$ with $f \notin \mathcal{A}nn(\mathbf{V})$.

Likewise $\mathbf{W} \subsetneq \mathbf{V}$, so $\mathcal{A}nn(\mathbf{V}) \subsetneq \mathcal{A}nn(\mathbf{W})$, so find $g \in \mathcal{A}nn(\mathbf{U})$ with $g \notin \mathcal{A}nn(\mathbf{V})$. I claim that $(f \cdot g) \in \mathcal{A}nn(\mathbf{V})$. To see this, let $v \in \mathbf{V}$. Then either $v \in \mathbf{U}$ (in which case $f(v) = 0$) or $v \in \mathbf{W}$ (in which case $g(v) = 0$). Either way, $(f \cdot g)(v) = 0$.

Thus, $f \notin \mathcal{A}nn(\mathbf{V})$ and $g \notin \mathcal{A}nn(\mathbf{V})$, but $(f \cdot g) \in \mathcal{A}nn(\mathbf{V})$. Hence, $\mathcal{A}nn(\mathbf{V})$ is not a prime ideal.

"(**c**) $\Longleftrightarrow$ (**d**)" **Exercise 123** Use Proposition 200 on page 165 and Proposition 140 on page 122. □

Corollary 151 on page 129 immediately implies:

Corollary 154 *There is a natural bijective correspondence:*

$$\left\{ \begin{array}{c} \textit{Isomorphism classes of irreducible} \\ \textit{algebraic varieties in } \mathbb{C}^n \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{c} \textit{Isomorphism classes of integral} \\ \textit{domain quotients of } \mathbb{C}[x_1, \ldots, x_n] \end{array} \right\}.$$

Irreducible varieties are the basic 'building blocks' out of which other varieties are made:

Proposition 155 *Any algebraic variety is a union of irreducible varieties.*

To be precise: let $\mathbf{V} \subset \mathbb{F}^n$ *be an algebraic variety with annihilator* $\mathcal{A}nn(\mathbf{V})$*. Consider the set of all prime ideals which contain* $\mathcal{A}nn(\mathbf{V})$*:*

$$\mathfrak{P} \quad = \quad \{ \mathcal{P} \triangleleft \mathbb{F}[x_1, \ldots, x_n] \; ; \; \mathcal{P} \textit{ prime, and } \mathcal{A}nn(\mathbf{V}) \subset \mathcal{P} \}.$$

For any $\mathcal{P} \in \mathfrak{P}$*, the corresponding variety* $\mathbb{V}(\mathcal{P})$ *is irreducible, and* $\mathbf{V} \;=\; \bigcup_{\mathcal{P} \in \mathfrak{P}} \mathbb{V}(\mathcal{P})$.

Proof: **Exercise 124** ______________________________ □

Because of this, algebraic geometers normally restrict their attention to irreducible varieties. Indeed, in many texts, (eg. [3]), an ‘algebraic variety’ is *defined* to be irreducible. .

Chapter 10

Ideal Theory

10.1 Principal Ideals and PIDs

Prerequisites: §??

Let $\mathcal{R}$ be a *commutative* ring and let $p \in \mathcal{R}$. The **principal ideal** generated by p is the set

$$p\mathcal{R} \quad = \quad \{pr \ ; \ r \in \mathcal{R}\}. \qquad \text{(see Example } \langle 117\text{k}\rangle\text{)} \tag{10.1}$$

We normally denote this ideal by '(p)'.

Example 156:

(a) If $\mathcal{R} = \mathbb{Z}$, then $(2) = 2\mathbb{Z} = \{2z \ ; \ z \in \mathbb{Z}\}$, the ideal of *even numbers* from Example ⟨117a⟩.

(b) If $\mathcal{R} = \mathbb{R}[x]$, then $(x) = x\mathbb{R}[x] = \{x \cdot q(x) \ ; \ q \in \mathcal{R}[x]\}$, the ideal from Example ⟨117e⟩. ◊

If $\mathcal{R}$ is a *noncommutative* ring, then in general, $p\mathcal{R}$ is not an ideal. Instead, we define the **principal ideal** generated by p to be

$$\mathcal{R}p\mathcal{R} \quad = \quad \{r_1 p s_1 + r_2 p s_2 + \ldots + r_n p s_n \ ; \ n \in \mathbb{N}, r_1, \ldots, r_n, s_1, \ldots, s_n \in \mathcal{R}\}. \tag{10.2}$$

(**Exercise 125** Verify this is an ideal.)

Again, we denote this ideal by '(p)'. Observe that, if $\mathcal{R}$ is commutative, then definition (10.2) agrees with definition (10.1) (**Exercise 126**).

A **principal ideal domain (PID)** is a commutative ring where *all ideals are principal ideals.*

Proposition 157 *The ring of integers is a principal ideal domain. To be precise:*

If $\mathcal{I} \lhd \mathbb{Z}$ is any ideal, then $\mathcal{I} = n\mathbb{Z}$, where n is the minimal positive element in $\mathcal{I}$.

Proof: Let $\mathcal{I} \lhd \mathbb{Z}$ be an ideal, and let n be the *smallest positive integer* in $\mathcal{I}$. That is: $n = \min(\mathcal{I} \cap \mathbb{N})$.

Our goal is to show that $\mathcal{I} = n\mathbb{Z}$.

Claim 1: $n\mathbb{Z} \subset \mathcal{I}$.

Proof: Since $n \in \mathcal{I}$ and $\mathcal{I}$ is an ideal, it follows that $nz \in \mathcal{I}$ for all $z \in \mathbb{Z}$. In other words, any element of $n\mathbb{Z}$ is in $\mathcal{I}$. .... □ `[Claim 1]`

Claim 2: $\mathcal{I} \subset n\mathbb{Z}$.

Proof: Suppose not. Find some $i \in \mathcal{I}$ with $i \notin n\mathbb{Z}$. Without loss of generality, we assume $i > 0$ (because if $i \in \mathcal{I}$, then $-i$ is also in $\mathcal{I}$). Since n is the minimal positive element of $\mathcal{I}$, we must have $n < i$. Apply the `Division algorithm` to write $i = qn + r$, where $0 < r < n$. By hypothesis, $i \notin n\mathbb{Z}$, so n doesn't divide i, so we must have $r \neq 0$. But then:

- $i \in \mathcal{I}$;
- $qn \in n\mathbb{Z} \subset \mathcal{I}$ thus, $qn \in \mathcal{I}$.
- $r = i - qn$; thus, $r \in \mathcal{I}$.

But $r < n$, and this is a contradiction, because n is the *minimal* positive element in $\mathcal{I}$. By contradiction, no such i can exist; hence $\mathcal{I} \subset (n)$. .... □ `[Claim 2]`

Claims 1 and 2 together imply that $\mathcal{I} = n\mathbb{Z}$. ________________________________ □

Principal Ideals in Polynomial Rings: The proof of Proposition 157 generalizes readily to the ring $\mathbb{R}[x]$ of polynomials. The key concepts used in the proof of Proposition 157 were

1. The existence of a *minimal element* in $\mathcal{I}$.

2. The use of the `Division Algorithm`.

We must generalize these ideas to polynomials. Recall that the *degree* of a polynomial $p(x)$ is the highest exponent appearing in $p(x)$ with nonzero coefficient. For example $\mathsf{degree}\,(()\,5x^3 - 7x^2 + 8x + 6) = 3$. In general, If $p(x) = p_n x^n + \ldots + p_1 x^1 + p_0$, (where $p_n \neq 0$), then $\mathsf{degree}\,(p) = n$.

Recall also the the `Polynomial Long Division` algorithm. Given any polynomials $p(x)$ and $P(x)$, where $\mathsf{degree}\,(P) \geq \mathsf{degree}\,(p)$, we can 'divide' $P(x)$ by $p(x)$ to get:

$$\frac{P(x)}{p(x)} \quad = \quad q(x) \; + \; \frac{r(x)}{p(x)}. \tag{10.3}$$

where $\mathsf{degree}\,(r) < \mathsf{degree}\,(p)$. If we multiply both sides of equation (10.3) by $p(x)$, we get:

$$P(x) \quad = \quad q(x) \cdot p(x) \; + \; r(x).$$

If the 'remainder' polynomial $r(x)$ equals zero, then we say that $p(x)$ **divides** $P(x)$.

Proposition 158 *The ring $\mathbb{R}[x]$ is a principal ideal domain. To be precise:*

If $\mathcal{I} \lhd \mathbb{R}[x]$ is any ideal, then $\mathcal{I} = (p)$, where p is a polynomial of minimal degree in $\mathcal{I}$.

Proof: Let $\mathcal{I} \lhd \mathbb{R}[x]$ be an ideal. Let p be a polynomial of minimal degree in $\mathcal{I}$. That is:

$$\mathsf{degree}\,(p) \quad = \quad \min\{\mathsf{degree}\,(q)\ ;\ q \in \mathcal{I}\}.$$

Our goal is to show that $\mathcal{I} = (p)$.

Claim 1: $(p) \subset \mathcal{I}$.

Proof: Since $p \in \mathcal{I}$ and $\mathcal{I}$ is an ideal, it follows that $p \cdot q \in \mathcal{I}$ for all $q \in \mathbb{R}[x]$. Hence any element of (p) is also in $\mathcal{I}$. .... □ `[Claim 1]`

Claim 2: $\mathcal{I} \subset (p)$.

Proof: Suppose not. Find some $i \in \mathcal{I}$ with $i \notin (p)$. Since p has minimal degree in $\mathcal{I}$, we must have $\mathsf{degree}\,(p) \leq \mathsf{degree}\,(i)$. Now apply `Polynomial Long Division` to write $i = qp + r$, where $0 < \mathsf{degree}\,(r) < \mathsf{degree}\,(p)$. By hypothesis, $i \notin (p)$, so p doesn't divide i, so we must have $r \neq 0$. But then:

- $i \in \mathcal{I}$;
- $qp \in (p) \subset \mathcal{I}$; thus, $qp \in \mathcal{I}$.
- $r = i - qp$; thus, $r \in \mathcal{I}$.

But $\mathsf{degree}\,(r) < \mathsf{degree}\,(p)$, and this is a contradiction, because p has *minimal* degree in $\mathcal{I}$. By contradiction, no such i can exist; hence $\mathcal{I} \subset (p)$. .... □ `[Claim 2]`

Claims 1 and 2 together imply that $\mathcal{I} = (p)$. □

There is nothing special about the real numbers or about $\mathbb{R}[x]$. The notion of *polynomial division* generalizes to any field, and yields the following theorem:

Proposition 159 *Let $\mathbb{F}$ be any field. Then $\mathbb{F}[x]$ is a principal ideal domain. To be precise:*

If $\mathcal{I} \lhd \mathbb{F}[x]$ is any ideal, then $\mathcal{I} = (p)$, where p is a polynomial of minimal degree in $\mathcal{I}$.

Proof: **Exercise 127** Literally just change the symbol '$\mathbb{R}$' to '$\mathbb{F}$' throughout the preceeding argument □

Do not get the idea that *every* polynomial ring is a PID....

Example 160:

(a) Let $\mathcal{R} = \mathbb{Z}[x]$ (Example ⟨98c⟩), and let $\mathcal{I}$ be the polynomial generated by the elements 2 and x. That is,

$$\mathcal{I} \quad = \quad \{2p(x) + xq(x) \; ; \; p(x) \in \mathbb{Z}[x] \text{ and } q(x) \in \mathbb{Z}[x]\}.$$

We claim that $\mathcal{I}$ is a nonprincipal ideal.

Claim 1: *$\mathcal{I}$ is an ideal.*

Proof: <u>**Exercise 128**</u> .... □ [Claim 1]

Claim 2: *$\mathcal{I}$ contains the polynomials 2 and x.*

Proof: $2 = 2 \cdot 1 + x \cdot 0$, and $2 = 2 \cdot 0 + x \cdot 1$. .... □ [Claim 2]

Claim 3: $\mathcal{I} \neq \mathbb{Z}[x]$.

Proof: We will show that *any element in $\mathcal{I}$ must have an even constant coefficient.* Hence, $\mathcal{I}$ cannot be all of $\mathbb{Z}[x]$.

Suppose $i(x) \in \mathcal{I}$. Then $i(x) = 2p(x) + xq(x)$ for some $p(x), q(x) \in \mathbb{Z}[x]$.

$$\begin{array}{llcllcl} \text{If} & p(x) & = & p_n x^n + \ldots + p_1 x + p_0, & \text{then} & p(x) & = & 2p_n x^n + \ldots + 2p_1 x + 2p_0. \\ \text{If} & q(x) & = & q_m x^m + \ldots + q_1 x + q_0, & \text{then} & x \cdot q(x) & = & q_m x^{m+1} + \ldots + q_1 x^2 + q_0 x. \end{array}$$

Hence, if $i(x) = i_\ell x^\ell + \ldots + i_1 x + i_0$, then $i_0 = 2p_0$ is even. .... □ [Claim 3]

Claim 4: *$\mathcal{I}$ is* not *a principal ideal.*

Proof: Suppose $\mathcal{I} = (g)$ for some polynomial $g(x) \in \mathbb{Z}[x]$.

Claim 4.1: *g must be a constant polynomial —ie. an integer.*

Proof: Claim 2 says $2 \in \mathcal{I}$, so $g(x)$ must divide 2. In other words, $2 = g(x) \cdot q(x)$ for some polynomial $q(x)$. But then

$$0 \quad = \quad \mathsf{degree}\,(2) \quad = \quad \mathsf{degree}\,(p(x) \cdot q(x)) \quad = \quad \mathsf{degree}\,(p(x)) + \mathsf{degree}\,(q(x))$$

Hence, we must have $\mathsf{degree}\,(p(x)) = 0 = \mathsf{degree}\,(q(x))$. Hence $g(x)$ must be a constant integer, say g_0. .... □ [Claim 4]

Now, since g_0 divides 2, we must have either $g_0 = 1$ or $g_0 = 2$.

Claim 4.2: $g_0 \neq 1$.

Proof: Claim 3 says $\mathcal{I} \neq \mathbb{Z}[x]$. Hence, 1 is not an element of $\mathcal{I}$, so $g \neq 1$. □ [Claim 4.2]

Thus, $g_0 = 2$. But g_0 must divide x, and 2 does not divide x in the ring $\mathbb{Z}[x]$. Hence, no such g could exist. Thus, $\mathcal{I}$ is not a principal ideal. .... □ [Claim 4]

As a consequence: $\mathbb{Z}[x]$ *is not a principal ideal domain.*

(b) Let $\mathcal{R} = \mathbb{R}[x, y]$ (Example ⟨98e⟩) and let $\mathcal{I}$ be the polynomial generated by the elements x and y. That is,

$$\mathcal{I} \quad = \quad \{x \cdot p(x, y) + y \cdot q(x, y) \ ; \ p(x, y) \in \mathbb{R}[x, y] \ \text{ and } \ q(x, y) \in \mathbb{R}[x, y]\}.$$

We claim that $\mathcal{I}$ is a nonprincipal ideal.

Claim 1: *$\mathcal{I}$ is an ideal.*

Claim 2: *$\mathcal{I}$ contains the polynomials x and y.*

Claim 3: *$\mathcal{I} \neq \mathbb{R}[x, y]$. In particular, every polynomial of $\mathcal{I}$ has zero constant term.*

The proofs of these claims are **Exercise 129** .

Claim 4: *$\mathcal{I}$ is* not *a principal ideal.*

Proof: To see this, suppose $\mathcal{I} = (g)$ for some polynomial $g(x, y) \in \mathbb{R}[x, y]$. Then $g(x)$ must divide x, so it must not contain *any* terms in y, or any powers of x greater than x^1. Hence, $g(x) = g_1 x + g_0$ for some $g_1, g_0 \in \mathbb{R}$. If $g(x)$ is to divide x, then we must have $g_0 = 0$. Hence, $g(x, y) = g_1 \cdot x$.

But $g(x, y)$ must divide y, and x does not divide y in the ring $\mathbb{R}[x, y]$. Hence, no such g could exist. Thus, $\mathcal{I}$ is not a principal ideal. . . . . □ `[Claim 4]`

As a consequence: $\mathbb{R}[x, y]$ *is not a principal ideal domain.*

Example ⟨160b⟩ generalizes to the following

Proposition 161 *Let $\mathbb{F}$ be a field, and let $n \geq 2$. If $\mathbb{F}[x_1, x_2, \ldots, x_n]$ is the ring of polynomials over $\mathbb{F}$ in n variables, then $\mathbb{F}[x_1, x_2, \ldots, x_n]$ is not a principal ideal domain.*

Proof: **Exercise 130** □

10.2 Maximal Ideals

Prerequisites: §8.6, §10.1

Let $\mathcal{R}$ be a ring and $\mathcal{M} \lhd \mathcal{R}$ an ideal. We say $\mathcal{M}$ is a **maximal** ideal if there exists no ideal $\mathcal{I} \lhd \mathcal{R}$ such that $\mathcal{M} \subsetneq \mathcal{I} \subsetneq \mathcal{R}$ (see Figure 10.1)

Example 162: Let $\mathcal{R} = \mathbb{Z}$ and consider the principal ideal $\mathcal{M} = 2\mathbb{Z}$ (Example ⟨117a⟩). I claim this ideal is maximal. To see this, suppose there was $\mathcal{I} \lhd \mathbb{Z}$ such that $2\mathbb{Z} \subsetneq \mathcal{I} \subsetneq \mathbb{Z}$. Thus, $\mathcal{I}$ contains least one odd number o (since $\mathcal{I} \not\subset 2\mathbb{Z}$); say $o = 2n + 1$. But $\mathcal{I}$ also contains all even numbers (since $2\mathbb{Z} \subset \mathcal{I}$); in particular, $\mathcal{I}$ contains $2n$. Thus, $\mathcal{I}$ contains $(2n + 1) - (2n) = 1$. Hence $\mathcal{I} = \mathbb{Z}$, by Lemma 120 on page 111.

Hence, any ideal properly containing $2\mathbb{Z}$ must be all of $\mathbb{Z}$. Hence, $2\mathbb{Z}$ is maximal.

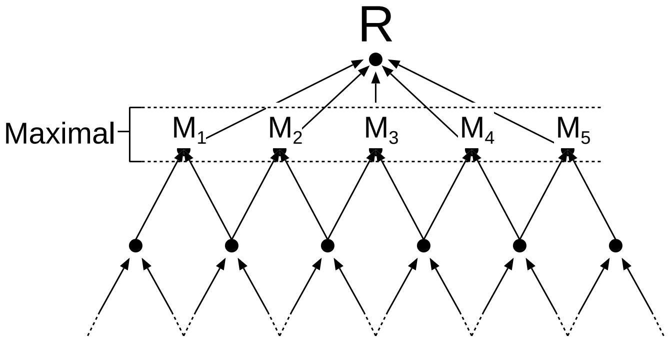


Figure 10.1: A schematic representation of the lattice of ideals of the ring $\mathcal{R}$. The *maximal* ideals form the 'top row' of the lattice, just below $\mathcal{R}$.

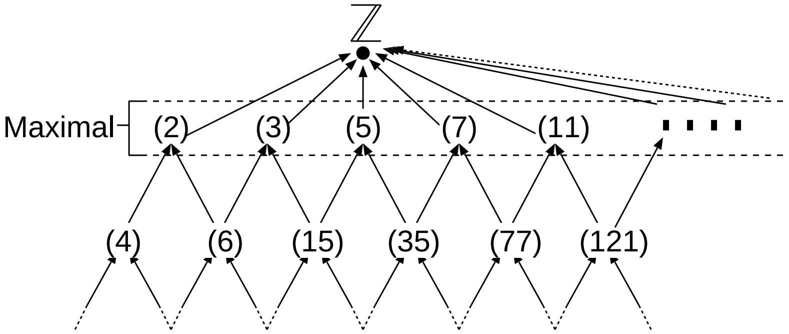


Figure 10.2: A schematic representation of the lattice of ideals of the ring $\mathbb{Z}$. The *maximal* ideals are the principal ideals (2), (3), (5), etc. generated by prime numbers.

Example ⟨162⟩ generalizes as follows:

Proposition 163 *The maximal ideals of $\mathbb{Z}$ are exactly the ideals $p\mathbb{Z}$, where p is prime.*

(see Figure 10.2.)

Proof: We want to show: $\Big(\ \mathcal{M}$ is a maximal ideal in $\mathbb{Z}\ \Big) \iff \Big(\ \mathcal{M} = p\mathbb{Z}$ for prime $p\ \Big)$.

Claim 1: *Let p be prime, and let $\mathcal{M} = p\mathbb{Z}$. Then $\mathcal{M}$ is a maximal ideal.*

Proof: Recall from Example ⟨117b⟩ that $\mathcal{M}$ is an ideal. We must show it is maximal.

To see this, suppose $\mathcal{I} \lhd \mathbb{Z}$ was an ideal such that $p\mathbb{Z} \subsetneq \mathcal{I} \subsetneq \mathbb{Z}$. Proposition 157 on page 133 says that $\mathcal{I} = (i)$, where i is the minimal positive element in $\mathcal{I}$. But if $\mathcal{M} \subset \mathcal{I}$, that means in particular that $p \in \mathcal{I}$ —ie. $p \in (i)$, so i must divide p. But p is prime, so....

either: $i = p$, in which case $(i) = (p)$ ie. $\mathcal{I} = \mathcal{M}$.

or: $i = 1$, in which case $(i) = \mathbb{Z}$ ie. $\mathcal{I} = \mathbb{Z}$.

It follows that $\mathcal{M}$ is maximal. . . . . ☐ `[Claim 1]`

Claim 2: *Let $\mathcal{M} \lhd \mathbb{Z}$ be a maximal ideal. Then $\mathcal{M} = p\mathbb{Z}$ for some prime p.*

Proof: $\mathcal{M}$ is an ideal of $\mathbb{Z}$, so Proposition 157 on page 133 says that $\mathcal{M} = p\mathbb{Z}$, where p is the smallest positive element in $\mathcal{M}$. Our goal is to show that p is prime.

Suppose p was not prime. Let q be a divisor of p, with $1 < q < p$. Let $p = q \cdot d$.

Claim 2.1: $\mathcal{M} \subset q\mathbb{Z}$.

Proof: Any element of $\mathcal{M}$ has the form pz for some $z \in \mathbb{Z}$. But $p = q \cdot d$, so $pz = q \cdot (dz)$ is also an element of $q\mathbb{Z}$. . . . . ☐ `[Claim 2.1]`

Claim 2.2: $q\mathbb{Z} \not\subset \mathcal{M}$.

Proof: Recall that $1 < q < p$. But p is the *smallest* positive element in $\mathcal{M}$; hence q cannot be in $\mathcal{M}$; hence $q\mathbb{Z} \not\subset \mathcal{M}$. . . . . ☐ `[Claim 2.2]`

Thus, $\mathcal{M} \subsetneq q\mathbb{Z} \subsetneq \mathbb{Z}$, so $\mathcal{M}$ cannot be maximal. Contradiction. ____☐`[Claim 2 & Theorem]`

10.2.1 Maximal ideals in polynomial rings

The maximal ideal structure of $\mathbb{Z}$ is closely mirrored by the maximal ideal structure of polynomial rings....

Example 164: Let $\mathcal{R} = \mathbb{R}[x]$ and consider the principal ideal $\mathcal{M} = (x)$ (Example ⟨156b⟩). I claim this ideal is maximal. To see this, suppose $\mathcal{I} \lhd \mathbb{R}[x]$ such that $(x) \subsetneq \mathcal{I} \subsetneq \mathbb{R}[x]$. Thus, $\mathcal{I}$ contains a polynomial p with $p \notin (x)$. Hence $p(x) = p_n x^n + \ldots + p_1 x + p_0$ with $p_0 \neq 0$. Let

$$q(x) \quad = \quad p_n x^{n-1} + \ldots + p_2 x + p_1.$$

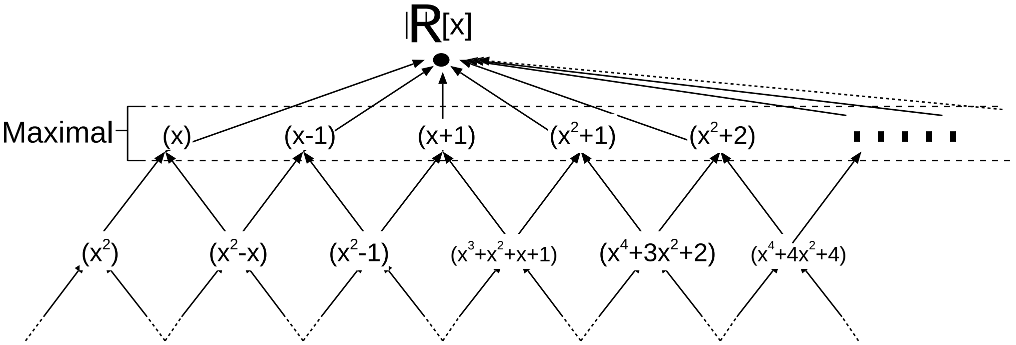


Figure 10.3: A schematic representation of the lattice of ideals of the ring $\mathbb{R}[x]$. The *maximal* ideals are the principal ideals (x), $(x-1)$, (x^2+1), etc. generated by irreducible polynomials.

Then $x \cdot q(x)$ is in (x), and thus, in $\mathcal{I}$. But

$$x \cdot q(x) \quad = \quad p_n x^n + \ldots + p_2 x^2 + p_1 x \quad = \quad p(x) - p_0.$$

Thus, $\mathcal{I}$ contains $p(x) - x \cdot q(x) \ = \ p_0$. Thus, $\mathcal{I}$ contains $p_0^{-1} \cdot p_0 = 1$. Thus, $\mathcal{I} = \mathbb{R}[x]$, by Lemma 120 on page 111.

Hence, any ideal properly containing (x) must be all of $\mathbb{R}[x]$. So (x) is maximal.

Example ⟨164⟩ generalizes as follows. Recall that a polynomial $p(x) \in \mathcal{F}[x]$ is **irreducible** if there exists no polynomials $q(x) \in \mathcal{F}[x]$ which divides $p(x)$, except for $q(x) = p(x)$ or $q(x) = 1$.

Proposition 165 *Let $\mathcal{F}$ be a field. The maximal ideals of $\mathcal{F}[x]$ are exactly the ideals (p), where p is an irreducible polynomial (see Figure 10.3).*

Proof: We want to show: $\Big(\ \mathcal{M}$ is a maximal ideal in $\mathcal{F}[x]\ \Big) \iff \Big(\ \mathcal{M} = (p)$ for irreducible $p\ \Big)$.

Claim 1: *Let p be an irreducible polynomial, and let $\mathcal{M} = (p)$. Then $\mathcal{M}$ is a maximal ideal.*

Proof: Recall from Example ⟨156b⟩ that $\mathcal{M}$ is an ideal. We must show it is maximal.

To see this, suppose $\mathcal{I} \lhd \mathcal{F}[x]$ such that $(p) \subsetneq \mathcal{I} \subsetneq \mathcal{F}[x]$. Proposition 159 on page 135 says that $\mathcal{I} = (i)$, where i is a polynomial of minimal degree in $\mathcal{I}$. But if $\mathcal{M} \subset \mathcal{I}$, that means in particular that $p \in \mathcal{I}$ —ie. $p \in (i)$, so i must divide p. But p is irreducible, so...

either: $i = p$, in which case $(i) = (p)$ ie. $\mathcal{I} = \mathcal{M}$.

or: $i = 1$, in which case $(i) = \mathcal{F}[x]$ ie. $\mathcal{I} = \mathcal{F}[x]$.

It follows that $\mathcal{M}$ is maximal. .... □ `[Claim 1]`

Claim 2: *If $\mathcal{M} \lhd \mathcal{F}[x]$ is a maximal ideal, then $\mathcal{M} = (p)$ for some irreducible polynomial p.*

Proof: $\mathcal{M}$ is an ideal of $\mathcal{F}[x]$, so Proposition 159 on page 135 says that $\mathcal{M} = (p)$, where p is an element in $\mathcal{M}$ of minimal degree. Our goal is to show that p is irreducible.

Suppose p was not irreducible. Let q be a divisor of p, with $0 \leq \mathsf{degree}\,(q) < \mathsf{degree}\,(p)$. Let $p = q \cdot d$.

Claim 2.1: $\mathcal{M} \subset (q)$.

Proof: Any element of $\mathcal{M}$ has the form pf for some $f \in \mathcal{F}[x]$. But $p = q \cdot d$, so $pz = q \cdot (df)$ is also an element of (q). .... □ `[Claim 2.1]`

Claim 2.2: $(q) \not\subset \mathcal{M}$.

Proof: Recall that $0 \leq \mathsf{degree}\,(q) < \mathsf{degree}\,(p)$. But p has *minimal* degree in $\mathcal{M}$; hence q cannot be in $\mathcal{M}$; hence $(q) \not\subset \mathcal{M}$. .... □ `[Claim 2.2]`

Thus, $\mathcal{M} \subsetneq (q) \subsetneq \mathcal{F}[x]$, so $\mathcal{M}$ cannot be maximal. Contradiction. □ `[Claim 2 & Theorem]`

10.2.2 Maximal Ideals have Simple Quotients

Recall that a ring $\mathcal{R}$ is *simple* if it contains no nontrivial ideals (ie. the only ideals in $\mathcal{R}$ are $\{0\}$ and $\mathcal{R}$ itself). For example, $\mathbb{Z}_{/2}$ is a simple ring. But $\mathbb{Z}_{/2}$ is just the quotient ring $\mathbb{Z}/2\mathbb{Z}$ (Example ⟨128⟩), and as we've seen, $2\mathbb{Z}$ is a maximal ideal. This exemplifies a general principle: If $\mathcal{M}$ *is a maximal ideal, then* $\mathcal{R}/\mathcal{M}$ *is a simple ring.* To be precise:

Proposition 166 *Let $\mathcal{R}$ be a ring and let $\mathcal{I} \lhd \mathcal{R}$ be an ideal. Then*

$$\Big(\ \mathcal{I} \text{ is maximal}\ \Big) \iff \Big(\ \mathcal{R}/\mathcal{I} \text{ is simple}\ \Big).$$

Proof: Let $\widetilde{\mathcal{R}} = \mathcal{R}/\mathcal{I}$. We apply the `Lattice Isomorphism Theorem` (Theorem 134 on page 116). Let $\mathcal{J} < \mathcal{R}$ be any subring with $\mathcal{I} < \mathcal{J}$. Let $\widetilde{\mathcal{J}} = \mathcal{J}/\mathcal{I}$, a subring of $\widetilde{\mathcal{R}}$. Then the `Lattice Isomorphism Theorem` says

$$\Big(\ \mathcal{J} \text{ is a proper ideal of } \mathcal{R}, \text{ and } \mathcal{I} \subsetneq \mathcal{J}\ \Big) \iff \Big(\ \widetilde{\mathcal{J}} \text{ is a proper ideal of } \widetilde{\mathcal{R}}, \text{ and } \widetilde{\mathcal{J}} \neq 0\ \Big).$$

Hence,

$$\begin{aligned}\Big(\ \mathcal{I} \text{ is maximal}\ \Big) &\iff \Big(\ \text{There are no proper ideals } \mathcal{J} \lhd \mathcal{R} \text{ with } \mathcal{I} \subsetneq \mathcal{J}\ \Big)\\ &\iff \Big(\ \text{There are no proper ideals } \widetilde{\mathcal{J}} \lhd \widetilde{\mathcal{R}} \text{ with } \mathcal{J} \neq \{0\}\ \Big)\\ &\iff \Big(\ \mathcal{R}/\mathcal{I} \text{ is simple}\ \Big).\end{aligned}$$

□

Recall that the only simple *commutative* rings are *fields*. Hence, we have the following:

Corollary 167 *Let $\mathcal{R}$ be a commutative ring and let $\mathcal{I} \lhd \mathcal{R}$ be an ideal. Then*

$$\Big(\mathcal{I} \text{ is maximal} \Big) \iff \Big(\mathcal{R}/\mathcal{I} \text{ is a field} \Big).$$

An equivalent formulation of Corollary 167 is:

Corollary 168 *Let $\mathcal{R}$ be a commutative ring and let $\mathcal{F}$ be a field.*

(a) *If $\phi : \mathcal{R} \longrightarrow \mathcal{F}$ is an epimorphism, then $\ker(\phi)$ is a maximal ideal.*

(b) *All maximal ideals of $\mathcal{R}$ arise in this fashion.* □

Example 169:

(a) Let $\mathcal{R} = \mathbb{R}[x]$, and consider the principal ideal $(x) = \{x \cdot q(x) \ ; \ q \in \mathcal{R}[x]$ any polynomial$\}$. We already saw (Example ⟨164⟩) that (x) is a maximal ideal in $\mathbb{R}[x]$, but now we can provide a second proof.

Recall the *evaluation* homomorphism $\epsilon_0 : \mathbb{R}[x] \longrightarrow \mathbb{R}$ from Example ⟨111g⟩, defined by $\epsilon_0(p_n x^n + \ldots + p_1 x + p_0) = p_0$. In Example ⟨113e⟩, we saw that $\ker(\epsilon_0) = (x)$. But $\mathsf{image}\,[\epsilon_0] = \mathbb{R}$ is a field, so Corollary 168 implies that (x) is a maximal ideal.

(b) Let $\mathcal{R} = \mathcal{C}(\mathbb{R})$, and let $r \in \mathbb{R}$ be any fixed number, and consider the ideal $\mathcal{M}_r = \{f \in \mathcal{C}(\mathbb{R}) \ ; \ f(r) = 0\}$ from Example ⟨117d⟩. I claim $\mathcal{M}_r$ is a maximal ideal in $\mathcal{C}(\mathbb{R})$.

To see this, let $\epsilon_r : \mathcal{C}(\mathbb{R}) \longrightarrow \mathbb{R}$ be the *evalutation map* from Example ⟨111e⟩, defined by $\epsilon_r(f) = f(r)$. In Example ⟨113d⟩, we saw that $\ker(\epsilon_r) = \mathcal{M}_r$. But $\mathsf{image}\,[\epsilon_r] = \mathbb{R}$ is a field, so Corollary 168 implies that $\mathcal{M}_r$ is a maximal ideal.

Thus, *every point in $\mathbb{R}$ defines a maximal ideal in $\mathcal{C}(\mathbb{R})$.* We will see in §10.3 that this correspondence goes in both directions.

(c) Let $\mathcal{R} = \mathbb{C}[x]$, and let $c \in \mathbb{C}$ be any complex number, and consider the ideal $\mathcal{M}_c = \{f \in \mathbb{C}[x] \ ; \ f(c) = 0\}$. Observe that $\mathcal{M}_c$ is the kernel of the evaluation homomorphism $\epsilon_c : \mathbb{C}[x] \longrightarrow \mathbb{C}$ defined by $\epsilon_c(f) = f(c)$. Since the image of ϵ_c is the field $\mathbb{C}$, it follows that $\mathcal{M}_c$ is maximal in $\mathbb{C}[x]$.

Corollary 167 also provides a powerful mechanism for constructing fields....

Proposition 170 *Let $\mathcal{F}$ be a field, and let $p(x) \in \mathcal{F}[x]$ be an irreducible polynomial. Then the quotient ring $\mathcal{F}[x]/(p)$ is also a field.*

Proof: Proposition 165 says that the principal ideal (p) is maximal in $\mathcal{F}[x]$. Hence, Corollary 167 says that the quotient ring $\mathcal{F}[x]/(p)$ is a field. □

To illustrate, we'll construct the complex numbers....

Example 171: *The field $\mathbb{C}$ is isomorphic to the quotient ring $\mathbb{R}[x]/(x^2+1)$.*

Proof: x^2+1 is irreducible because you can't factor x^2+1 in $\mathbb{R}[x]$. Thus, Proposition 170 says that $\mathcal{C} = \mathbb{R}[x]/(x^2+1)$ is a field. We want to show that $\mathcal{C}$ is isomorphic to $\mathbb{C}$.

For any polynomial $p(x)$ in $\mathbb{R}[x]$, let $\overline{p}$ be the corresponding element in $\mathcal{C}$. In particular, let $\mathbf{j} = \overline{x}$.

Claim 1: *$\mathbf{j}$ is a square root of -1.*

Proof: Observe that $\mathbf{j}^2 + \overline{1} \;=\; \overline{x}^2 + \overline{1} \;=\; \overline{x^2+1} \;=\; 0$. Hence, $\mathbf{j}^2 = -1$. □ `[Claim 1]`

Claim 2: *Every element of $\mathcal{C}$ can be written in a unique way as $\overline{r}_1 + \overline{r}_2\mathbf{j}$, where r_1 and r_2 are real numbers.*

Proof: Any element of $\mathcal{C}$ has the form $\overline{p(x)}$ for some polynomial $p \in \mathbb{R}[x]$. To illustrate, we'll use the polynomial $p(x) = 8x^8 + 7x^7 + 6x^6 + 5x^5 + 4x^4 + 3x^3 + 2x^2 + x + \frac{1}{2}$. Observe

$$\begin{aligned}
\overline{p(x)} &= \overline{8x^8 + 7x^7 + 6x^6 + 5x^5 + 4x^4 + 3x^3 + 2x^2 + x + 1/2}\\
&= \bar{8}\overline{x}^8 + \bar{7}\overline{x}^7 + \bar{6}\overline{x}^6 + \bar{5}\overline{x}^5 + \bar{4}\overline{x}^4 + \bar{3}\overline{x}^3 + \bar{2}\overline{x}^2 + \overline{x} + \overline{1/2}\\
&= \bar{8}\mathbf{j}^8 + \bar{7}\mathbf{j}^7 + \bar{6}\mathbf{j}^6 + \bar{5}\mathbf{j}^5 + \bar{4}\mathbf{j}^4 + \bar{3}\mathbf{j}^3 + \bar{2}\mathbf{j}^2 + \mathbf{j} + \overline{1/2}\\
&\underset{\text{Clm.1}}{=\!=} \bar{8}(-\bar{1})^4 + \bar{7}(-\bar{1})^3\mathbf{j} + \bar{6}(-\bar{1})^3 + \bar{5}(-\bar{1})^2\mathbf{j} + \bar{4}(-\bar{1})^2\\
&\qquad\qquad + \bar{3}(-\bar{1})\mathbf{j} + \bar{2}(-\bar{1}) + \mathbf{j} + \overline{1/2}\\
&= \bar{8} - \bar{7}\mathbf{j} - \bar{6} + \bar{5}\mathbf{j} + \bar{4} - \bar{3}\mathbf{j} - \bar{2} + \mathbf{j} + \overline{1/2}\\
&= (\bar{8} - \bar{6} + \bar{4} - \bar{2} + \overline{1/2}) + (-\bar{7} + \bar{5} - \bar{3} + \bar{1})\mathbf{j} \;=\; \overline{1/2} - \bar{4}\mathbf{j}.
\end{aligned}$$

The same reduction will clearly work for any polynomial. .... □ `[Claim 2]`

Now, define the map $\phi : \mathbb{C} \longrightarrow \mathcal{C}$ by $\phi(r_1 + r_2\mathbf{i}) \;=\; \overline{r}_1 + \overline{r}_2\mathbf{j}$. It is left as **Exercise 131** to verify that this is a ring isomorphism. □

10.3 The Maximal Spectrum

Prerequisites: §10.2

Alg.Geo.

Let $\mathcal{R}$ be a ring. The **maximal spectrum** of $\mathcal{R}$ is the set of all maximal ideals of $\mathcal{R}$:

$$\overline{\mathsf{Spec}}\,(\mathcal{R}) \quad=\quad \{\mathcal{M}\ ;\ \mathcal{M} \lhd \mathcal{R} \text{ is a maximal ideal}\}. \qquad \text{(see Figure 10.4)}$$

Example 172: Let $\mathcal{R} = \mathbb{Z}$. Then Proposition 163 on page 139 says that the maximal ideals of $\mathbb{Z}$ are exactly the principal ideals $p\mathbb{Z}$, where p is prime. In other words,

$$\overline{\mathsf{Spec}}\,(\mathbb{Z}) \quad=\quad \{p\mathbb{Z}\ ;\ p \in \mathbb{N} \text{ a prime number}\}$$

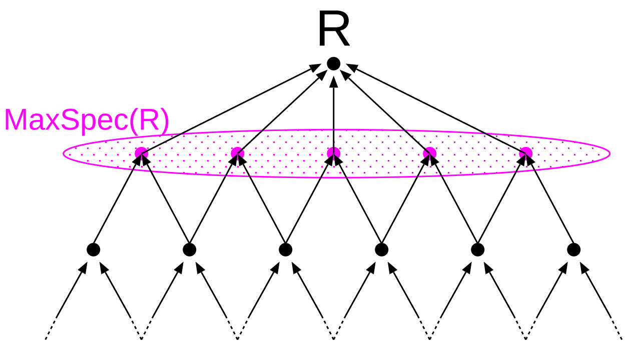


Figure 10.4: A schematic representation of the maximal spectrum of the ring $\mathcal{R}$.

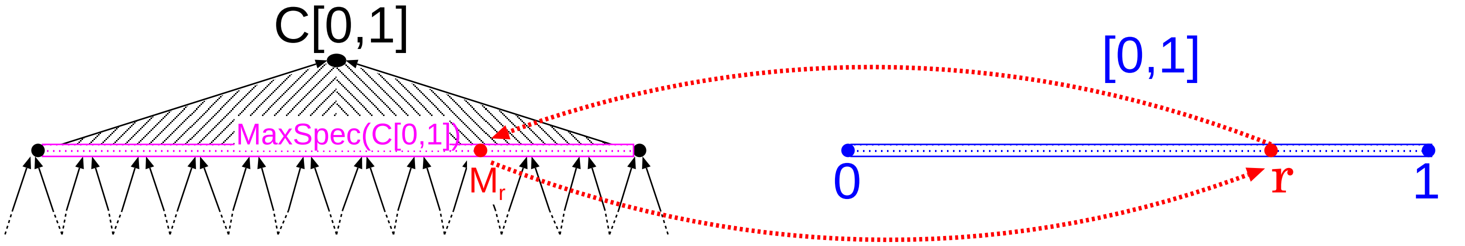


Figure 10.5: A schematic representation of the maximal spectrum of the ring $\mathcal{C}[0,1]$, showing the bijective correspondence between points in $[0,1]$ and maximal ideals in $\mathcal{C}[0,1]$.

On of the fundamental concepts of algebraic geometry is the following

> **Correspondence Principle:** *If* $\mathbf{X}$ *is a space, and* $\mathcal{R}$ *is a ring of functions over* $\mathbf{X}$*, then the* maximal ideals *of* $\mathcal{R}$ *correspond to the* points *of* $\mathbf{X}$*. Hence,* $\overline{\mathsf{Spec}}(\mathcal{R})$ *is a sort of 'image' of* $\mathbf{X}$*.*

Thus, the maximal spectrum has a natural 'geometric' interpretation. The basic idea of *commutative ideal theory* is to extend this interpretation to *arbitrary* rings; given *any* ring $\mathcal{R}$, we interpret $\overline{\mathsf{Spec}}(\mathcal{R})$ as a kind of abstract 'space', and $\mathcal{R}$ as a ring of 'functions' on this space.

10.3.1 The Maximal Spectrum of a Continuous Function Ring

Prerequisites: §10.2 **Recommended:** §A

Our first illustration of the `Correspondence Principle` is the following:

Proposition 173 *Let* $\mathcal{C}[0,1]$ *be the ring of continuous functions from* $[0,1]$ *into* $\mathbb{R}$*. Then:*

(a) *For every $r \in [0,1]$, there is a maximal ideal $\mathcal{M}_r = \{f \in \mathcal{C}[0,1]\ ;\ f(r) = 0\}$, which is the kernel of the evaluation map $\epsilon_r : \mathcal{C}[0,1] \ni f \mapsto f(r) \in \mathbb{R}$.*

(b) *Every maximal ideal of $\mathcal{C}[0,1]$ arises in this fashion.*

(c) *Thus, there is a natural bijective correspondence $[0,1] \leftrightarrow \overline{\mathsf{Spec}}\left(\mathcal{C}[0,1]\right)$, defined by the map $r \mapsto \mathcal{M}_r$ (see Figure 10.5).*

Proof: We have already seen that $\mathcal{M}_r$ is a maximal ideal in Example ⟨169b⟩ (there we were in the ring $\mathcal{C}(\mathbb{R})$, but the argument for $\mathcal{C}[0,1]$ is identical). It remains to show that the map $[0,1] \ni r \mapsto \mathcal{M}_r \in \overline{\mathsf{Spec}}\left(\mathcal{C}[0,1]\right)$ is bijective.

Injective: Suppose $r, s \in [0,1]$ with $r \neq s$. We want to show that $\mathcal{M}_r \neq \mathcal{M}_s$. To see this, consider the continuous function $f(x) = |x - r|$. Clearly, $f \in \mathcal{C}[0,1]$, and $f(r) = 0$, so $f \in \mathcal{M}_r$. On the other hand, $f(s) \neq 0$, so $f \notin \mathcal{M}_s$. Hence, $\mathcal{M}_r \neq \mathcal{M}_s$.

Surjective: Let $\mathcal{M} \lhd \mathcal{C}[0,1]$ be some maximal ideal; we want to show that $\mathcal{M} = \mathcal{M}_r$ for some $r \in [0,1]$. To do this, for any $f \in \mathcal{M}$, define

$$\mathbb{V}(f) \quad = \quad \{v \in [0,1]\ ;\ f(v) = 0\}.$$

Claim 1: *$\mathbb{V}(f) \neq \emptyset$, and is a closed subset of $[0,1]$.*

Proof: *Nonempty:* Suppose $\mathbb{V}(f) = \emptyset$. This means that $f(x) \neq 0$ for all $x \in [0,1]$. Thus, f is a *unit* element of $\mathcal{C}[0,1]$. To see this, define $g(x) = \frac{1}{f(x)}$. Then $g \in \mathcal{C}[0,1]$ also, and $f \cdot g = \mathbb{1}$.

But if f is a unit element, then $\mathcal{M} = \mathcal{C}[0,1]$, by Lemma 120 on page 111. By contradiction, $\mathbb{V}(f) \neq \emptyset$

Closed: To see that $\mathbb{V}(f)$ is *closed*, suppose that $v_1, v_2, \ldots,$ was a sequence of points in $\mathbb{V}(f)$. Thus, $f(v_n) = 0$ for all $n \in \mathbb{N}$. Suppose $v = \lim_{n\to\infty} v_n$. Since f is continuous, we have:

$$f(v) \quad = \quad \lim_{n\to\infty} f(v_n) \quad = \quad \lim_{n\to\infty} 0 \quad = \quad 0.$$

Hence, $v \in \mathbb{V}(f)$ also. .... □ [Claim 1]

Claim 2: *If $f_1, f_2, \ldots, f_N \in \mathcal{M}$, then $\bigcap_{n=1}^{N} \mathbb{V}(f_n) \neq \emptyset$, and is a closed subset of $[0,1]$.*

Proof: Define $F(x) = f_1^2(x) + f_2(x)^2 + \ldots f_n(x)^2$. Then F is also in $\mathcal{M}$ (because $\mathcal{M}$ is a subring), so Claim 1 says $\mathbb{V}(F) \neq \emptyset$. But $\mathbb{V}(F) = \bigcap_{n=1}^{N} \mathbb{V}(f_n)$ (**Exercise 132**).

To see that $\bigcap_{n=1}^{N} \mathbb{V}(f_n)$ is *closed*, recall that the intersection of any number of closed subsets is also closed. (Lemma 267 on page 224) .... □ [Claim 2]

Now, define

$$\mathbb{V}(\mathcal{M}) \quad = \quad \{v \in [0,1] \ ; \ f(v) = 0 \text{ for all } f \in \mathcal{M}\} \quad = \quad \bigcap_{f \in \mathcal{M}} \mathbb{V}(f).$$

Claim 3: $\mathbb{V}(\mathcal{M}) \neq \emptyset$.

Proof: The interval $[0,1]$ is compact, so we apply the `Finite Intersection Property` (230). Now Claim 3 follow from Claim 2. .... □ `[Claim 3]`

Now, let $x \in \mathbb{V}(\mathcal{M})$.

Claim 4: $\mathcal{M} \subset \mathcal{M}_x$.

Proof: If $f \in \mathcal{M}$, then $x \in \mathbb{V}(f)$; hence $f(x) = 0$, hence $f \in \mathcal{M}_x$. ..... □ `[Claim 4]`

But $\mathcal{M}$ is a *maximal* ideal, so we conclude that $\mathcal{M} = \mathcal{M}_x$. □

There is nothing special about the interval $[0,1]$; the same result holds for any compact[1] subset $\mathbf{X} \subset \mathbb{R}^N$:

Proposition 174 *Let $n \in \mathbb{N}$, and let $\mathbf{X} \subset \mathbb{R}^n$ be a compact subset. Let $\mathcal{C}(\mathbf{X})$ be the ring of continuous functions from $\mathbf{X}$ into $\mathbb{R}$. Then:*

(a) *For every $\mathbf{x} \in \mathbf{X}$, there is a maximal ideal $\mathcal{M}_{\mathbf{x}} = \{f \in \mathcal{C}(\mathbf{X}) \ ; \ f(\mathbf{x}) = 0\}$, which is the kernel of the evaluation map $\epsilon_{\mathbf{x}} : \mathcal{C}(\mathbf{X}) \ni f \mapsto f(\mathbf{x}) \in \mathbb{R}$.*

(b) Every *maximal ideal of $\mathcal{C}(\mathbf{X})$ arises in this fashion.*

(c) *Thus, there is a natural bijective correspondence $\mathbf{X} \leftrightarrow \overline{\mathsf{Spec}}\left(\mathcal{C}(\mathbf{X})\right)$, defined by the map $\mathbf{x} \mapsto \mathcal{M}_{\mathbf{x}}$.*

Proof: **<u>Exercise 133</u>** Generalize the proof of Proposition 173 □

It is important that $\mathbf{X}$ be *compact*. Proposition 174 is *not* true if we let $\mathbf{X} = \mathbb{R}$ or $\mathbb{R}^n$.

Example 175: Let $\mathcal{R} = \mathcal{C}(\mathbb{R})$, and let $\mathcal{C}_0(\mathbb{R})$ be the ring of continuous functions with *compact support* from Example ⟨3k⟩. Then $\mathcal{C}_0(\mathbb{R})$ is an ideal within $\mathcal{C}(\mathbb{R})$ (see Example ⟨117l⟩), and indeed, a *maximal* ideal (**<u>Exercise 134</u>**). However, there is no point $r \in \mathbb{R}$ such that $\mathcal{C}_0(\mathbb{R}) = \mathcal{M}_r$. Hence, the `Correspondence Principle` *fails* for $\mathcal{C}(\mathbb{R})$.

To see how this is related to compactness, note that, in a sense, $\mathcal{C}_0(\mathbb{R})$ is the set of continuous functions which 'vanish at infinity'. Metaphorically we could write: "$\mathcal{C}_0(\mathbb{R}) = \mathcal{M}_\infty$." The `Correspondence Principle` fails because '∞' is not an element of $\mathbb{R}$. The solution to this problem is to *compactify* $\mathbb{R}$ by adding a 'point at ∞'. This is called the *Stone-Čech compactification.*

[1] $\mathbf{X}$ is *compact* if $\mathbf{X}$ is both *closed* and *bounded* in $\mathbb{R}^n$. See Appendix §A.4.

We can generalize Proposition 174 much further:

Proposition 176 *Let $\mathbf{X}$ be a compact metric space, and let $\mathcal{C}(\mathbf{X})$ be the ring of continuous functions from $\mathbf{X}$ into $\mathbb{R}$. Then:*

(a) *For every $x \in \mathbf{X}$, there is a maximal ideal $\mathcal{M}_x = \{f \in \mathcal{C}(\mathbf{X})\ ;\ f(x) = 0\}$, which is the kernel of the evaluation map $\epsilon_x : \mathcal{C}(\mathbf{X}) \ni f \mapsto f(x) \in \mathbb{R}$.*

(b) Every *maximal ideal of $\mathcal{C}(\mathbf{X})$ arises in this fashion.*

(c) *Thus, there is a natural bijective correspondence $\mathbf{X} \leftrightarrow \overline{\mathsf{Spec}}\left(\mathcal{C}(\mathbf{X})\right)$, defined by the map $x \mapsto \mathcal{M}_x$.*

Proof: **<u>Exercise 135</u>** Hint: Use the fact that metric spaces are *regular.* □

The `Correspondence Principle` also holds if we restrict ourselves to *differentiable* functions...

Proposition 177 *Let $\mathbf{X} \subset \mathbb{R}^n$ be a compact subset of $\mathbb{R}^n$ (or let $\mathbf{X}$ be a compact differentiable manifold). Let $\mathcal{C}$ be one of the following rings of functions from $\mathbf{X}$ into $\mathbb{R}$:*

$$\mathcal{C} = \mathcal{C}^k(\mathbf{X}) \text{ (for some } k \in \mathbb{N}) \qquad \text{or} \qquad \mathcal{C} = \mathcal{C}^\infty(\mathbf{X}), \qquad \text{or} \qquad \mathcal{C} = \mathcal{C}^\omega(\mathbf{X}).$$

Then:

(a) *For every $\mathbf{x} \in \mathbf{X}$, there is a maximal ideal $\mathcal{M}_\mathbf{x} = \{f \in \mathcal{C}\ ;\ f(\mathbf{x}) = 0\}$, which is the kernel of the evaluation map $\epsilon_\mathbf{x} : \mathcal{C} \ni f \mapsto f(\mathbf{x}) \in \mathbb{R}$.*

(b) Every *maximal ideal of $\mathcal{C}$ arises in this fashion.*

(c) *Thus, there is a natural bijective correspondence $\mathbf{X} \leftrightarrow \overline{\mathsf{Spec}}(\mathcal{C})$, defined by the map $\mathbf{x} \mapsto \mathcal{M}_\mathbf{x}$.*

Proof: **<u>Exercise 136</u>** □

10.3.2 The Maximal Spectrum of a Polynomial Ring

Most important for classical algebraic geometry is a `Correspondence Principle` for the ring of *polynomial* functions. To make this work, however, we must pass to the ring of *complex* polynomials....

Proposition 178 *Let $\mathbb{C}[x]$ be the ring of complex polynomials. Then:*

(a) *For every $c \in \mathbb{C}$, there is a maximal ideal $\mathcal{M}_c = \{f \in \mathbb{C}[x]\ ;\ f(c) = 0\}$, which is the kernel of the evaluation map $\epsilon_c : \mathbb{C}[x] \ni f \mapsto f(c) \in \mathbb{C}$.*

(b) *$\mathcal{M}_c$ is the principal ideal generated by the linear polynomial $\ell(x) = x - c$. That is,*

$$\mathcal{M}_c \quad = \quad \Big\{(x-c)\cdot q(x)\ ;\ \ q(x) \in \mathbb{C}[x]\Big\}.$$

(c) *<u>Every</u> maximal ideal of $\mathbb{C}[x]$ arises in this fashion.*

(d) *Thus, there is a natural bijective correspondence $\mathbb{C} \leftrightarrow \overline{\mathsf{Spec}}\Big(\mathbb{C}[x]\Big)$, defined by the map $c \mapsto \mathcal{M}_c$.*

The proof of Proposition 178 requires the following result, which we state without proof:

Fundamental Theorem of Algebra: *Any polynomial in $\mathbb{C}[x]$ can be factored into* linear *polynomials. In other words, if $p(x) \in \mathbb{C}[x]$, then there are complex constants $c_1, c_2, \ldots, c_n \in \mathbb{C}$ (possibly not distinct) so that* $p(x) = (x-c_1)\cdot(x-c_2)\cdots(x-c_n)$. □

(The proof of the `Fundamental Theorem` does not involve the theory of maximal ideals, so there is no circularity here)

Proof of Proposition 178: **(a)** This was Example ⟨169c⟩.

(b) **<u>Exercise 137</u>** .

(c) Let $\mathcal{M} \triangleleft \mathbb{C}[x]$ be a maximal ideal. The ring $\mathbb{C}[x]$ is a principal ideal domain (Proposition 159 on page 135) so $\mathcal{M}$ is the principal ideal generated by some polynomial $p(x) \in \mathbb{C}[x]$. Our goal is to show that $p(x) = x - c$ for some $c \in \mathbb{C}$.

The `Fundamental Theorem of Algebra` says that $p(x)$ is divisible by some linear polynomial $\ell(x) = (x - c)$. That is, $p(x) = (x-c)\cdot q(x)$ for some polynomial $q(x)$. Thus, p is in the principal ideal $\mathcal{I}$ generated by $(x-c)$. But then $\mathcal{M} \subset \mathcal{I}$. Since $\mathcal{M}$ is maximal, it follows that $\mathcal{M} = \mathcal{I}$ —that is, $\mathcal{M}$ is the principal ideal generated by $(x-c)$. □

The generalization of Proposition 178 to polynomials of n variables is one of the most important theorems in algebraic geometry:

Theorem 179 Hilbert's Nullstellensatz (Complex Version)

Let $\mathbb{C}[x_1, \ldots, x_n]$ be the ring of complex polynomials in n variables.

(a) *For every $\mathbf{c} \in \mathbb{C}^n$, there is a maximal ideal $\mathcal{M}_{\mathbf{c}} = \{f \in \mathbb{C}[x_1, \ldots, x_n] \ ; \ f(\mathbf{c}) = 0\}$, which is the kernel of the evaluation map $\epsilon_{\mathbf{c}} : \mathbb{C}[x_1, \ldots, x_n] \ni f \mapsto f(\mathbf{c}) \in \mathbb{C}$.*

(b) *$\mathcal{M}_{\mathbf{c}}$ is the ideal generated by the linear polynomials*

$$\ell_1(\mathbf{x}) = x_1 - c_1; \qquad \ell_2(\mathbf{x}) = x_2 - c_2; \quad \ldots \quad \ell_n(\mathbf{x}) = x_n - c_n;$$

That is,

$$\mathcal{M}_c \quad = \quad \Big\{(x_1 - c_1) \cdot q_1(\mathbf{x}) \ + \ldots \ + (x_n - c_n) \cdot q_n(\mathbf{x}) \ ; \ \ q_1(\mathbf{x}), \ldots, q_n(\mathbf{x}) \in \mathbb{C}[x_1, \ldots, x_n]\Big\}.$$

(c) *<u>Every</u> maximal ideal of $\mathbb{C}[x_1, \ldots, x_n]$ arises in this fashion.*

(d) *Thus, there is a natural bijective correspondence $\mathbb{C}^n \leftrightarrow \overline{\mathsf{Spec}}\Big(\mathbb{C}[x_1, \ldots, x_n]\Big)$, defined by the map $\mathbf{c} \mapsto \mathcal{M}_{\mathbf{c}}$.*

Proof: (a) <u>**Exercise 138**</u> .

(b) Our strategy is to perform a 'change of coordinates' on $\mathbb{C}^n$, so that the point $\mathbf{c} = (c_1, \ldots, c_n)$ acts as the 'origin'. It may therefore be helpful, when first reading this argument, to pretend that $\mathbf{c} = (0, 0, \ldots, 0)$.

Claim 1: *Let $p(\mathbf{x}) \in \mathbb{C}[x_1, \ldots, x_n]$ be any polynomial. Then $p(\mathbf{x})$ can be written as a polynomial in the variables $(x_1 - c_1)$, $(x_2 - c_2)$, $\ldots$, $(x_n - c_n)$. In other words, there are complex coefficients $q^*, q_1, \ldots, q_n$, $q_{11}, \ldots, q_{nn}$, $q_{111}, \ldots, q_{nnn}$, etc. so that*

$$\begin{aligned} p(\mathbf{x}) \quad = \quad & q^* + q_1(x_1 - c_1) \ + \ \ldots \ + \ q_n(x_n - c_n) \\ & + q_{11}(x_1 - c_1)^2 \ + \ q_{12}(x_1 - c_1)(x_2 - c_2) + \ldots + q_{1n}(x_1 - c_1)(x_n - c_n) \\ & + q_{22}(x_2 - c_2)^2 \ + \ q_{23}(x_2 - c_2)(x_3 - c_3) + \ldots + q_{2n}(x_2 - c_2)(x_n - c_n) \ + \ \ldots \ + \\ & + q_{nn}(x_n - c_n)^2 \ + \ q_{111}(x_1 - c_1)^3 \ + \ q_{112}(x_1 - c_1)^2(x_2 - c_2) + \ldots \end{aligned}$$

Here, $q^* = p(\mathbf{c})$.

Proof: <u>**Exercise 139**</u> Hint: There are two approaches:

Calculus Approach: Compute the Taylor series of the function p around the point $\mathbf{c} = (c_1, \ldots, c_n)$. (*The advantage of this approach is that it is simple.*)

Purely Algebraic Approach: Expand out the expression on the right hand side and collect like terms, to yield a system of linear equations for the coefficients $q_1, q_2, \ldots$ in terms of the coefficients of p. Now solve the resulting system. (*The advantage of this approach is that it will work for any field, not just $\mathbb{C}$.*) .... □ `[Claim 1]`

Note that every term on the right hand side —except for the constant q^* —is divisible by at least one of the polynomials $(x_1 - c_1)$, $\ldots$, $(x_n - c_n)$. Thus, every term except q^* is in the ideal generated by $(x_1 - c_1)$, $\ldots$, $(x_n - c_n)$. It follows:

$$\begin{aligned}\Big(p \in \mathcal{M}_{\mathbf{c}} \Big) \quad &\Longleftrightarrow \quad \Big(p(\mathbf{c}) = 0 \Big) \quad \Longleftrightarrow \quad \Big(q^* = 0 \Big) \\ &\Longleftrightarrow \quad \Big(p \text{ is in the ideal generated by } (x_1 - c_1),\ \ldots,\ (x_n - c_n) \Big).\end{aligned}$$

Hence, $\mathcal{M}_{\mathbf{c}}$ *is* the ideal generated by $(x_1 - c_1)$, $\ldots$, $(x_n - c_n)$.

(c) Let $\mathcal{M} \lhd \mathbb{C}[x_1, \ldots, x_n]$ be a maximal ideal, and let $\mathcal{F} = \mathbb{C}[x_1, \ldots, x_n]/\mathcal{M}$ be the corresponding quotient field (Corollary 167). Let $\pi : \mathbb{C}[x_1, \ldots, x_n] \longrightarrow \mathcal{F}$ be the quotient map.

Let $\mathbb{C}[x_1]$ be the ring of polynomials containing only the x_1 variable; thus, $\mathbb{C}[x_1]$ is a subring of $\mathbb{C}[x_1, \ldots, x_n]$. Let $\pi_1 : \mathbb{C}[x_1] \longrightarrow \mathcal{F}$ be the restriction of π to $\mathbb{C}[x_1]$.

Claim 2: $\ker(\pi_1) \subset \mathcal{M}$.

Proof: By construction, $\mathcal{M} = \ker(\pi)$. Since π_1 is just the restriction of π to the smaller domain $\mathbb{C}[x_1]$, it follows that $\ker(\pi_1) \subset \ker(\pi)$. . . . . □ [Claim 2]

We want to show that $\mathcal{M}$ contains a linear polynomial of the form $\ell(\mathbf{x}) = (x_1 - c_1)$. By Claim 2, it will be sufficient to show that $\ker(\pi_1)$ contains such a polynomial.

Claim 3: $\ker(\pi_1) \neq \{0\}$.

Proof: Recall the field of *rational functions:*

$$\mathbb{C}(x_1) \quad = \quad \left\{ \frac{p(x_1)}{q(x_1)} \ ; \ p, q \in \mathbb{C}[x_1] \right\} \qquad \text{(Example } \langle 98h \rangle \text{ on page 84)}$$

Claim 3.1: *If* $\ker(\pi_1) = \{0\}$*, then we can extend* π_1 *to a monomorphism* $\widetilde{\pi}_1 : \mathbb{C}(x_1) \longrightarrow \mathcal{F}$*. Thus,* $\mathcal{F}$ *contains an isomorphic copy of the field* $\mathbb{C}(x_1)$*.*

Proof: Suppose $p \in \mathbb{C}[x_1]$ has nonzero image —ie. $\pi_1(p) \neq 0$ in $\mathcal{F}$. Since $\mathcal{F}$ is a field, this means that $\pi_1(p)$ has an inverse, which we will denote by $\frac{1}{\pi_1(p)}$.

If $\ker(\pi_1) = \{0\}$, then $\pi_1(p)$ is invertible for *any* nonzero polynomial $p \in \mathbb{C}[x_1]$. Thus, then we can extend π_1 to a map $\widetilde{\pi}_1 : \mathbb{C}(x_1) \longrightarrow \mathcal{F}$ as follows:

$$\widetilde{\pi}_1 \left(\frac{p}{q} \right) \quad = \quad \frac{\pi_1(p)}{\pi_2(p)}, \qquad \text{for any } p, q \in \mathbb{C}[x_1].$$

Exercise 140 Check that $\widetilde{\pi}_1$ is well-defined, and a ring monomorphism. □ [Claim 3.1]

We will now show that the conclusion of Claim 3.1 is impossible, because $\mathbb{C}(x_1)$ is 'too big' to fit inside of $\mathcal{F}$. To do this, it will be helpful to treat $\mathbb{C}(x_1)$ and $\mathcal{F}$ as complex vector spaces.

Recall that $\mathbb{C}[x_1, x_2, \ldots, x_n]$ is a *vector space* over the field $\mathbb{C}$ of complex numbers. In other words, if p and q are two polynomials, then $p + q$ is also a polynomial, and if $c \in \mathbb{C}$ is any scalar, then $c \cdot p$ is also a polynomial.

Claim 3.2: *$\mathbb{C}[x_1, x_2, \ldots, x_n]$ is a countable-dimensional complex vector space.*

Proof: Consider the monomials $x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}$, where $k_1, \ldots, k_n \in \mathbb{N}$. Clearly, any element of $\mathbb{C}[x_1, \ldots, x_n]$ can be written as $\mathbb{C}$-linear combination of monomials of this form. Thus, the (countable) set $\left\{x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n} \ ; \ k_1, \ldots, k_n \in \mathbb{N}\right\}$ spans $\mathbb{C}[x_1, \ldots, x_n]$ as a complex vector space. . . . . □ `[Claim 3.2]`

It follows that $\mathcal{F}$ is also a complex vector space, and $\pi : \mathbb{C}[x_1, \ldots, x_n] \longrightarrow \mathcal{F}$ is a surjective $\mathbb{C}$-linear map. To be precise: If $p, q \in \mathbb{C}[x_1, \ldots, x_n]$, let $\overline{p}$ and $\overline{q}$ denote the corresponding elements of $\mathcal{F}$; then $\overline{p} + \overline{q} = \overline{p+q}$, and, if if $c \in \mathbb{C}$ is any scalar, then $c \cdot \overline{p} = \overline{c \cdot p}$.

Claim 3.3: *$\mathcal{F}$ is a countable-dimensional complex vector space.*

Proof: Since $\mathcal{F}$ is the image of $\mathbb{C}[x_1, \ldots, x_n]$, it follows that $\dim(\mathcal{F}) \leq \dim\Big(\mathbb{C}[x_1, \ldots, x_n]\Big)$. To be more concrete, observe that the (countable) set $\left\{\overline{x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}} \ ; \ k_1, \ldots, k_n \in \mathbb{N}\right\}$ spans $\mathcal{F}$ as a complex vector space. . . . . □ `[Claim 3.3]`

Observe that $\mathbb{C}(x_1)$ can also be treated as a complex vector space (the sum of two rational functions is rational, etc.).

Claim 3.4: *$\mathbb{C}(x_1)$ is a $\mathbb{C}$-vector space of <u>uncountable</u> dimension.*

Proof: For any $c \in \mathbb{C}$, consider the rational function $\rho_c(x) \quad = \quad \dfrac{1}{x - c}$. Observe that $\rho_c(x)$ is finite everywhere except at $x = c$, and $\rho_c(c) = \infty$. It follows that, if $c_1, \ldots, c_n$ are distinct complex numbers, then any finite $\mathbb{C}$-linear combination

$$a_1 \rho_{c_1} \ + \ a_2 \rho_{c_2} \ + \ \ldots \ + \ a_n \rho_{c_2} \tag{10.4}$$

(where $a_1, \ldots, a_n \in \mathbb{C}$ are nonzero complex numbers) is equal to infinity on the set $\{c_1, \ldots, c_n\}$. In particular, the linear combination (10.4) can never be zero. In other words, the set of rational functions $\{\rho_c \ ; \ c \in \mathbb{C}\}$ is linearly independent in $\mathbb{C}(x)$. But $\mathbb{C}$ is uncountable, so the set $\{\rho_c \ ; \ c \in \mathbb{C}\}$ is also uncountable. Hence, $\mathbb{C}(x)$ has uncountable dimension. . . . . □ `[Claim 3.4]`

Thus, Claims 3.1, 3.2, and 3.3 together amount to an embedding of an uncountable dimensional $\mathbb{C}$-vector space within a countable-dimensional space. This is impossible. By contradiction, $\ker(\pi_1)$ cannot be $\{0\}$. . . . . □ `[Claim 3]`

Claim 4: *$\ker(\pi_1)$ contains a linear polynomial $\ell_1(x_1) = (x_1 - c_1)$ for some $c_1 \in \mathbb{C}$.*

Proof: Claim 3 says $\ker(\pi_1) \neq \{0\}$, so let $k(x_1) \in \ker(\pi_1)$ be some nonzero polynomial. The `Fundamental Theorem of Algebra` (p. 148) says that there are constants $a_1, a_2, \ldots, a_m \in \mathbb{C}$ so that

$$k(x_1) \quad = \quad (x_1 - a_1) \cdot (x_1 - a_2) \cdots (x_1 - a_m).$$

Hence,

$$0 \quad = \quad \pi_1(k) \quad = \quad \pi_1(x_1 - a_1) \cdot \pi_1(x_1 - a_2) \cdots \pi_1(x_1 - a_m).$$

Now, $\mathcal{F}$ is a field, so it has no zero divisors. So at least *one* of the factors $\pi_1(x_1 - a_1)$, $\pi_1(x_1-a_2)$, ..., $\pi_1(x_1-a_m)$ must be zero. By reordering, we can assume $\pi_1(x_1-a_1) \;=\; 0$. In other words, the polynomial $\ell(x_1) = (x_1 - a_1)$ is in $\ker(\pi_1)$. so, set $c_1 = a_1$, and we're done. .... □ [Claim 4]

It follows from Claims 2 and 4 that:

Claim 5: *$\mathcal{M}$ contains a linear polynomial $\ell_1(\mathbf{x}) = (x_1 - c_1)$ for some $c_1 \in \mathbb{C}$.*

Now, we can repeat the argument of Claim 5 for $x_2, \ldots, x_n$ to conclude:

$\mathcal{M}$ contains linear polynomials $\ell_1(\mathbf{x}) = (x_1 - c_1)$, $\quad \ell_2(\mathbf{x}) = (x_2 - c_2), \ldots, \ell_n(\mathbf{x}) = (x_n - c_n)$, for some $c_1, \ldots, c_n \in \mathbb{C}$.

Thus, $\mathcal{M}$ contains the ideal generated by $\ell_1, \ldots, \ell_n$. But from **(b)**, we know that the ideal generated by $\ell_1, \ldots, \ell_n$ is the maximal ideal $\mathcal{M}_{\mathbf{c}}$ (where $\mathbf{c} = (c_1, \ldots, c_n)$). Thus $\mathcal{M}$ contains $\mathcal{M}_{\mathbf{c}}$, which means $\mathcal{M}$ must *equal* $\mathcal{M}_{\mathbf{c}}$. □

Further Reading: A good, elementary introduction to the Nullstellensatz and its relation to algebraic geometry is [1, §10.7]. More advanced discussions of the maximal spectrum can be found in [4, §7.5] or [6, II.5]. A thorough development of the subject is [7]. For applications to algebraic geometry, see [8, 3]. For applications to differential geometry, see [9]. The closely related spectral theory of abelian Banach algebras is discussed in [2, VII.8]

10.3.3 The Maximal Spectrum of a Coordinate Ring

Prerequisites: §8.6, §9.2, §10.3.2

Recall that the `Nullstellensatz`[2] says there is a bijective correspondence between the points in $\mathbb{C}^n$ and the maximal ideals of $\mathbb{C}[x_1,..., x_n]$ given by

$$\mathbb{C}^n \ni \mathbf{c} \mapsto \mathcal{M}_{\mathbf{c}} \in \overline{\mathsf{Spec}}\left(\mathbb{C}[x_1,..., x_n]\right) \qquad (\text{where } \mathcal{M}_{\mathbf{c}} \;=\; \{p \in \mathbb{C}[x_1,..., x_n]\ ;\ p(\mathbf{c}) \;=\; 0\}.)$$

We will now extend this to any complex algebraic variety. If $\mathbf{V} \subset \mathbb{C}^n$ is an algebraic variety, then we define

$$\overline{\mathsf{Spec}}_{\mathbf{V}}\left(\mathbb{C}[x_1,..., x_n]\right) \quad = \quad \{\mathcal{M}_v\ ;\ v \in \mathbf{V}\} \quad \subset \quad \overline{\mathsf{Spec}}\left(\mathbb{C}[x_1,..., x_n]\right).$$

[2] p. 149

Lemma 180 *Let $\mathbf{V} \subset \mathbb{C}^n$, with annihilator $\mathcal{A}nn(\mathbf{V})$. Then $\overline{\mathsf{Spec}}_{\mathbf{V}}(\mathbb{C}[x_1,\ldots,x_n])$ is just the set of maximal ideals in $\mathbb{C}[x_1,\ldots,x_n]$ which contain $\mathcal{A}nn(\mathbf{V})$. Formally:*

$$\overline{\mathsf{Spec}}_{\mathbf{V}}(\mathbb{C}[x_1,\ldots,x_n]) \quad = \quad \left\{\mathcal{M} \in \overline{\mathsf{Spec}}\left(\mathbb{C}[x_1,\ldots,x_n]\right) \ ; \ \mathcal{A}nn(\mathbf{V}) \subset \mathcal{M}\right\}.$$

Proof: **<u>Exercise 141</u>** ______ □

Corollary 181 Nullstellensatz for Complex Algebraic Varieties

Let $\mathbf{V} \subset \mathbb{C}^n$ be an algebraic variety, with coordinate ring $\mathcal{C}oord(\mathbf{V})$. Then there is a natural bijection between the points in $\mathbf{V}$ and the maximal ideals $\mathcal{C}oord(\mathbf{V})$, given by

$$\mathbb{C}^n \ni \mathbf{c} \mapsto \overline{\mathcal{M}}_{\mathbf{c}} \in \overline{\mathsf{Spec}}\left(\mathcal{C}oord(\mathbf{V})\right),$$

where $\overline{\mathcal{M}}_{\mathbf{c}} = \{f \in \mathcal{C}oord(\mathbf{V}) \ ; \ f(\mathbf{c}) = 0\}$.

Proof: Let $\mathcal{A}nn(\mathbf{V})$ be the annihilator of $\mathbf{V}$; Recall that Proposition 140 on page 122 says that

$$\mathcal{C}oord(\mathbf{V}) \quad \cong \quad \frac{\mathbb{C}[x_1,\ldots,x_n]}{\mathcal{A}nn(\mathbf{V})}.$$

Hence, there is a natural bijection $\quad \overline{\mathsf{Spec}}\left(\mathcal{C}oord(\mathbf{V})\right) \longleftrightarrow \overline{\mathsf{Spec}}\left(\dfrac{\mathbb{C}[x_1,\ldots,x_n]}{\mathcal{A}nn(\mathbf{V})}\right).$

The `Lattice Isomorphism Theorem` (p. 116) yields an order-preserving bijection:

$$\left\{\text{ideals of } \mathbb{C}[x_1,\ldots,x_n] \text{ containing } \mathcal{A}nn(\mathbf{V})\right\} \ni \mathcal{I} \quad \mapsto \quad \overline{\mathcal{I}} \in \left\{\text{ideals of } \ \frac{\mathbb{C}[x_1,\ldots,x_n]}{\mathcal{A}nn(\mathbf{V})}\right\}.$$

It follows that

$$\left(\ \mathcal{I} \text{ is maximal in } \mathbb{C}[x_1,\ldots,x_n]\ \right) \quad \Longleftrightarrow \quad \left(\ \overline{\mathcal{I}} \text{ is maximal in } \ \frac{\mathbb{C}[x_1,\ldots,x_n]}{\mathcal{A}nn(\mathbf{V})}\ \right)$$

(because the map is order-preserving). Hence, we get a bijection:

$$\left\{\mathcal{M} \in \overline{\mathsf{Spec}}\left(\mathbb{C}[x_1,\ldots,x_n]\right) \ ; \ \mathcal{A}nn(\mathbf{V}) \subset \mathcal{M}\right\} \ni \mathcal{M} \ \mapsto \ \overline{\mathcal{M}} \in \overline{\mathsf{Spec}}\left(\frac{\mathbb{C}[x_1,\ldots,x_n]}{\mathcal{A}nn(\mathbf{V})}\right)$$

But Lemma 180 says $\left\{\mathcal{M} \in \overline{\mathsf{Spec}}\left(\mathbb{C}[x_1,\ldots,x_n]\right) \ ; \ \mathcal{A}nn(\mathbf{V}) \subset \mathcal{M}\right\}$ is just $\overline{\mathsf{Spec}}_{\mathbf{V}}(\mathbb{C}[x_1,\ldots,x_n])$. Finally, there is a natural bijection between $\mathbf{V}$ and $\overline{\mathsf{Spec}}_{\mathbf{V}}(\mathbb{C}[x_1,\ldots,x_n])$, given by $v \mapsto \mathcal{M}_v$. □

10.4 The Zariski Topology

Prerequisites: §10.3

Let $\mathbf{X}$ be some kind of 'space' (eg. $\mathbf{X} = [0,1]$ or $\mathbf{X} = \mathbb{C}^n$), and let $\mathcal{R}$ be some ring of functions on $\mathbf{X}$ (eg. $\mathcal{R} = \mathcal{C}[0,1]$, or $\mathcal{R} = \mathbb{C}[x_1,\ldots,x_n]$). The `Correspondence Principle` (page 144) sets up a natural bijection between the *maximal ideals* in $\mathcal{R}$ and the *points* in $\mathbf{X}$. We will now see how we can turn this bijection into a *homeomorphism*, by endowing the maximal spectrum of $\mathcal{R}$ with a natural topological structure called the *Zariski topology*.

10.4.1 The Zariski Topology (Continuous Function Rings)

Prerequisites: §10.3.1 **Recommended:** §A

Throughout this section, let $\mathbf{X}$ be one of the following:

1. $\mathbf{X} = [0,1]$.

2. $\mathbf{X} \subset \mathbb{R}^n$ any compact subset.

3. $\mathbf{X}$ any compact metric space.

(whichever you're most comfortable with).

Let $\mathcal{C}(\mathbf{X})$ be the ring of continuous functions from $\mathbf{X}$ into $\mathbb{R}$. Recall Propositions 173 (for $\mathbf{X} = [0,1]$), 174 (for $\mathbf{X} \subset \mathbb{R}^n$) or 176 (for $\mathbf{X}$ a metric space), all of which say:

> *There is a natural bijective correspondence between the points in $\mathbf{X}$ and the maximal ideals in $\overline{\mathsf{Spec}}\left(\mathcal{C}(\mathbf{X})\right)$, given by $x \mapsto \mathcal{M}_x$, where $\mathcal{M}_x = \{f \in \mathcal{C}(\mathbf{X})\ ;\ f(x) = 0\}$.*

We will now see how this bijection also preserves the *topological* structure of $\mathbf{X}$.

If $\mathbf{C} \subset \mathbf{X}$ is any closed set, then we define a corresponding subset of $\overline{\mathsf{Spec}}\left(\mathcal{C}(\mathbf{X})\right)$ as follows:

$$\mathfrak{Zar}(\mathbf{C}) \quad = \quad \{\mathcal{M}_c\ ;\ c \in \mathbf{C}\}.$$

We call $\mathfrak{Zar}(\mathbf{C})$ the **Zariski closed set** corresponding to $\mathbf{C}$.

The **Zariski topology** on $\overline{\mathsf{Spec}}\left(\mathcal{C}(\mathbf{X})\right)$ is the topology where a subset $\mathfrak{C} \subset \overline{\mathsf{Spec}}\left(\mathcal{C}(\mathbf{X})\right)$ is considered closed if and only if $\mathfrak{C}$ is a Zariski closed set. It follows:

Corollary 182 *Endow $\overline{\mathsf{Spec}}\left(\mathcal{C}(\mathbf{X})\right)$ with the Zariski topology. Then the map $\mathbf{X} \ni x \mapsto \mathcal{M}_x \in \overline{\mathsf{Spec}}\left(\mathcal{C}(\mathbf{X})\right)$ is a homeomorphism.* ______ □

This wouldn't be very interesting if the Zariski topology on $\overline{\mathsf{Spec}}\left(\mathcal{C}(\mathbf{X})\right)$ was merely obtained by 'importing' the topology of $\mathbf{X}$ in the obvious way. However, we'll now construct the Zariski topology using only the *intrinsic* algebraic structure of $\mathcal{C}(\mathbf{X})$.

If $\mathcal{S} \subset \mathcal{C}(\mathbf{X})$ is any subset, then the **envelope** of $\mathcal{S}$ is the set of all maximal ideals in $\mathcal{C}(\mathbf{X})$ which contain $\mathcal{S}$:

$$\mathfrak{Env}\,(\mathcal{S}) \quad = \quad \left\{\mathcal{M} \in \overline{\mathsf{Spec}}\left(\mathcal{C}(\mathbf{X})\right) \ ; \ \mathcal{S} \subset \mathcal{M}\right\}.$$

Let $\mathfrak{C} \subset \overline{\mathsf{Spec}}\left(\mathcal{C}(\mathbf{X})\right)$ be some collection of maximal ideals. We call $\mathfrak{C}$ an **envelope** if $\mathfrak{C} = \mathfrak{Env}\,(\mathcal{S})$ for some subset $\mathcal{S} \subset \mathcal{C}(\mathbf{X})$.

Proposition 183 *The envelope subsets of* $\overline{\mathsf{Spec}}\left(\mathcal{C}(\mathbf{X})\right)$ *are exactly the Zariski closed sets.*

To be specific:

(a) *If* $\mathcal{S} \subset \mathcal{C}(\mathbf{X})$ *is any subset, then* $\mathfrak{Env}\,(\mathcal{S}) \;=\; \mathfrak{Zar}\left(\mathbb{V}\,(\mathcal{S})\right)$, *where*

$$\mathbb{V}\,(\mathcal{S}) \;=\; \{x \in \mathbf{X}\ ;\ f(x) = 0 \textit{ for all } f \in \mathcal{S}\}.$$

is the **variety** *of* $\mathcal{S}$ *(see §10.7.2). Hence, every envelope is a Zariski closed set.*

(b) *Conversely, if* $\mathbf{C} \subset \mathbf{X}$ *is a closed set, and* $\mathfrak{Zar}\,(\mathbf{C})$ *is the corresponding Zariski closed set, then* $\mathfrak{Zar}\,(\mathbf{C}) \;=\; \mathfrak{Env}\left(\mathcal{A}nn\,(\mathbf{C})\right)$, *where*

$$\mathcal{A}nn\,(\mathbf{C}) \quad = \quad \{f \in \mathcal{C}(\mathbf{X})\ ;\ f(c) = 0 \textit{ for all } c \in \mathbf{C}\}.$$

is the **annihilator** *of* $\mathbf{C}$ *(see §10.7.2). Hence, every Zariski closed set is an envelope.*

Proof: **Exercise 142** □

Thus, Zariski topology on $\overline{\mathsf{Spec}}\left(\mathcal{C}(\mathbf{X})\right)$ is a completely *algebraic* construction. In other words, if you just encountered $\mathcal{R} = \mathcal{C}(\mathbf{X})$ as an 'abstract ring', and you had no idea that it was the ring of continuous functions on some space, you could still construct the Zariski topology, based on purely algebraic information. Thus, the topology of $\mathbf{X}$ is *completely encoded* by the algebraic structure of $\mathcal{C}(\mathbf{X})$.

This encoding extends to morphisms: we will see that *continuous functions* between spaces $\mathbf{X}$ and $\mathbf{Y}$ correspond to *ring homomorphisms* between $\mathcal{C}(\mathbf{X})$ and $\mathcal{C}(\mathbf{Y})$.

Proposition 184 *Suppose* $\mathbf{X}$ *and* $\mathbf{Y}$ *are compact metric spaces*[3], *and let* $\phi : \mathbf{X} \longrightarrow \mathbf{Y}$ *be a continuous map.*

[3] For example, suppose $\mathbf{X} = [0, 1] = \mathbf{Y}$.

(a) *If $f : \mathbf{Y} \longrightarrow \mathbb{R}$ is continuous, then $f \circ \phi : \mathbf{X} \longrightarrow \mathbb{R}$ is also continuous.*

Thus, if $f \in \mathcal{C}(\mathbf{Y})$, then $f \circ \phi \in \mathcal{C}(\mathbf{X})$.

(b) *Define $\Phi : \mathcal{C}(\mathbf{Y}) \longrightarrow \mathcal{C}(\mathbf{X})$ by $\Phi(f) = f \circ \phi$. Then:*

1. *Φ is a ring homomorphism, and $\Phi(\mathbb{1}_{\mathbf{Y}}) = \mathbb{1}_{\mathbf{X}}$.*
2. $\Big(\,$*ϕ is injective*$\,\Big) \iff \Big(\,$*Φ is surjective*$\,\Big)$.
3. $\Big(\,$*ϕ is surjective*$\,\Big) \iff \Big(\,$*Φ is injective*$\,\Big)$.
4. $\Big(\,$*ϕ is a homeomorphism*$\,\Big) \iff \Big(\,$*Φ is a ring isomorphism*$\,\Big)$.

(c) *Suppose $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ are compact metric spaces, and $\phi : \mathbf{X} \longrightarrow \mathbf{Y}$ and $\psi : \mathbf{Y} \longrightarrow \mathbf{Z}$ are continuous functions. Let $\gamma = \psi \circ \phi : \mathbf{X} \longrightarrow \mathbf{Z}$, so that diagram* **(A)** *below commutes.*

Define homomorphisms $\Phi : \mathcal{C}(\mathbf{Y}) \longrightarrow \mathcal{C}(\mathbf{X})$, $\Psi : \mathcal{C}(\mathbf{Z}) \longrightarrow \mathcal{C}(\mathbf{Y})$, and $\Gamma : \mathcal{C}(\mathbf{Z}) \longrightarrow \mathcal{C}(\mathbf{X})$ as in part **(b)**. *Then $\Gamma = \Phi \circ \Psi$. In other words, diagram* **(B)** *commutes:*

$$\begin{tikzcd}
\mathbf{X} \arrow[r, "\phi"] \arrow[dr, "\gamma"'] & \mathbf{Y} \arrow[d, "\psi"] & & \mathcal{C}(\mathbf{X}) & \mathcal{C}(\mathbf{Y}) \arrow[l, "\Phi"'] \\
\textbf{(A)} & \mathbf{Z} & \Longrightarrow & \textbf{(B)} & \mathcal{C}(\mathbf{Z}) \arrow[u, "\Psi"'] \arrow[ul, "\Gamma"]
\end{tikzcd}$$

Proof: <u>**Exercise 143**</u> Hint: See the proof of Proposition 149 on page 126 □

The map Φ described in Proposition 184**(b)** is called the **pullback** of ϕ, and is usually writted as $\phi^\sharp$. Thus, in part **(c)**, $\Phi = \phi^\sharp$, $\Psi = \psi^\sharp$, and $\Gamma = \gamma^\sharp$, and part **(c)** could be reformulated:

$$(\psi \circ \phi)^\sharp \quad = \quad \phi^\sharp \circ \psi^\sharp.$$

This correspondence also goes the other way:

Proposition 185 *Suppose $\mathbf{X}$ and $\mathbf{Y}$ are compact metric spaces*[4], *and let $\Phi : \mathcal{C}(\mathbf{Y}) \longrightarrow \mathcal{C}(\mathbf{X})$ be a ring homomorphism, such that $\Phi(\mathbb{1}_{\mathbf{Y}}) = \mathbb{1}_{\mathbf{X}}$.*

(a) *If $\mathcal{M} \lhd \mathcal{C}(\mathbf{X})$ is any maximal ideal in $\mathcal{C}(\mathbf{X})$, then $\Phi^{-1}(\mathcal{M})$ is a maximal ideal in $\mathcal{C}(\mathbf{Y})$.*

(b) *Thus, we can define a map $\widetilde{\phi} : \overline{\mathsf{Spec}}\left(\mathcal{C}(\mathbf{X})\right) \longrightarrow \overline{\mathsf{Spec}}\left(\mathcal{C}(\mathbf{Y})\right)$ by $\widetilde{\phi}(\mathcal{M}) = \Phi^{-1}(\mathcal{M})$.*

[4] For example, suppose $\mathbf{X} = [0,1] = \mathbf{Y}$.

1. *$\widetilde{\phi}$ is a continuous function from $\overline{\mathsf{Spec}}\Big(\mathcal{C}(\mathbf{X})\Big)$ to $\overline{\mathsf{Spec}}\Big(\mathcal{C}(\mathbf{Y})\Big)$.*
2. $\Big($ *Φ is a ring monomorphism* $\Big)$ $\Longleftrightarrow$ $\Big($ *$\widetilde{\phi}$ is surjective* $\Big)$.
3. $\Big($ *Φ is a ring epimorphism* $\Big)$ $\Longleftrightarrow$ $\Big($ *$\widetilde{\phi}$ is injective* $\Big)$.
4. $\Big($ *Φ is a ring isomorphism* $\Big)$ $\Longleftrightarrow$ $\Big($ *$\widetilde{\phi}$ is a homeomorphism* $\Big)$.

(c) *Recall that the map $\mathbf{X} \ni x \mapsto \mathcal{M}_x \in \overline{\mathsf{Spec}}\Big(\mathcal{C}(\mathbf{X})\Big)$ is a homeomorphism. Define $\phi : \mathbf{X} \longrightarrow \mathbf{Y}$ by $\phi(x) = y$, where $y \in \mathbf{Y}$ is the unique point such that $\mathcal{M}_y = \widetilde{\phi}(\mathcal{M}_x)$. In other words, define ϕ so that the following diagram commutes:*

$$\begin{tikzcd}
\mathbf{X} \ni x \arrow[r, dashed, "\phi"] \arrow[d, leftrightarrow] & \phi(x) \in \mathbf{Y} \arrow[d, leftrightarrow] \\
\overline{\mathsf{Spec}}\Big(\mathcal{C}(\mathbf{X})\Big) \ni \mathcal{M}_x \arrow[r, dashed, "\widetilde{\phi}"'] & \Phi^{-1}(\mathcal{M}_x) \in \overline{\mathsf{Spec}}\Big(\mathcal{C}(\mathbf{Y})\Big)
\end{tikzcd}$$

Then:

1. *ϕ is well-defined, and a continuous function from $\mathbf{X}$ to $\mathbf{Y}$.*
2. $\Big($ *Φ is a ring monomorphism* $\Big)$ $\Longleftrightarrow$ $\Big($ *ϕ is surjective* $\Big)$.
3. $\Big($ *Φ is a ring epimorphism* $\Big)$ $\Longleftrightarrow$ $\Big($ *ϕ is injective* $\Big)$.
4. $\Big($ *Φ is a ring isomorphism* $\Big)$ $\Longleftrightarrow$ $\Big($ *ϕ is a homeomorphism* $\Big)$.
5. *For any $f \in \mathcal{C}(\mathbf{Y})$, $\Phi(f) = f \circ \phi$. In other words, $\Phi = \phi^\sharp$ is the* **pullback** *of ϕ, as described in Proposition 184.*

(d) *Suppose $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ are compact metric spaces, and that $\Phi : \mathcal{C}(\mathbf{Y}) \longrightarrow \mathcal{C}(\mathbf{X})$ and $\Psi : \mathcal{C}(\mathbf{Z}) \longrightarrow \mathcal{C}(\mathbf{Y})$ are ring homomorphisms. Let $\Gamma = \Phi \circ \Psi : \mathcal{C}(\mathbf{Z}) \longrightarrow \mathcal{C}(\mathbf{X})$, so that diagram* **(A)** *below commutes.*

Define $\phi : \mathbf{X} \longrightarrow \mathbf{Y}$, $\psi : \mathbf{Y} \longrightarrow \mathbf{Z}$, and $\gamma : \mathbf{X} \longrightarrow \mathbf{Z}$ as in part **(c)**. *Then $\gamma = \psi \circ \phi$. In other words, diagram* **(B)** *commutes:*

$$\begin{tikzcd}
\mathcal{C}(\mathbf{X}) & & \mathcal{C}(\mathbf{Y}) \arrow[ll, "\Phi"'] & & \mathbf{X} \arrow[rr, "\phi"] \arrow[rrd, "\gamma"'] & & \mathbf{Y} \arrow[d, "\psi"] \\
\textbf{(A)} & & \mathcal{C}(\mathbf{Z}) \arrow[u, "\Psi"'] \arrow[llu, "\Gamma"] & \Longrightarrow & \textbf{(B)} & & \mathbf{Z}
\end{tikzcd}$$

Proof: **Exercise 144** □

The map ϕ described in Proposition 185**(c)** is called the **push forward** of Φ, and is usually writted as $\Phi_\flat$. Thus, in part **(d)**, $\phi = \Phi_\flat$, $\psi = \Psi_\flat$, and $\gamma = \Gamma_\flat$, and part **(d)** could be reformulated:

$$(\Psi \circ \Phi)_\flat \quad = \quad \Phi_\flat \circ \Psi_\flat.$$

Part **(c5)** of Proposition 185 says that the operations of 'pulling back' and 'pushing forward' are inverse to each other. In other words:

For any continous function $\phi : \mathbf{X} \longrightarrow \mathbf{Y}$, and any ring homomorphism $\Phi : \mathcal{C}(\mathbf{Y}) \longrightarrow \mathcal{C}(\mathbf{X})$,

$$\Big(\ \Phi = \phi^\sharp\ \Big) \iff \Big(\ \phi = \Phi_\flat\ \Big)$$

10.4.2 (∗) The Zariski Topology (for an arbitrary Ring)

Prerequisites: §10.3, §A **Recommended:** §10.4.1

Let $\mathcal{R}$ be a ring. If $\mathcal{S} \subset \mathcal{R}$ is any subset, then the **Zariski envelope** of $\mathcal{S}$ is the set of all maximal ideals in $\mathcal{R}$ which contain $\mathcal{S}$:

$$\mathfrak{Env}(\mathcal{S}) \quad = \quad \left\{\mathcal{M} \in \overline{\mathsf{Spec}}(\mathcal{R})\ ;\ \mathcal{S} \subset \mathcal{M}\right\}$$

Example 186:

(a) If $\mathcal{R} = \mathbb{Z}$ and $\mathcal{S} = \{5\}$, then $\mathfrak{Env}(\mathcal{S}) = \{5\mathbb{Z}\}$. Also, if $\mathcal{S} = \{5, 10, 35\}$, then $\mathfrak{Env}(\mathcal{S}) = \{5\mathbb{Z}\}$. Finally, if $\mathcal{S} = 5\mathbb{Z}$, then $\mathfrak{Env}(\mathcal{S}) = \{5\mathbb{Z}\}$.

(b) If $\mathcal{R} = \mathbb{Z}$ and $\mathcal{S} = \{12\}$, then $\mathfrak{Env}(\mathcal{S}) = \{2\mathbb{Z},\ 3\mathbb{Z}\}$. If $\mathcal{S} = \{12, 24, 72\}$, then $\mathfrak{Env}(\mathcal{S}) = \{2\mathbb{Z},\ 3\mathbb{Z}\}$. If $\mathcal{S} = 12\mathbb{Z}$, then $\mathfrak{Env}(\mathcal{S}) = \{2\mathbb{Z},\ 3\mathbb{Z}\}$.

(c) In general, let $\mathcal{R} = \mathbb{Z}$, and suppose $n \in \mathbb{N}$ has prime factorization: $n = p_1^{\nu_1} \cdot p_2^{\nu_2} \cdots p_k^{\nu_k}$. If $\mathcal{S} = \{n\}$, or $\mathcal{S} = n\mathbb{Z}$, then $\mathfrak{Env}(\mathcal{S}) \ = \ \Big\{(p_1),\ (p_2),\ \ldots, (p_n)\Big\}$. (**Exercise 145**)

(d) Let $\mathcal{R} = \mathbb{R}[x]$, and let $p(x) \ = \ x^2 - 3x + 2 \ = \ (x-2)(x-1)$. If $\mathcal{S} = \{p(x)\}$, then $\mathfrak{Env}(\mathcal{S}) = \{\mathcal{M}_1,\ \mathcal{M}_2\}$, where $\mathcal{M}_1 = \{q(x) \in \mathbb{C}[x]\ ;\ q(1) = 0\}$ and $\mathcal{M}_2 = \{q(x) \in \mathbb{C}[x]\ ;\ q(2) = 0\}$.

(e) Let $\mathcal{R} = \mathbb{C}[x]$, and suppose $p(x) \in \mathbb{C}[x]$ factors into a product of linear polynomials:

$$p(x) \quad = \quad (x - c_1)^{k_1} \cdot (x - c_2)^{k_2} \cdots (x - c_n)^{k_n},$$

for some $c_1, \ldots, c_n \in \mathbb{C}$. If $\mathcal{S} = \{p\}$, then $\mathfrak{Env}(\mathcal{S}) \ = \ \Big\{\mathcal{M}_{c_1},\ \mathcal{M}_{c_2},\ \ldots, \mathcal{M}_{c_n}\Big\}$, where $\mathcal{M}_c = \{q(x) \in \mathbb{C}[x]\ ;\ q(c) = 0\}$. (**Exercise 146**)

(f) If $\mathcal{M} \lhd \mathcal{R}$ is a maximal ideal, then $\mathfrak{Env}(\{\mathcal{M}\}) = \{\mathcal{M}\}$.

(g) $\mathfrak{Env}(\mathcal{R}) = \emptyset$ (because no maximal ideal can contain the whole ring).

(h) $\mathfrak{Env}(\{0\}) = \overline{\mathsf{Spec}}(\mathcal{R})$ (because every maximal ideal contains 0).

Proposition 187 *Let $\mathcal{R}$ be any ring. Then*

(a) *The map $\mathfrak{Env}(\bullet)$ is inclusion reversing. In other words, if $\mathcal{S} \subset \mathcal{T}$, then $\mathfrak{Env}(\mathcal{T}) \subset \mathfrak{Env}(\mathcal{S})$.*

(b) *If $\mathcal{S} \subset \mathcal{R}$ and $\mathcal{T} \subset \mathcal{R}$, then $\mathfrak{Env}(\mathcal{S} \cap \mathcal{T}) = \mathfrak{Env}(\mathcal{S}) \cup \mathfrak{Env}(\mathcal{T})$.*

(c) *If $\mathcal{S}_1, \mathcal{S}_2, \ldots$ is any collection of subsets of $\mathcal{R}$, then $\mathfrak{Env}\left(\bigcup_{n=1}^{\infty} \mathcal{S}_n\right) = \bigcap_{n=1}^{\infty} \mathfrak{Env}(\mathcal{S}_n)$.*

Proof: Exercise 147 □

Let $\mathfrak{C} \subset \overline{\mathsf{Spec}}(\mathcal{R})$ be some collection of maximal ideals. We call $\mathfrak{C}$ a **Zariski subset** if $\mathfrak{C} = \mathfrak{Env}(\mathcal{S})$ for some subset $\mathcal{S} \subset \mathcal{R}$. It follows from Proposition 187:

Corollary 188

(a) *$\emptyset$ and $\overline{\mathsf{Spec}}(\mathcal{R})$ are Zariski subsets of $\overline{\mathsf{Spec}}(\mathcal{R})$.*

(b) *If $\mathfrak{C}_1$ and $\mathfrak{C}_2$ are Zariski sets, then so is $\mathfrak{C}_1 \cup \mathfrak{C}_2$.*

(c) *The intersection of any number of Zariski subsets is a Zariski subset.* □

In other words, Corollary 188 says that the collection of Zariski subsets obeys all the axioms required for the closed sets in a topological space (page 226). Hence, we define **Zariski topology** on $\overline{\mathsf{Spec}}(\mathcal{R})$ by the condition:

A subset $\mathfrak{C} \subset \overline{\mathsf{Spec}}(\mathcal{R})$ is considered closed if and only if $\mathfrak{C}$ is a Zariski set.

In §10.4.1, we saw that a ring homomomorphism from $\mathcal{C}(\mathbf{Y})$ to $\mathcal{C}(\mathbf{X})$ induced a continuous map from $\mathbf{X}$ to $\mathbf{Y}$, and vice versa. (Propositions 184 and 185). Since the `Correspondence Principle` identifies $\mathbf{X}$ with the maximal spectrum of $\mathcal{C}(\mathbf{X})$, we could also interpret this as a continuous map from $\overline{\mathsf{Spec}}\left(\mathcal{C}(\mathbf{X})\right)$ to $\overline{\mathsf{Spec}}\left(\mathcal{C}(\mathbf{Y})\right)$. The same principle holds for arbitrary rings...

Proposition 189 *Let $\Phi : \mathcal{R} \longrightarrow \mathcal{S}$ be a ring homomorphism, such that $\Phi(1_{\mathcal{R}}) = 1_{\mathcal{S}}$.*

(a) *If $\mathcal{M} \lhd \mathcal{S}$ is any maximal ideal in $\mathcal{S}$, then $\Phi^{-1}(\mathcal{M})$ is a maximal ideal in $\mathcal{R}$.*

(b) *Define* $\phi : \overline{\mathsf{Spec}}\,(\mathcal{S}) \longrightarrow \overline{\mathsf{Spec}}\,(\mathcal{R})$ *by* $\phi(\mathcal{M}) = \Phi^{-1}(\mathcal{M})$. *Then:*

1. ϕ *is a continuous function from* $\overline{\mathsf{Spec}}\,(\mathcal{S})$ *to* $\overline{\mathsf{Spec}}\,(\mathcal{R})$.
2. $\Big(\ \Phi$ *is a ring monomorphism* $\Big) \Longrightarrow \Big(\ \phi$ *is surjective* $\Big)$.
3. $\Big(\ \Phi$ *is a ring epimorphism* $\Big) \Longrightarrow \Big(\ \phi$ *is injective* $\Big)$.
4. $\Big(\ \Phi$ *is a ring isomorphism* $\Big) \Longrightarrow \Big(\ \phi$ *is a homeomorphism* $\Big)$.

(c) *Suppose* $\Phi : \mathcal{R} \longrightarrow \mathcal{S}$ *and* $\Psi : \mathcal{S} \longrightarrow \mathcal{T}$ *are ring homomorphisms. Let* $\Gamma = \Phi \circ \Psi : \mathcal{R} \longrightarrow \mathcal{T}$, *so that diagram* **(A)** *below commutes.*

Define $\phi : \overline{\mathsf{Spec}}\,(\mathcal{S}) \longrightarrow \overline{\mathsf{Spec}}\,(\mathcal{R})$, $\quad \psi : \overline{\mathsf{Spec}}\,(\mathcal{T}) \longrightarrow \overline{\mathsf{Spec}}\,(\mathcal{S})$, *and* $\gamma : \overline{\mathsf{Spec}}\,(\mathcal{T}) \longrightarrow \overline{\mathsf{Spec}}\,(\mathcal{R})$ *as above. Then* $\gamma = \psi \circ \phi$. *In other words, diagram* **(B)** *commutes:*

$$\begin{tikzcd}
\mathcal{R} \arrow[r, "\Phi"] \arrow[dr, "\Gamma"'] & \mathcal{S} \arrow[d, "\Psi"] & & \overline{\mathsf{Spec}}\,(\mathcal{R}) & & \overline{\mathsf{Spec}}\,(\mathcal{S}) \arrow[ll, "\phi"'] \\
\textbf{(A)} & \mathcal{T} & \Longrightarrow & \textbf{(B)} & & \overline{\mathsf{Spec}}\,(\mathcal{T}) \arrow[u, "\psi"'] \arrow[ull, "\gamma"]
\end{tikzcd}$$

Proof: <u>**Exercise 148**</u> □

The map ϕ described in Proposition 189**(b)** is called the **push forward** of Φ, and is usually writted as $\Phi_\flat$. Thus, in part **(c)**, $\phi = \Phi_\flat$, $\ \psi = \Psi_\flat$, and $\gamma = \Gamma_\flat$, and part **(c)** could be reformulated:

$$(\Psi \circ \Phi)_\flat \quad = \quad \Phi_\flat \circ \Psi_\flat$$

10.4.3 $(*)$ The Zariski Topology (on a field)

Prerequisites: §10.4.2, §10.3.2 **Recommended:** §10.4.1

In §10.4.1, we saw that that the Zariski topology on the ring of continuous functions $\mathcal{C}[0,1]$ completely encodes the topology of $[0,1]$. What if we instead consider a ring of *polynomial* functions?

Let $\mathcal{F}$ be a field (eg. either $\mathcal{F} = \mathbb{R}$ or $\mathcal{F} = \mathbb{C}$), and consider the ring $\mathcal{F}[x]$ of polynomials over $\mathcal{F}$. If $\mathcal{P} \subset \mathcal{F}[x]$ is any subset of $\mathcal{F}[x]$, then the **(algebraic) variety** defined by $\mathcal{P}$ is the subset of $\mathcal{F}$ defined:

$$\mathbb{V}\,(\mathcal{P}) \quad = \quad \{x \in \mathcal{F}\ ;\ p(x) = 0 \text{ for all } p \in \mathcal{P}\}.$$

Example 190:

(a) Suppose $\mathcal{F} = \mathbb{R}$ and $p(x) = x^2 - 3x + 2 = (x-2)(x-1)$. If $\mathcal{P} = \{p(x)\}$, then $\mathbb{V}(\mathcal{P}) = \{1,2\}$.

(b) More generally, suppose $\mathcal{P} = \{p(x)\}$ is a singleton set, where $p(x) \in \mathcal{F}[x]$ is a product of linear factors:

$$p(x) \quad = \quad (x-c_1)^{k_1} \cdot (x-c_2)^{k_2} \cdots (x-c_n)^{k_n},$$

Then $\mathbb{V}(\{p\}) = \{c_1, c_2, \ldots, c_n\}$ is just the set of *roots* of $p(x)$. ______________

The **Zariski topology** on $\mathcal{F}$ is the topology where a set is consider 'closed' if and only if it is an algebraic variety. That is:

$$\textit{For any } \mathcal{C} \subset \mathcal{F}, \quad \Big(\mathcal{C} \textit{ is closed} \Big) \iff \Big(\mathcal{C} = \mathbb{V}(\mathcal{S}) \textit{ for some subset } \mathcal{S} \subset \mathcal{F}[x] \Big).$$

Suppose that $\mathcal{F} = \mathbb{C}$; recall that Proposition 178 says:

> *There is a natural bijective correspondence between the points in $\mathbb{C}$ and the maximal ideals in $\overline{\mathsf{Spec}}\left(\mathbb{C}[x]\right)$, given by $c \mapsto \mathcal{M}_c$, where $\mathcal{M}_c = \{f \in \mathbb{C}[x]\ ;\ f(c) = 0\}$.*

We'll now see that this bijection is actually a *homeomorphism* between Zariski topologies:

Proposition 191

(a) *If $\mathcal{S} \subset \mathbb{C}[x]$ is any subset, then $\mathfrak{Env}(\mathcal{S}) = \{\mathcal{M}_c\ ;\ c \in \mathbb{V}(\mathcal{S})\}$.*

(b) *Endow $\mathbb{C}$ and $\overline{\mathsf{Spec}}\left(\mathbb{C}[x]\right)$ with their respective Zariski topologies. Then the map $\mathbb{C} \ni c \mapsto \mathcal{M}_c \in \overline{\mathsf{Spec}}\left(\mathbb{C}[x]\right)$ is a homeomorphism.*

Proof: **Exercise 149** ______________ □

The Zariski topology on $\mathbb{R}$ or $\mathbb{C}$ is much different from the familiar Euclidean topology:

Proposition 192 *The Zariski topology on $\mathcal{F}$ is just the* **cofinite** *topology. That is, for any $\mathcal{C} \subset \mathcal{F}$,*

$$\Big(\mathcal{C} \textit{ is (Zariski) closed} \Big) \iff \Big(\mathcal{C} \textit{ is finite} \Big).$$

To be precise:

(a) *For any $\mathcal{S} \subset \mathbb{C}[x]$, $\mathbb{V}(\mathcal{S})$ is a finite subset of $\mathbb{C}$.*

(b) *Conversely, every finite subset of $\mathbb{C}$ arises in this fashion. Given any finite subset $\{c_1, c_2, \ldots, c_n\} \subset \mathbb{C}$, let $p(x) = (x-c_1)\cdot(x-c_2)\cdots(x-c_n)$. Then $\mathbb{V}(\{p\}) = \{c_1, c_2, \ldots, c_n\}$.*

Proof: **Exercise 150** □

Thus, the Zariski topology has far fewer closed or open sets than the Euclidean topology. This has two consequences:

- It is 'easier' for a sequence to converge in the Zariski topology. In other words, if the sequence $\{x_1, x_2, \ldots\}$ converges to x in the Euclidean topology, then it automatically converges in the Zariski topology. However, the converse is not true.

- It is 'harder' for a function to be continuous in the Zariski topology. If $f : \mathcal{F} \longrightarrow \mathcal{F}$, then

$$\Big(\ f \text{ is continuous}\ \Big) \iff \Big(\ \text{either } f \text{ is constant, or } f \text{ is everywhere finite-to-one.}\ \Big)$$

  In particular, any polynomial function is continous.

Exercise 151 Verify these statements.

10.4.4 (∗) The Zariski Topology (on affine n-space)

Prerequisites: §10.4.2, §10.3.2 **Recommended:** §10.4.1, §10.4.3

Suppose $\mathbf{X} \subset \mathbb{R}^n$ was some compact subset. In §10.4.1, we saw that that the Zariski topology on the ring of continuous functions $\mathcal{C}(\mathbf{X})$ completely encodes the topology of $\mathbf{X}$. What if we instead consider a ring of *polynomial* functions?

Let $\mathcal{F}$ be a field (eg. either $\mathcal{F} = \mathbb{R}$ or $\mathcal{F} = \mathbb{C}$), and consider the ring $\mathcal{F}[x_1, \ldots, x_n]$ of polynomials in n variables over $\mathcal{F}$. If $\mathcal{P} \subset \mathcal{F}[x_1, \ldots, x_n]$ be any subset of $\mathcal{F}[x_1, \ldots, x_n]$, then the **(algebraic) variety** defined by $\mathcal{P}$ is the subset of $\mathcal{F}^n$ defined:

$$\mathbb{V}(\mathcal{P}) \quad = \quad \{\mathbf{x} \in \mathcal{F}^n \ ; \ p(\mathbf{x}) = 0 \text{ for all } p \in \mathcal{P}\}$$

Example 193: Suppose $\mathcal{F} = \mathbb{R}$, and consider $\mathbb{R}[x, y]$. Suppose $\mathcal{P} = \{p(x, y)\}$ is a singleton set, where $p(x, y) = x^2 + y^2 - 1$. Then

$$\mathbb{V}(\{p\}) \ = \ \big\{(x, y) \in \mathbb{R}^2 \ ; \ x^2 + y^2 = 1\big\}$$

is just the circle of radius 1 around zero.

The **Zariski topology** on $\mathcal{F}^n$ is the topology where a set is consider 'closed' if and only if it is an algebraic variety. That is:

$$\textit{For any } \mathcal{C} \subset \mathcal{F}^n, \quad \Big(\ \mathcal{C} \textit{ is closed}\ \Big) \iff \Big(\ \mathcal{C} = \mathbb{V}(\mathcal{S}) \textit{ for some subset } \mathcal{S} \subset \mathcal{F}[x_1, \ldots, x_n]\ \Big).$$

Suppose that $\mathcal{F} = \mathbb{C}$; recall that `Hilbert's Nullstellensatz` (Proposition 179) says:

There is a natural bijection between the points in $\mathbb{C}^n$ *and the maximal ideals in* $\overline{\mathsf{Spec}}\left(\mathbb{C}[x_1,\ldots,x_n]\right)$*, given by* $\mathbf{c} \mapsto \mathcal{M}_{\mathbf{c}}$*, where* $\mathcal{M}_{\mathbf{c}} = \{f \in \mathbb{C}[x_1,\ldots,x_n]\ ;\ f(\mathbf{c}) = 0\}$*.*

We'll now see that this bijection is actually a *homeomorphism* between Zariski topologies:

Proposition 194

(a) *If* $\mathcal{S} \subset \mathbb{C}[x_1,\ldots,x_n]$ *is any subset, then* $\mathfrak{Env}(\mathcal{S}) = \{\mathcal{M}_{\mathbf{c}}\ ;\ \mathbf{c} \in \mathbb{V}(\mathcal{S})\}$*.*

(b) *Endow* $\mathbb{C}^n$ *and* $\overline{\mathsf{Spec}}\left(\mathbb{C}[x_1,\ldots,x_n]\right)$ *with their respective Zariski topologies. Then the map* $\mathbb{C}^n \ni \mathbf{c} \mapsto \mathcal{M}_{\mathbf{c}} \in \overline{\mathsf{Spec}}\left(\mathbb{C}[x_1,\ldots,x_n]\right)$ *is a homeomorphism.*

Proof: **Exercise 152** □

10.4.5 (∗) More about the Zariski Topology

Prerequisites: §10.4.2, §10.7.1 **Recommended:** §10.4.1

Recall, when constructing the Zariski topology in §10.4.1, we specifically worked with compact spaces. Indeed, we saw in §10.3.1 that the `Correspondence Principle` actually fails if $\mathbf{X}$ is not compact (Example 175). One reason for this is that the Zariski topology itself is *always* compact. Hence, if $\mathbf{X}$ was a noncompact space, we could hardly expect a homemomorphism between $\mathbf{X}$ and $\overline{\mathsf{Spec}}(\mathcal{C}(\mathbf{X}))$...

Proposition 195 *Let* $\mathcal{R}$ *be any ring. Then* $\overline{\mathsf{Spec}}(\mathcal{R})$ *is compact in the Zariski topology.* □

If $\mathbf{X}$ and $\mathbf{Y}$ are compact metric spaces, then Propositions 184 and 185 showed that

$$\Big(\ \mathbf{X} \text{ and } \mathbf{Y} \text{ are homeomorphic}\ \Big) \iff \Big(\ \mathcal{C}(\mathbf{X}) \text{ and } \mathcal{C}(\mathbf{Y}) \text{ are isomorphic as rings}\ \Big)$$
$$\iff \Big(\ \overline{\mathsf{Spec}}(\mathcal{C}(\mathbf{X})) \text{ and } \overline{\mathsf{Spec}}(\mathcal{C}(\mathbf{Y})) \text{ are homeomorphic}\ \Big)$$

Thus, the maximal spectrum is a *complete invariant* of the ring $\mathcal{C}(\mathbf{X})$. It is natural to wonder whether a similar conclusion holds for the maximal spectrum of arbitrary rings. Unfortunately, it does not.

Proposition 196 *Let* $\mathcal{R}$ *be a ring, and let* $\mathcal{J} = \mathcal{J}\,0_{\mathcal{R}}$ *be its Jacobson radical (p. 167). Then* $\overline{\mathsf{Spec}}(\mathcal{R})$ *is homeomorphic to* $\overline{\mathsf{Spec}}(\mathcal{R}/\mathcal{J})$*.*

To be precise: let $\phi : \mathcal{R} \longrightarrow \mathcal{R}/\mathcal{J}$ *be the quotient homomorphism, and let* $\phi_\flat : \overline{\mathsf{Spec}}(\mathcal{R}/\mathcal{J}) \longrightarrow \overline{\mathsf{Spec}}(\mathcal{R})$ *be its push-forward*[5]*. Then* $\phi_\flat$ *is a homeomorphism.* □

[5] See Proposition 189 on page 159**(b)**.

Nevertheless, we can relate algebraic properties of $\mathcal{R}$ to topological properties of its maximal spectrum.

Proposition 197 *Suppose $\mathcal{R} = \mathcal{R}_1 \oplus \mathcal{R}_2 \oplus \ldots \oplus \mathcal{R}_n$. Then $\overline{\mathsf{Spec}}(\mathcal{R})$ is a disconnected union of n components:*

$$\overline{\mathsf{Spec}}(\mathcal{R}) \quad = \quad \overline{\mathsf{Spec}}(\mathcal{R}_1) \sqcup \overline{\mathsf{Spec}}(\mathcal{R}_2) \sqcup \ldots \sqcup \overline{\mathsf{Spec}}(\mathcal{R}_n).$$

Conversely, if $\overline{\mathsf{Spec}}(\mathcal{R})$ is connected, then $\mathcal{R}$ cannot be decomposed as a direct sum of two rings. □

Further Reading: The Zariski topology is discussed in [4, §7.5]; a thorough development is [7]. For applications to algebraic geometry, see [8, 3].

10.5 Prime Ideals

Prerequisites: §10.2

Recall from § 8.3 on page 108 that the set $2\mathbb{Z}$ of *even numbers* is an ideal in the ring $\mathbb{Z}$. This ideal has an interesting additional property. For any integers $n, m \in \mathbb{Z}$

$$\Big(\ n \cdot m \text{ is even}\ \Big) \iff \Big(\ \text{either } n \text{ is even or } m \text{ is even}\ \Big).$$

In other words, the product of two *odd* numbers *cannot* be even. We say that $2\mathbb{Z}$ is a *prime* ideal.

Let $\mathcal{R}$ be a ring and let $\mathcal{P} \triangleleft \mathcal{R}$ be an ideal. We say $\mathcal{P}$ is a **prime** ideal if, for any $r, s \in \mathcal{R}$,

$$\Big(\ r \cdot s \text{ is in } \mathcal{P}\ \Big) \iff \Big(\ \text{either } r \in \mathcal{P} \text{ or } s \in \mathcal{P}\ \Big).$$

Equivalently, $\mathcal{P}$ is prime if its complement $\mathcal{R} \setminus \mathcal{P}$ is *multiplicatively closed*. That is:

$$\text{for any } r, s \in \mathcal{R}, \qquad \Big(\ r \not\in \mathcal{P} \text{ and } s \not\in \mathcal{P}\ \Big) \Longrightarrow \Big(\ r \cdot s \not\in \mathcal{P}\ \Big).$$

Prime ideals get their name from their prototypical example: principal ideals of the integers generated by prime numbers.

Example 198: Let $p \in \mathbb{N}$ be prime. Then the principal ideal[6] $p\mathbb{Z} = \{pz\ ;\ z \in \mathbb{Z}\}$ is a prime ideal in $\mathbb{Z}$. To see this, suppose $n, m \in \mathbb{Z}$. Then:

$$\begin{aligned} \Big(\ n \cdot m \in p\mathbb{Z}\ \Big) &\iff \Big(\ p \text{ divides } n \cdot m\ \Big) \\ &\iff \Big(\ \text{either } p \text{ divides } n \text{ or } p \text{ divides } m\ \Big) \\ &\iff \Big(\ \text{either } n \in p\mathbb{Z} \text{ or } m \in p\mathbb{Z}\ \Big). \end{aligned}$$

[6] See Example ⟨117b⟩ on page 108.

These are not the *only* prime ideals in $\mathbb{Z}$. The zero ideal $\{0\}$ is also prime:

Example 199: Let $\mathcal{D}$ be any integral domain. Then the zero ideal $\{0\}$ is a prime ideal. To see this, recall that, if $d_1, d_2 \in \mathcal{D}$, then $\Big(d_1 \cdot d_2 = 0 \Big) \iff \Big(d_1 = 0 \text{ or } d_2 = 0 \Big)$. ______

Indeed, Example ⟨199⟩ is prototypical....

Proposition 200 *Let $\mathcal{R}$ be a commutative ring and let $\mathcal{I} \lhd \mathcal{R}$ be an ideal. Then*

$$\Big(\mathcal{I} \text{ is a prime ideal} \Big) \iff \Big(\mathcal{R}/\mathcal{I} \text{ is an integral domain} \Big).$$

Proof: Let $\widetilde{\mathcal{R}} = \mathcal{R}/\mathcal{I}$. For any elements r and s in $\mathcal{R}$, let the corresponding elements of $\widetilde{\mathcal{R}}$ be $\widetilde{r}$ and $\widetilde{s}$. Then

$$\Big(r \in \mathcal{I} \Big) \iff \Big(\widetilde{r} = \widetilde{0} \Big); \qquad \Big(s \in \mathcal{I} \Big) \iff \Big(\widetilde{s} = \widetilde{0} \Big);$$
$$\text{and} \qquad \Big(r \cdot s \in \mathcal{I} \Big) \iff \Big(\widetilde{r} \cdot \widetilde{s} = \widetilde{0} \Big).$$

$$\begin{aligned} \text{Hence,} \quad \Big(\mathcal{I} \text{ is a prime ideal} \Big) &\iff \Big(\text{If } r \cdot s \in \mathcal{I}, \text{ then } r \in \mathcal{I} \text{ or } s \in \mathcal{I} \Big) \\ &\iff \Big(\text{If } \widetilde{r} \cdot \widetilde{s} = \widetilde{0}, \text{ then } \widetilde{r} = \widetilde{0} \text{ or } \widetilde{s} = \widetilde{0} \Big) \\ &\iff \Big(\mathcal{R}/\mathcal{I} \text{ has no zero divisors.} \Big) \\ &\iff \Big(\mathcal{R}/\mathcal{I} \text{ is an integral domain.} \Big) \end{aligned}$$

______□

Corollary 201 *Let $\mathcal{R}$ be a commutative ring and let $\mathcal{I} \lhd \mathcal{R}$ be an ideal. Then*

$$\Big(\mathcal{I} \text{ is a maximal ideal} \Big) \Longrightarrow \Big(\mathcal{I} \text{ is a prime ideal} \Big).$$

Proof:
$$\begin{aligned} \Big(\mathcal{I} \text{ is a maximal ideal} \Big) &=\!{\scriptstyle (C167)}\!\Rightarrow \Big(\mathcal{R}/\mathcal{I} \text{ is a field.} \Big) \\ &=\!{\scriptstyle (X109c)}\!\Rightarrow \Big(\mathcal{R}/\mathcal{I} \text{ is an integral domain.} \Big) \\ &=\!{\scriptstyle (C200)}\!\Rightarrow \Big(\mathcal{I} \text{ is prime.} \Big) \end{aligned}$$

Here, (C167) is by Corollary 167 on page 142; (X109c) is by Example ⟨109c⟩ on page 101, and (C200) is by Corollary 200 above. ______□

Example 202: (Not *all* prime ideals are maximal.)

⟨a⟩ The *zero* ideal is prime in any integral domain (Example 199), but it is not maximal.

⟨b⟩ Let $\mathcal{R} = \mathbb{Z}[x]$ (Example ⟨98c⟩). Then the principal ideal (x) is prime, because $\mathbb{Z}[x]/(x) \cong \mathbb{Z}$ is an integral domain. However, (x) is not maximal, because it is contained in the ideal

$$(2, x) \quad = \quad \{2 \cdot p(x) + x \cdot q(x) \; ; \; p, q \in \mathbb{Z}[x]\}$$

(**Exercise 153**) ______

An equivalent formulation of Proposition 200 is:

Corollary 203 *Let $\mathcal{R}$ be a commutative ring and let $\mathcal{D}$ be an integral domain*

(a) *If $\phi : \mathcal{R} \longrightarrow \mathcal{D}$ is an epimorphism, then $\ker(\phi)$ is a prime ideal.*

(b) *All prime ideals of $\mathcal{R}$ arise in this fashion.* ______ □

Corollary 204 *Let $\phi : \mathcal{R} \longrightarrow \mathcal{S}$ be a ring homomorphism, and let $\mathcal{P} \lhd \mathcal{R}$ be an ideal such that $\ker(\Phi) \subset \mathcal{P}$. Then $\Big(\mathcal{P}$ is a prime ideal in $\mathcal{R} \Big) \iff \Big(\Phi(\mathcal{P})$ is a prime ideal in $\mathcal{S} \Big)$.*

Proof: **Exercise 154** Hint: Combine two applications of Proposition 200 with the `Chain Isomorphism Theorem` (Theorem 133 on page 116). ______ □

Examples 198 and 199 exhaust the prime ideals of $\mathbb{Z}$:

Proposition 205 *The prime ideals of $\mathbb{Z}$ are exactly:*

1. *The zero ideal $\{0\}$.*

2. *The principal ideals $p\mathbb{Z}$, where p is prime.*

Proof: We've already seen that the zero ideal and prime principal ideals are prime. Now we show there are no others.

Suppose $\mathcal{I} \lhd \mathbb{Z}$ was a *nonzero* ideal. Proposition 157 on page 133 says that $\mathcal{I} = (i)$, where i is the minimal positive element in $\mathcal{I}$. Thus, $\mathcal{R}/\mathcal{I} = \mathbb{Z}_{/i}$. But from Example ⟨109f⟩ on page 101, we know:

$$\Big(\mathbb{Z}_{/i} \text{ is an integral domain} \Big) \iff \Big(\mathbb{Z}_{/i} \text{ is a field} \Big).$$

Hence, combining this with Proposition 200, we obtain:

$$\Big((i) \text{ is prime} \Big) \iff \Big((i) \text{ is is maximal} \Big)$$

But Proposition 163 (p. 139) says: $\Big((i) \text{ is is maximal} \Big) \iff \Big(i \text{ is prime.} \Big)$ ______ □

10.6 The Prime Spectrum

Prerequisites: §10.5 **Recommended:** 10.3

Let $\mathcal{R}$ be a ring. The **prime spectrum** of $\mathcal{R}$ is the set of all prime ideals of $\mathcal{R}$:

$$\mathsf{Spec}'(\mathcal{R}) \quad = \quad \{\mathcal{P}\ ;\ \mathcal{P} \lhd \mathcal{R} \text{ is a prime ideal}\}.$$

Example 206: Let $\mathcal{R} = \mathbb{Z}$. Then Proposition 205 on the facing page says that the prime ideals of $\mathbb{Z}$ are the zero ideal, and the principal ideals $p\mathbb{Z}$, where p is prime. In other words,

$$\mathsf{Spec}'(\mathbb{Z}) \quad = \quad \{p\mathbb{Z}\ ;\ p \in \mathbb{N} \text{ a prime number, or } p = 0\}.$$

10.7 Radical, Variety, Annihilator

Let $\mathbf{X}$ be some kind of 'space' (eg. $\mathbf{X} = [0,1]$ or $\mathbf{X} = \mathbb{C}^n$), and let $\mathcal{R}$ be some ring of functions on $\mathbf{X}$ (eg. $\mathcal{R} = \mathcal{C}[0,1]$, or $\mathcal{R} = \mathbb{C}[x_1, \ldots, x_n]$). In this section, we will investigate an interesting duality between *subspaces* of $\mathbf{X}$ (eg. closed subsets or algebraic varieties), and certain *ideals* of $\mathcal{R}$. A key concept linking the two is that of the *radical*, an algebraic device for 'closing' an ideal.

10.7.1 The Jacobson Radical

Prerequisites: §10.2 **Recommended:** §10.7.3

Let $\mathcal{R}$ be a ring, and let $\mathcal{I} \subset \mathcal{R}$ be any subset (usually, $\mathcal{I}$ is an ideal). The **Jacobson radical** of $\mathcal{I}$ in $\mathcal{R}$ is the intersection of all maximal ideals in $\mathcal{R}$ which contain $\mathcal{I}$:

$$\mathcal{J}\mathcal{I} \quad = \quad \bigcap_{\substack{\mathcal{I}\subset\mathcal{M}\lhd\mathcal{R} \\ \mathcal{M} \text{ maximal}}} \mathcal{M}.$$

Example 207:

(a) If $\mathcal{M} \lhd \mathcal{R}$ is a maximal ideal, then $\mathcal{J}\mathcal{M} = \mathcal{M}$.

(b) Let $\mathcal{R} = \mathbb{Z}$, and let $\mathcal{I} = (12)$. Then $\mathcal{J}\mathcal{I} \ = \ (3) \cap (2) \ = \ (6)$.

(c) More generally, let $\mathcal{R} = \mathbb{Z}$, and let $i \in \mathbb{Z}$. Suppose i has prime factorization: $i = p_1^{\iota_1} \cdot p_2^{\iota_2} \cdots p_k^{\iota_k}$. Let $\mathcal{I} = (i)$ be the principal ideal generated by i. Then

$$\mathcal{J}\mathcal{I} \quad = \quad (p_1) \cap (p_2) \cap \ldots \cap (p_n) \quad = \quad (P), \text{ where } P \ = \ p_1 \cdot p_2 \cdots p_n.$$ (**Exercise 155**)

(d) Let $\mathcal{R} = \mathbb{R}[x]$, and let $p(x) \ = \ x^3 - 4x^2 + 5x - 2 \ = \ (x-1)^2 \cdot (x-2)$. Let $\mathcal{I} = (p)$ be the principal ideal generated by $p(x)$. Let $\mathcal{M}_1$ be the principal ideal $(x-1)$, and $\mathcal{M}_2$ be the principal ideal $(x-2)$. Then

$$\mathcal{J}\mathcal{I} \quad = \quad \mathcal{M}_1 \cap \mathcal{M}_2 \ = \ (x^2 - 3x + 2).$$

Also, observe that $\mathcal{M}_1 = \{q(x) \in \mathbb{R}[x]\ ;\ q(1) = 0\}$ and $\mathcal{M}_2 = \{q(x) \in \mathbb{R}[x]\ ;\ q(2) = 0\}$.

(e) More generally, suppose $\mathbb{F}$ is a field and $\mathcal{R} = \mathbb{F}[x]$. Let $p(x) \in \mathbb{F}[x]$ be a product of linear polynomials:

$$p(x) \quad = \quad (x - c_1)^{k_1} \cdot (x - c_2)^{k_2} \cdots (x - c_n)^{k_n},$$

for some $c_1, \ldots, c_n \in \mathbb{F}$. Let $\mathcal{I} = (p)$ be the principal ideal generated by $p(x)$. Then $\mathcal{J}\overline{\mathcal{I}} \;=\; \mathcal{M}_{c_1} \cap \mathcal{M}_{c_2} \cap \ldots \mathcal{M}_{c_n}$, where $\mathcal{M}_c \;=\; (x - c) \;=\; \{q(x) \in \mathbb{F}[x] \; ; \; q(c) = 0\}$ is the principal ideal generated by $x - c$. (**Exercise 156**)

We say that the ideal $\mathcal{I}$ is a **Jacobson ideal** if $\mathcal{J}\overline{\mathcal{I}} = \mathcal{I}$.

Example 208:

(a) Any maximal ideal is a Jacobson ideal.

(b) If $\mathcal{J} = \mathcal{J}\overline{\mathcal{I}}$, then $\mathcal{J}$ is a Jacobson ideal. (**Exercise 157**)

The **Jacobson radical** of the ring $\mathcal{R}$ itself is the Jacobson radical of the *zero* ideal —in other words, the intersection of *all* maximal ideals in $\mathcal{R}$:

$$\mathcal{J}\overline{0_{\mathcal{R}}} \quad = \quad \bigcap_{\substack{\mathcal{M} \lhd \mathcal{R} \\ \text{maximal}}} \mathcal{M}.$$

Example 209:

(a) Let $\mathcal{R} = \mathbb{Z}$. Then $\mathcal{J}\overline{0_{\mathbb{Z}}} \;=\; \{0\}$.

(b) Let $\mathcal{R} \;=\; \mathbb{Z}_{/12}$. Then $\mathcal{J}\overline{0_{\mathbb{Z}_{/12}}} \;=\; \{\bar{0},\ \bar{6}\}$.

(c) More generallly, let $\mathcal{R} = \mathbb{Z}_{/n}$, where n has prime factorization: $n = p_1^{\nu_1} \cdot p_2^{\nu_2} \cdots p_k^{\nu_k}$. Then

$$\mathcal{J}\overline{0_{\mathbb{Z}_{/n}}} \quad = \quad \{\overline{m} \; ; \; m = p_1^{\mu_1} \cdots p_k^{\mu_k}, \text{ where } 0 < \mu_1 < \nu_1, \ \ldots, \ 0 < \mu_k < \nu_k\}.$$

(**Exercise 158**)

Lemma 210 *Let $\mathcal{I} \lhd \mathcal{R}$, and let $\mathcal{Q} = \mathcal{R}/\mathcal{I}$ be the quotient ring, with quotient map $\pi : \mathcal{R} \longrightarrow \mathcal{Q}$. Then $\mathcal{J}\overline{\mathcal{I}} \;=\; \pi^{-1}\left(\mathcal{J}\overline{0_{\mathcal{Q}}}\right)$.*

Proof: **Exercise 159** □

Further Reading: [4, §4.2].

10.7.2 Variety and Annihilator (Continuous Function Rings)

Prerequisites: §10.7.1 **Recommended:** §10.7.4

Throughout the following discussion, let $\mathbf{I}$ be one of the following (whichever you are most comfortable with).

1. $\mathbf{I} = [0, 1]$.

2. $\mathbf{I} \subset \mathbb{R}^n$ any compact subset.

3. $\mathbf{I}$ any compact metric space.

Consider the ring $\mathcal{C}(\mathbf{I})$ of continuous functions from $\mathbf{I}$ into $\mathbb{R}$. For any subset $\mathbf{X} \subset \mathbf{I}$, we define the **annihilator** of $\mathbf{X}$ to be the set

$$\mathcal{A}nn\,(\mathbf{X}) \quad = \quad \{f \in \mathcal{C}(\mathbf{I})\ ;\ f(x) = 0 \text{ for all } x \in \mathbf{X}\}.$$

For example, if $\mathbf{X} = \{r\}$ is a singleton set, then

$$\mathcal{A}nn\,(\{r\}) \quad = \quad \{f \in \mathcal{C}(\mathbf{I})\ ;\ f(r) = 0\} \quad = \quad \mathcal{M}_r,$$

where $\mathcal{M}_r$ is the maximal ideal from Example ⟨169b⟩.

Lemma 211

(a) *For any* $\mathbf{X} \subset \mathbf{I}$, $\mathcal{A}nn\,(\mathbf{X})$ *is an ideal of* $\mathcal{C}(\mathbf{I})$.

(b) *The operation* $\mathcal{A}nn\,(\bullet)$ *is inclusion-reversing. That is: If* $\mathbf{X} \subset \mathbf{Y}$, *then* $\mathcal{A}nn\,(\mathbf{Y}) \subset \mathcal{A}nn\,(\mathbf{X})$.

(c) *For any* $\mathbf{X} \subset \mathbf{I}$ *and* $\mathbf{Y} \subset \mathbf{I}$, $\mathcal{A}nn\,(\mathbf{X} \cup \mathbf{Y}) = \mathcal{A}nn\,(\mathbf{X}) \cap \mathcal{A}nn\,(\mathbf{Y})$.

Proof: **Exercise 160** □

Let $\mathcal{S} \subset \mathcal{C}(\mathbf{I})$ be any subset. Recall that the **variety** of $\mathcal{S}$ is the subset $\mathbb{V}\,(\mathcal{S}) = \{i \in \mathbf{I}\ ;\ f(i) = 0 \text{ for all } f \in \mathcal{S}\}$.

Lemma 212

(a) *For any* $\mathcal{S} \subset \mathcal{C}(\mathbf{I})$, $\mathbb{V}\,(\mathcal{S})$ *is a closed subset of* $\mathbf{I}$.

(b) *The operation* $\mathbb{V}\,(\bullet)$ *is inclusion-reversing. That is: If* $\mathcal{S} \subset \mathcal{T}$, *then* $\mathbb{V}\,(\mathcal{T}) \subset \mathbb{V}\,(\mathcal{S})$.

Proof: **Exercise 161** □

Thus, we have defined two operations:

$$\begin{aligned} \mathcal{A} : \Big\{\text{subsets of } \mathbf{I}\Big\} &\longrightarrow \Big\{\text{ideals of } \mathcal{C}(\mathbf{I})\Big\} \\ \mathbb{V} : \Big\{\text{subsets of } \mathcal{C}(\mathbf{I})\Big\} &\longrightarrow \Big\{\text{closed subsets of } \mathbf{I}\Big\} \end{aligned}$$

These operations are 'inverses' of each other, in the following sense:

Lemma 213

(a) *Let* $\mathbf{X} \subset \mathbf{I}$. *Then* $\mathbb{V}\Big(\mathcal{A}m\,(\mathbf{X})\Big) \;=\; \overline{\mathbf{X}}$ *is the* topological closure *of* $\mathbf{X}$ *in* $\mathbf{I}$.

(b) *In particular,* $\Big(\ \mathbf{X}$ *is a closed subset* $\Big) \iff \Big(\ \mathbb{V}\Big(\mathcal{A}m\,(\mathbf{X})\Big) \;=\; \mathbf{X}\ \Big)$.

(c) *Let* $\mathcal{I} \lhd \mathcal{C}(\mathbf{I})$. *Then* $\mathcal{A}m\Big(\mathbb{V}\,(\mathcal{I})\Big) \;=\; \mathcal{J}\mathcal{I}$ *is the* Jacobson radical *of* $\mathcal{I}$ *in* $\mathcal{C}(\mathbf{I})$.

(d) *In particular,* $\Big(\ \mathcal{I}$ *is a Jacobson ideal* $\Big) \iff \Big(\ \mathcal{A}m\Big(\mathbb{V}\,(\mathcal{I})\Big) \;=\; \mathcal{I}\ \Big)$.

(e) *Let* $\mathfrak{C} \;=\; \{$*closed subsets* $\mathbf{X} \subset \mathbf{I}\}$ *and* $\mathfrak{J} \;=\; \{$*Jacobson ideals* $\mathcal{I} \lhd \mathcal{C}(\mathbf{I})\}$. *Then the maps*

$$\mathcal{A} : \mathfrak{C} \longrightarrow \mathfrak{J} \qquad \textit{and} \qquad \mathbb{V} : \mathfrak{J} \longrightarrow \mathfrak{C}$$

are bijections, and inverse to one another.

Proof: **Exercise 162** □

10.7.3 The Nil Radical

Prerequisites: §10.5 **Recommended:** §10.7.1

Let $\mathcal{R}$ be a ring. An element $r \in \mathcal{R}$ is called **nilpotent** if $r^n = 0_{\mathcal{R}}$ for some $n \in \mathbb{N}$.

Example 214:

(a) Let $\mathcal{R} = \mathbb{Z}_{/72}$ and let $r = \bar{6}$. Then r is nilpotent, because $(\bar{6})^3 \;=\; \overline{216} \;=\; \bar{0}$.

(b) More generally, let $\mathcal{R} = \mathbb{Z}_{/m}$, and suppose that m has prime factorization $m = p_1^{\mu_1} \cdot p_2^{\mu_2} \cdots p_k^{\mu_k}$. Let $r = \overline{p}_1 \cdot \overline{p}_2 \ldots \overline{p}_k$ in $\mathcal{R}$. Then r is nilpotent in $\mathbb{Z}_{/m}$. To see this, let $M = \max\{\mu_1, \ldots, \mu_k\}$. Then $r^M = \overline{p}_1^M \cdots \overline{p}_k^M \;=\; \bar{0}$, because $p_1^M \cdots p_k^M$ is divisible by m.

(c) Let $\mathcal{R} = \mathcal{M}_2(\mathbb{R})$, and let $\mathbf{M} = \left[\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right]$. Then $\mathbf{M}$ is nilpotent in $\mathcal{M}_2(\mathbb{R})$, because $\mathbf{M}^2 = \mathbf{0}$.

The **nilradical** of $\mathcal{R}$ is the set of all nilpotent elements:

$$\sqrt[*]{0_{\mathcal{R}}} \quad = \quad \{r \in \mathcal{R}\ ;\ r^n = 0_{\mathcal{R}} \text{ for some } n \in \mathbb{N}\}.$$

Proposition 215 (Krull) *Let $\mathcal{R}$ be a commutative ring.*

(a) $\sqrt[*]{0_{\mathcal{R}}}$ *is a prime ideal in $\mathcal{R}$.*

(b) $\sqrt[*]{0_{\mathcal{R}}}$ *is the intersection of all prime ideals in $\mathcal{R}$:* $\sqrt[*]{0_{\mathcal{R}}} \;=\; \bigcap_{\substack{\mathcal{P} \triangleleft \mathcal{R} \\ \mathcal{P} \text{ prime}}} \mathcal{P}.$

Proof: **(a)** $\sqrt[*]{0_{\mathcal{R}}}$ **is closed under addition:** **<u>Exercise 163</u>** .

$\sqrt[*]{0_{\mathcal{R}}}$ **is an ideal:** Suppose $z \in \sqrt[*]{0_{\mathcal{R}}}$, and let $r \in \mathcal{R}$ be arbitrary. We claim that $r \cdot z \in \sqrt[*]{0_{\mathcal{R}}}$ also. To see this, find $n \in \mathbb{N}$ such that $z^n = 0$. Then $(r \cdot z)^n \underset{(c)}{=} r^n z^n \;=\; r^n 0 \;=\; 0$ (where (c) is because $\mathcal{R}$ is commutative.)

$\sqrt[*]{0_{\mathcal{R}}}$ **is a prime ideal:** We'll prove this in the case when $\mathcal{R}$ is an integral domain (the general proof is more complicated and involves Zorn's lemma).

Suppose $ab \in \sqrt[*]{0_{\mathcal{R}}}$. Then $(ab)^n = 0$ for some $n \in \mathbb{N}$. But $(ab)^n = a^n b^n$, and $\mathcal{R}$ has no zero divisors, so either $a^n = 0$ or $b^n = 0$; hence, either $a \in \sqrt[*]{0_{\mathcal{R}}}$ or $b \in \sqrt[*]{0_{\mathcal{R}}}$.

(b) "$\supset$" Since $\sqrt[*]{0_{\mathcal{R}}}$ itself is a prime ideal, it is clear that $\bigcap_{\substack{\mathcal{P} \triangleleft \mathcal{R} \\ \mathcal{P} \text{ prime}}} \mathcal{P} \;\subset\; \sqrt[*]{0_{\mathcal{R}}}$.

"$\subset$" It suffices to show that $\sqrt[*]{0_{\mathcal{R}}} \;\subset\; \mathcal{P}$ for every prime ideal $\mathcal{P} \triangleleft \mathcal{R}$. To show this, suppose $z \in \sqrt[*]{0_{\mathcal{R}}}$. Then $z^n = 0$ for some $n \in \mathbb{N}$. But then $z^n \in \mathcal{P}$ (because $0 \in \mathcal{P}$). Hence, since $\mathcal{P}$ is prime, it follows that $z \in \mathcal{P}$. □

If $\mathcal{I} \triangleleft \mathcal{R}$ is any ideal, then the **nilradical** of $\mathcal{I}$ in $\mathcal{R}$ is defined:

$$\sqrt[*]{\mathcal{I}} \quad = \quad \{r \in \mathcal{R} \,;\; r^n \in \mathcal{I} \text{ for some } n \in \mathbb{N}\}.$$

Example 216:

(a) If $\mathcal{I} = \{0\}$ is the zero ideal, then the nilradical of $\mathcal{I}$ is just the nilradical of $\mathcal{R}$.

(b) If $\mathcal{P} \triangleleft \mathcal{R}$ is a prime ideal, then $\sqrt[*]{\mathcal{P}} = \mathcal{P}$. (**<u>Exercise 164</u>**)

(c) Let $\mathcal{R} = \mathbb{Z}$, and let $\mathcal{I} = (12)$. Then $\sqrt[*]{\mathcal{I}} \;=\; (6) \;=\; (3) \cap (2)$.

(d) Let $\mathcal{R} = \mathbb{Z}$, and suppose $i \in \mathbb{N}$ has prime factorization: $i = p_1^{\iota_1} \cdot p_2^{\iota_2} \cdots p_k^{\iota_k}$. Let $\mathcal{I} = (i)$ be the principal ideal generated by i. Then:

$$\begin{aligned} \sqrt[*]{\mathcal{I}} \;&=\; \left\{\text{all numbers of the form } p_1^{j_1} p_2^{j_2} \cdots p_n^{j_n} \;;\; \text{for some } j_1, j_2, \ldots, j_n > 0\right\} \\ &=\; (P) \qquad \text{where } P = p_1 \cdot p_2 \cdots p_k \\ &=\; (p_1) \cap (p_2) \cap \ldots \cap (p_k) \qquad (\textbf{\underline{Exercise 165}}) \end{aligned}$$

(e) Let $\mathcal{R} = \mathbb{C}[x]$, and let $\mathcal{I} = (x^2 - 1)$ be the principal ideal generated by $p(x) = x^2 - 1$. Then

$$\sqrt[*]{\mathcal{I}} \quad = \quad (x+1) \cap (x-1) \quad = \quad \mathcal{M}_{(-1)} \cap \mathcal{M}_1,$$

where $\mathcal{M}_1 \quad = \quad (x-1) \quad = \quad \{q(x) \in \mathbb{C}[x]\ ;\ q(1) = 0\}$, and $\mathcal{M}_{(-1)} \quad = \quad (x+1) \quad = \quad \{q(x) \in \mathbb{C}[x]\ ;\ q(-1) = 0\}$.

(f) Let $\mathcal{R} = \mathbb{C}[x]$, and suppose $p(x) \in \mathbb{C}[x]$ factors into a product of linear polynomials:

$$p(x) \quad = \quad (x - c_1)^{k_1} \cdot (x - c_2)^{k_2} \cdots (x - c_n)^{k_n},$$

for some $c_1, \ldots, c_n \in \mathbb{C}$. Let $\mathcal{I} = (p)$ be the principal ideal generated by $p(x)$. Then

$$\sqrt[*]{\mathcal{I}} \quad = \quad (x - c_1) \cap (x - c_2) \cap (x - c_n) \quad = \quad \mathcal{M}_{c_1} \cap \mathcal{M}_{c_2} \cap \ldots \mathcal{M}_{c_n},$$

where $\mathcal{M}_c \ = \ (x - c_1) \ = \ \{q(x) \in \mathbb{C}[x]\ ;\ q(c) = 0\}$. (**Exercise 166**)

Proposition 217 *Let $\mathcal{R}$ be a commutative ring, and let $\mathcal{I} \lhd \mathcal{R}$.*

(a) *$\sqrt[*]{\mathcal{I}}$ is a prime ideal in $\mathcal{R}$.*

(b) *$\sqrt[*]{\mathcal{I}}$ is the intersection of all prime ideals in $\mathcal{R}$ which contain $\mathcal{I}$:* $\quad \sqrt[*]{\mathcal{I}} \ = \ \bigcap_{\substack{\mathcal{I} \subset \mathcal{P} \lhd \mathcal{R} \\ \mathcal{P} \text{ prime}}} \mathcal{P}.$

(c) *Let $\mathcal{Q} = \mathcal{R}/\mathcal{I}$ be the quotient ring, with quotient map $\pi : \mathcal{R} \longrightarrow \mathcal{Q}$. Then* $\sqrt[*]{\mathcal{I}} \ = \ \pi^{-1}\left(\sqrt[*]{0_{\mathcal{Q}}}\right)$.

Proof: **Exercise 167** Hint: Prove part **(c)** first; then prove parts **(a)** and **(b)** by combining part **(c)** with Proposition 215, Corollary 204 on page 166 and the `Lattice Isomorphism Theorem` (Theorem 134 on page 116) □

We say that an ideal $\mathcal{I}$ is **radical** if $\sqrt[*]{\mathcal{I}} = \mathcal{I}$.

Example 218:

(a) Any prime ideal is radical.

(b) If $\mathcal{N} = \sqrt[*]{\mathcal{I}}$, then $\mathcal{N}$ is radical. (**Exercise 168**)

We say that the ring $\mathcal{R}$ is **perfect** if $\mathcal{R}$ has no nilpotent elements —ie. $\sqrt[*]{0_{\mathcal{R}}} \ = \ \{0\}$. In other words, $\mathcal{R}$ is perfect if the zero ideal $\{0_{\mathcal{R}}\}$ is radical.

Lemma 219 *Let $\mathcal{I} \lhd \mathcal{R}$. Then:* $\Big(\ \mathcal{I}$ *is radical* $\Big) \iff \Big(\ \mathcal{R}/\mathcal{I}$ *is perfect* $\Big)$.

Proof: **Exercise 169** □

Further Reading: [4, §7.1].

10.7.4 Variety and Annihilator (Polynomial Rings)

Prerequisites: §10.7.3 **Recommended:** §10.7.2

Consider the ring $\mathbb{C}[x_1, \ldots, x_n]$ of polynomial functions on $\mathbb{C}^n$. For any subset $\mathbf{X} \subset \mathbb{C}^n$, we define the **annihilator** of $\mathbf{X}$ to be the set

$$\mathcal{A}nn\,(\mathbf{X}) \quad = \quad \{p \in \mathbb{C}[x_1, \ldots, x_n] \; ; \; p(x) = 0 \text{ for all } x \in \mathbf{X}\}.$$

For example, if $\mathbf{X} = \{c\}$ is a singleton set, then

$$\mathcal{A}nn\,(\{c\}) \quad = \quad \{p \in \mathbb{C}[x_1, \ldots, x_n] \; ; \; p(c) = 0\} \quad = \quad \mathcal{M}_c,$$

where $\mathcal{M}_c$ is the maximal ideal from Example ⟨169b⟩.

Lemma 220

(a) *For any* $\mathbf{X} \subset \mathbb{C}^n$, $\mathcal{A}nn\,(\mathbf{X})$ *is an ideal of* $\mathbb{C}[x_1, \ldots, x_n]$.

(b) *The operation* $\mathcal{A}nn\,(\bullet)$ *is inclusion-reversing. That is: If* $\mathbf{X} \subset \mathbf{Y}$*, then* $\mathcal{A}nn\,(\mathbf{Y}) \subset \mathcal{A}nn\,(\mathbf{X})$.

(c) *For any* $\mathbf{X} \subset \mathbb{C}^n$ *and* $\mathbf{Y} \subset \mathbb{C}^n$, $\mathcal{A}nn\,(\mathbf{X} \cup \mathbf{Y}) = \mathcal{A}nn\,(\mathbf{X}) \cap \mathcal{A}nn\,(\mathbf{Y})$.

Proof: <u>**Exercise 170**</u> □

Let $\mathcal{S} \subset \mathbb{C}[x_1, \ldots, x_n]$ be any subset. Recall that the **variety** of $\mathcal{S}$ is the subset $\mathbb{V}(\mathcal{S}) = \{\mathbf{c} \in \mathbb{C}^n \; ; \; p(\mathbf{c}) = 0 \text{ for all } p \in \mathcal{S}\}$.

Lemma 221 *The operation* $\mathbb{V}(\bullet)$ *is inclusion-reversing. That is: If* $\mathcal{S} \subset \mathcal{T}$*, then* $\mathbb{V}(\mathcal{T}) \subset \mathbb{V}(\mathcal{S})$.

Proof: <u>**Exercise 171**</u> □

Thus, we have defined two operations:

$$\begin{aligned} \mathcal{A} : \Big\{\text{subsets of } \mathbb{C}^n\Big\} &\longrightarrow \Big\{\text{ideals of } \mathbb{C}[x_1, \ldots, x_n]\Big\} \\ \mathbb{V} : \Big\{\text{subsets of } \mathbb{C}[x_1, \ldots, x_n]\Big\} &\longrightarrow \Big\{\text{varieties of } \mathbb{C}^n\Big\} \end{aligned}$$

In a certain sense, the operation $\mathbb{V}$ is the 'inverse' of $\mathcal{A}$:

Lemma 222

(a) *Let* $\mathbf{X} \subset \mathbb{C}^n$*. Then* $\mathbb{V}\left(\mathcal{A}{\scriptstyle nn}\,(\mathbf{X})\right)$ *is the smallest algebraic variety in* $\mathbb{C}^n$ *which contains* $\mathbf{X}$ *(we call this the 'Zariski closure' of* $\mathbf{X}$*).*

(b) *In particular,* $\Big(\ \mathbf{X}$ *is an algebraic variety* $\Big) \iff \Big(\ \mathbb{V}\left(\mathcal{A}{\scriptstyle nn}\,(\mathbf{X})\right) = \mathbf{X}\ \Big).$

Proof: <u>**Exercise 172**</u> □

It would be nice if we could also say that the operation $\mathcal{A}$ was the inverse of $\mathbb{V}$. The precise statement of this is:

Theorem 223 Hilbert's Nullstellensatz (Radical Version)[7]

Let $\mathcal{I} \lhd \mathbb{C}[x_1, \ldots, x_n]$*. Then*

(a) $\mathcal{A}{\scriptstyle nn}\left(\mathbb{V}(\mathcal{I})\right) = \sqrt[*]{\mathcal{I}}$ *is the* nil radical *of* $\mathcal{I}$ *in* $\mathbb{C}[x_1, \ldots, x_n]$.

(b) *In particular,* $\Big(\ \mathcal{I}$ *is a radical ideal* $\Big) \iff \Big(\ \mathcal{A}{\scriptstyle nn}\left(\mathbb{V}(\mathcal{I})\right) = \mathcal{I}\ \Big).$ □

Corollary 224 *Let* $\mathfrak{Z} = \{$*algebraic varieties* $\mathbf{X} \subset \mathbb{C}^n\}$ *and* $\mathfrak{N} = \{$*radical ideals* $\mathcal{I} \lhd \mathbb{C}[x_1, \ldots, x_n]\}$*. Then the maps*

$$\mathcal{A} : \mathfrak{Z} \longrightarrow \mathfrak{N} \qquad \text{and} \qquad \mathbb{V} : \mathfrak{N} \longrightarrow \mathfrak{Z}$$

are bijections, and inverse to one another. □

Further Reading: [4, §7.12] or [6, §X.2]

[7] This is actually the original version of the Nullstellensatz, so it is sometimes called the 'classical' Nullstellensatz.

Part III

Fields

Chapter 11

Field Theory

11.1 Compass & Straight-Edge Constructions I

In Mesopotamia, India, and especially Greece, ancient mathematicians perfected the art of constructing precise geometric figures in the plane, using only a compass and straight-edge. A *compass* is a length of string with a stylus at one end; using it, you can draw a circle of any radius around any point in the plane. A *straight-edge* is simply a long, perfectly straight object (eg. a tightly stretched string); using it, you can draw a straight line passing through any chosen pair of points. In societies lacking computers or precision-engineered tools, these were the only devices available for drafting and engineering.

We will provide three illustrations of these techniques: bisecting a line segment, bisecting an angle, and constructing a regular hexagon.

Example A *Bisecting a Line Segment*

Consider a line segment $\overline{x\ y}$ between two points x and y in the plane (Figure 11.1A). The construction proceeds as follows:

1. Construct a circle of radius r about x, as in Figure 11.1B. Here, r is any value greater than half the distance from x to y.

2. Construct another circle with the same radius r, but centred at y, as in Figure 11.1C.

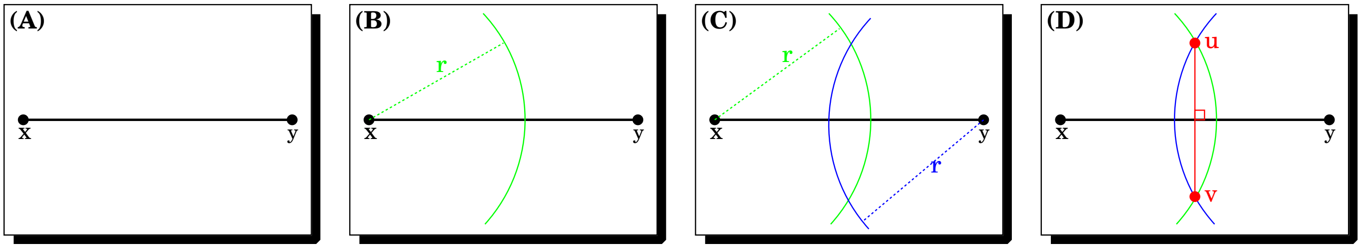


Figure 11.1: Bisecting a line segment.

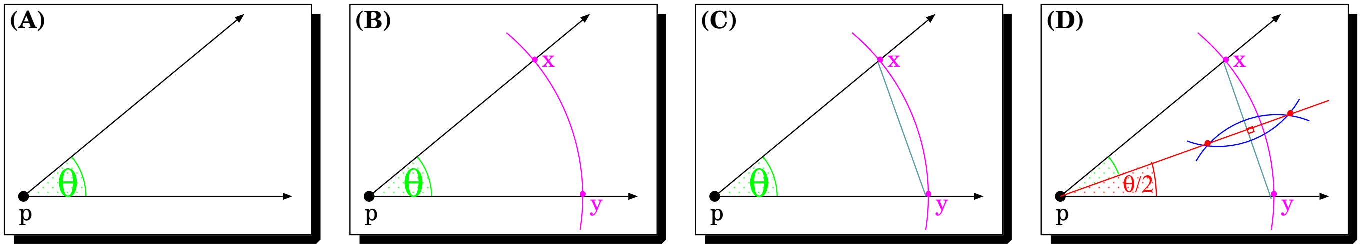


Figure 11.2: Bisecting an angle

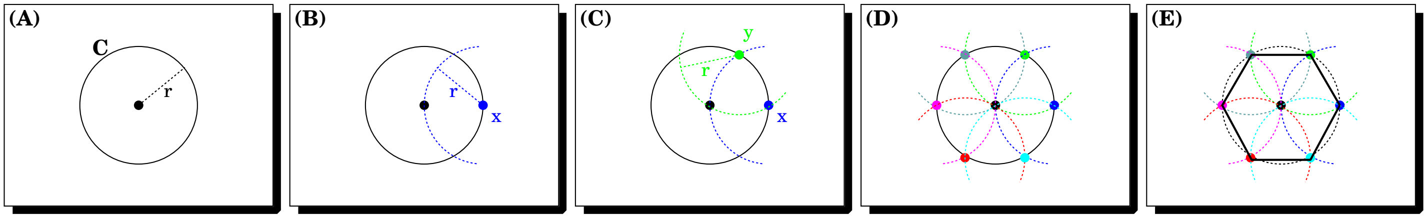


Figure 11.3: Constructing a regular hexagon

3. Let u and v be the points where the two circles intersect, as in Figure 11.1D. Then the line $\overline{u\ v}$ is perpendicular to $\overline{x\ y}$, and intersects it exactly halfway between x and y.

Example B *Bisecting an Angle*
Consider an angle θ formed by two intersecting lines in the plane, as in Figure 11.2A. The construction proceeds as follows:

1. Let p be the place where the two lines intersect. Construct a circle of any radius about p, as in Figure 11.2B.

2. Let x and y be the places where this circle intersects the two lines forming θ. Form the line segment $\overline{x\ y}$, as in Figure 11.2C.

3. Using the previous construction, bisect the line $\overline{x\ y}$. Draw a line from p to the midpoint of $\overline{x\ y}$, as in Figure 11.2D. The angle formed by this line is $\theta/2$

Example C *Constructing a Regular Hexagon*
We will construct a regular hexagon whose sides all have length r.

1. Draw a circle **C** of radius r as in Figure 11.3A.

2. Let x be an arbitrary point on the circle, and draw another circle of radius r around x, as in 11.3B.

3. Let y be a point where the new circle intersects **C**; draw another circle of radius r around y, as in 11.3C.

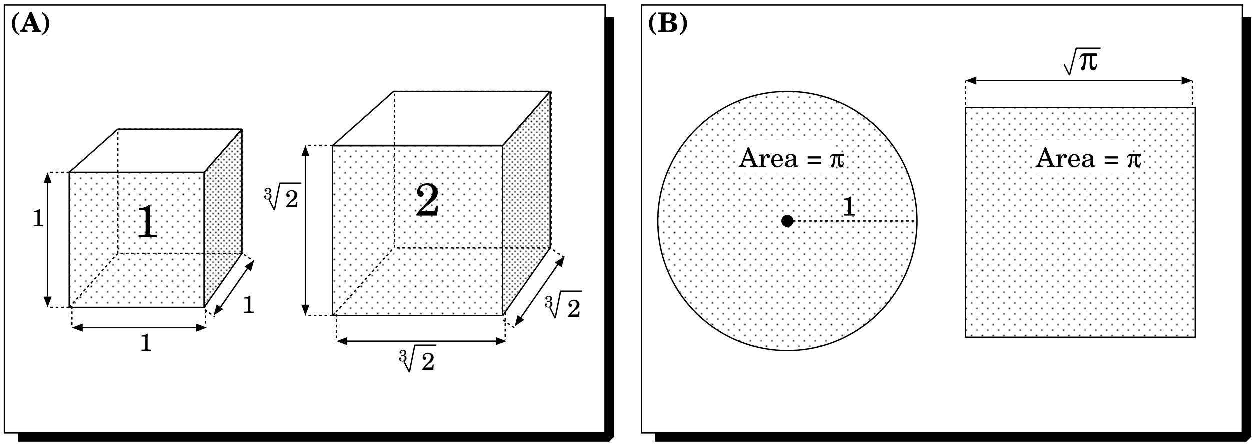


Figure 11.4: **(A)** Doubling the cube. **(B)** Squaring the circle

4. Proceeding in this fashion, draw four more circles of radius r, each centered at the place where the last circle intersects **C** (Figure 11.3D)

5. We now have six equally spaced circles of radius r around **C**. Their centres are six equally spaced points on **C**. Connect these points to get a regular hexagon, as in Figure 11.3E.

Four Outstanding problems: The Greeks developed a vast and sophisticated technology of geometric constructions. However, for more than two thousand years, four problems have remained unsolved:

I. Trisecting an angle: We saw how to bisect an angle in Example B. Is there a similar construction to divide an angle into three equal parts?

II. Constructing arbitrary regular polygons: We saw how to construct a regular hexagon in Example C. By using only three of the six vertices, we also obtain a regular (ie. equilateral) triangle. By bisecting the angles of the hexagon (as in Example B), we can also construct a dodecagon (12 sides), a 24-gon, etc.

By constructing perpendicular bisectors as in Example A, we can also construct a regular square; by bisecting its angles, we get an octagon, a 16-gon, etc.

A fairly elaborate construction yields a regular pentagon. Angle bisection yields a regular decagon (10 sides), a 20-gon, etc.

If you could trisect angles, then you could get a regular nonagon (9 sides) from an equilateral triangle. Is there another way? Can you construct a heptagon (7 sides)? An 11-gon? In general, is there a way to construct a regular N-gon for any N?

III. Doubling the Cube: If we build a cube with sides 1 metre long, then the interior of cube has a volume of 1 cubic metre. Is it possible to build a cube whose volume is exactly

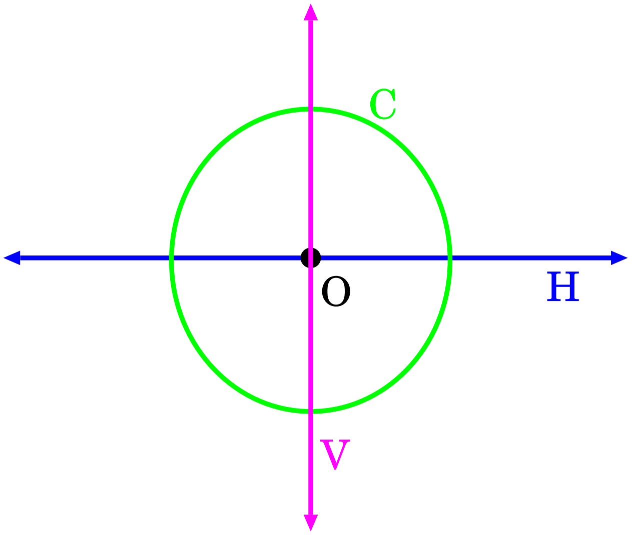


Figure 11.5: The starting point: an origin point **O**, a unit circle **C**, and lines **H** and **V**.

2 cubic metres, as in Figure 11.4A? That is: can you construct a cube whose sides have length $\sqrt[3]{2}$?

IV. Squaring the Circle: If we draw a circle with a radius of 1 metre, then its interior has an area of π square metres. Is it possible to build a square with exactly the same area, as in Figure 11.4B? That is: can you construct a square whose sides have length $\sqrt{\pi}$?

We will see that the answer to these four questions is "No". This does not mean we don't yet *know* how to do these constructions. It is a much more sweeping statement: we can prove that it is simply *impossible* to perform these constructions using only a compass and straight-edge. While this may a depressing and negative result, it has one important consequence: we no longer need waste mental effort trying to solve these problems.

To show that problems **I** to **IV** are unsolvable, we will use field theory. We will show that the set of all possible lengths constructable using compass and straight-edge forms a field, $\mathbb{K}$. We will then show that lengths such as $\sqrt[3]{2}$ or $\sqrt{\pi}$ simply cannot be elements of $\mathbb{K}$. To do this will require understanding the theory of fields and field extensions.

To begin with, let's more precisely define what a compass and straightedge (C&S) construction consists of. We begin with an unmarked Euclidean plane. As shown in Figure 11.5, We set down a single point **O**, which we consider the 'origin', and we decide on a unit of length measurement by drawing a circle **C** around **O**, and *defining* this to be the 'unit circle'. We then draw a single line **H** through **O**, which we *define* to be 'horizontal'. We can then construct the perpendicular line **V**, which we define as 'vertical'. At this point, we can identify the Euclidean plane with the set $\mathbb{R}^2$ in the obvious way.

A C&S construction is any sequence of the following five basic operations, illustrated in Figure 11.6.

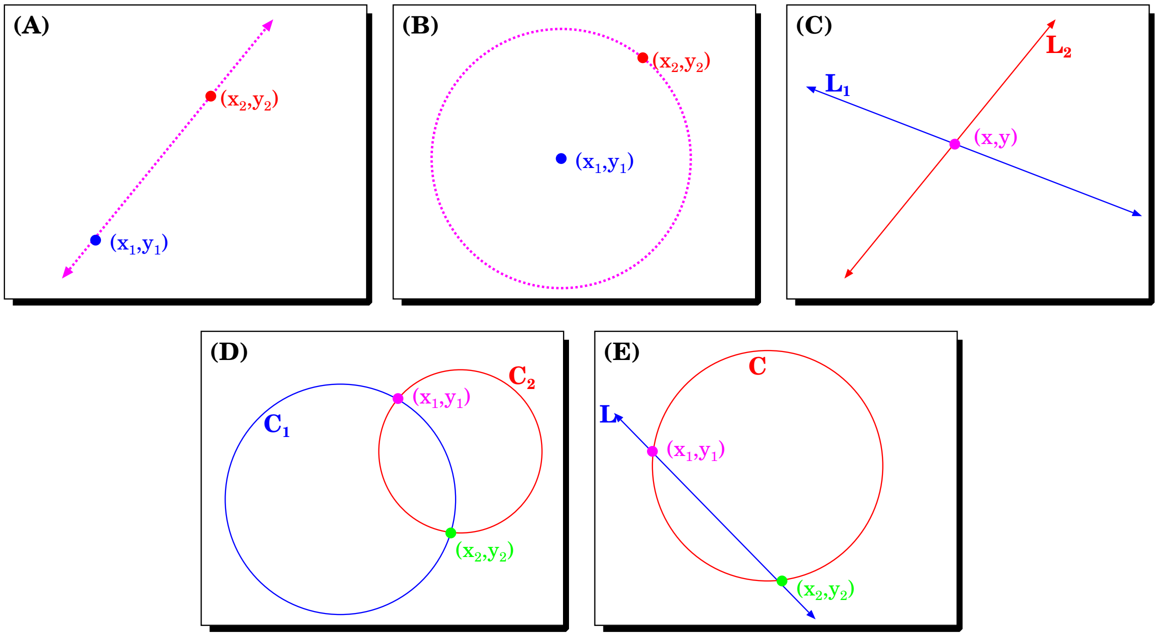


Figure 11.6: The five basic operations with compass and straight-edge.

(Figure 11.6A) Given two points (x_1, y_1) and (x_2, y_2) in $\mathbb{R}^2$, we can draw a unique **line** passing through these points.

(Figure 11.6B) Given two points (x_1, y_1) and (x_2, y_2) in $\mathbb{R}^2$, we can draw a unique **circle** centered at (x_1, y_2) and passing through (x_2, y_2).

(Figure 11.6C) Given two (nonparallel) lines $\mathbf{L}_1$ and $\mathbf{L}_2$, we can find the unique **point** where they intersect.

(Figure 11.6D) Given two circles $\mathbf{C}_1$ and $\mathbf{C}_2$, we can find the unique one or two **points** where they intersect.

(Figure 11.6E) Given a circle $\mathbf{C}$ and a line $\mathbf{L}$, we can find the unique one or two **points** where they intersect.

We will say that a point, line, or circle is **constructable** if it can be obtained through some sequence of these five operations, starting from nothing more than the origin $\mathbf{O}$, the unit circle $\mathbf{C}$, and the lines $\mathbf{H}$ and $\mathbf{V}$.

A **constructable length** is any real number r which appears as the distance between two constructable points x and y. Let $\mathbb{K}$ be the set of all constructable lengths; hence $\mathbb{K} \subset \mathbb{R}$.

Proposition 225 *$\mathbb{K}$ satisfies the following properties:*

(a) *$\mathbb{K}$ is a field.*

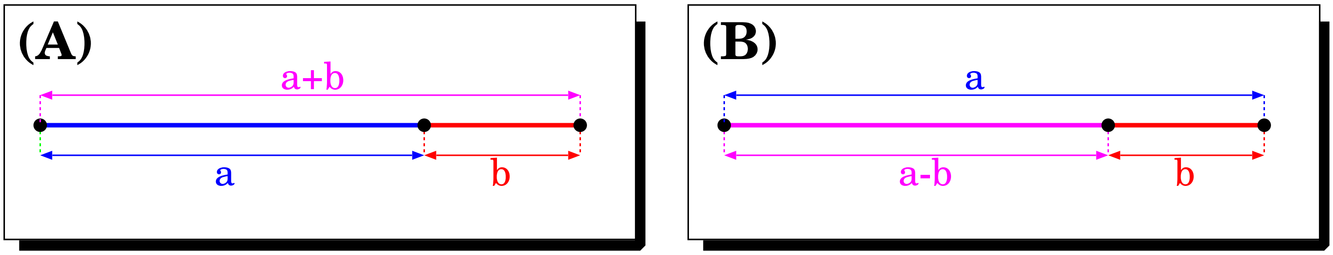


Figure 11.7: **(A)** Addition of lengths. **(B)** Subtraction of lengths.

(b) $\mathbb{Q} \subset \mathbb{K}$.

(c) *$\mathbb{K}$ is closed under square roots. That is: if $k \in \mathbb{K}$, then $\sqrt{k}$ is also in $\mathbb{K}$.*

Furthermore:

1. *$\mathbb{K}$ is the* smallest *field in $\mathbb{R}$ satisfying properties* **(A)**, **(B)**, *and* **(C)**. *That is: all elements of $\mathbb{K}$ can be obtained by starting with $\mathbb{Q}$, and iteratively applying operations of addition, subtraction, multiplication, division, and square root.*

2. *The set of all constructable points in the plane is exactly $\mathbb{K}^2 \subset \mathbb{R}^2$.*

3. *A line $\mathbf{L}$ is constructable if and only if $\mathbf{L} = \{(x, y) \in \mathbb{R}^2 \; ; \; ax + by = c\}$, where $a, b, c \in \mathbb{K}$.*

4. *A circle $\mathbf{C}$ is constructable if and only if $\mathbf{C} = \{(x, y) \in \mathbb{R}^2 \; ; \; (x - a)^2 + (y - b)^2 = c\}$, where $a, b, c \in \mathbb{K}$.*

Proof: **(A)** We must show that $\mathbb{K}$ is closed under addition, subtraction, multiplication, and division.

Closure under addition/subtraction: Let $a, b \in \mathbb{K}$. To construct the length $a + b$, we proceed as in Figure 11.7A: we construct parallel, adjoining line segments of lengths a and b; the length of the combination is $a + b$. To get $a - b$, we construct a line segment of length a, and then remove a segment of length b, as in Figure 11.7B.

Closure under multiplication: Let $a, b \in \mathbb{K}$. Construct a right-angle triangle whose base has length 1 and whose height has length a as in Figure 11.8A. Now extend this to a similar triangle whose base has length b, as in Figure 11.8B. Then the height of the new triangle is $a \cdot b$.

Closure under division: Let $a, b \in \mathbb{K}$. Construct a right-angle triangle whose base has length b and whose height has length a as in Figure 11.9A. Now construct a similar triangle, whose base has length 1, as in Figure 11.9B. Then the height of the new triangle is a/b.

Proof of (B) The set $\mathbb{K}$ contains the length 1 by definition. Since $\mathbb{K}$ is closed under addition/subtraction, we then have $\mathbb{Z} \subset \mathbb{K}$. Since $\mathbb{K}$ is closed under multiplication/division, we get $\mathbb{Q} \subset \mathbb{K}$.

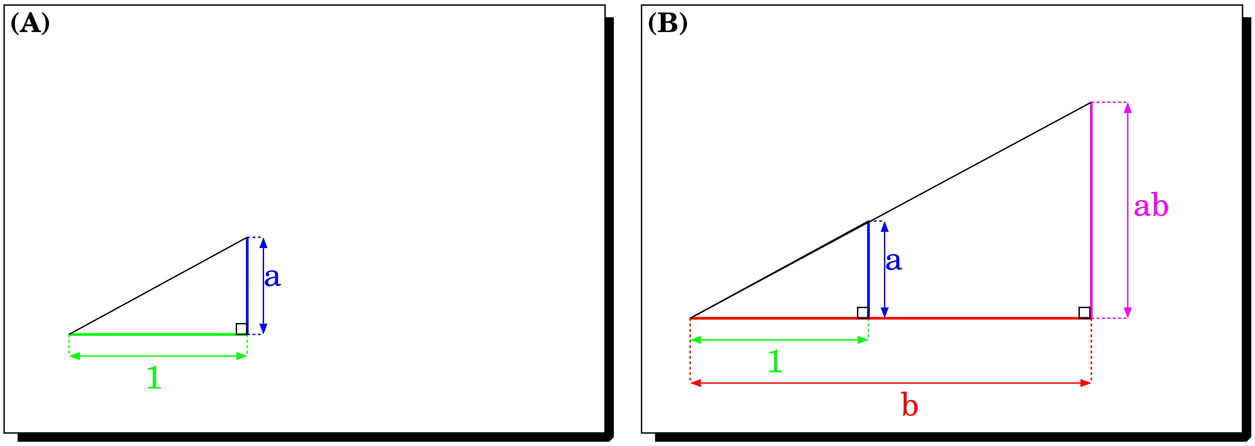


Figure 11.8: Multiplication of lengths.

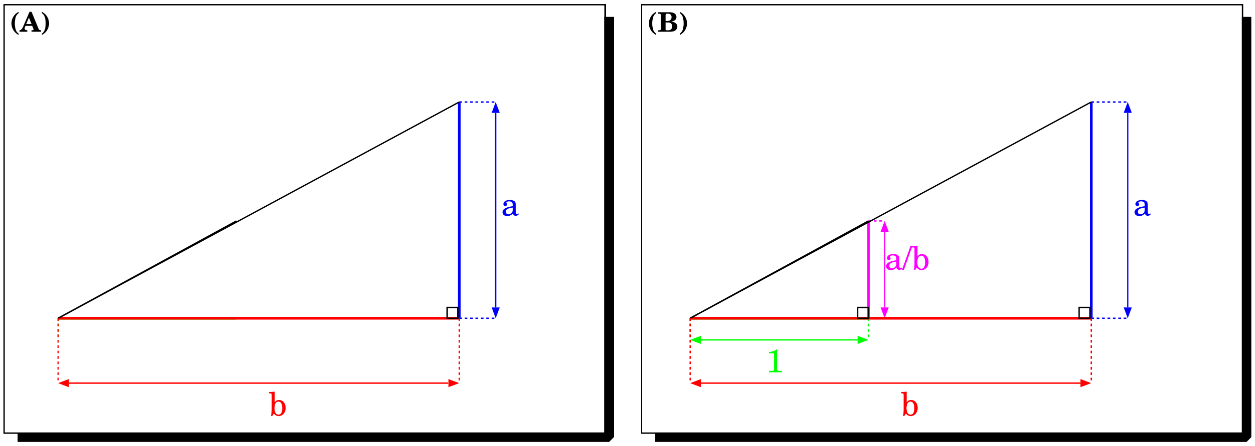


Figure 11.9: Division of lengths.

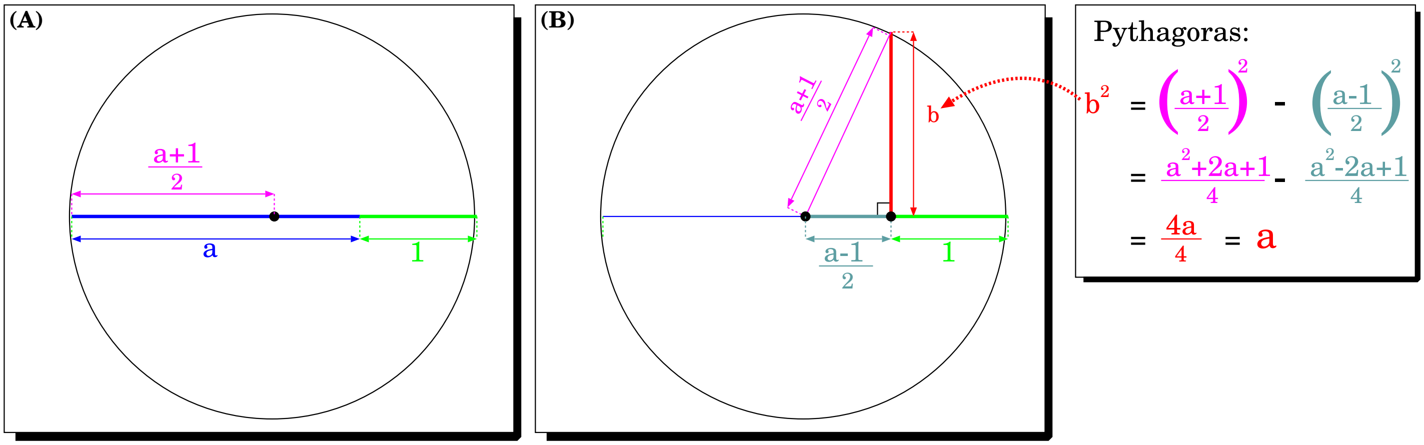

(A)
$\frac{a+1}{2}$
a
1
(B)
$\frac{a+1}{2}$
b
$\frac{a-1}{2}$
1
Pythagoras:
$b^2 = \left(\frac{a+1}{2}\right)^2 - \left(\frac{a-1}{2}\right)^2$
$= \frac{a^2+2a+1}{4} - \frac{a^2-2a+1}{4}$
$= \frac{4a}{4} = a$

Figure 11.10: Computing the square root of a length

Proof of (C) Let $a \in \mathbb{K}$. We want to construct the length $\sqrt{a}$. To do this, first build a circle of diameter $a+1$, as in Figure 11.10A. Thus, the radius of this circle is $\frac{a+1}{2}$. Now draw a perpendicular to the diameter, as in Figure 11.10B. A simple computation shows that this perpendicular has length $b = \sqrt{a}$.

Proof of 2: Suppose $(x, y) \in \mathbb{R}^2$ was a constructable point. We can draw a perpendicular line from **H** to (x, y); the length of this line is x, so $x \in \mathbb{K}$. We can draw a perpendicular line from **V** to (x, y); the length of this line is y, so $y \in \mathbb{K}$.

Proof of 3: A line **L** is constructable if it is the unique line passing through two constructable points (x_1, y_1) and (x_2, y_2), as in Figure 11.6A. In this case, the coefficients a, b, c can be obtained by solving the following system of linear equations:

$$\begin{array}{ccccccc} x_1 a & + & y_1 b & + & c & = & 0; \\ x_2 a & + & y_2 b & + & c & = & 0. \end{array}$$

If (x_1, y_1) is directly over (x_2, y_2) (ie. the line **L** is parallel to the vertical **V**) then $x_1 = x_2$, and we set $a = 1$, $b = 0$, and $c = x_1 = x_2$. Otherwise, set $b = -1$, to get the simpler system:

$$\begin{aligned} x_1 a - c &= y_1; \\ x_2 a - c &= y_2. \end{aligned}$$

We then solve this system to obtain:

$$\begin{bmatrix} a \\ c \end{bmatrix} = \begin{bmatrix} x_1 & -1 \\ x_2 & -1 \end{bmatrix}^{-1} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.$$

Since x_1, x_2, y_1, y_2 and 1 are in $\mathbb{K}$, and $\mathbb{K}$ is a field, we conclude that a and c are also in $\mathbb{K}$.

Proof of 4: A circle **C** is constructable if it is the unique circle centred at a constructable point (x_1, y_1) and passing through another constructable point (x_2, y_2), as in Figure 11.6C. In this case, set $a = x_1$ and $b = y_1$, both of which are in $\mathbb{K}$. Now let $c = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$. Since $\mathbb{K}$ is closed under square roots, it follows that $c \in \mathbb{K}$ also.

Proof of 1: Let $\widetilde{\mathbb{K}}$ be the smallest field in $\mathbb{R}$ satisfying **(A)**, **(B)**, and **(C)**. Clearly, $\widetilde{\mathbb{K}} \subset \mathbb{K}$; we'll show $\mathbb{K} \subset \widetilde{\mathbb{K}}$.

The arguments from the proofs of parts **2**, **3**, and **4** show:

- If $\mathbb{L}$ is the line between two points in $\widetilde{\mathbb{K}}^2$, then $\mathbf{L} = \{(x, y) \in \mathbb{R}^2 \ ; \ ax + by = c\}$, where $a, b, c \in \widetilde{\mathbb{K}}$. We say that $\mathbb{L}$ is a **$\widetilde{\mathbb{K}}$-line**.
- If (x, y) is the point of intersection of two $\widetilde{\mathbb{K}}$-lines, then $(x, y) \in \widetilde{\mathbb{K}}^2$.
- If $\mathbb{C}$ is the circle defined by two points in $\widetilde{\mathbb{K}}^2$, then $\mathbf{C} = \{(x, y) \in \mathbb{R}^2 \ ; \ (x - a)^2 + (y - b)^2 = c\}$, where $a, b, c \in \widetilde{\mathbb{K}}$. We say that $\mathbb{C}$ is a **$\widetilde{\mathbb{K}}$-circle**.

It remains to show:

1. If (x, y) is a point of intersection of two $\widetilde{\mathbb{K}}$-circles then $(x, y) \in \widetilde{\mathbb{K}}^2$.
2. If (x, y) is a point of intersection of a $\widetilde{\mathbb{K}}$-line and a $\widetilde{\mathbb{K}}$-circle then $(x, y) \in \widetilde{\mathbb{K}}^2$.

It is **Exercise 173** to show:

Claim 1: *If (x, y) is a point of intersection of a $\widetilde{\mathbb{K}}$-circle with a $\widetilde{\mathbb{K}}$-line or another $\widetilde{\mathbb{K}}$-circle, then x satisfies a quadratic equation $ax^2 + bx + c = 0$, where $a, b, c \in \widetilde{\mathbb{K}}$. Likewise, y satisfies a quadratic equation $dx^2 + ex + f = 0$, where $d, e, f \in \widetilde{\mathbb{K}}$.*

It follows that $x = \frac{-b \pm \sqrt{b^2-4ac}}{2a}$ and $y = \frac{-e \pm \sqrt{e^2-4df}}{2d}$. Since $\widetilde{\mathbb{K}}$ is a field closed under square roots, it follows that x and y are also in $\widetilde{\mathbb{K}}$. □

To show the impossibility of constructions **I** to **IV**, we'll show certain lengths are not in $\mathbb{K}$:

I. Trisecting an angle: Angle θ is trisectable iff $\sin(\theta/3)$ is in $\mathbb{K}$. This is obviously the case for some values of θ (eg. $\theta = \pi/2$), but we'll show it is not the case for most θ.

II. Constructing arbitrary regular polygons: A regular N-gon is constructable iff $\sin(\frac{2\pi}{N})$ is in $\mathbb{K}$. We'll show this is not true for some N.

III. Doubling the Cube: We can double a cube iff $\sqrt[3]{2} \in \mathbb{K}$. We'll show this is false.

IV. Squaring the Circle: We can double a circle iff $\sqrt{\pi} \in \mathbb{K}$. We'll show this is false.

11.2 Field Extensions

Prerequisites: §7.5 **Recommended:** §11.1

Let $\mathbb{F}$ be a field. An **extension** of $\mathbb{F}$ is another field $\mathbb{E}$ such that $\mathbb{F} \subset \mathbb{E}$.

Example 226:

(a) $\mathbb{R}$ is an extension of $\mathbb{Q}$.

(b) $\mathbb{C}$ is an extension of $\mathbb{R}$.

(c) Let $\mathbb{K}$ be the field of *constructable lengths* from Proposition 225 in §11.1. Then $\mathbb{K}$ is an extension of $\mathbb{Q}$.

(d) Let $\mathbb{R}(x)$ be the field of rational functions with real coefficients (Example ⟨98h⟩ on page 84), and let $\mathbb{C}(x)$ be the field of rational functions with complex coefficients. Then $\mathbb{C}(x)$ is an extension of $\mathbb{R}(x)$.

We consider the pair $\mathbb{F}$ and $\mathbb{E}$ as a single structure, called a **field extension**. We will denote this structure by "$(\mathbb{E} \supset \mathbb{F})$". Sometimes the notation "$\mathbb{E}/\mathbb{F}$" is used (read "$\mathbb{E}$ over $\mathbb{F}$"). We will also portray this field extension diagramatically by:

$$\begin{tikzcd} \mathbb{E} \arrow[d, dash] \\ \mathbb{F} \end{tikzcd}$$

11.2.1 Extensions as Vector Spaces:

It is natural to think of $\mathbb{C}$ as a 'two-dimensional real vector space'. We often speak of the 'complex plane' (as opposed to the 'real line'). Every element of $c \in \mathbb{C}$ can be written in a unique way as $c = r_1 + r_2\mathbf{i}$, so the set $\{1, \mathbf{i}\}$ is a *basis* for $\mathbb{C}$.

This idea generalizes to any field extension. If $\mathbb{E}$ is a field extension of $\mathbb{F}$, then $\mathbb{E}$ is automatically a **vector space** over $\mathbb{F}$. That is:

1. $\mathbb{E}$ is an *abelian group* under the operation of addition:
   - $e_1 + e_2 = e_2 + e_1$, for all $e_1, e_2 \in \mathbb{E}$
   - $e_1 + (e_2 + e_3) = (e_1 + e_2) + e_3$, for all $e_1, e_2, e_3 \in \mathbb{E}$
   - $e + 0 = e$ for all $e \in \mathbb{E}$.
   - $e + (-e) = 0$ for all $e \in \mathbb{E}$.
2. $\mathbb{F}$ acts linearly on $\mathbb{E}$ by *scalar multiplication*:
   - For any $f \in \mathbb{F}$ and $e \in \mathbb{E}$, the product $f \cdot e$ is in $\mathbb{E}$.
   - $f \cdot (e_1 + e_2) = fe_1 + fe_2$, for any $f \in \mathbb{F}$ and $e_1, e_2 \in \mathbb{E}$.
   - $(f_1 \cdot f_2) \cdot e = f_1 \cdot (f_2 \cdot e)$ for any $f_1, f_2 \in \mathbb{F}$ and $e \in \mathbb{E}$.

The **degree** of the extension $(\mathbb{E} \supset \mathbb{F})$ is the *dimension* of $\mathbb{E}$ as an $\mathbb{F}$-vector space. That is:

$$\begin{aligned}
\mathsf{deg}\left(\mathbb{E} \supset \mathbb{F}\right) &= \min\left\{|\mathcal{S}| \;;\; \text{where } \mathcal{S} \subset \mathbb{E} \text{ is an } \mathbb{F}\text{-spanning set for } \mathbb{E}\right\} \\
&= \max\left\{|\mathcal{I}| \;;\; \text{where } \mathcal{I} \subset \mathbb{E} \text{ is an } \mathbb{F}\text{-linearly independent subset of } \mathbb{E}\right\} \\
&= |\mathcal{B}|, \quad \text{where } \mathcal{B} \subset \mathbb{E} \text{ is any } \mathbb{F}\text{-basis for } \mathbb{E}.
\end{aligned}$$

Sometimes the notation "$[\mathbb{E} : \mathbb{F}]$" is used to indicate degree. We say $\mathbb{E}$ is a **finite** extension if $\mathsf{deg}\left(\mathbb{E} \supset \mathbb{F}\right)$ is finite, and an **infinite** extension if $\mathsf{deg}\left(\mathbb{E} \supset \mathbb{F}\right)$ is infinite.

Example 227:

(a) $\mathbb{C}$ is a 2-dimensional vector space over $\mathbb{R}$, with basis $\{1, \mathbf{i}\}$. Thus, $\mathsf{deg}\left(\mathbb{C} \supset \mathbb{R}\right) = 2$.

(b) $\mathbb{R}$ is an *infinite*-dimensional vector space over $\mathbb{Q}$. One way to see this is to observe that $\mathsf{card}\left[\mathbb{Q}^N\right] = \mathsf{card}\left[\mathbb{Q}\right] = \aleph_0$ for any finite N. In other words, any finite-dimensional rational vector space is *countable*. But $\mathbb{R}$ is *uncountable*; hence its dimension (as a rational vector space) must be infinite (and indeed, uncountable). Thus, $\mathsf{deg}\left(\mathbb{R} \supset \mathbb{Q}\right) = \infty$. ———

11.2.2 Simple Extensions

Let $\mathbb{E} \supset \mathbb{F}$ be any extension of $\mathbb{F}$, and let $\alpha \in \mathbb{E}$. We define $\mathbb{F}(\alpha)$ to be the *smallest subfield* of $\mathbb{E}$ containing both α and $\mathbb{F}$. That is:

$$\mathbb{F}(\alpha) \quad = \quad \bigcap_{\substack{\mathbb{F}\subset\mathbb{L}\subset\mathbb{E} \\ \alpha\in\mathbb{L}}} \mathbb{L}.$$

We say $\mathbb{F}(\alpha)$ is a **simple extension** of $\mathbb{F}$.

Example 228:

(a) $\mathbb{Q}(\sqrt{2})$ is the smallest subfield of $\mathbb{R}$ containing $\mathbb{Q}$ and $\sqrt{2}$. We'll see later (Example ⟨233a⟩ on page 191) that:

Every element of $\mathbb{Q}(\sqrt{2})$ *has the form* $a + b\sqrt{2}$ *for unique rational numbers* $a, b \in \mathbb{Q}$.

Addition in $\mathbb{Q}(\sqrt{2})$ takes the obvious form:

$$\left(a_1 + b_1\sqrt{2}\right) \; + \; \left(a_2 + b_2\sqrt{2}\right) \quad = \quad (a_1 + a_2) + (b_1 + b_2)\sqrt{2}.$$

Thus, $\mathbb{Q}(\sqrt{2})$ is a 2-dimensional vector space over $\mathbb{Q}$, with basis $\{1, \sqrt{2}\}$. Hence $\mathbb{Q}(\sqrt{2})$ is a degree-2 extension of $\mathbb{Q}$.

Multiplication in $\mathbb{Q}(\sqrt{2})$ takes the obvious form:

$$\left(a_1 + b_1\sqrt{2}\right) \cdot \left(a_2 + b_2\sqrt{2}\right) \quad = \quad (a_1a_2 + 2b_1b_2) \; + \; (a_1b_2 + a_2b_1)\sqrt{2}.$$

For example: $\left(1 + 3\sqrt{2}\right) \cdot \left(5 + 7\sqrt{2}\right) \; = \; (1\cdot5 + 2\cdot3\cdot7) \; + \; (1\cdot7 + 3\cdot5)\sqrt{2} \; = \; 47 \; + \; 22\sqrt{2}$.

(b) $\mathbb{Q}(\pi)$ is the smallest subfield of $\mathbb{R}$ containing $\mathbb{Q}$ and π. We'll see later that elements of $\mathbb{Q}(\pi)$ take the form of 'rational functions in π' ie. they have the form

$$\frac{q_0 + q_1\pi + q_2\pi^2 + \ldots + q_n\pi^n}{r_0 + r_1\pi + r_2\pi^2 + \ldots + r_m\pi^m},$$

where $n < m$, and $q_0, \ldots, q_n, r_0, \ldots, r_m \in \mathbb{Q}$.

Addition and multiplication proceed exactly as for rational functions. For example:

$$\left(\frac{1 + 2\pi}{3 - 4\pi}\right) \cdot \left(\frac{5 + 1\pi}{2 - 1\pi}\right) \quad = \quad \frac{(1 + 2\pi)\cdot(5 + 1\pi)}{(3 - 4\pi)\cdot(2 - 1\pi)} \quad = \quad \frac{5 + 7\pi + 2\pi^2}{6 - 5\pi + 4\pi^2}.$$

The elements

$$\left\{\ldots, \frac{1}{\pi^2}, \frac{1}{\pi}, 1, \pi, \pi^2, \pi^3, \ldots\right\}$$

are all $\mathbb{Q}$-linearly independent (this is because π is *transcendental* over $\mathbb{Q}$, a fact which we will not prove). Thus, $\mathbb{Q}(\pi)$ is an *infinite*-degree extension of $\mathbb{Q}$.

11.2.3 Finitely Generated Extensions

Let $\mathbb{E} \supset \mathbb{F}$ be any extension of $\mathbb{F}$, and let $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{E}$. We define $\mathbb{F}(\alpha_1, \ldots, \alpha_n)$ to be the *smallest subfield* of $\mathbb{E}$ containing $\{\alpha_1, \ldots, \alpha_n\}$ and $\mathbb{F}$. That is:

$$\mathbb{F}(\alpha_1, \ldots, \alpha_n) \quad = \quad \bigcap_{\substack{\mathbb{F}\subset\mathbb{L}\subset\mathbb{E} \\ \alpha_1,\ldots,\alpha_n\in\mathbb{L}}} \mathbb{L}.$$

We say $\mathbb{F}(\alpha_1, \ldots, \alpha_n)$ is a **finitely generated extension** of $\mathbb{F}$.

Example 229:

(a) $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ is the smallest subfield of $\mathbb{R}$ containing $\mathbb{Q}$ and $\sqrt{2}$. We'll see later (in Example ⟨243a⟩ on page 200) that:

> *Every element of* $\mathbb{Q}(\sqrt{2}, \sqrt{3})$ *has the form* $a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6}$*, for unique rational numbers* $a, b, c, d \in \mathbb{Q}$.

(b) $\mathbb{Q}(\pi, e)$ is the smallest subfield of $\mathbb{R}$ containing $\mathbb{Q}$, π, and e. Elements of $\mathbb{Q}(\pi, e)$ take the form of 'rational functions in π and e' ie. they have the form

$$\frac{\displaystyle\sum_{j=1}^{J}\sum_{k=0}^{K} q_{jk}\pi^n e^m}{\displaystyle\sum_{n=1}^{N}\sum_{m=0}^{M} r_{nm}\pi^n e^m}$$

where $\{q_{jk}\} \subset \mathbb{Q}$ and $\{r_{nm}\} \subset \mathbb{Q}$.

(This is because π and e are both is *transcendental* over $\mathbb{Q}$, and further are *algebraically independent* of one another. We will not prove these facts) Thus, $\mathbb{Q}(\pi, e)$ is an *infinite*-degree extension of $\mathbb{Q}$.

Lemma 230 *Let* $\mathbb{E} \supset \mathbb{F}$.

(a) *If* $\alpha, \beta \in \mathbb{E}$, *and* $\mathbb{A} = \mathbb{F}(\alpha)$, *then* $\mathbb{F}(\alpha, \beta) = \mathbb{A}(\beta)$.

(b) *More generally, if* $\alpha_1, \alpha_2, \ldots, \alpha_n, \beta \in \mathbb{E}$, *then* $\mathbb{F}(\alpha_1, \ldots, \alpha_n, \beta) = \mathbb{F}(\alpha_1, \ldots, \alpha_n)(\beta)$.

Proof: **Exercise 174** □

11.3 Finite and Algebraic Extensions

Prerequisites: §11.2

Let $\mathbb{E}$ be an extension of $\mathbb{F}$, and let $\alpha \in \mathbb{E}$. If $\mathbf{p}(\alpha) = 0$ for some $\mathbf{p} \in \mathbb{F}[x]$, then we say α is **algebraic** over $\mathbb{F}$; if not, we say α is a **transcendental** over $\mathbb{F}$.

Example 231:

(a) $\alpha = \sqrt[3]{2}$ is algebraic over over $\mathbb{Q}$, because $\mathbf{p}(\alpha) = 0$, where $\mathbf{p}(x) = x^3 - 2$.

(b) $\alpha = \mathbf{i}$ is algebraic over over $\mathbb{Q}$, because $\mathbf{p}(\alpha) = 0$, where $\mathbf{p}(x) = x^2 + 1$.

(c) $\alpha = \mathbf{i}$ is algebraic over over $\mathbb{R}$ for the same reason.

(d) π and e are transcendental over $\mathbb{Q}$. There is no polynomial $\mathbf{p}(x) \in \mathbb{Q}[x]$ such that $\mathbf{p}(\pi) = 0$ or $\mathbf{p}(e) = 0$. (This is actually very hard to prove).

If $\mathbb{E}$ is an extension of $\mathbb{F}$, then then we say $\mathbb{E}$ is an **algebraic extension** if ϵ is algebraic over $\mathbb{F}$ for every $\epsilon \in \mathbb{E}$. Otherwise, we say $\mathbb{E}$ is a **transcendental extension** of $\mathbb{F}$. For example, we'll show:

- $\mathbb{Q}(\sqrt{3})$ is an algebraic extension of $\mathbb{Q}$.
- $\mathbb{Q}(\mathbf{i})$ is an algebraic extension of $\mathbb{Q}$.
- $\mathbb{Q}(\sqrt{3},\ \mathbf{i})$ is an algebraic extension of $\mathbb{Q}$.
- $\mathbb{Q}(\pi)$ is a transcendental extension of $\mathbb{Q}$.
- $\mathbb{Q}(\sqrt{3},\ \pi)$ is a transcendental extension of $\mathbb{Q}$.

These claims follow from the following result:

Proposition 232 *Let $\mathbb{E}$ be an extension of $\mathbb{F}$, and let $\alpha \in \mathbb{F}$. The following are equivalent:*

(a) *α is algebraic over $\mathbb{F}$.*

(b) *$\mathbb{F}(\alpha)$ is a finite extension of $\mathbb{F}$.*

(c) *There exists a unique monic, irreducible polynomial $\mathbf{m}(x) \in \mathbb{F}[x]$ so that α is a root of $\mathbf{m}(x)$. In this case,*

1. *If $\mathbf{p}(x) \in \mathbb{F}[x]$ is any polynomial such that $\mathbf{p}(\alpha) = 0$, then $\mathbf{m}(x)$ divides $\mathbf{p}(x)$.*
2. *$\mathbb{F}(\alpha) \ \cong\ \mathbb{F}[x]/(\mathbf{m})$, via the isomorphism*

$$\Psi : \frac{\mathbb{F}[x]}{(\mathbf{m})} \ni \overline{\mathbf{p}(x)} \ \mapsto\ \mathbf{p}(\alpha) \in\ \mathbb{F}(\alpha).$$

(Note that this does not depend on the field $\mathbb{E}$).

3. $\mathsf{deg}\left(\mathbb{F}(\alpha) \supset \mathbb{F}\right) = \mathsf{degree}\,(\mathbf{m}(x))$.
4. $\mathbb{F}(\alpha)$ *has a* $\mathbb{F}$*-basis:* $\{1, \alpha, \alpha^2, \ldots, \alpha^N\}$.

Proof: (**a**)$\Longrightarrow$(**c**) Let $\mathcal{A} = \{\mathbf{p} \in \mathbb{F}[x]\ ;\ \mathbf{p}(\alpha) = 0\}$, and let $\mathbf{m}(x)$ be a nonzero element of minimal degree in $\mathcal{A}$. By multiplying $\mathbf{m}(x)$ by some element of $\mathbb{F}$, we can assume $\mathbf{m}$ is monic.

Claim 1: $\mathbf{m}(x)$ *is irreducible.*

Proof: Suppose $\mathbf{m}(x) = \mathbf{p}(x) \cdot \mathbf{q}(x)$. Then $\mathbf{p}(\alpha) \cdot \mathbf{q}(\alpha) = \mathbf{m}(\alpha) = 0$, so either $\mathbf{p}(\alpha) = 0$ or $\mathbf{q}(\alpha) = 0$. Hence, either $\mathbf{p} \in \mathcal{A}$ or $\mathbf{q} \in \mathcal{A}$. Suppose $\mathbf{p} \in \mathcal{A}$. But if $\mathbf{m}(x) = \mathbf{p}(x) \cdot \mathbf{q}(x)$, then $\mathsf{degree}\,(\mathbf{p}) < \mathsf{degree}\,(\mathbf{m})$, contradicting the minimality of $\mathbf{m}$. By contradiction, $\mathbf{m}$ must be irreducible. .... □ `[Claim 1]`

Proof of (c1) Let $\mathbf{p}(x) \in \mathcal{A}$; we want to show that $\mathbf{m}(x)$ divides $\mathbf{p}(x)$. Apply `Polynomial Long-Division` to write $\mathbf{p}(x) = \mathbf{q}(x)\mathbf{m}(x) + \mathbf{r}(x)$, where $\mathsf{degree}\,(\mathbf{r}) < \mathsf{degree}\,(\mathbf{q})$. We want to show $\mathbf{r} = 0$. To see this, observe that

$$0 \quad = \quad \mathbf{p}(\alpha) \quad = \quad \mathbf{q}(\alpha) \cdot \mathbf{m}(\alpha) + \mathbf{r}(x) \quad = \quad \mathbf{q}(\alpha) \cdot 0 + \mathbf{r}(x) \quad = \quad \mathbf{r}(x).$$

Hence, $\mathbf{r} \in \mathcal{A}$. But $\mathbf{q}$ is the nonzero element of minimal degree in $\mathcal{A}$, and $\mathsf{degree}\,(\mathbf{r}) < \mathsf{degree}\,(\mathbf{q})$, so we must have $\mathbf{r} = 0$.

Proof of (c2) Define $\Psi : \mathbb{F}[x] \longrightarrow \mathbb{F}(\alpha)$ by $\mathbf{p}(x) \mapsto \mathbf{p}(\alpha)$. Let $\mathbb{I} = \mathsf{image}\,[\Psi] \subset \mathbb{F}(\alpha)$.

It is **Exercise 175** to verify that

$$\ker(\Psi) \quad = \quad (\mathbf{m}).$$

It follows from the `Fundamental Isomorphism Theorem` that there is an isomorphism

$$\mathbb{I} \quad \cong \quad \frac{\mathbb{F}[x]}{\ker(\Psi)} \quad = \quad \frac{\mathbb{F}[x]}{(\mathbf{m})}.$$

It remains to show that $\mathbb{I} = \mathbb{F}(\alpha)$. To see this, observe that

$$\Big(\ \mathbf{m}(x) \text{ is irreducible.}\ \Big) =_{\text{Prop.165}}\Rightarrow \Big(\ (\mathbf{m}) \text{ is a maximal ideal.}\ \Big) =_{\text{Cor.167}}\Rightarrow \Big(\ \mathbb{I} \text{ is a field.}\ \Big).$$

Also note that $\mathbb{I}$ contains α and $\mathbb{F}$. Thus, $\mathbb{I}$ must contain $\mathbb{F}(\alpha)$. But $\mathbb{I} \subset \mathbb{F}(\alpha)$, so we conclude that $\mathbb{I} = \mathbb{F}(\alpha)$.

Proof of (c4) We must show that the set $\{1, \alpha, \alpha^2, \ldots, \alpha^{n-1}\}$ is both *linearly independent* and a *spanning set* for $\mathbb{F}(\alpha)$ as a $\mathbb{F}$-vector space.

Spanning Set: For any $\mathbf{f}(x) \in \mathbb{F}[x]$, let $\overline{\mathbf{f}(x)} = \Psi\left(\mathbf{f}(x)\right)$.

Claim 2: *Every element of* $\mathbb{F}(\alpha)$ *has the form* $f_0 + f_1\alpha + f_2\alpha^2 + \ldots + f_m\alpha^m$ *for some* $m \in \mathbb{N}$ *and* $f_0, \ldots, f_m \in \mathbb{F}$.

Proof: Let $e \in \mathbb{F}(\alpha)$. It follows from **(c2)** that $e = \overline{\mathbf{f}(x)}$, for some polynomial $\mathbf{f}(x) \in \mathbb{F}[x]$. Suppose $\mathbf{f}(x) = f_0 + f_1 x + f_2 x^2 + \ldots + f_m x^m$; then $\overline{\mathbf{f}(x)} = \overline{f}_0 + \overline{f}_1 \overline{x} + \overline{f}_2 \overline{x}^2 + \ldots + \overline{f}_m \overline{x}^m = f_0 + f_1\alpha + f_2\alpha^2 + \ldots + f_m\alpha^m$. .... □ `[Claim 2]`

Claim 3: *If $m \geq n$, then $\alpha^m = r_0 + r_1\alpha + r_2\alpha^2 + \ldots + r_{n-1}\alpha^{n-1}$ for some coefficients $r_0, \ldots, r_{n-1} \in \mathbb{F}$.*

Proof: Apply `Polynomial Long Division` to write $x^m = \mathbf{p}(x)\mathbf{q}(x) + \mathbf{r}(x)$ for some $\mathbf{q}, \mathbf{r} \in \mathbb{F}[x]$, where $\mathbf{r}(x)$ is a polynomial of degree less than n. Then

$$\alpha^m \quad = \quad \overline{x^m} \quad = \quad \overline{\mathbf{p}(x)\mathbf{q}(x)} + \overline{\mathbf{r}(x)} \quad = \quad \mathbf{r}(\alpha).$$

Thus, if $\mathbf{r}(x) = r_0 + r_1 x + r_2 x^2 + \ldots + r_{n-1}x^{n-1}$, then $\alpha^m = r_0 + r_1\alpha + r_2\alpha^2 + \ldots + r_{n-1}\alpha^{n-1}$. □ `[Claim 3]`

Combine Claims 4 and 5 to conclude that $\{1, \alpha, \alpha^2, \ldots, \alpha^{n-1}\}$ spans $\mathbb{F}(\alpha)$.

Linearly Independent: Suppose that $f_0 + f_1\alpha + f_2\alpha^2 + \ldots + f_{n-1}\alpha^{n-1} = 0$ for some $f_0, \ldots, f_{n-1} \in \mathbb{F}$; we want to show that $f_0 = f_1 = \ldots = f_{n-1} = 0$.

Let $\mathbf{f}(x) = f_0 + f_1 x + f_2 x^2 + \ldots + f_{n-1}x^{n-1}$, a polynomial in $\mathbb{F}[x]$. Then we have $\overline{\mathbf{f}(x)} = 0$, which means that $\mathbf{f}(x) \in \mathcal{I}$, which means that $\mathbf{p}(x)$ divides $\mathbf{f}(x)$. Thus, either $\mathbf{f} = 0$, or $\mathsf{degree}\,(\mathbf{f}) \geq \mathsf{degree}\,(\mathbf{p}) = n$. But $\mathsf{degree}\,(\mathbf{f}) \leq n - 1$ by construction, so this is impossible. We conclude that $\mathbf{f} = 0$ —in other words, $f_0 = f_1 = \ldots = f_{n-1} = 0$.

Proof of (c4) It follows from **(c3)** that $\mathsf{deg}\left(\mathbb{F}(\alpha) \supset \mathbb{F}\right) = \mathsf{degree}\,(\mathbf{m}(x))$.

(c)$\Longrightarrow$(b) This is immediate.

(b)$\Longrightarrow$(a) Suppose $\mathbb{F}(\alpha)$ has degree n. Then the $n+1$ elements $\{1, \alpha, \alpha^2, \ldots, \alpha^n\}$ must be $\mathbb{F}$-linearly dependent. In other words, there are elements $f_0, f_1, \ldots, f_n \in \mathbb{F}$ (not all zero) so that $f_0 + f_1\alpha + \ldots + f_n\alpha^n = 0$. But this means that $\mathbf{f}(\alpha) = 0$, where $\mathbf{f}(x) = f_n x^n + \ldots + f_1 x + f_0$. Hence, α is a root of the polynomial $\mathbf{f}$, so α is algebraic. □

The polynomial $\mathbf{m}(x)$ in Proposition 232 is called the **minimal polynomial** of α, and is usually denoted $\mathbf{m}_{\alpha,\mathbb{F}}(x)$, to make its dependence on α and $\mathbb{F}$ explicit. The **degree** of α (over $\mathbb{F}$) is the degree of $\mathbf{m}_{\alpha,\mathbb{F}}(x)$, which we denote by $\mathsf{degree}\,(\alpha; \mathbb{F})$. It follows that $\mathsf{deg}\left(\mathbb{F}(\alpha) \supset \alpha\right) = \mathsf{degree}\,(\alpha; \mathbb{F})$.

Example 233:

(a) $\sqrt{2}$ is algebraic over $\mathbb{Q}$, with minimal polynomial $\mathbf{m}_{\sqrt{2};\mathbb{Q}}(x) = x^2 - 2$. Thus, Proposition 232 part **(c2)** says $\mathbb{Q}(\sqrt{2})$ is isomorphic to $\mathbb{Q}[x]/(x^2 - 2)$. Part **(c3)** says that $\mathbb{Q}(\sqrt{2})$ is an extension of degree 2 over $\mathbb{Q}$, and Part **(c4)** says that $\mathbb{Q}(\sqrt{2})$ has $\mathbb{Q}$-basis $\{1, \sqrt{2}\}$.

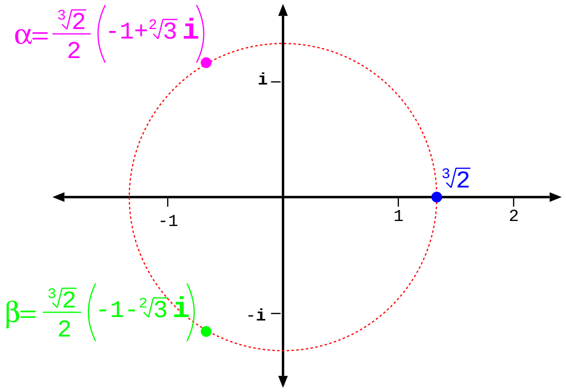

α= $\frac{\sqrt[3]{2}}{2}$ (-1+$\sqrt[2]{3}$ i)
i
$\sqrt[3]{2}$
-1
1
2
β= $\frac{\sqrt[3]{2}}{2}$ (-1-$\sqrt[2]{3}$ i)
-i

Figure 11.11: The three cube roots of 3 in the complex plane.

(b) Let $\mathbf{p}(x) = x^3 - 2$. Then $\mathbf{p}(x)$ is irreducible over $\mathbb{Q}$; its three roots in $\mathbb{C}$ are:

$$\sqrt[3]{2}; \qquad \alpha \;=\; \frac{\sqrt[3]{2}}{2}\cdot\left(-1+\sqrt{3}\cdot\mathbf{i}\right); \qquad \text{and} \qquad \beta \;=\; \frac{\sqrt[3]{2}}{2}\cdot\left(-1-\sqrt{3}\cdot\mathbf{i}\right) \qquad \text{(see Figure 11.11)}$$

Thus, $\sqrt[3]{2}$ is algebraic over $\mathbb{Q}$, with minimal polynomial $\mathbf{m}_{\sqrt[3]{2};\mathbb{Q}}(x) = \mathbf{p}(x)$. Thus, Proposition 232**(c2)** says that $\mathbb{Q}(\sqrt[3]{2})$ is isomorphic to $\mathbb{Q}[x]/(\mathbf{p}(x))$. Part **(c3)** says that $\mathbb{Q}(\sqrt[3]{2})$ is an extension of degree 3. Part **(c4)** says that $\mathbb{Q}(\sqrt[3]{2})$ has $\mathbb{Q}$-basis $\{1,\ \sqrt[3]{2},\ \sqrt[3]{4}\}$.

(c) However, $\sqrt[3]{2} \in \mathbb{R}$, so $\mathbb{R}(\sqrt[3]{2}) = \mathbb{R}$ –in other words, $\mathbb{R}(\sqrt[3]{2})$ is a trivial extension of $\sqrt[3]{2}$. The polynomial $\mathbf{p}(x) = x^3 - 2$ is *not* irreducible over $\mathbb{R}$, because it factors:

$$x^3 - 2 \quad = \quad (x - \sqrt[3]{2})(x^2 + \ \sqrt[3]{2}\cdot x + \sqrt[3]{4}) \qquad \text{(check this)}$$

Thus, $\mathbf{p}(x)$ is *not* a minimal polynomial for $\sqrt[3]{2}$ over $\mathbb{R}$. Indeed, the minimal polynomial for $\sqrt[3]{2}$ over $\mathbb{R}$ is just $\mathbf{m}_{\sqrt[3]{2};\mathbb{R}}(x) \;=\; x - \sqrt[3]{2}$.

(d) Let $\mathbf{p}(x) = x^3 - 2$ as in Example ⟨233b⟩, and now let $\alpha = \frac{\sqrt[3]{2}}{2}\cdot\left(-1+\sqrt{3}\cdot\mathbf{i}\right)$ (see Figure 11.11). As in Example ⟨233b⟩, α is algebraic over $\mathbb{Q}$, with minimal polynomial $\mathbf{p}(x)$. Thus, Proposition 232 part **(c2)** says that $\mathbb{Q}(\alpha)$ is isomorphic to $\mathbb{Q}[x]/(\mathbf{p}(x))$; part **(c3)** says $\mathbb{Q}(\alpha)$ is an extension of degree 3, and part **(c4)** says $\mathbb{Q}(\alpha)$ has basis $\{1,\ \alpha,\ \alpha^2\}$. This is portrayed on the *left* side of Figure 11.12**(A)**.

Observe that this means $\mathbb{Q}(\alpha)$ is isomorphic to $\mathbb{Q}(\sqrt[3]{2})$ (from Example ⟨233b⟩). However, $\mathbb{Q}(\alpha)$ is *not equal* to $\mathbb{Q}(\sqrt[3]{2})$, because $\mathbb{Q}(\sqrt[3]{2}) \subset \mathbb{R}$, whereas $\mathbb{Q}(\alpha)$ extends into the complex plane.

(e) Let α be as in Example ⟨233d⟩. Then $\alpha \not\in \mathbb{R}$, so $\mathbb{R}(\alpha)$ is a nontrivial extension of $\mathbb{R}$. However, $\mathbf{p}(x) = x^3 - 2$ is *not* the minimal polynomial for α over $\mathbb{R}$, because $\mathbf{p}(x)$ is not irreducible, as we saw in Example ⟨233c⟩. Instead, the minimal polynomial for α over $\mathbb{R}$ is given by:

$$\mathbf{m}_{\alpha;\mathbb{R}}(x) \quad = \quad (x^2 + \ \sqrt[3]{2}\cdot x + \sqrt[3]{4})$$

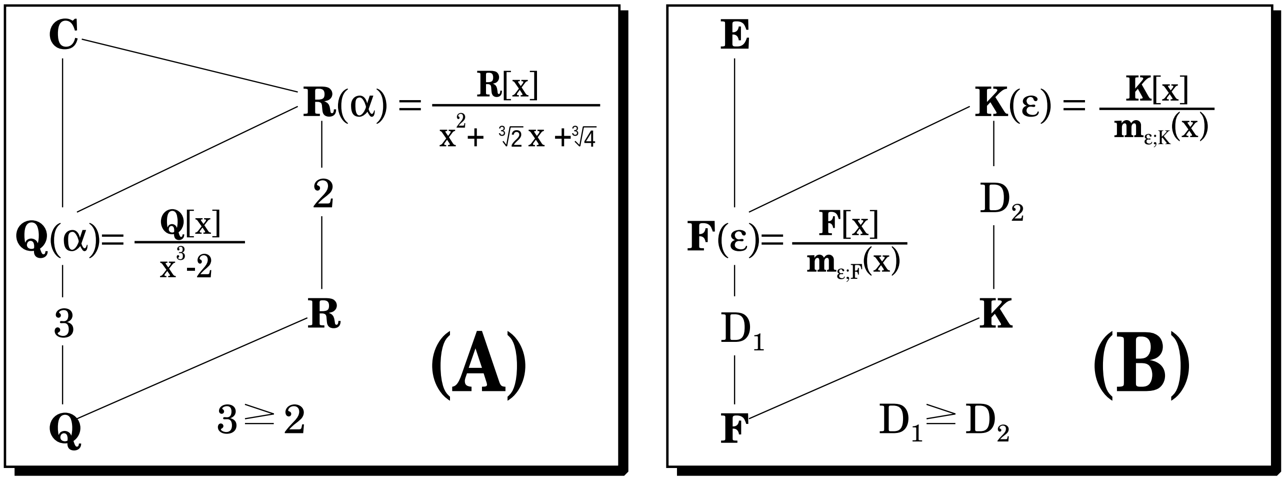


Figure 11.12: **(A)** $\mathbb{Q}(\alpha)$ has degree 3 over $\mathbb{Q}$, but $\mathbb{R}(\alpha)$ only has degree 2 over $\mathbb{R}$. This is because $\mathbf{m}_{\alpha;\mathbb{Q}}(x) = x^3 - 2$, but $\mathbf{m}_{\alpha;\mathbb{R}}(x) = x^2 + \sqrt[3]{2}x + \sqrt[3]{4}$. **(B)** $\mathbb{F}(\epsilon)$ has degree D_1 over $\mathbb{F}$, while $\mathbb{K}(\epsilon)$ has degree $D_2 \leq D_1$ over $\mathbb{K}$.

Thus, Proposition 232**(c)** says that

$$\mathbb{R}(\alpha) \quad \cong \quad \mathbb{R}[x]/(x^2 + \sqrt[3]{2} \cdot x + \sqrt[3]{4})$$

is an extension of degree 2, with $\mathbb{R}$-basis $\{1, \alpha\}$.

This is portrayed on the *right* side of Figure 11.12**(A)**.

Compare Example ⟨233b⟩ and ⟨233c⟩. Compare Examples ⟨233d⟩ and ⟨233e⟩. It appears that the minimal polynomial of an element α depends on the field we are comparing α to. The formal statement is as follows:

Proposition 234 *Let $\mathbb{E} \supset \mathbb{F}$ be an extension of $\mathbb{F}$, and let $\epsilon \in \mathbb{E}$ be algebraic over $\mathbb{F}$. Let $\mathbb{K} \supset \mathbb{F}$ be another extension of $\mathbb{F}$. Then, as shown in Figure 11.12***(B)***:*

(a) *ϵ is algebraic over $\mathbb{K}$.*

(b) *If we treat $\mathbf{m}_{\epsilon;\mathbb{F}}(x)$ as element of $\mathbb{K}[x]$, then $\mathbf{m}_{\epsilon;\mathbb{F}}(x)$ may no longer be irreducible. However:*

(c) *$\mathbf{m}_{\epsilon;\mathbb{K}}(x)$ divides $\mathbf{m}_{\epsilon;\mathbb{F}}(x)$ (in $\mathbb{K}[x]$).*

(d) *Thus, $\mathsf{deg}\left(\mathbb{K}(\epsilon) \supset \mathbb{K}\right) \ \leq\ \mathsf{deg}\left(\mathbb{F}(\epsilon) \supset \mathbb{F}\right)$.*

Proof: **(a)** Clearly, $\mathbb{F}[x] \subset \mathbb{K}[x]$. Thus, $\mathbf{m}_{\epsilon;\mathbb{F}}(x)$ is an element of $\mathbb{K}[x]$, so since $\mathbf{m}_{\epsilon;\mathbb{F}}(\epsilon) = 0$ by definition, it follows that ϵ is algebraic over $\mathbb{K}$.

(c) Let $\mathbf{m}_{\epsilon;\mathbb{K}}(x)$ be the minimal polynomial of ϵ over $\mathbb{K}$, then Proposition 232**(c1)** says $\mathbf{m}_{\epsilon;\mathbb{K}}(x)$ divides $\mathbf{m}_{\epsilon;\mathbb{F}}(x)$.

(b) and **(d)** follow immediately from **(c)**. □

Corollary 235 *Let $\mathbb{E} \supset \mathbb{F}$. Then*

$$\Big(\mathbb{E} \text{ is a finite extension of } \mathbb{F} \Big) \Longrightarrow \Big(\mathbb{E} \text{ is an algebraic extension of } \mathbb{F} \Big).$$

Proof: Let $\epsilon \in \mathbb{E}$. We want to show ϵ is algebraic.

Consider the subfield $\mathbb{F}(\epsilon)$. Observe that $\mathsf{deg}\Big(\mathbb{F}(\epsilon) \supset \mathbb{F}\Big) \leq \mathsf{deg}\Big(\mathbb{E} \supset \mathbb{F}\Big)$, so it is finite. Say $\mathsf{deg}\Big(\mathbb{F}(\epsilon) \supset \mathbb{F}\Big) = N$. Thus, the elements $1, \epsilon, \epsilon^2, \ldots, \epsilon^N$ cannot be linearly independant; hence there are some $f_0, f_1, \ldots, f_N \in \mathbb{F}$ (not all zero) so that

$$f_0 + f_1\epsilon + \ldots + f_N\epsilon^N \quad = \quad 0.$$

But now define polynomial $\mathbf{f}(x) \in \mathbb{F}[x]$ by

$$\mathbf{f}(x) = f_0 + f_1 x + \ldots + f_N x^N.$$

It follows that $\mathbf{f}(\epsilon) = 0$. Hence, ϵ is algebraic.

This holds for any $\epsilon \in \mathbb{E}$, so $\mathbb{E}$ is finite. □

Note: The converse of Corollary 235 is false. While it is true that every finite extension of $\mathbb{F}$ is algebraic, it is *not* true that every algebraic extension is finite.

The next theorem can be summarized: *A finite extension of a finite extension is finite; an algebraic extension of an algebraic extension is algebraic.*

Theorem 236 *Let $\mathbb{F} \subset \mathbb{E} \subset \mathbb{D}$ be a chain of field extensions.*

(a) $\mathsf{deg}\Big(\mathbb{D} \supset \mathbb{F}\Big) \quad = \quad \mathsf{deg}\Big(\mathbb{D} \supset \mathbb{E}\Big) \cdot \mathsf{deg}\Big(\mathbb{E} \supset \mathbb{F}\Big)$ *(whether these degrees are finite or infinite).*

(b) *Thus, if $\mathbb{D}$ has finite degree over $\mathbb{E}$, and $\mathbb{E}$ has finite degree over $\mathbb{F}$, and then $\mathbb{D}$ also has finite degree over $\mathbb{F}$ See Figure 11.13(A).*

(c) *In particular, suppose* $\mathsf{deg}\Big(\mathbb{E} \supset \mathbb{F}\Big) = N$*, and $\{\epsilon_1, \ldots, \epsilon_N\}$ is a $\mathbb{F}$-basis for $\mathbb{E}$. Suppose* $\mathsf{deg}\Big(\mathbb{D} \supset \mathbb{E}\Big) = M$*, and $\{\delta_1, \ldots, \delta_N\}$ is a $\mathbb{E}$-basis for $\mathbb{D}$. Then*

$$\left\{ \begin{array}{cccc} \epsilon_1\delta_1, & \epsilon_1\delta_2, & \ldots, & \epsilon_1\delta_M, \\ \epsilon_2\delta_1, & \epsilon_2\delta_2, & \ldots, & \epsilon_2\delta_M, \ldots \\ \vdots & \vdots & \ddots & \vdots \\ \ldots\epsilon_N\delta_1, & \epsilon_N\delta_2, & \ldots, & \epsilon_N\delta_M \end{array} \right\}$$

is an $\mathbb{F}$-basis for $\mathbb{D}$.

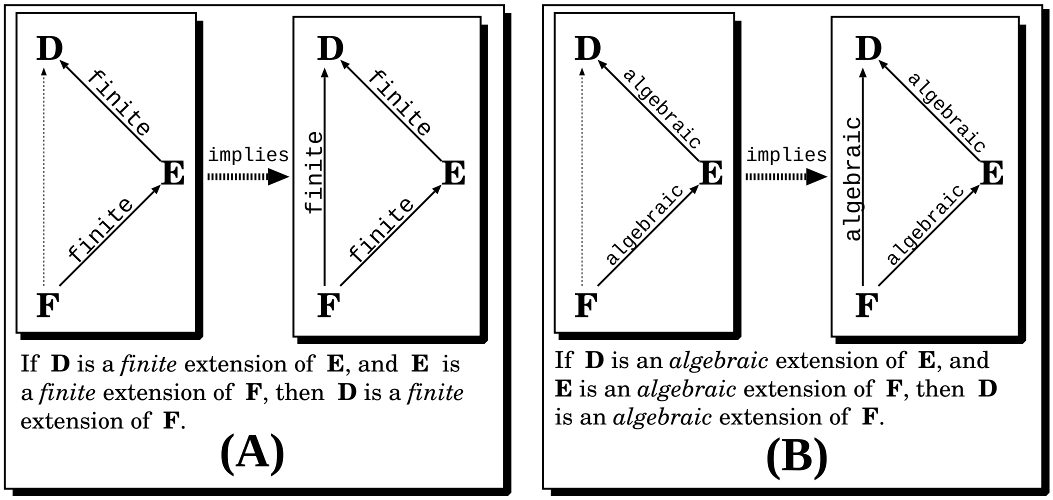

D
finite
E
finite
F
implies
D
finite
E
finite
finite
F
If D is a finite extension of E, and E is a finite extension of F, then D is a finite extension of F.
(A)
D
algebraic
E
algebraic
F
implies
D
algebraic
E
algebraic
algebraic
F
If D is an algebraic extension of E, and E is an algebraic extension of F, then D is an algebraic extension of F.
(B)

Figure 11.13:

(d) *If $\mathbb{D}$ is algebraic over $\mathbb{E}$, and $\mathbb{E}$ is algebraic over $\mathbb{F}$, and then $\mathbb{D}$ is also algebraic over $\mathbb{F}$ (whether these extensions are finite or infinite). See Figure 11.13(B).*

Proof: Clearly **(a)** and **(b)** follows from **(c)**, so we'll prove **(b)**. We must show that the set $\{\epsilon_n \delta_m\}_{n=1\ldots N}^{m=1\ldots M}$ is $\mathbb{F}$-*linearly independent* and an $\mathbb{F}$-*spanning set* for $\mathbb{D}$.

Spanning Set: Let $d \in \mathbb{D}$ be arbitrary. Since $\{\delta_1, \ldots, \delta_M\}$ is a $\mathbb{E}$-basis for $\mathbb{D}$, we have:

$$d \quad = \quad e_1\delta_1 + \ldots + e_M\delta_M \tag{11.1}$$

for some $e_1, \ldots, e_M \in \mathbb{E}$. But $\{\epsilon_1, \ldots, \epsilon_N\}$ is a $\mathbb{F}$-basis for $\mathbb{E}$. Hence, we can write:

$$\begin{array}{rcl} e_1 & = & f_{11}\epsilon_1 + \ldots + f_{N1}\epsilon_N \\ \vdots & \vdots & \vdots \\ e_M & = & f_{1M}\epsilon_1 + \ldots + f_{NM}\epsilon_N \end{array} \tag{11.2}$$

Combining equations (11.1) and (11.2) yields:

$$d \quad = \quad \sum_{m=1}^{M} e_m\delta_m \quad = \quad \sum_{m=1}^{M}\left(\sum_{n=1}^{N} f_{nm}\epsilon_n\right)\delta_m \quad = \quad \sum_{m=1}^{M}\sum_{n=1}^{N} f_{nm}\epsilon_n\delta_m.$$

Linearly Independent: Suppose we had an equation $\displaystyle\sum_{n=1}^{N}\sum_{m=1}^{M} f_{nm}\epsilon_n\delta_m \quad = \quad 0$ for some coefficients $f_{nm} \in \mathbb{F}$ not all zero. We can rewrite this as:

$$0 \quad = \quad \sum_{m=1}^{M}\left(\sum_{n=1}^{N} f_{nm}\epsilon_n\right)\delta_m \quad = \quad \sum_{m=1}^{M} e_m\delta_m$$

where $e_1 = f_{11}\epsilon_1 + \ldots + f_{N1}\epsilon_N$ etc. Since $\{\delta_1, \ldots, \delta_m\}$ are $\mathbb{E}$-linearly independent, we must conclude that $e_1 = \ldots = e_M = 0$. But then, since $\epsilon_1, \ldots, \epsilon_n$ are $\mathbb{F}$-linearly independent, we must have $f_{nm} = 0$ for all n and m.

(d) Let $\delta \in \mathbb{D}$; we want to show that δ is algebraic over $\mathbb{F}$.

Since δ is algebraic over $\mathbb{E}$, there is some polynomial $\mathbf{p}(x) \in \mathbb{E}[x]$ so that $\mathbf{p}(\delta) = 0$. Suppose $\mathbf{p}(x) = \pi_n x^n + \ldots + \pi_1 x + \pi_0$, where $\pi_0, \pi_1, \ldots, \pi_n \in \mathbb{E}$. Let $\mathbb{P} = \mathbb{F}(\pi_1, \ldots, \pi_n) \subset \mathbb{E}$, as in the figure on the left.

Claim 1: *$\mathbb{P}$ is finite over $\mathbb{F}$.*

Proof: Let $\mathbb{P}_0 = \mathbb{F}(\pi_0)$. Then $\mathbb{P}_0$ is a finite extension of $\mathbb{F}$ by Proposition 232**(b)**, because π_0 is algebraic over $\mathbb{F}$.

Next, define:

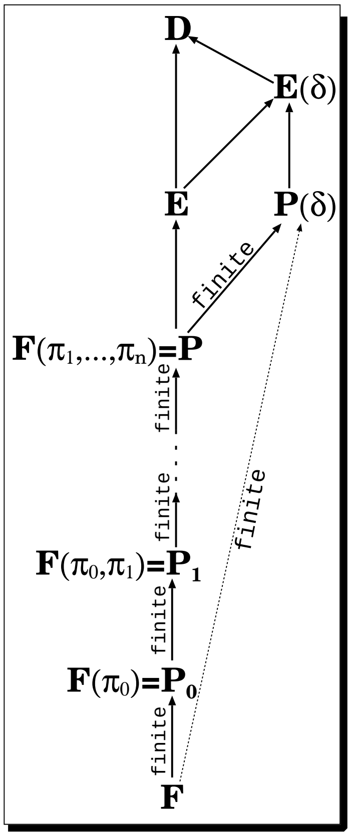
$$\begin{array}{ccccccc}
\mathbb{P}_1 & = & \mathbb{P}_0(\pi_1) & = & \mathbb{F}(\pi_0, \pi_1) & & \\
\mathbb{P}_2 & = & \mathbb{P}_1(\pi_2) & = & \mathbb{F}(\pi_0, \pi_1, \pi_2) & & \\
\mathbb{P}_3 & = & \mathbb{P}_2(\pi_3) & = & \mathbb{F}(\pi_0, \pi_1, \pi_2, \pi_3) & & \\
\vdots & \vdots & \vdots & \vdots & \vdots & & \\
\mathbb{P}_n & = & \mathbb{P}_{n-1}(\pi_n) & = & \mathbb{F}(\pi_0, \pi_1, \pi_2, \pi_3, \ldots, \pi_n) & = & \mathbb{P}
\end{array}$$

Then $\mathbb{P}_1$ is a finite extension of $\mathbb{P}_0$ by Proposition 232**(b)**, because π_1 is algebraic over $\mathbb{F}$ (and thus, over $\mathbb{P}_1$). Likewise, $\mathbb{P}_2$ is finite over $\mathbb{P}_1$; $\mathbb{P}_3$ is finite over $\mathbb{P}_2$, and so on.

Thus, we have a chain of finite extensions:

$$\mathbb{F} \subset \mathbb{P}_0 \subset \mathbb{P}_1 \subset \mathbb{P}_2 \subset \ldots \subset \mathbb{P}_n = \mathbb{P}$$

Hence, by iterating part **(b)** of this theorem, we conclude that $\mathbb{P}$ is finite over $\mathbb{F}$. .... □ [Claim 1]



Claim 2: *δ is algebraic over $\mathbb{P}$.*

Proof: Recall that $\mathbf{p}(x) = \pi_n x^n + \ldots + \pi_1 x + \pi_0$. Now, $\pi_1, \ldots, \pi_n \in \mathbb{P}$ by construction, so $\mathbf{p}(x) \in \mathbb{P}[x]$. But $\mathbf{p}(\delta) = 0$; hence δ is algebraic over $\mathbb{P}$. .... □ [Claim 2]

Claim 3: *$\mathbb{P}(\delta)$ is finite over $\mathbb{P}$.*

Proof: δ is algebraic over $\mathbb{P}$, so this follows from Proposition 232. ..... □ [Claim 3]

Claim 4: *$\mathbb{P}(\delta)$ is finite over $\mathbb{F}$.*

Proof: Combine Claims 1 and 3 with part **(b)** of this theorem. .... □ [Claim 4]

Claim 4 and Corollary 235**(b)** imply that $\mathbb{P}(\delta)$ is algebraic over $\mathbb{F}$. Hence, δ is algebraic over $\mathbb{F}$. This holds for any $\delta \in \mathbb{D}$. Hence, $\mathbb{D}$ is algebraic over $\mathbb{F}$. □

Corollary 237 *Suppose $\mathbb{D} \supset \mathbb{E} \supset \mathbb{F}$. Then:*

(a) $\mathsf{deg}\left(\mathbb{E} \supset \mathbb{F}\right)$ *divides* $\mathsf{deg}\left(\mathbb{D} \supset \mathbb{F}\right)$.

(b) *Thus,* $\mathsf{deg}\left(\mathbb{E} \supset \mathbb{F}\right) \leq \mathsf{deg}\left(\mathbb{D} \supset \mathbb{F}\right)$, *and*

(c) $\left(\ \mathsf{deg}\left(\mathbb{E} \supset \mathbb{F}\right) = \mathsf{deg}\left(\mathbb{D} \supset \mathbb{F}\right)\ \right) \iff \left(\ \mathbb{E} = \mathbb{D}\ \right)$.

(d) *Hence, if* $\mathsf{deg}\left(\mathbb{D} \supset \mathbb{F}\right)$ *is prime, then there is no subfield $\mathbb{E}$ such that $\mathbb{D} \supsetneq \mathbb{E} \supsetneq \mathbb{F}$.* □

Corollary 238 *Let $\mathbb{E} \supset \mathbb{F}$. The following are equivalent:*

(a) *$\mathbb{E}$ is a finite extension of $\mathbb{F}$*

(b) *$\mathbb{E} = \mathbb{F}(\alpha_1, \alpha_2, \ldots, \alpha_n)$, where $\alpha_1, \ldots, \alpha_n$ are elements algebraic over $\mathbb{F}$.*

(c) $\mathsf{deg}\left(\mathbb{E} \supset \mathbb{F}\right) \ \leq\ d_1 \cdot d_2 \cdots d_n,$ *where* $d_k = \mathsf{degree}\,(\alpha_k; \mathbb{F})$.

Proof: **(a)**$\Longrightarrow$**(b)** (by induction on $\mathsf{deg}\left(\mathbb{E} \supset \mathbb{F}\right)$)

Base Case: Suppose $\mathsf{deg}\left(\mathbb{E} \supset \mathbb{F}\right) = 2$. Let $\alpha \in \mathbb{E}$ be any element not in $\mathbb{F}$. Then $\mathbb{F}(\epsilon)$ is an extension of $\mathbb{F}$ of degree 2 or greater, so Corollary 237**(c)** says $\mathbb{F}(\alpha) = \mathbb{E}$.

Induction: Suppose that "**(a)**$\Longrightarrow$**(b)**" is true for *all* fields $\mathbb{F}'$ and all extensions $\mathbb{E}' \supset \mathbb{F}'$ of degree less than D. Suppose that $\mathsf{deg}\left(\mathbb{E} \supset \mathbb{F}\right) = D$.

Let $\alpha_1 \in \mathbb{E}$. Then α_1 is algebraic over $\mathbb{F}$, so Corollary 232**(b)** says $\mathsf{deg}\left(\mathbb{F}(\alpha_1) \supset \mathbb{F}\right) = d_1$, where $d_1 = \mathsf{degree}\,(\alpha_k; \mathbb{F}) > 1$.

Now, let $\mathbb{F}' = \mathbb{F}(\alpha_1)$; then $\mathbb{E} \supset \mathbb{F}' \supset \mathbb{F}$, so Theorem 236**(a)** says that

$$\mathsf{deg}\left(\mathbb{E} \supset \mathbb{F}'\right) \quad = \quad \frac{\mathsf{deg}\left(\mathbb{E} \supset \mathbb{F}\right)}{\mathsf{deg}\left(\mathbb{F}' \supset \mathbb{F}\right)} \quad = \quad \frac{D}{d_1} \quad < \quad D.$$

Hence, by induction, $\mathbb{E} = \mathbb{F}'(\alpha_2, \ldots, \alpha_n)$ for some elements $\alpha_2, \ldots, \alpha_n \in \mathbb{E}$. But then

$$\mathbb{E} \quad = \quad \mathbb{F}'(\alpha_2, \ldots, \alpha_n) \quad = \quad \mathbb{F}(\alpha_1)(\alpha_2, \ldots, \alpha_n) \quad = \quad \mathbb{F}(\alpha_1, \alpha_2, \ldots, \alpha_n).$$

Since $\mathbb{E}$ is algebraic over $\mathbb{F}$, we know that $\alpha_2, \ldots, \alpha_n$ are algebraic over $\mathbb{F}$.

(b)$\Longrightarrow$**(c)** (by induction on n)

Base Case $(n = 1)$ If $\mathbb{E} = \mathbb{F}(\alpha)$, then $\mathsf{deg}\left(\mathbb{E} \supset \mathbb{F}\right) = \mathsf{degree}\,(\alpha; \mathbb{F})$ by definition.

Induction: Let $\mathbb{K} = \mathbb{F}(\alpha_1, \ldots, \alpha_{n-1})$. By induction, we suppose

$$\mathsf{deg}\left(\mathbb{K} \supset \mathbb{F}\right) \quad \leq \quad d_1 \cdot d_2 \cdots d_{n-1}. \tag{11.3}$$

Note that $\mathbb{E} = \mathbb{K}(\alpha_n)$. Thus,

$$\mathsf{deg}\left(\mathbb{E} \supset \mathbb{K}\right) \underset{(*)}{=} \mathsf{degree}\,(\alpha_n; \mathbb{K}) \quad \leq_{(\dagger)} \quad \mathsf{degree}\,(\alpha_n; \mathbb{F}) \quad = \quad d_n. \tag{11.4}$$

Here, $(*)$ is by the **Base Case**, and $(\dagger)$ is by Proposition 234**(d)**. Thus,

$$\mathsf{deg}\left(\mathbb{E} \supset \mathbb{F}\right) \underset{(**)}{=} \mathsf{deg}\left(\mathbb{E} \supset \mathbb{K}\right) \cdot \mathsf{deg}\left(\mathbb{K} \supset \mathbb{F}\right) \quad \leq_{(\ddagger)} \quad d_n \cdot (d_1 \cdot d_2 \cdots d_{n-1}).$$

Here, $(**)$ is by Theorem 236**(a)**, and $(\ddagger)$ follows from equations (11.3) and (11.4).

(c)$\Longrightarrow$**(a)** If $\mathsf{deg}\left(\mathbb{E} \supset \mathbb{F}\right) \leq d_1 \cdot d_2 \cdots d_n$, then $\mathsf{deg}\left(\mathbb{E} \supset \mathbb{F}\right)$ is finite, so $\mathbb{E}$ is algebraic over $\mathbb{F}$ by Corollary 235. □

Corollary 239 *Let $\mathbb{E} \supset \mathbb{F}$, and let* $\mathbb{A} = \{\alpha \in \mathbb{E}\ ;\ \alpha$ *is algebraic over* $\mathbb{F}\}$.

Then $\mathbb{A}$ is a subfield of $\mathbb{E}$, and is an algebraic extension of $\mathbb{F}$.

Proof: Clearly, $\mathbb{F} \subset \mathbb{A} \subset \mathbb{E}$. We must show that $\mathbb{A}$ is a field —ie. that it is closed under addition, subtraction, multiplication, and division.

To see this, let $\alpha, \beta \in \mathbb{A}$. Then α, β are algebraic over $\mathbb{F}$, so the field $\mathbb{F}(\alpha, \beta)$ has finite degree over $\mathbb{F}$, by Corollary 238. Hence, $\mathbb{F}(\alpha, \beta)$ is algebraic over $\mathbb{F}$, by Corollary 235. Hence, all elements of $\mathbb{F}(\alpha, \beta)$ are algebraic —ie. $\mathbb{F}(\alpha, \beta) \subset \mathbb{A}$.

But $\alpha + \beta$, $\alpha - \beta$, $\alpha \cdot \beta$, and α/β are all elements of $\mathbb{F}(\alpha, \beta)$, and thus, of $\mathbb{A}$. □

11.4 (multi)Quadratic Extensions

Prerequisites: §11.3

Let $\mathbb{E}$ be an extension of $\mathbb{F}$, and let $\alpha \in \mathbb{E}$. If $\mathbf{q}(\alpha) = 0$ for some quadratic polynomial $\mathbf{q}(x) = x^2 + q_1 x + q_0$ in $\mathbb{F}[x]$, then we say α is a **quadratic root** over $\mathbb{F}$. We call $\mathbb{F}(\alpha)$ a **quadratic extension** of $\mathbb{F}$.

Example 240: $\alpha = \sqrt{2}$ is a quadratic root over $\mathbb{Q}$, because $\mathbf{q}(\alpha) = 0$, where $\mathbf{q}(x) = x^2 - 2$. Thus, $\mathbb{Q}(\sqrt{2})$ is a quadratic extension of $\mathbb{Q}$.

Recall that $\mathbb{F}$ has **characteristic 2** if $1 + 1 = 0$ in $\mathbb{F}$. For example, $\mathbb{Z}_{/2}$ has characteristic 2. However, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ do not have characteristic 2; nor does $\mathbb{Z}_{/p}$, for $p > 2$.

Proposition 241 *Suppose $\mathbb{F}$ does not have characteristic 2. Let $\mathbb{E} \supset \mathbb{F}$. The following are equivalent:*

(a) *$\mathbb{E}$ is a quadratic extension of $\mathbb{F}$.*

(b) *$\mathbb{E} = \mathbb{F}(\sqrt{D})$ for some $D \in \mathbb{F}$ which is not the square of any other element in $\mathbb{F}$.*

(c) $\mathsf{deg}\left(\mathbb{E} \supset \mathbb{F}\right) = 2$.

Proof: (**a**)$\Longrightarrow$(**b**) Suppose $\mathbb{E} = \mathbb{F}(\alpha)$, where α is a quadratic root. Thus, $\mathbf{q}(\alpha) = 0$, or some quadratic polynomial $\mathbf{q}(x) = x^2 + bx + c$.

The roots of $\mathbf{q}$ are given by the `Quadratic Formula`:

$$\alpha \quad = \quad \frac{b \pm \sqrt{b^2 - 4c}}{2}.$$

(It is elementary to verify that this holds in any field not of characteristic 2; we need characteristic $\neq 2$ because the formula calls for division by 2).

Let $D = b^2 - 4c$; then $D \in \mathbb{F}$. I claim that $\mathbb{F}(\alpha) = \mathbb{F}(\sqrt{D})$. To see this, it suffices to show that $\alpha \in \mathbb{F}(\sqrt{D})$ and that $\sqrt{D} \in \mathbb{F}(\alpha)$.

To see that $\alpha \in \mathbb{F}(\sqrt{D})$, observe that $\alpha = \frac{1}{2}(b \pm \sqrt{D})$.

To see that $\sqrt{D} \in \mathbb{F}(\alpha)$, observe that $\sqrt{D} = 2\alpha - b$.

(**b**)$\Longrightarrow$(**a**) This is immediate.

(**a**)$\Longrightarrow$(**c**) Suppose $\mathbb{E} = \mathbb{F}(\alpha)$, where α is a quadratic root. Then Proposition 232 implies that $\mathsf{deg}\left(\mathbb{F}(\alpha) \supset \mathbb{F}\right) = 2$.

(**c**)$\Longrightarrow$(**a**) Let $\alpha \in \mathbb{E}$ such that $\alpha \notin \mathbb{F}$.

Claim 1: *α is a quadratic root over $\mathbb{F}$*

Proof: Note that $\mathbb{E}$ is a 2-dimensional $\mathbb{F}$-vector space. Hence, the elements $\{1, \ \alpha, \ \alpha^2\}$ cannot be $\mathbb{F}$-linearly independent. Thus, there are $f_0, f_1, f_2 \in \mathbb{F}$ (not all zero) such that

$$f_0 + f_1\alpha + f_2\alpha^2 \ = \ 0. \tag{11.5}$$

We know that $f_2 \neq 0$, because

$$\left(f_2 = 0 \right) \quad \Longrightarrow \quad \left(f_0 + f_1\alpha = 0 \right) \quad \Longrightarrow \quad \left(\alpha \ = \ -f_0/f_1 \ \in \ \mathbb{F} \right),$$

and we specified $\alpha \not\in \mathbb{F}$. Hence, we can divide equation (11.5) by f_2, to obtain:

$$q_0 + q_1\alpha + \alpha^2 \quad = \quad 0,$$

where $q_0 = f_0/f_2$ and $q_1 = f_1/f_2$. In other words, $\mathbf{q}(\alpha) = 0$, where $\mathbf{q}(x) = x^2 + q_1 x + x_0$ is a quadratic polynomial. . . . . □ `[Claim 1]`

It remains to show that $\mathbb{E} = \mathbb{F}(\alpha)$. To see this, observe that

$$2 \quad = \quad \mathsf{deg}\left(\mathbb{E} \supset \mathbb{F}\right) \quad \geq \quad \mathsf{deg}\left(\mathbb{F}(\alpha) \supset \mathbb{F}\right) \quad > \quad 1.$$

Hence, we must have $\mathsf{deg}\left(\mathbb{F}(\alpha) \supset \mathbb{F}\right) = 2$, which means $\mathbb{F}(\alpha) = \mathbb{E}$. ________________□

We say that $\mathbb{E}$ is a **multiquadratic extension** of $\mathbb{F}$ if we have a sequence of extensions

$$\mathbb{F} \quad = \quad \mathbb{E}_0 \quad \subset \quad \mathbb{E}_1 \quad \subset \quad \mathbb{E}_2 \quad \subset \quad \cdots \quad \subset \quad \mathbb{E}_N \quad = \quad \mathbb{E},$$

where $\mathbb{E}_n$ is a quadratic extension of $\mathbb{E}_{n-1}$ for all $n \geq 1$.

Corollary 242 *If $\mathbb{E} \supset \mathbb{F}$ is a multiquadratic extension, then* $\mathsf{deg}\left(\mathbb{E} \supset \mathbb{F}\right) = 2^n$ *for some n.*

Proof: **<u>Exercise 176</u>** Hint: Combine Proposition 241(c) with Theorem 236(a) on page 194. □

Note that the converse of this is false: if $\mathsf{deg}\left(\mathbb{E} \supset \mathbb{F}\right) = 2^n$, it is not necessarily true that $\mathbb{E}$ is multiquadratic over $\mathbb{F}$.

Example 243:

(a) Consider $\mathbb{Q}(\sqrt{2}, \sqrt{3})$. Clearly, $\mathbb{Q} \subsetneq \mathbb{Q}(\sqrt{2}) \subseteq \mathbb{Q}(\sqrt{2}, \sqrt{3})$.

Claim 1: $\sqrt{3} \not\in \mathbb{Q}(\sqrt{2})$.

Proof: If $\sqrt{3} \in \mathbb{Q}(\sqrt{2})$, then there are rational numbers $a, b \in \mathbb{Q}$ such that

$$\sqrt{3} \quad = \quad a + b\sqrt{2}. \tag{11.6}$$

First observe that $a \neq 0 \neq b$. To see this, suppose $b = 0$; then equation (11.6) becomes $\sqrt{3} = a$ —that is, $\sqrt{3} \in \mathbb{Q}$, which we know is false.

Now suppose $a = 0$; then equation (11.6) becomes $\sqrt{3} = b\sqrt{2}$. Multiply by $\sqrt{2}$ to get: $\sqrt{6} = 2b$, meaning , $\sqrt{6} \in \mathbb{Q}$, which we know is false.

Thus, $ab \neq 0$. Square both sides of (11.6) to get:

$$3 \quad = \quad (a + b\sqrt{2})^2 \quad = \quad a^2 + 2b^2 - 2ab\sqrt{2}.$$

Since $ab \neq 0$, this implies:

$$\sqrt{2} \quad = \quad \frac{3 - a^2 - 2b^2}{-2ab},$$

which means $\sqrt{2} \in \mathbb{Q}$, which we know is false. . . . . □ `[Claim 1]`

Claim 2: *$\mathbb{Q}(\sqrt{2},\sqrt{3})$ is a quadratic extension of $\mathbb{Q}(\sqrt{2})$.*

> **Proof:** Claim 1 implies that $\mathbb{Q}(\sqrt{2}) \subsetneq \mathbb{Q}(\sqrt{2},\sqrt{3})$. Also the minimal polynomial of $\sqrt{3}$ over $\mathbb{Q}(\sqrt{3})$ is $\mathbf{m}_{\sqrt{3}}(x) = x^2 - 3$ (because $\mathbf{m}_{\sqrt{3}}(x)$ is irreducible over $\mathbb{Q}(\sqrt{2})$, by Claim 1). .... □ `[Claim 2]`

But $\mathbb{Q}(\sqrt{2})$ itself is a quadratic extension of $\mathbb{Q}$. Thus, $\mathbb{Q}(\sqrt{2},\sqrt{3})$ is a *multiquadratic* extension of $\mathbb{Q}$.

Furthermore, Theorem 236**(a)** implies:

$$\begin{aligned}\mathsf{deg}\left(\mathbb{Q}(\sqrt{2},\sqrt{3}) \supset \mathbb{Q}\right) \quad &= \quad \mathsf{deg}\left(\mathbb{Q}(\sqrt{2},\sqrt{3}) \supset \mathbb{Q}(\sqrt{2})\right) \cdot \mathsf{deg}\left(\mathbb{Q}(\sqrt{2}) \supset \mathbb{Q}\right) \quad = \quad 2\cdot 2 \\ &= \quad 4.\end{aligned}$$

This means that $\mathbb{Q}(\sqrt{2},\sqrt{3})$ is a 4-dimensional vector space over $\mathbb{Q}$. Claim 1 shows that $\{1,\sqrt{2},\sqrt{3}\}$ is a linearly independent set. Similar reasoning shows that $\{1,\sqrt{2},\sqrt{3},\sqrt{6}\}$ is also linearly independent, hence, a basis. We conclude:

> *Every element of $\mathbb{Q}(\sqrt{2},\sqrt{3})$ has the form $a+b\sqrt{2}+c\sqrt{3}+d\sqrt{6}$, for unique rational numbers $a,b,c,d \in \mathbb{Q}$.*

(b) Consider $\mathbb{Q}(\sqrt[4]{2})$. Since $\sqrt{2} = (\sqrt[4]{2})^2$, it follows that $\sqrt{2} \in \mathbb{Q}(\sqrt[4]{2})$; hence

$$\mathbb{Q} \quad \subsetneq \quad \mathbb{Q}(\sqrt{2}) \quad \subseteq \quad \mathbb{Q}(\sqrt[4]{2}).$$

Claim 3: $\sqrt[4]{2} \;\not\in\; \mathbb{Q}(\sqrt{2})$.

> **Proof:** **<u>Exercise 177</u>** Hint: Imitate the proof of Claim 1 above. ...... □ `[Claim 3]`

Thus, $\mathbb{Q}(\sqrt{2}) \subsetneq \mathbb{Q}(\sqrt[4]{2})$. Since $\sqrt[4]{2} = \sqrt{\sqrt{2}}$, it follows that $\mathbb{Q}(\sqrt[4]{2})$ is a quadratic extension of $\mathbb{Q}(\sqrt{2})$, and thus, a *multiquadratic* extension of $\mathbb{Q}$.

Example ⟨243a⟩ generalizes as follows:

Proposition 244 *Let $\mathbb{F} \subset \mathbb{C}$, and let $\alpha_1,\alpha_2,\ldots,\alpha_n \in \mathbb{F}$. Then:*

(a) *$\mathbb{F}(\sqrt{\alpha_1},\ \sqrt{\alpha_2},\ \ldots,\sqrt{\alpha_n})$ is a multiquadratic extension of $\mathbb{F}$.*

(b) $\mathsf{deg}\left(\mathbb{F}(\sqrt{\alpha_1},\ \sqrt{\alpha_2},\ \ldots,\sqrt{\alpha_n}) \supset \mathbb{F}\right) \;=\; 2^m$, *for some $m \leq n$.*

Proof: **<u>Exercise 178</u>** Hint: proceed by induction on n. □

The following result is reminiscent of Theorem 236 on page 194:

Proposition 245 *If $\mathbb{D}$ is a multiquadratic extension of $\mathbb{E}$, and $\mathbb{E}$ is a multiquadratic extension of $\mathbb{F}$, then $\mathbb{D}$ is a multiquadratic extension of $\mathbb{F}$.*

Proof: **Exercise 179** □

A 'subextension' of a multiquadratic extension is also multiquadratic:

Proposition 246 *If $\mathbb{D}$ is a multiquadratic extension of $\mathbb{F}$, and $\mathbb{F} \subset \mathbb{E} \subset \mathbb{D}$, then $\mathbb{E}$ is also a multiquadratic extension of $\mathbb{F}$.*

Proof: **Exercise 180** □

11.5 Compass & Straight-Edge Constructions II

Prerequisites: §11.1, §11.4

Let $\mathbb{K} \subset \mathbb{R}$ be the field of constructable lengths, from Proposition 225 on page 181 of §11.1.

Proposition 247 *Let $\kappa \in \mathbb{K}$. Then:*

(a) *$\mathbb{Q}(\kappa)$ is a multiquadratic extension of $\mathbb{Q}$.*

(b) *Thus, $\mathsf{deg}\left(\mathbb{Q}(\kappa) \supset \mathbb{Q}\right) = 2^n$ for some $n \in \mathbb{N}$.*

Proof: Part **1** of Proposition 225 says that $\mathbb{K}$ is the *smallest* subfield of $\mathbb{R}$ which contains $\mathbb{Q}$ and is closed under square roots.

We will first construct another subfield $\mathbb{K}_\infty \subset \mathbb{R}$ so that $\mathbb{Q} \subset \mathbb{K}_\infty$ and $\mathbb{K}_\infty$ is closed under square roots. We proceed as follows:

- Let $\mathbb{K}_1 \subset \mathbb{R}$ be the smallest subfield containing $\mathbb{Q}$ and the square roots of all elements in $\mathbb{Q}$ (thus, $\mathbb{K}_1$ contains $\sqrt{2}$, $\sqrt{3}$, etc.)
- Let $\mathbb{K}_2 \subset \mathbb{R}$ be the smallest subfield containing $\mathbb{K}_1$ and the square roots of all elements in $\mathbb{K}_1$ (thus, $\mathbb{K}_2$ contains $\sqrt[4]{2} = \sqrt{\sqrt{2}}$, $\sqrt{\sqrt{2}+\sqrt{3}}$, etc.)
- ....inductively, let $\mathbb{K}_{n+1}$ be the smallest subfield containing $\mathbb{K}_n$ and the square roots of all elements in $\mathbb{K}_n$.

We now have an ascending sequence of field extensions:

$$\mathbb{Q} \quad \subset \quad \mathbb{K}_1 \quad \subset \quad \mathbb{K}_2 \quad \subset \quad \mathbb{K}_3 \quad \subset \quad \cdots$$

Now, let $\mathbb{K}_\infty \;=\; \bigcup_{n=1}^{\infty} \mathbb{K}_n$.

Claim 1: *$\mathbb{K}_\infty$ is a field, and is closed under square roots.*

Proof: Exercise 181 .... □ [Claim 1]

Clearly, $\mathbb{Q} \subset \mathbb{K}_\infty \subset \mathbb{R}$. It follows from Claim 1 and Proposition 225 (Part **1**) that $\mathbb{K} \subset \mathbb{K}_\infty$. (Actually, $\mathbb{K} = \mathbb{K}_\infty$, but this is not important for our purposes.)

Now suppose that $\kappa \in \mathbb{K}$. It follows that $\kappa \in \mathbb{K}_\infty$. But this means that $\kappa \in \mathbb{K}_n$ for some $n \in \mathbb{N}$. It thus suffices to show:

Claim 2: *If $\kappa \in \mathbb{K}_n$ for some n, then $\mathbb{Q}(\kappa)$ is a multiquadratic extension of $\mathbb{Q}$.*

Proof: (by induction on n)

Base Case $(n = 1)$**:** If $\kappa \in \mathbb{K}_1$, this means that

$$\kappa \quad = \quad a_0 + a_1\sqrt{b_1} + a_2\sqrt{b_2} + \ldots + a_m\sqrt{b_m},$$

for some $a_0, a_1, \ldots, a_m$ and $b_1, \ldots, b_m \in \mathbb{Q}$. Thus, $\kappa \in \mathbb{Q}(\sqrt{b_1}, \ldots, \sqrt{b_n})$, and thus, $\mathbb{Q}(\kappa) \subset \mathbb{Q}(\sqrt{b_1}, \ldots, \sqrt{b_n})$.

But Proposition 244(a) says $\mathbb{Q}(\sqrt{b_1}, \ldots, \sqrt{b_n})$ is a multiquadratic extension of $\mathbb{Q}$. Since $\mathbb{Q}(\kappa) \subset \mathbb{Q}(\sqrt{b_1}, \ldots, \sqrt{b_n})$, it follows from Proposition 246 that $\mathbb{Q}(\kappa)$ is also a multiquadratic extension of $\mathbb{Q}$.

Induction: If $\kappa \in \mathbb{K}_n$, then $\kappa = \alpha_0 + \alpha_1\sqrt{\beta_1} + \alpha_2\sqrt{\beta_2} + \ldots + \alpha_m\sqrt{\beta_m}$, for some $\alpha_0, \alpha_1, \ldots, \alpha_m$ and $\beta_1, \ldots, \beta_m \in \mathbb{K}_{n-1}$. Let $\mathbb{F} = \mathbb{Q}(\alpha_0, \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_m)$. Then:

- By identical reasoning to the `Base Case`, $\mathbb{F}(\kappa)$ is a multiquadratic extension of $\mathbb{F}$.
- By induction hypothesis, $\mathbb{F}$ is a multiquadratic extension of $\mathbb{Q}$.
- Thus, Proposition 245 implies that $\mathbb{F}(\kappa)$ is a multiquadratic extension of $\mathbb{Q}$.

But $\mathbb{Q}(\kappa) \subset \mathbb{F}(\kappa)$, so Proposition 246 implies that $\mathbb{Q}(\kappa)$ is a multiquadratic extension of $\mathbb{Q}$. .... □ [Claim 2]

Part **(b)** of our theorem follows from part **(a)** and Corollary 242. □

Corollary 248 *It is impossible to do any of the following using only compass and straight-edge:*

I *Trisect an arbitrary angle.*

II *Square the circle.*

III *Double the Cube.*

Proof: **(II)** Doubling the Cube is equivalent to constructing $\sqrt[3]{2}$. But $\mathbb{Q}(\sqrt[3]{2})$ is a degree-3 extension of $\mathbb{Q}$, contradicting Proposition 247(b) (because 3 is not a power of 2). Thus, $\sqrt[3]{2}$ cannot be in $\mathbb{K}$.

(III) Squaring the circle is equivalent to constructing $\sqrt{\pi}$, and hence constructing π. But π is transcendental over $\mathbb{Q}$, so $\mathbb{Q}(\pi)$ is an infinite extension of $\mathbb{Q}$, contradicting Proposition 247(b) (because ∞ is not a power of 2). Thus, π cannot be in $\mathbb{K}$.

(I) Let Θ be some angle and let $\theta = \frac{1}{3}\Theta$. Thus, $\Theta = 3\theta$. Let $\alpha = \cos(\theta)$. We'll show that, in general, $\mathbb{Q}(\alpha)$ is *not* a multiquadratic extension of $\mathbb{Q}$; hence α cannot be in $\mathbb{K}$.

Claim 1: $\cos(\Theta) = 4\alpha^3 \;-\; 3\alpha.$

Proof: We employ some standard trigonometric identities...

$$\begin{aligned}
\cos(\Theta) &= \cos(3\theta) \quad = \quad \cos(\theta + 2\theta) \\
&= \cos(\theta)\cdot\cos(2\theta) - \sin(\theta)\cdot\sin(2\theta) \\
&= \cos(\theta)\cdot\Big(2\cos^2(\theta) - 1\Big) \;-\; \sin(\theta)\cdot 2\sin(\theta)\cos(\theta) \\
&= 2\cos^3(\theta) - \cos(\theta) \;-\; 2\sin^2(\theta)\cdot\cos(\theta) \\
&= 2\cos^3(\theta) - \cos(\theta) \;-\; 2\Big(1 - \cos^2(\theta)\Big)\cdot\cos(\theta) \\
&= 2\cos^3(\theta) - \cos(\theta) \;-\; 2\cos(\theta) + 2\cos^3(\theta) \\
&= 4\cos^3(\theta) \;-\; 3\cos(\theta) \quad = \quad 4\alpha^3 - 3\alpha\cdot \quad \ldots\ldots\ldots\ldots\ldots\ldots \; \Box \; \texttt{[Claim 1]}
\end{aligned}$$

Let $\Theta = 60^o$. Then Θ is a constructable angle, and $\cos(\Theta) = \frac{1}{2}$. Thus Claim 1 says:

$$\frac{1}{2} \quad = \quad 4\alpha^3 - 3\alpha.$$

In other words, $\mathbf{p}(\alpha) = 0$, where $\mathbf{p}(x) = 4x^3 - 3x - \frac{1}{2}$. It can be checked that $\mathbf{p}$ is irreducible over $\mathbb{Q}$. Thus, $\mathbf{p}(x)$ is the minimal polynomial of α over $\mathbb{Q}$, and Proposition 232**(c3)** says that $\mathbb{Q}(\alpha)$ is an extension of degree 3 over $\mathbb{Q}$. But 3 is not a power of 2, so Proposition 247(b) says that α cannot be in $\mathbb{K}$. ______ □

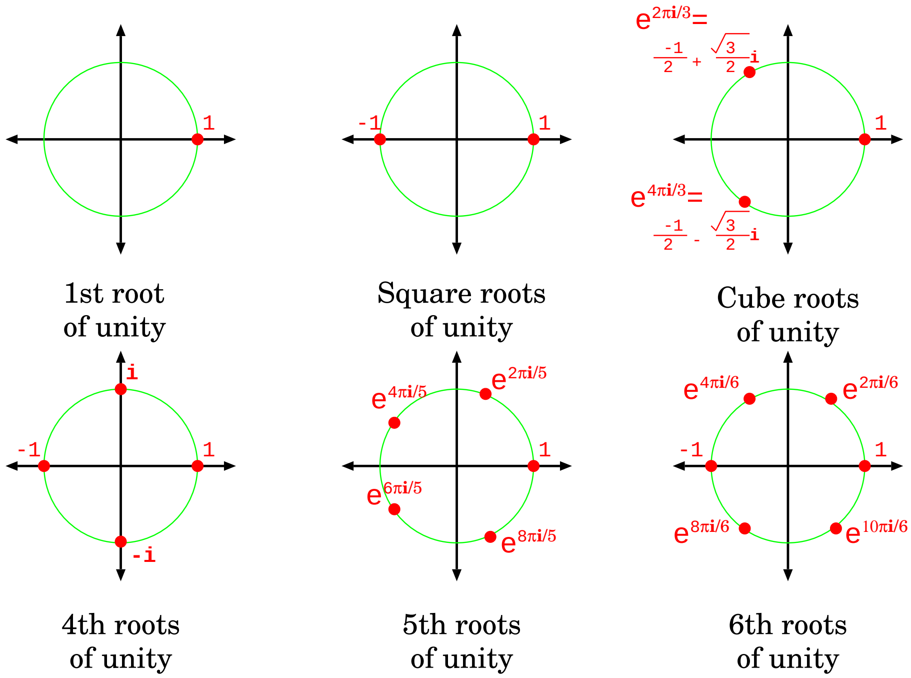
11.6 Cyclotomic Extensions

Prerequisites: §11.3

Let $\mathbb{F}$ be a field. If $n \in \mathbb{N}$, then an ***n*th root of unity** is a number $\zeta \in \mathbb{F}$ so that $\zeta^n = 1$.

Example 249:

(a) Let $\mathbb{F} = \mathbb{C}$. The Nth roots of unity (for $N = 1, 2, 3, 4, 5, 6$) are shown in Figure 11.14 and listed in the following table:



1st root of unity | Square roots of unity | Cube roots of unity

4th roots of unity | 5th roots of unity | 6th roots of unity

Figure 11.14: Roots of unity in the complex plane.

N	Nth roots of unity
$N = 1$	1.
$N = 2$	1 and -1.
$N = 3$	1, $e^{2\pi\mathbf{i}/3} = \dfrac{-1}{2} + \dfrac{\sqrt{3}}{2}\mathbf{i}$, and $e^{4\pi\mathbf{i}/3} = \dfrac{-1}{2} - \dfrac{\sqrt{3}}{2}\mathbf{i}$.
$N = 4$	1, $\mathbf{i}$, -1, and $-\mathbf{i}$.
$N = 5$	1, $e^{2\pi\mathbf{i}/5}$, $e^{4\pi\mathbf{i}/5}$, $e^{6\pi\mathbf{i}/5}$, and $e^{8\pi\mathbf{i}/5}$.
$N = 6$	1, $e^{\pi\mathbf{i}/3}$, $e^{2\pi\mathbf{i}/3}$, $e^{3\pi\mathbf{i}/6} = -1$, $e^{4\pi\mathbf{i}/3}$, and $e^{5\pi\mathbf{i}/6}$.

In general, for any $n \in \mathbb{N}$, the nth roots of unity are: 1, $e^{2\pi\mathbf{i}/n}$, $e^{4\pi\mathbf{i}/n}, \ldots, e^{\frac{2(n-1)\pi\mathbf{i}}{n}}$ (see Lemma 250(a) below).

As Figure 11.14 suggests, the nth roots of unity form a set of n equally spaced points around the unit circle in the complex plane. There is thus a relationship between nth roots of unity and the construction of a regular n-gon (see §11.7)

(b) Let $\mathbb{F} = \mathbb{Z}_{/5}$. Then $1, 2, 3$ and 4 are *all* 4th roots of unity:

$$\begin{aligned}
1^4 &= 1; \\
2^4 &= 16 \equiv 1 \pmod{5}; \\
3^4 &= 81 \equiv 1 \pmod{5}; \\
4^4 &= 256 \equiv 1 \pmod{5}.
\end{aligned}$$

(c) Let $\mathbb{F} = \mathbb{Z}_{/p}$, for some prime p. Then *all* nonzero elements of $\mathbb{Z}_{/p}$ are $(p-1)$th roots of unity. In other words, we have:

Fermat's Little Theorem: $z^{p-1} \equiv 1 \pmod{p}$, *whenever* $z \not\equiv 0 \pmod{p}$.

Proof: Let $\mathbb{Z}_{/p}{}^{\times} = \{1, 2, \ldots, (p-1)\}$, and note that $\mathbb{Z}_{/p}{}^{\times}$ forms a group of order $(p-1)$ under multiplication. Thus, if $z \in \mathbb{Z}_{/p}{}^{\times}$, then the (multiplicative) order of z divides $p-1$, so that $z^{p-1} = 1$. ______

Lemma 250 *Let $\mathcal{Z}_n$ be the set of all nth roots of unity in $\mathbb{C}$. Then*

(a) $\mathcal{Z}_n = \left\{1,\ e^{2\pi\mathbf{i}/n},\ e^{4\pi\mathbf{i}/n}, \ldots, e^{\frac{2(n-1)\pi\mathbf{i}}{n}}\right\}$. *Thus,* $\mathsf{card}\,[\mathcal{Z}_n] = n$.

(b) *$\mathcal{Z}_n$ is a cyclic group under multiplication, and is generated by $e^{2\pi\mathbf{i}/n}$.*

(c) *$\mathcal{Z}_n$ is isomorphic to $\mathbb{Z}_{/n}$ via the map* $\psi : \mathbb{Z}_{/n} \ni k \mapsto e^{\frac{2k\pi\mathbf{i}}{n}} \in \mathcal{Z}_n$.

(d) *If $d \in \{1, 2, \ldots, n\}$, then* $\Big(d \text{ divides } n\Big) \iff \Big(\mathcal{Z}_d \subset \mathcal{Z}_n\Big)$.

Proof: **(a,c)** **Exercise 182**.

(b) First note that $\mathcal{Z}_n$ is *closed under multiplication:* if $\zeta_1, \zeta_2 \in \mathcal{Z}_n$, then

$$(\zeta_1 \cdot \zeta_2)^n = \zeta_1^n \cdot \zeta_2^n = 1 \cdot 1 = 1; \qquad \text{hence } (\zeta_1 \cdot \zeta_2) \in \mathcal{Z}_n \text{ also.}$$

Thus, $\mathcal{Z}_n$ forms a group. It is clear that all elements of $\mathcal{Z}_n$ are powers of $e^{2\pi\mathbf{i}/n}$, so $e^{2\pi\mathbf{i}/n}$ generates $\mathcal{Z}_n$, so $\mathcal{Z}_n$ is cyclic.

(d) Recall that $\zeta = e^{\frac{2\pi\mathbf{i}}{d}}$ generates $\mathcal{Z}_d$. Thus,

$$\Big(\mathcal{Z}_d \subset \mathcal{Z}_n\Big) \iff \Big(\zeta \in \mathcal{Z}_n\Big) \iff \Big(\zeta^n = 1\Big) \iff \Big(e^{\frac{2n\pi\mathbf{i}}{d}} = 1\Big)$$
$$\iff \Big(\tfrac{2n\pi}{d} \text{ is an integer multiple of } 2\pi\Big) \iff \Big(\tfrac{n}{d} \text{ is an integer}\Big)$$
$$\iff \Big(d \text{ divides } n\Big).$$ ______ □

The nth **cyclotomic field**[1] is the field

$$\mathbb{CF}_n = \mathbb{Q}\left(e^{2\pi\mathbf{i}/5},\ e^{4\pi\mathbf{i}/n}, \ldots, e^{\frac{2(n-1)\pi\mathbf{i}}{n}}\right),$$

generated by all nth roots of unity over the rational numbers. Our first task is to show that $\mathbb{CF}_n$ is in fact a *simple* extension of $\mathbb{Q}$, generated by a single element: a 'primitive' root of unity.

Lemma 250(a) says that $\mathcal{Z}_n$ is a group under multiplication. A **primitive nth root of unit** is an element of $\mathcal{Z}_n$ which *generates* $\mathcal{Z}_n$ as a group. The following table lists the primitive Nth roots of unity for $N = 2, 3, 4, 5, 6$:

[1] The term 'cyclotomic' means 'circle-cutting' ('cyclo-tomic') in Greek, and refers to the fact that the nth roots of unity 'cut' the circle into n equal peices.

N	Primitive Nth roots of unity
$N=2$	-1.
$N=3$	$e^{2\pi\mathbf{i}/3}$ and $e^{4\pi\mathbf{i}/3}$.
$N=4$	$\mathbf{i}$, and $-\mathbf{i}$.
$N=5$	$e^{2\pi\mathbf{i}/5}$, $e^{4\pi\mathbf{i}/5}$, $e^{6\pi\mathbf{i}/5}$, and $e^{8\pi\mathbf{i}/5}$.
$N=6$	$e^{2\pi\mathbf{i}/6}$ and $e^{10\pi\mathbf{i}/6}$.
$N=7$	$e^{2\pi\mathbf{i}/7}$, $e^{4\pi\mathbf{i}/7}$, $e^{6\pi\mathbf{i}/7}$, $e^{8\pi\mathbf{i}/7}$, $e^{10\pi\mathbf{i}/7}$, and $e^{12\pi\mathbf{i}/7}$.
$N=8$	$e^{\pi\mathbf{i}/4}$, $e^{3\pi\mathbf{i}/4}$, $e^{5\pi\mathbf{i}/4}$, and $e^{7\pi\mathbf{i}/4}$.
$N=9$	$e^{2\pi\mathbf{i}/9}$, $e^{4\pi\mathbf{i}/9}$, $e^{8\pi\mathbf{i}/9}$, $e^{10\pi\mathbf{i}/9}$, $e^{14\pi\mathbf{i}/9}$, and $e^{16\pi\mathbf{i}/9}$.

Let $\mathcal{Z}_n^* = \{\zeta \in \mathcal{Z}_n \,;\, \zeta \text{ a primitive root}\}$. The primitive roots allow us to express $\mathbb{CF}_n$ as a *simple* extension of $\mathbb{Q}$:

Lemma 251 *If $\zeta \in \mathcal{Z}_n^*$ be any primitive nth root of unity, then $\mathbb{CF}_n = \mathbb{Q}(\zeta)$.*

Proof: First note that $\mathbb{Q}(\zeta) \subset \mathbb{CF}_n$. We claim that also $\mathbb{CF}_n \subset \mathbb{Q}(\zeta)$. To see this, note that $\mathbb{Q}(\zeta)$ contains all powers of ζ; since ζ generates $\mathcal{Z}_n$, it follows that $\mathbb{Q}(\zeta)$ contains all nth roots of unity, so $\mathbb{Q}(\zeta)$ contains $\mathbb{CF}_n$. □

Our next task is to count $\mathcal{Z}_n^*$. We define **Euler's φ-function:**

$$\boldsymbol{\varphi}(n) \quad = \quad \#\text{ of numbers } j \in \{1,2,...,n-1\} \text{ which are relatively prime to } n.$$

Example 252:

(a) $\boldsymbol{\varphi}(5) = 4$, because all of the numbers $\{1,2,3,4\}$ are relatively prime to 5.

(b) $\boldsymbol{\varphi}(16) = 8$, because the numbers $\{1,3,5,7,9,11,13,15\}$ are relatively prime to 16.

See the `Appendix` at the end of this section for more about the φ-function.

Lemma 253

(a) *If $\zeta \in \mathcal{Z}_n$, then* $\Big(\zeta \in \mathcal{Z}_n^*\Big) \iff \Big(k = n-1 \text{ is the smallest } k > 0 \text{ for which } \zeta^k = 1\Big)$.

(b) $\mathcal{Z}_n^* = \left\{e^{2p\pi\mathbf{i}/n} \,;\, \text{where } p \text{ is relatively prime to } n\right\}$.

(c) *More generally, if $\zeta \in \mathcal{Z}_n^*$ is any fixed element, then $\mathcal{Z}_n^* = \{\zeta^p \,;\; p \text{ is relatively prime to } n\}$.*

(d) *Thus,* $\mathsf{card}\,[\mathcal{Z}_n^*] = \boldsymbol{\varphi}(n)$.

(e) $\mathcal{Z}_n = \bigsqcup_{d \text{ divides } n} \mathcal{Z}_d^*$.

Proof: **(a)** If $\zeta \in \mathcal{Z}_n$, then $\zeta^{n-1} = 1$, because $|\mathcal{Z}_n| = (n-1)$. Thus, $\zeta \in \mathcal{Z}_n^*$ if and only if $\mathcal{Z}_n = \{1, \zeta, \zeta^2, \ldots, \zeta^{n-1}\}$, and this occurs only if $k = n-1$ is the smallest $k > 0$ for which $\zeta^k = 1$.

(b,c) follow from **(a)** (**Exercise 183**). **(d)** follows from the definition of $\boldsymbol{\varphi}(n)$.

(e) For all $d \leq n$, we define

$$\mathcal{Z}'_d \quad = \quad \left\{\zeta \in \mathcal{Z}_n \; ; \; k = d-1 \text{ is the smallest } k > 0 \text{ for which } \zeta^k = 1\right\}.$$

Claim 1: $\mathcal{Z}_n \;=\; \bigsqcup_{d \text{ divides } n} \mathcal{Z}'_d.$

Proof: **Exercise 184** . First check the following:

- If $d_1 \neq d_2$, then $\mathcal{Z}'_{d_1}$ and $\mathcal{Z}'_{d_2}$ are disjoint.
- Every element of $\mathcal{Z}_n$ must belong to $\mathcal{Z}'_d$ for some d.
- $\mathcal{Z}'_d = \emptyset$ unless d divides n.

Now combine these to get the Claim. .... □ [Claim 1]

Claim 2: *For every d which divides n,* $\mathcal{Z}'_d = \mathcal{Z}^*_d$.

Proof: Observe that:

- $\mathcal{Z}_d \subset \mathcal{Z}_n$; this follows from Lemma 250**(d)**.
- $\mathcal{Z}'_d \subset \mathcal{Z}_d$; to see this, note that, if $\zeta \in \mathcal{Z}'_d$, then $\zeta^d = 1$ by definition.
- Part **(a)** of the present theorem says: for any $\zeta \in \mathcal{Z}_d$, $\left(\zeta \in \mathcal{Z}^*_n\right) \iff \left(\zeta \in \mathcal{Z}'_n\right)$.

It follows that $\mathcal{Z}'_d = \mathcal{Z}^*_d$. .... □ [Claim 2]

(d) follows from Claims 1 and 2. ____________________________________ □

Remark: Note that part **(d)** of the theorem yields the surprising identity:

$$(n-1) \quad = \quad \sum_{d \text{ divides } n} \boldsymbol{\varphi}(n)\,.$$

Our next goal is to represent $\mathbb{CF}_n$ as the splitting field of a single irreducible polynomial. We define the n**th cyclotomic polynomial**:

$$\mathbf{cp}_n(x) \quad = \quad \prod_{\zeta \in \mathcal{Z}^*_n} (x - \zeta). \tag{11.7}$$

This is the 'smallest' polynomial having all elements of $\mathcal{Z}_n^*$ as roots. For example:

$$\begin{array}{rcccl}
x-(1) & = & \mathbf{cp}_1(x) & = & x-1; \\
x-(-1) & = & \mathbf{cp}_2(x) & = & x+1; \\
\left(x-e^{\frac{2\pi\mathbf{i}}{3}}\right)\cdot\left(x-e^{\frac{4\pi\mathbf{i}}{3}}\right) & = & \mathbf{cp}_3(x) & = & x^2+x+1; \\
(x+\mathbf{i})\cdot(x-\mathbf{i}) & = & \mathbf{cp}_4(x) & = & x^2+1; \\
\left(x-e^{\frac{2\pi\mathbf{i}}{5}}\right)\cdot\left(x-e^{\frac{4\pi\mathbf{i}}{5}}\right)\cdot\left(x-e^{\frac{6\pi\mathbf{i}}{5}}\right)\cdot\left(x-e^{\frac{8\pi\mathbf{i}}{5}}\right) & = & \mathbf{cp}_5(x) & = & x^4+x^3+x^2+x+1; \\
\left(x-e^{\frac{\pi\mathbf{i}}{3}}\right)\cdot\left(x-e^{\frac{5\pi\mathbf{i}}{3}}\right) & = & \mathbf{cp}_6(x) & = & x^2-x+1; \\
\left(x-e^{\frac{2\pi\mathbf{i}}{7}}\right)\cdot\left(x-e^{\frac{4\pi\mathbf{i}}{7}}\right)\cdot\left(x-e^{\frac{6\pi\mathbf{i}}{7}}\right) & & & & \\
\cdot\left(x-e^{\frac{8\pi\mathbf{i}}{7}}\right)\cdot\left(x-e^{\frac{10\pi\mathbf{i}}{7}}\right)\cdot\left(x-e^{\frac{12\pi\mathbf{i}}{7}}\right) & = & \mathbf{cp}_7(x) & = & x^6+x^5+x^4+x^3+x^2+x+1; \\
\left(x-e^{\frac{\pi\mathbf{i}}{4}}\right)\cdot\left(x-e^{\frac{3\pi\mathbf{i}}{4}}\right)\cdot\left(x-e^{\frac{5\pi\mathbf{i}}{4}}\right)\cdot\left(x-e^{\frac{7\pi\mathbf{i}}{4}}\right) & = & \mathbf{cp}_8(x) & = & x^4+1; \\
\left(x-e^{\frac{2\pi\mathbf{i}}{9}}\right)\cdot\left(x-e^{\frac{4\pi\mathbf{i}}{9}}\right)\cdot\left(x-e^{\frac{8\pi\mathbf{i}}{9}}\right) & & & & \\
\cdot\left(x-e^{\frac{10\pi\mathbf{i}}{9}}\right)\cdot\left(x-e^{\frac{14\pi\mathbf{i}}{9}}\right)\cdot\left(x-e^{\frac{16\pi\mathbf{i}}{9}}\right) & = & \mathbf{cp}_9(x) & = & x^6+x^3+1; \\
 & \vdots & \vdots & \vdots &
\end{array}$$

This partial list suggests that $\mathbf{cp}_n(x)$ always has *integer* coefficients, something which is not obvious from equation (11.7).

Proposition 254

(a) $\mathbf{cp}_n(x)$ *is a monic, irreducible polynomial with integer coefficients.*

(b) $\mathsf{degree}\left(\mathbf{cp}_n(x)\right) \;=\; \boldsymbol{\varphi}(n)$.

(c) *For any* $n\in\mathbb{N}$, $\quad (x^n-1) \;=\; \prod_{d \text{ divides } n} \mathbf{cp}_d(x)$.

Proof: **(b)** $\mathbf{cp}_n(x)$ is a product of $\boldsymbol{\varphi}(n)$ linear factors. Thus, $\mathsf{degree}(\mathbf{cp}_n(x)) \;=\; \boldsymbol{\varphi}(n)$.

(c) The roots of (x^n-1) are exactly the nth roots of unity. Hence, (x^n-1) factors completely over $\mathbb{CF}_n$:

$$x^n-1 \;=\; \prod_{\zeta\in\mathcal{Z}_n}(x-\zeta). \tag{11.8}$$

But Lemma 253**(d)** says $\mathcal{Z}_n \;=\; \bigsqcup_{d \text{ divides } n} \mathcal{Z}_d^*$. Hence, we can rewrite equation (11.8) as:

$$x^n-1 \;=\; \prod_{d \text{ divides } n}\ \prod_{\zeta\in\mathcal{Z}_n^*}(x-\zeta) \;=\; \prod_{d \text{ divides } n} \mathbf{cp}_d(x),$$

where the last equality is just the equation (11.7) which defines $\mathbf{cp}_d(x)$. □

The proof of part **(a)** will require three preliminary lemmas.

Lemma 255 *Let $\mathbb{E} \supset \mathbb{Q}$ be a field extension. Let $\mathbf{q}(x) \in \mathbb{Q}[x]$ and $\mathbf{e}(x) \in \mathbb{E}[x]$, and let $\mathbf{p}(x) = \mathbf{q}(x) \cdot \mathbf{e}(x)$. If $\mathbf{p}(x) \in \mathbb{Q}[x]$, then $\mathbf{e}(x) \in \mathbb{Q}[x]$ also.*

Proof: Apply the `Division Algorithm` to divide $\mathbf{p}(x)$ by $\mathbf{q}(x)$ in $\mathbb{Q}[x]$. We get unique polynomials $\mathbf{s}(x)$ and $\mathbf{r}(x)$ in $\mathbb{Q}[x]$ such that

$$\mathbf{p}(x) \quad = \quad \mathbf{s}(x) \cdot \mathbf{q}(x) \; + \; \mathbf{r}(x), \text{ and } \; \mathsf{degree}\,(\mathbf{r}(x)) \; < \; \mathsf{degree}\,(\mathbf{q}(x))\,. \tag{11.9}$$

Now, the `Division Algorithm` yields a *unique* solution in any polynomial ring, and equation (11.9) is *also* valid in the ring $\mathbb{E}[x]$; hence, $\mathbf{q}(x)$ and $\mathbf{r}(x)$ are the *unique* elements in $\mathbb{E}[x]$ such that (11.9) is true.

However, we already have a division equation in $\mathbb{E}[x]$, namely:

$$\mathbf{p}(x) \quad = \quad \mathbf{q}(x) \cdot \mathbf{e}(x).$$

Hence, we conclude: $\mathbf{e}(x) = \mathbf{s}(x)$ (and $\mathbf{r}(x) = 0$). In other words, $\mathbf{e}(x) \in \mathbb{Q}[x]$. ________□

(**Remark:** Lemma 255 holds for any field, not just $\mathbb{Q}$.)

Lemma 256 (Gauss' Lemma)

Suppose that $\mathbf{q}(x) \in \mathbb{Z}[x]$ is monic, and let $\mathbf{e}(x) \in \mathbb{Q}[x]$. Let $\mathbf{p}(x) = \mathbf{q}(x) \cdot \mathbf{e}(x)$.
If $\mathbf{p}(x) \in \mathbb{Z}[x]$, then $\mathbf{e}(x) \in \mathbb{Z}[x]$ also.

Proof: Suppose $\mathbf{e}(x) = e_0 + e_1 x + e_2 x^2 + \ldots + e_n x^n$, where $e_0, \ldots, e_n \in \mathbb{Q}$. Let D be the lowest common multiple of the denominators of $e_0, \ldots, e_n$. I claim $D = 1$ —ie. all of $e_0, \ldots, e_n$ are actually integers.

To see this, let

$$\mathbf{e}'(x) \quad = \quad D \cdot \mathbf{e}(x) \quad = \quad e'_0 + e'_1 x + e'_2 x^2 + \ldots + e'_n x^n,$$

where $e'_0 = D \cdot e_0$, $e'_1 = D \cdot e_1$, etc. Thus, $\mathbf{e}'(x)$ has all integer coefficients, and we have:

$$D \cdot \mathbf{p}(x) \quad = \quad \mathbf{q}(x) \cdot \mathbf{e}'(x). \tag{11.10}$$

Let p be a prime divisor of D.

Claim 1: *p divides all of $e'_0, \ldots, e'_n$.*

Proof: Reduce all coefficients in $\mathbf{p}(x)$, $\mathbf{q}(x)$, and $\mathbf{e}'(x)$ modulo p. Then equation (11.10) becomes:

$$0 \quad = \quad \overline{\mathbf{q}}(x) \cdot \overline{\mathbf{e}}'(x), \tag{11.11}$$

an equation in the polynomial ring $\mathbb{Z}_{/p}[x]$. But:

$$\Big(\ p \text{ is prime } \Big) =_{\text{Exmpl}\langle 107\rangle}\Rightarrow \Big(\ \mathbb{Z}_{/p} \text{ is a field } \Big) =_{\text{Exmpl}\langle 109\text{g}\rangle}\Rightarrow \Big(\ \mathbb{Z}_{/p}[x] \text{ is an integral domain } \Big).$$

Thus, equation (11.11) implies that either $\overline{\mathbf{q}}(x) = 0$ or $\overline{\mathbf{e}}'(x) = 0$. In other words, either all coefficients in $\mathbf{q}$ are congruent to zero mod p, or all coefficients in $\mathbf{e}'$ are zero mod p.

Now, $\mathbf{q}(x)$ is monic. That is, $\mathbf{q}(x) = x^n + q_{n-1}x^{n-1} + \ldots + q_1 x + q_0$. Thus, $\overline{\mathbf{q}}(x) = x^n + \overline{q}_{n-1}x^{n-1} + \ldots + \overline{q}_1 x + \overline{q}_0$ has a nonzero leading coefficient, so $\overline{\mathbf{q}}(x)$ cannot be zero.

It follows that $e'_0 \equiv \ldots \equiv e'_n \equiv 0 \pmod p$. Thus, all coefficients $e'_0, \ldots, e'_n$ of $\mathbf{e}'$ must be divisible by p. . . . . ☐ `[Claim 1]`

But D is the *lowest* common multiple of the denominators of $e_0, \ldots, e_n$, so $e'_0, \ldots, e'_n$ cannot have any common factors. Contradiction.

We conclude that $D = 1$, which means that $e_0, \ldots, e_n$ were integers all along. ☐

Lemma 257 *If $\zeta, \xi \in \mathcal{Z}_n^*$ are any two primitive roots of unity, then there is a field automorphism $\Phi : \mathbb{CF}_n \longrightarrow \mathbb{CF}_n$ such that:*

(a) $\Phi(\zeta) = \xi$.

(b) *Φ acts as the identity on $\mathbb{Q}$.*

Proof: Lemma 251 says that $\mathbb{Q}(\zeta) = \mathbb{CF}_n = \mathbb{Q}(\xi)$. Thus, every element of $\mathbb{CF}_n$ has the form $q_0 + q_1\zeta + q_2\zeta^2 + \ldots + q_{n-1}\zeta^{n-1}$ for some (not necessarily unique) $q_1, \ldots, q_{n-1} \in \mathbb{Q}$. Define the map $\Phi : \mathbb{CF}_n \longrightarrow \mathbb{CF}_n$ by

$$\Phi(q_0 + q_1\zeta + q_2\zeta^2 + \ldots + q_{n-1}\zeta^{n-1}) \quad = \quad q_0 + q_1\xi + q_2\xi^2 + \ldots + q_{n-1}\xi^{n-1}.$$

Clearly, $\Phi(\zeta) = \xi$, and thus, $\Phi\left(\mathbb{Q}(\zeta)\right) = \mathbb{Q}(\xi)$ —in other words, $\Phi(\mathbb{CF}_n) = \mathbb{CF}_n$. Also, Φ acts as the identity on $\mathbb{Q}$ (because $\Phi(q_0) = q_0$). It is **Exercise 185** to verify that Φ is a field automorphism. ☐

Proof of Proposition 254(a) $\mathbf{cp}_n(x)$ is a product of monic linear polynomials, so it is monic.

Claim 1: *$\mathbf{cp}_n(x)$ has integer coefficients —ie. $\mathbf{cp}_n(x) \in \mathbb{Z}[x]$.*

Proof: (by induction on n).

Assume that $\mathbf{cp}_d(x) \in \mathbb{Z}[x]$ for all $d < n$.

Let $\mathbf{p}(x) = x^n - 1$. From **(c)**, we know that

$$\mathbf{p}(x) \quad = \quad \prod_{d \text{ divides } n} \mathbf{cp}_d(x) \quad = \quad \mathbf{q}(x) \cdot \mathbf{cp}_n(x), \qquad \text{where } \mathbf{q}(x) \;=\; \prod_{\substack{d<n \\ d \text{ divides } n}} \mathbf{cp}_d(x).$$

By induction, $\mathbf{cp}_d(x) \in \mathbb{Z}[x]$ for all $d < n$; hence, $\mathbf{q}(x) \in \mathbb{Q}[x]$. Also, clearly, $\mathbf{p}(x) \in \mathbb{Q}[x]$. Apply Lemma 255 (with $\mathbb{E} = \mathbb{CF}_n$ and $\mathbf{e}(x) = \mathbf{cp}_n(x)$) to conclude that $\mathbf{cp}_n(x) \in \mathbb{Q}[x]$ also.

Now observe that $\mathbf{p}(x) \in \mathbb{Z}[x]$ and $\mathbf{q}(x) \in \mathbb{Z}[x]$, and that $\mathbf{q}(x)$ is monic. It follows from `Gauss' Lemma` (Lemma 256) that $\mathbf{cp}_n(x) \in \mathbb{Z}[x]$. .... □ `[Claim 1]`

$\mathbf{cp}_n(x)$ *is irreducible:* Suppose not. Then $\mathbf{cp}_n(x) = \mathbf{p}(x) \cdot \mathbf{q}(x)$ for some $\mathbf{p}(x)$ and $\mathbf{q}(x)$ in $\mathbb{Q}[x]$. But equation (11.7) says $\mathbf{cp}_n(x) = \prod_{\zeta \in \mathcal{Z}_n^*} (x - \zeta)$, so this means that there are subsets $\mathcal{P}, \mathcal{Q} \subset \mathcal{Z}_n^*$ so that $\mathcal{Z}_n^* = \mathcal{P} \sqcup \mathcal{Q}$, and $\mathbf{p}(x) = \prod_{\zeta \in \mathcal{P}} (x - \zeta)$, and $\mathbf{q}(x) = \prod_{\zeta \in \mathcal{Q}} (x - \zeta)$.

Now, let $\zeta \in \mathcal{P}$ and $\xi \in \mathcal{Q}$. By Lemma 255 we can define an automorphism $\Phi : \mathbb{CF}_n \longrightarrow \mathbb{CF}_n$ so that $\Phi(\zeta) = \xi$ and Φ acts as the identity on $\mathbb{Q}$.

Since, Φ acts as the identity on $\mathbb{Q}$, it follows that $\Phi\left(\mathbf{p}(x)\right) = \mathbf{p}(x)$. But $\mathbf{p}(x) = \prod_{\zeta \in \mathcal{P}} (x - \zeta)$, so this means that $\Phi(\mathcal{P}) = \mathcal{P}$. But $\Phi(\zeta) = \xi \in \mathcal{Q}$; a contradiction.

By contradiction, the hypothetical factorization cannot exist: $\mathbf{cp}_n(x)$ is irreducible. □

Remark: In this proof, we built an automorphism of $\mathbb{CF}_n$ which permuted the roots of $\mathbf{cp}_n(x)$, and used this to deduce information about the structure of $\mathbf{cp}_n(x)$. This use of automorphisms of field extensions (especially those which permute the roots of some polynomial) is the key concept of `Galois Theory`.

Corollary 258

(a) $\mathbf{cp}_n(x)$ *is the minimal polynomial (over $\mathbb{Q}$) for any primitive nth root of unity.*

(b) $\mathsf{deg}\left(\mathbb{CF}_n \supset \mathbb{Q}\right) = \varphi(n)$.

Proof: **(a)** Let ζ be a primitive nth root of unity. Then $\mathbf{cp}_n(x)$ is irreducible, monic, and has ζ as a root, so $\mathbf{cp}_n(x)$ is the minimal polynomial for ζ, according to Proposition 232**(c)**.

(b) This follows from part **(a)**, Proposition254**(b)**, and Proposition 232**(c3)**. □

Appendix: Euler's φ-function

We define **Euler's φ-function:**

$$\boldsymbol{\varphi}(n) \quad = \quad \#\text{ of numbers } j \in \{1, 2, ..., n-1\} \text{ which are relatively prime to } n.$$

Example 259:

(a) $\boldsymbol{\varphi}(7) = 6$, because *all* of the numbers $\{1, 2, 3, 4, 5, 6\}$ are relatively prime to 7.

(b) In general, if p is prime, then $\boldsymbol{\varphi}(p) = p - 1$, because *all* of the numbers $\{1, 2, ..., p-1\}$ are relatively prime to p.

(c) $\boldsymbol{\varphi}(25) = 20$, because the numbers $\{1, 2, 3, 4,\ 6, 7, 8, 9,\ 11, 12, 13, 14,\ 16, 17, 18, 19,\ 21, 22, 23, 24\}$ are relatively prime to 25. The only numbers *not* relatively prime to 25 are the multiples of 5, namely $\{5, 10, 15, 20\}$.

(d) $\boldsymbol{\varphi}(15) = 8$, because the numbers $\{1, 2,\ 4, 7, 8,\ 11, 13, 14\}$ are relatively prime to 25. Note that $15 = 3 \cdot 5$, and $8 = 2 \cdot 4 = \boldsymbol{\varphi}(3) \cdot \boldsymbol{\varphi}(5)$.

These examples illustrate the general rule:

Proposition 260 Properties of Euler's φ-function

(a) *If p is prime, and $n \in \mathbb{N}$, then $\boldsymbol{\varphi}(p^n) = p^{(n-1)} \cdot (p-1)$.*

(b) *If n and m are relatively prime, then $\boldsymbol{\varphi}(n \cdot m) \;=\; \boldsymbol{\varphi}(n) \cdot \boldsymbol{\varphi}(m)$.*

(c) *If n has prime factorization $n = p_1^{\nu_1} \cdot_2^{\nu_2} \cdots p_k^{\nu_k}$, then*

$$\boldsymbol{\varphi}(n) \;=\; p_1^{(\nu_1-1)} \cdot (p_1 - 1) \cdot p_2^{(\nu_2-1)} \cdot (p_2 - 1) \cdots p_k^{(\nu_k-1)} \cdot (p_k - 1).$$

Proof: **(a)** Note that the only numbers in $\{1, 2, \ldots, p^n - 1\}$ which are *not* relatively prime to p^n are the multiples of p, namely $\{p, 2p, \ldots, p^n - p\}$. There are exactly p^{n-1} of these, leaving $p^n - p^{n-1} \;=\; p^{n-1}(p-1)$ relatively prime elements.

(b) **Exercise 186**. **(c)** follows by combining **(a)** and **(b)**. □

11.7 Compass & Straight-Edge Constructions III

Prerequisites: §11.1, §11.4, §11.6 **Recommended:** §11.5

We will now characterize the regular polygons can be constructed with compass and straight-edge. Let $\mathbb{K} \subset \mathbb{R}$ be the field of constructable lengths, from Proposition 225 in §11.1. Recall Proposition 247 from §11.5:

Proposition 247 *Let $\kappa \in \mathbb{K}$. Then:*

(a) *$\mathbb{Q}(\kappa)$ is a multiquadratic extension of $\mathbb{Q}$.*

(b) *Thus, $\mathsf{deg}\left(\mathbb{Q}(\kappa) \supset \mathbb{Q}\right) = 2^n$ for some $n \in \mathbb{N}$.* □

A **Fermat prime** is a prime number of the form $p = 2^{(2^n)} + 1$ for some n. For example:

$$\begin{array}{rclcll} 3 &=& 2^1+1 &=& 2^{(2^0)}+1 & \text{is prime.} \\ 5 &=& 2^2+1 &=& 2^{(2^1)}+1 & \text{is prime.} \\ 17 &=& 2^4+1 &=& 2^{(2^2)}+1 & \text{is prime.} \\ 257 &=& 2^8+1 &=& 2^{(2^3)}+1 & \text{is prime.} \\ 65537 &=& 2^{16}+1 &=& 2^{(2^4)}+1 & \text{is prime.} \end{array}$$

(However, $2^{32} + 1$ is not prime, because it is divisible by 641.) It is an open problem whether there are infinitely many Fermat primes.

Proposition 261 $\left(\textit{The regular } N\textit{-gon is constructable with compass \& straight-edge}\right) \iff$

$\left(N = 2^n \cdot p_1 \cdot p_2 \cdots p_k, \textit{ where } p_1, p_2, \ldots, p_n \textit{ are distinct Fermat primes}\right)$.

Thus, a regular N-gon is constructable if and only if N is one of the following values:

2	4	8	16	32	64	...
3	6	12	24	48	92	...
5	10	20	40	80	160	...
15	30	60	120	240	480	...
17	34	68	136	272	544	...
51	102	204	408	816	1632	...
85	170	340	680	1360	2720	...
257	514	1028	2056	4112	8224	...
⋮	⋮	⋮	⋮	⋮	⋮	⋱

However, a regular N-gon is *not* constructible if N is one of:

$7, 9, 11, 13, 14, 18, 19, 21, 22, 23, 25, 26, 27, 28, 29, 31, 33, 35, 36, 37, 38, 39, 41, 42, 43, 44, 45, 46, 47, 49, \ldots$

Proof of Proposition 261: We will only prove "$\Longrightarrow$". Let $\zeta = e^{2\pi \mathbf{i}/N}$, a primitive Nth root of unity. Write $\zeta = z_r + z_i\mathbf{i}$, where $z_r, z_i \in \mathbb{R}$.

Claim 1: $\left(\textit{The regular } N\textit{-gon is constructable with compass \& straight-edge}\right) \iff$

$\left(z_r \textit{ and } z_i \textit{ are in } \mathbb{K}.\right)$

Proof: Constructing the N-gon is equivalent to cutting the circle into N equal segments, which is equivalent to constructing the angle $2\pi/N$, which is equivalent to constructing the lengths $\cos(2\pi/N)$ and $\sin(2\pi/N)$. But $z_r = \cos(2\pi/N)$ and $z_i = \sin(2\pi/N)$. □ `[Claim 1]`

Claim 2: $\Big(z_r$ *and* z_i *are in* $\mathbb{K}.\Big) \Longrightarrow \Big(\boldsymbol{\varphi}(N)$ *is a power of 2* $\Big)$.

Proof: Let $\mathbb{E} = \mathbb{Q}(z_r, z_i)$.

Claim 2.1: $\mathsf{deg}\Big(\mathbb{Q}(\zeta) \supset \mathbb{E}\Big) = 2.$

Proof: First note that $\mathbb{Q}(\zeta) \subset \mathbb{E}(\mathbf{i})$, because $\zeta = z_r + z_i\mathbf{i}$. Thus, $\mathsf{deg}\Big(\mathbb{Q}(\zeta) \supset \mathbb{E}\Big) \leq \mathsf{deg}\Big(\mathbb{E}(\mathbf{i}) \supset \mathbb{E}\Big) = 2$. But $\mathsf{deg}\Big(\mathbb{Q}(\zeta) \supset \mathbb{E}\Big)$ can't be 1 (because $\mathbb{E}$ is a subset of $\mathbb{R}$, whereas $\mathbb{Q}(\zeta)$ contains complex numbers). So $\mathsf{deg}\Big(\mathbb{Q}(\zeta) \supset \mathbb{E}\Big)$ must be 2. ... □ `[Claim 2.1]`

. Next, observe that:

- $\mathsf{deg}\Big(\mathbb{Q}(\zeta) \supset \mathbb{Q}\Big) = \boldsymbol{\varphi}(N)$, by Corollary 258**(b)** (because ζ is a primitive Nth root of unity).
- $\mathsf{deg}\Big(\mathbb{Q}(\zeta) \supset \mathbb{Q}\Big) \;=\; \mathsf{deg}\Big(\mathbb{Q}(\zeta) \supset \mathbb{E}\Big) \cdot \mathsf{deg}\Big(\mathbb{E} \supset \mathbb{Q}\Big)$, by Theorem 236**(a)**.

It follows that $\boldsymbol{\varphi}(N) \;=\; 2 \cdot \mathsf{deg}\Big(\mathbb{E} \supset \mathbb{Q}\Big)$. But Proposition 247**(b)** says:

$$\Big(z_r \text{ and } z_i \text{ are in } \mathbb{K}. \Big) \Longrightarrow \Big(\mathsf{deg}\Big(\mathbb{E} \supset \mathbb{Q}\Big) \text{ is a power of 2 } \Big).$$

It follows that $\boldsymbol{\varphi}(N)$ is also a power of 2. .... □ `[Claim 2]`

Claim 3: $\Big(\boldsymbol{\varphi}(N)$ *is also a power of* $2 \Big) \iff$
$\Big(N = 2^n \cdot p_1 \cdot p_2 \cdots p_k$, *where* $p_1, p_2, \ldots, p_n$ *are distinct Fermat primes* $\Big)$.

Proof: Suppose N has prime factorization: $N \;=\; 2^n \cdot p_1^{\nu_1} \cdot p_2^{\nu_2} \cdots p_k^{\nu_k}$.
Then Proposition 260**(c)** says that

$$\boldsymbol{\varphi}(N) \;=\; 2^{n-1} p_1^{(\nu_1-1)} \cdot (p_1 - 1) \cdot p_2^{(\nu_2-1)} \cdot (p_2 - 1) \cdots p_k^{(\nu_k-1)} \cdot (p_k - 1).$$

Thus, $\boldsymbol{\varphi}(N)$ is a power of 2 if and only if:

- $\nu_1 = \nu_2 = \ldots = \nu_k = 1$,
- $(p_1 - 1)$, $(p_2 - 1)$, ..., and $(p_k - 1)$ are powers of 2.

In other words, $p_1 = 2^{n_1} + 1$, ..., $p_k = 2^{n_k} + 1$, for some $n_1, \ldots, n_k \in \mathbb{N}$. It remains to show that, $p_1, \ldots, p_k$ must be Fermat primes —ie. that $n_1, \ldots, n_k$ must themselves be powers of 2. This follows from Claim 4 below. .... □ `[Claim 3]`

Claim 4: *If $p = 2^n + 1$ is a prime number, then n must be a power of 2.*

Proof: **Exercise 187** . ... □ `[Claim 4]`

The proof of the converse in the theorem requires Galois theory, and is beyond our scope. □

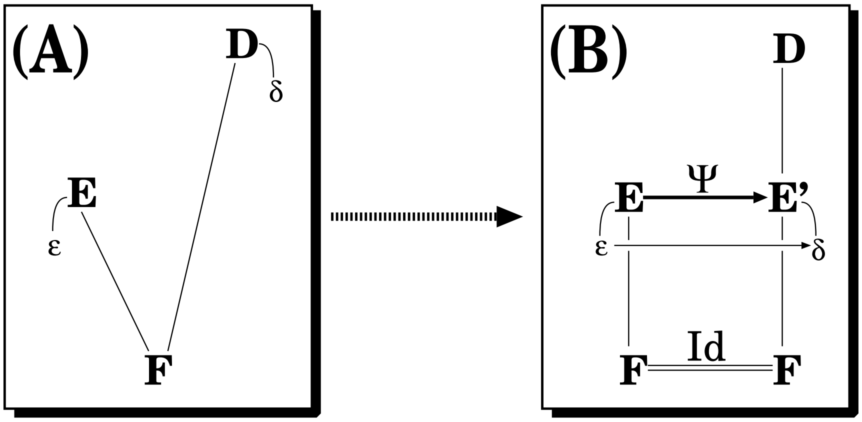


Figure 11.15: $\mathbb{E}$ is the unique minimal root extension of $\mathbb{F}$ for the polynomial $p(x)$.

11.8 Minimal Root Extensions

Prerequisites: §11.2 **Recommended:** §??

Normally, the reason we 'extend' the field $\mathbb{F}$ is to find solutions to polynomial equations. For example, it is irritating that the equation

$$x^2 + 1 \;=\; 0 \tag{11.12}$$

has no solutions $x \in \mathbb{R}$. Thus, we extend $\mathbb{R}$ to the larger field $\mathbb{C}$, which is specifically designed so that the equation (11.12) has solutions (namely $x = \pm\mathbf{i}$).

Recall that a **root** of the polynomial $\mathbf{p}(x)$ is some $\epsilon \in \mathbb{E}$ so that $\mathbf{p}(\epsilon) = 0$. Thus, the polynomial $\mathbf{p}(x) = x^2+1$ has roots $\epsilon = \pm\mathbf{i}$. We often extend fields by 'adjoining' roots to them; for example, we get $\mathbb{C}$ be 'adjoining' the root $\mathbf{i}$ to $\mathbb{R}$. The generalization of this procedure is the following proposition. On first reading, imagine $\mathbb{F} = \mathbb{R}$, $\mathbf{p}(x) = x^2 + 1$, $\mathbb{E} = \mathbb{C}$, and $\epsilon = \mathbf{i}$ (see Example 171 on page 143)

Proposition 262 (Minimal Root Extension) *Let $\mathbb{F}$ be any field and let $\mathbf{p}(x) \in \mathbb{F}[x]$ be an irreducible polynomial.*

There is a unique field extension $\mathbb{E} \supset \mathbb{F}$ such that:

(a) *$\mathbf{p}(x)$ has a root in $\mathbb{E}$ —in other words, there is some $\epsilon \in \mathbb{E}$ so that $\mathbf{p}(\epsilon) = 0$.*

(b) *$\mathbb{E}$ is the minimal extension with this property. In other words, suppose $\mathbb{D} \supset \mathbb{F}$ is another extension of $\mathbb{F}$, and $\delta \in \mathbb{D}$ is a root of $\mathbf{p}(x)$ (see Figure 11.15A). Then there is a subfield $\mathbb{E}' \subset \mathbb{D}$ (see Figure 11.15B) so that*

- *$\mathbb{F} \subset \mathbb{E}'$ and $\delta \in \mathbb{E}'$;*
- *$\mathbb{E} \cong \mathbb{E}'$, via an isomorphism $\Psi : \mathbb{E} \longrightarrow \mathbb{E}'$ such that:*

- $\Psi(\epsilon) = \delta$;
- $\Psi|_{\mathbb{F}} = \mathbf{Id}$.

(c) $\mathsf{deg}\left(\mathbb{E} \supset \mathbb{F}\right) = \mathsf{degree}\left(\mathbf{p}(x)\right)$.

(d) *An $\mathbb{F}$-basis for $\mathbb{E}$ is given by the elements $\{1, \epsilon, \epsilon^2, \ldots, \epsilon^{n-1}\}$, where $n = \mathsf{degree}\left(\mathbf{p}(x)\right)$.*

Proof: **(a)** Let $\mathcal{I} \subset \mathbb{F}[x]$ be the principal ideal generated by $\mathbf{p}(x)$, and let $\mathbb{E} = \mathbb{F}[x]/\mathcal{I}$.

Claim 1: *$\mathbb{E}$ is a field.*

Proof: This follows from Proposition 170, whose proof we recapitulate:
$\Big(\ \mathbf{p}(x)\text{ is irreducible.}\ \Big) =_{\text{Prop.}165}\Rightarrow \Big(\ \mathcal{I}\text{ is a maximal ideal.}\ \Big) =_{\text{Cor.}167}\Rightarrow \Big(\ \mathbb{E}\text{ is a field.}\ \Big)$.
□ [Claim 1]

Let $\Phi : \mathbb{F}[x] \longrightarrow \mathbb{E}$ be the quotient map. For any polynomial $\mathbf{q} \in \mathbb{F}[x]$, let $\overline{\mathbf{q}} = \Phi(\mathbf{q})$ be the corresponding element of $\mathbb{E}$. Recall that $\mathbb{F} \subset \mathbb{F}[x]$ (since elements of $\mathbb{F}$ can be treated as constant polynomials). Let $\overline{\mathbb{F}} \subset \mathbb{E}$ be the image of $\mathbb{F}$ in $\mathbb{E}$.

Claim 2: $\overline{\mathbb{F}} \cong \mathbb{F}$.

Proof: <u>**Exercise 188**</u> .... □ [Claim 2]

Thus, we will identify $\mathbb{F}$ with $\overline{\mathbb{F}}$, and think of it $\mathbb{F}$ as a subfield of $\mathbb{E}$.

Now, consider the polynomial $x \in \mathbb{F}[x]$. Let $\epsilon = \overline{x}$

Claim 3: $\mathbf{p}(\epsilon) = 0$.

Proof: $\mathbf{p}(\epsilon) = \mathbf{p}(\overline{x}) = \overline{\mathbf{p}(x)} = \Phi(\mathbf{p}) = 0$, by definition. .... □ [Claim 3]

Proof of (d) (*This is very similar to the proof of* Proposition 232**(b4)**).

We must show that the set $\{1, \epsilon, \epsilon^2, \ldots, \epsilon^{n-1}\}$ is both *linearly independent* and a *spanning set* for $\mathbb{E}$ as an $\mathbb{F}$-vector space.

Spanning Set:

Claim 4: *Every element of $\mathbb{E}$ has the form $f_0 + f_1\epsilon + f_2\epsilon^2 + \ldots + f_m\epsilon^m$ for some $m \in \mathbb{N}$ and $f_0, \ldots, f_m \in \mathbb{F}$.*

Proof: If $e \in \mathbb{E}$, then $e = \overline{\mathbf{f}(x)}$, for some polynomial $\mathbf{f}(x) \in \mathbb{F}[x]$. Suppose $\mathbf{f}(x) = f_0 + f_1x + f_2x^2 + \ldots + f_mx^m$; then $\overline{\mathbf{f}(x)} = \overline{f}_0 + \overline{f}_1\overline{x} + \overline{f}_2\overline{x}^2 + \ldots + \overline{f}_m\overline{x}^m = f_0 + f_1\epsilon + f_2\epsilon^2 + \ldots + f_m\epsilon^m$.
□ [Claim 4]

Claim 5: *If $m \geq n$, then $\epsilon^m = r_0 + r_1\epsilon + r_2\epsilon^2 + \ldots + r_{n-1}\epsilon^{n-1}$ for some coefficients $r_0, \ldots, r_{n-1} \in \mathbb{F}$.*

Proof: Apply Polynomial Long Division to write $x^m = \mathbf{p}(x)\mathbf{q}(x) + \mathbf{r}(x)$ for some $\mathbf{q}, \mathbf{r} \in \mathbb{F}[x]$, where $\mathbf{r}(x)$ is a polynomial of degree less than n. Then

$$\epsilon^m \quad = \quad \overline{x^m} \quad = \quad \overline{\mathbf{p}(x)\mathbf{q}(x)} + \overline{\mathbf{r}(x)} \quad = \quad \mathbf{r}(\epsilon). \qquad \text{(because } \overline{\mathbf{p}(x)} = 0\text{)}$$

Thus, if $\mathbf{r}(x) = r_0 + r_1 x + r_2 x^2 + \ldots + r_{n-1}x^{n-1}$, then $\epsilon^m = r_0 + r_1\epsilon + r_2\epsilon^2 + \ldots + r_{n-1}\epsilon^{n-1}$. □ [Claim 5]

Combine Claims 4 and 5 to conclude that $\{1, \epsilon, \epsilon^2, \ldots, \epsilon^{n-1}\}$ spans $\mathbb{E}$.

Linearly Independent: Suppose that $f_0 + f_1\epsilon + f_2\epsilon^2 + \ldots + f_{n-1}\epsilon^{n-1} = 0$ for some $f_0, \ldots, f_{n-1} \in \mathbb{F}$; we want to show that $f_0 = f_1 = \ldots = f_{n-1} = 0$.

Let $\mathbf{f}(x) = f_0 + f_1 x + f_2 x^2 + \ldots + f_{n-1}x^{n-1}$, a polynomial in $\mathbb{F}[x]$. Then we have $\overline{\mathbf{f}(x)} = \mathbf{f}(\epsilon) = 0$, which means that $\mathbf{f}(x) \in \mathcal{I}$, which means that $\mathbf{p}(x)$ divides $\mathbf{f}(x)$. Thus, either $\mathbf{f} = 0$, or degree $(\mathbf{f}) \geq$ degree $(\mathbf{p}) = n$. But degree $(\mathbf{f}) \leq n - 1$ by construction, so this is impossible. We conclude that $\mathbf{f} = 0$ —in other words, $f_0 = f_1 = \ldots = f_{n-1} = 0$.

Proof of (c) This follows immediately from **(d)**.

Proof of (b) Let $\mathbb{D}$ be some other extension of $\mathbb{F}$, and suppose $\delta \in \mathbb{D}$ was a root of $\mathbf{p}(x)$ —ie. $\mathbf{p}(\delta) = 0$. From **(d)**, we know that every element of $\mathbb{E}$ has the form $f_0 + f_1\epsilon + f_2\epsilon^2 + \ldots + f_{n-1}\epsilon^{n-1}$ for some $f_0, \ldots, f_{n-1} \in \mathbb{F}$. Define $\Psi : \mathbb{E} \longrightarrow \mathbb{D}$ as follows:

$$\Psi\left(f_0 + f_1\epsilon + f_2\epsilon^2 + \ldots + f_{n-1}\epsilon^{n-1}\right) \quad = \quad f_0 + f_1\delta + f_2\delta^2 + \ldots + f_{n-1}\delta^{n-1}$$

It is **Exercise 189** to check that Ψ is a field monomorphism. Thus, if $\mathbb{E}' = \Psi(\mathbb{E})$, then $\mathbb{E}'$ is isomorphic to $\mathbb{E}$, and is a subfield of $\mathbb{D}$. Clearly, Ψ acts as the identity on $\mathbb{F}$, so $\mathbb{F} \subset \mathbb{E}'$. □

We call $\mathbb{E}$ the **minimal root extension** of $\mathbb{F}$ induced by the irreducibe polynomial $\mathbf{p}(x)$.

It follows from the proof of Proposition 262**(d)** that elements of $\mathbb{E}$ can be treated as 'polynomials' in the symbol ϵ, of degree less than n. The multiplication of two elements of $\mathbb{E}$ is then equivalent to multiplying these polynomials, subject to the reduction provided by Claim 5.

Since $\mathbb{E}$ is obtained by 'adjoining' the root ϵ to $\mathbb{F}$, and since elements of $\mathbb{E}$ are 'polynomials in ϵ', we often write "$\mathbb{E} = \mathbb{F}(\epsilon)$".

Example 263:

(a) Let $\mathbb{F} = \mathbb{R}$ and let $\mathbf{p}(x) = x^2 + 1$. Then the root extension $\mathbb{E} = \mathbb{R}[x]/(x^2+1)$ is isomorphic to $\mathbb{C}$, via the identification $\epsilon = \mathbf{i}$. Any element of $c \in \mathbb{C}$ is a 'polynomial in $\mathbf{i}$' of degree 1 —namely, $c = r_1 + r_2\mathbf{i}$, where $r_1, r_2 \in \mathbb{R}$. The multiplication of two complex numbers then has the form:

$$\begin{aligned} (r_1 + r_2\mathbf{i}) \cdot (s_1 + s_2\mathbf{i}) &= r_1 s_1 + r_1 s_2\mathbf{i} + r_2 s_1\mathbf{i} + r_2 s_2\mathbf{i}^2 \underset{(*)}{=} r_1 s_1 + (r_1 s_2 + r_2 s_1)\mathbf{i} - r_2 s_2 \\ &= (r_1 s_1 - r_2 s_2) + (r_1 s_2 + r_2 s_1)\mathbf{i}. \end{aligned}$$

Here, $(*)$ follows from the reduction formula: $\mathbf{i}^2 = -1$, which is the form which Claim 5 takes here.

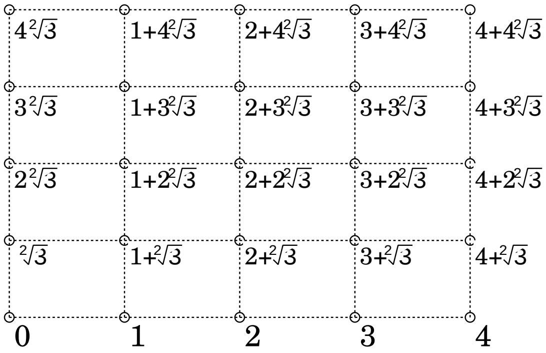


Figure 11.16: $\mathbb{E} = \mathbb{Z}_{/5}(\sqrt{3})$ has 25 elements.

(b) Let $\mathbb{F} = \mathbb{Z}_{/5}$. Observe that the polynomial $\mathbf{p}(x) = x^2 - 3$ has no roots in $\mathbb{Z}_{/5}$ (because $1^2 = 1 = 4^2$ and $2^2 = 4 = 3^2$). Let $\mathbb{E} = \mathbb{Z}_{/5}[x]/(\mathbf{p})$. Then $\mathbb{E}$ is an extension of $\mathbb{Z}_{/5}$ of degree 2, generated by an element ϵ which satisfies $\epsilon^2 - 3 = 0$. Thus, we could write $\epsilon = \sqrt{3}$, so that $\mathbb{E} = \mathbb{Z}_{/5}(\sqrt{3})$. As shown in Figure 11.16, field $\mathbb{E}$ contains exactly 25 elements, each of the form $a + b\sqrt{3}$, where $a, b \in \mathbb{Z}_{/5}$.

Addition in $\mathbb{Z}_{/5}(\sqrt{3})$ takes the obvious form:

$$\left(a_1 + b_1\sqrt{3}\right) + \left(a_2 + b_2\sqrt{3}\right) \quad = \quad (a_1 + a_2) + (b_1 + b_2)\sqrt{3}.$$

Thus, $\mathbb{Z}_{/5}(\sqrt{3})$ is a 2-dimensional vector space over $\mathbb{Z}_{/5}$, with basis $\{1, \sqrt{3}\}$.

Multiplication in $\mathbb{Z}_{/5}(\sqrt{3})$ takes the obvious form:

$$\left(a_1 + b_1\sqrt{3}\right) \cdot \left(a_2 + b_2\sqrt{3}\right) \quad = \quad (a_1a_2 + 3b_1b_2) \;+\; (a_1b_2 + a_2b_1)\,\sqrt{3}.$$

For example:

$$\left(1 + 2\sqrt{3}\right) \cdot \left(1 + 4\sqrt{3}\right) \;=\; (1 \cdot 1 + 3 \cdot 2 \cdot 4) \;+\; (1 \cdot 4 + 2 \cdot 1)\,\sqrt{3} \;=\; 25 + 6\sqrt{3} \;=\; \sqrt{3}.$$

Note that $\mathbb{E}$ contains exactly $25 = 5^2$ elements. We'll see later that a similar construction yields a field of cardinality p^n for any $n \in \mathbb{N}$ and any prime p.

The notation $\mathbb{E} = \mathbb{F}(\epsilon)$ may seem ambiguous, because it suggests the simple extensions introduced on page 187. However, Proposition 262**(b)** can be restated:

Corollary 264 *Let $\mathbb{F} \subset \mathbb{E}$ and $\mathbf{p}(x)$ be as in Proposition 262. If $\mathbb{K}$ is any extension of $\mathbb{F}$ containing a root δ of $\mathbf{p}(x)$, then there is a natural isomorphism $\mathbb{E} \cong \mathbb{F}(\delta)$.* □.

Appendix A

Background: Topology

A.1 Introduction

There are several mathematical approaches to studying space and its properties. Each approach uses a different mathematical structure to represent spatial structure...

Linear Algebra studies the geometry of vector spaces and their subspaces, and is concerned primarily with 'flat' objects like lines and planes.

Differential Geometry uses the tools of differential calculus to study the geometry of curves, surfaces, and other *differentiable manifolds.*

Algebraic Geometry uses the algebraic structure of function rings to study the geometry of curves, surfaces, and other *algebraic varieties.*

Metric Space Theory studies geometry of a space endowed with a concept of distance (a 'metric').

Functional Analysis combines methods of linear algebra and metric spaces to study the geometry of infinite dimensional vector spaces, such as *Banach spaces* and *Hilbert spaces.*

Of all of these, **Topology** is the most abstract and fundamental. Topology studies those properties of space which are *independent* of any particular metric or coordinate structure. Topology is primarily concerned with four issues:

- Convergence.
- Closure.
- Connectedness.
- Continuity.

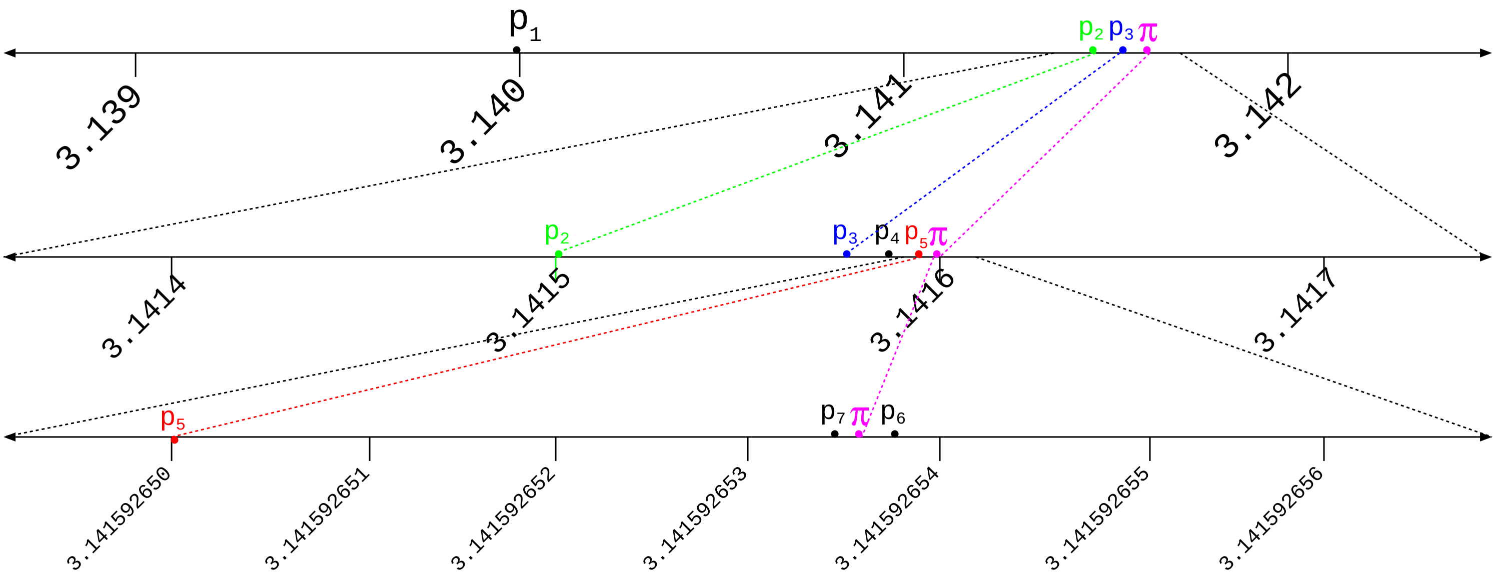


Figure A.1: The sequence $\{p_1, p_2, p_3, \ldots\}$ converges to π.

Convergence: Approximation is an ancient and familiar idea. For example, consider the number π. We can never exactly express π, but for practical purposes it suffices to *approximate* π, for example, by:

$$\pi \;\approx\; 3.14159265358979323854626400000\ldots$$

So, consider this sequence of numbers:

$$\begin{array}{rcl} p_1 & = & 3.14000000000000000\ldots \\ p_2 & = & 3.14160000000000000\ldots \\ p_3 & = & 3.14159300000000000\ldots \\ p_4 & = & 3.14159265000000000\ldots \\ p_5 & = & 3.14159265360000000\ldots \\ p_6 & = & 3.14159265359000000\ldots \\ p_7 & = & 3.14159265358979000\ldots \\ \vdots & \vdots & \ddots \end{array} \qquad \text{(Figure A.1)} \tag{A.1}$$

Clearly, this sequence of numbers is getting closer and closer to π. We say that the sequence $\{p_1, p_2, p_3, \ldots\}$ is **converging** to π. We can justify this as follows:

- p_1 agrees with π up to *two* decimal places.

- p_2 agrees with π up to *four* decimal places.

- p_3 agrees with π up to *six* decimal places.

- ....and in general, for any $L > 0$, there is some N so that p_N agrees with π up to L decimal places. Furthermore, p_n also agrees with π up to L decimal places, for any $n > N$.

Observe that $\Big(p_n \text{ agrees with } \pi \text{ up to } L \text{ decimal places} \Big) \iff \Big(|p_n - \pi| \ < \ \frac{1}{10^L} \Big)$.

Thus, we could just say:

For any $L > 0$, there is some N so that $|p_N - \pi| \ < \ \frac{1}{10^L}$.
Furthermore, $|p_n - \pi| \ < \ \frac{1}{10^L}$, for any $n > N$.

or, more succinctly,

For any $L > 0$, there is some N so that, for any $n \geq N$, $\quad |p_n - \pi| \ < \ \frac{1}{10^L}$.

Of course, there is nothing special about powers of 10. Instead of talking about $\frac{1}{10^L}$, we could measure convergence using any number $\epsilon > 0$. We have arrived at the definition of convergence:

$$\boxed{\text{For any } \epsilon > 0, \text{ there is some } N \text{ so that, for any } n \geq N, \quad |p_n - \pi| \ < \ \epsilon.}$$

We say π is the **limit** of the sequence $\{p_1, p_2, p_3, \ldots\}$, and write:

$$\lim_{n\to\infty} p_n \quad = \quad \pi.$$

If we are looking at sequences of real numbers, then the concept of convergence is pretty straightforward. But in an abstract space $\mathbf{X}$, it is not so transparent. An important question in topology is, *When do sequences in $\mathbf{X}$ converge, and what does this mean?*

Closure: We say that a subset $\mathbf{U} \subset \mathbb{R}$ is **closed** if *no sequence of elements in $\mathbf{U}$ can converge to a point outside of $\mathbf{U}$.*

Example 265:

(a) The open interval $\mathbf{U} = (0, \infty) = \{r \in \mathbb{R} \ ; \ 0 < r\}$ is *not* closed. To see this, consider the sequence

$$\begin{array}{rcl} u_1 & = & 0.1 \\ u_2 & = & 0.01 \\ u_3 & = & 0.001 \\ u_4 & = & 0.0001 \\ \vdots & \vdots & \ddots \end{array} \qquad (A.2)$$

Clearly, the points $u_1, u_2, u_3, \ldots$ are all in $(0, \infty)$, but they converge to 0, which is *not* in $(0, \infty)$. Hence, $(0, \infty)$ is not closed.

(b) The interval $\mathbf{U} = [0,\infty) = \{r \in \mathbb{R} \; ; \; 0 \leq r\}$ *is* closed. Any sequence of numbers greater than or equal to zero must converge to a limit greater than or equal to zero. So no sequence in $[0,\infty)$ can converge to a point outside of $[0,\infty)$. Hence, $[0,\infty)$ is closed.

(c) The interval $[0,1] = \{r \in \mathbb{R} \; ; \; 0 \leq r \leq 1\}$ is closed.

(d) In general, if $-\infty < a < b < \infty$, then (as the name suggests):

- The *closed interval* $[a,b] = \{r \in \mathbb{R} \; ; \; a \leq r \leq b\}$ is closed.
- The *open interval* $(a,b) = \{r \in \mathbb{R} \; ; \; a < r < b\}$ is *not* closed. (a and b are the limits of sequences in (a,b), but a and b are not themselves in (a,b).)
- The *left half-open interval* $(a,b] = \{r \in \mathbb{R} \; ; \; a < r \leq b\}$ is *not* closed. (a is a limit of a sequence in (a,b).)
- The *right half-open interval* $[a,b) = \{r \in \mathbb{R} \; ; \; a \leq r < b\}$ is *not* closed. (b is a limit of a sequence in (a,b).)

(e) The set $\mathbb{Q}$ of rational numbers is *not* closed. To see this, observe that the sequence (A.1) above can be rewritten:

$$p_1 \;=\; \frac{314}{100}; \qquad p_2 \;=\; \frac{31416}{10000}; \qquad p_3 \;=\; \frac{3141593}{1000000}; \qquad \ldots\ldots$$

Thus, $\{p_1, p_2, p_3, \ldots\}$ a sequence of rational numbers, but converges to π, which is *irrational.* Hence $\mathbb{Q}$ is not closed.

The **closure** of a set $\mathbf{U}$ in $\mathbb{R}$ is the *smallest closed subset* of $\mathbb{R}$ which contains $\mathbf{U}$.

Example 266:

(a) The closure of $(0,\infty)$ is $[0,\infty)$.

(b) The closure of $\mathbb{Q}$ is all of $\mathbb{R}$.

The collection of all closed subsets of $\mathbb{R}$ has the following properties:

Lemma 267

(a) *The empty set $\emptyset$ is closed.*

(b) *The set $\mathbb{R}$ of* all *real numbers is closed.*

(c) *If $\mathbf{C}_1$ and $\mathbf{C}_2$ are two closed sets, then their union $\mathbf{C}_1 \cup \mathbf{C}_2$ is also closed.*

(d) *If $\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3, \ldots$ is any collection of closed sets, then the intersection $\bigcap_{n=1}^{\infty} \mathbf{C}_n$ is closed (even if it is empty —see* **(a)***).* □

It is these four properties which motivate the definition of an abstract topological space (see Page 226.)

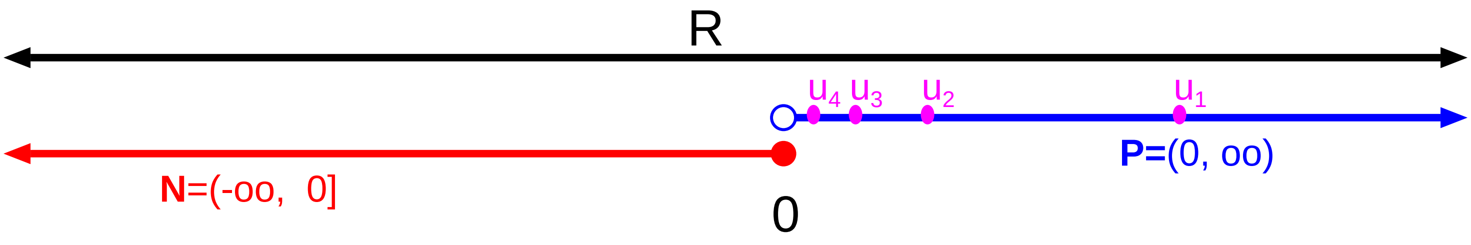


Figure A.2: The subsets $\mathbf{N} = (-\infty, 0]$ and $\mathbf{P} = (0, \infty)$ are disjoint, but there is a sequence in $\mathbf{P}$ converging to a point in $\mathbf{N}$.

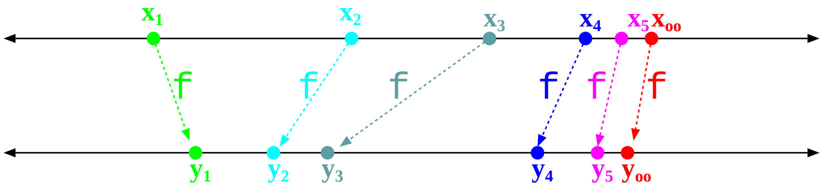


Figure A.3: If $\{x_1, x_2, x_3, \ldots\}$ is a sequence converging to x_∞, and $y_n = f(x_n)$ for all n, then the sequence $\{y_1, y_2, y_3, \ldots\}$ must converge to $y = f(x_\infty)$.

Connectedness: Imagine 'cutting' the real line into two halves: $\mathbf{N} = (-\infty, 0]$ and $\mathbf{P} = (0, \infty)$. Clearly, $\mathbb{R} = \mathbf{N} \sqcup \mathbf{P}$, and the two sets are disjoint. And yet, somehow they seem 'glued together'. The reason is that there is a sequence in $\mathbf{P}$ converging to a point in $\mathbf{N}$. Specifically, the sequence $\{u_1, u_2, u_3, \ldots\}$ defined by (A.2) above is a sequence in $\mathbf{P}$ which converges to 0, a point in $\mathbf{N}$.

Because of this, it seems 'unnatural' to cut $\mathbb{R}$ into $\mathbf{N}$ and $\mathbf{P}$. In doing so, we are 'disrupting' the topological structure of $\mathbb{R}$, by separating sequences from their limits.

The property we are encountering here is *connectedness*. We say that the real number line $\mathbb{R}$ is **connected** because

> *It is impossible to separate $\mathbb{R}$ into two disjoint sets $\mathbf{A}$ and $\mathbf{B}$, such that $\mathbb{R} = \mathbf{A} \sqcup \mathbf{B}$, and so that no sequence in $\mathbf{A}$ converges to a point in $\mathbf{B}$ or vice versa.*

We can express this more succinctly using the concept of closed sets:

> *It is impossible to separate $\mathbb{R}$ into two disjoint $\underline{\text{closed}}$ sets $\mathbf{A}$ and $\mathbf{B}$, such that $\mathbb{R} = \mathbf{A} \sqcup \mathbf{B}$.*

On the other hand, consider the integers $\mathbb{Z}$. Let $\mathbf{N} = \{\ldots, -3, -2, -1, 0\}$, and let $\mathbf{P} = \{1, 2, 3, \ldots\}$. Then $\mathbb{Z} = \mathbf{N} \sqcup \mathbf{P}$, and no sequence in $\mathbf{N}$ can converge to a point in $\mathbf{P}$ or vice versa. To see this, note that $\mathbf{N}$ and $\mathbf{P}$ are separated from one another by the interval $(0, 1)$. No act of sequence convergence can leap across this chasm. Thus, $\mathbb{Z}$ is **disconnected**.

Continuity A function $f : \mathbb{R} \longrightarrow \mathbb{R}$ is *continuous* if it 'preserves' the topological structure of $\mathbb{R}$. For example:

(i) If $\{x_1, x_2, x_3, \ldots\}$ is a sequence converging to x_∞, and $y_n = f(x_n)$ for all n, then the sequence $\{y_1, y_2, y_3, \ldots\}$ must converge to $y_\infty = f(x_\infty)$ (see Figure A.3). In other words,

$$\lim_{n\to\infty} f(x_n) \quad = \quad f\left(\lim_{n\to\infty} x_n\right).$$

(ii) If $\mathbf{U} \subset \mathbb{R}$ is a connected subset, then $f(\mathbf{U})$ is also connected.

(iii) If $\mathbf{U} \subset \mathbb{R}$ is *not* closed, then $f(\mathbf{U})$ is not closed either.

Thus, continuous functions are the 'homomorphisms' of topological structure.

All three properties are contained within the formal definition of continuity. If $f : \mathbb{R} \longrightarrow \mathbb{R}$ is any function, then f is **continuous** if, for any subset $\mathbf{C} \subset \mathbb{R}$,

$$\Big(\ \mathbf{C} \text{ is closed}\ \Big) \Longrightarrow \Big(\ f^{-1}(\mathbf{C}) \text{ is also closed}\ \Big). \tag{A.3}$$

(**Exercise 190** Suppose that the function f satisfies definition (A.3). Show that f satisfies properties **(i)**, **(ii)**, and **(iii)**.)

A.2 Abstract Topological Spaces

So far we've looked at the topology of $\mathbb{R}$. Now we'll develop a mathematical abstraction of these properties. A **topological space** is a set $\mathbf{X}$, along with a collection of subsets $\mathfrak{C}$ of $\mathbf{X}$ (called *closed subsets*), satisfying the following **Closed Set Axioms:**

(C1) The empty set $\emptyset$ is in $\mathfrak{C}$.

(C2) The set $\mathbf{X}$ is in $\mathfrak{C}$.

(C3) If subsets $\mathbf{C}_1$ and $\mathbf{C}_2$ are in $\mathfrak{C}$, then their union $\mathbf{C}_1 \cup \mathbf{C}_2$ is also in $\mathfrak{C}$.

(C4) If $\mathfrak{K} \subset \mathfrak{C}$ is any collection of closed sets, then the intersection $\bigcap_{\mathbf{K}\in\mathfrak{K}} \mathbf{K}$ is also in $\mathfrak{C}$.

Observe that these properties simply recapitulate the properties of Lemma 267 (where, in **(d)**, we set $\mathfrak{K} = \{\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3, \ldots\ldots\}$).

We extend the concepts of *convergence*, *closure*, *connectedness*, and *continuity* to this abstract setting:

Closure: If $\mathbf{U} \subset \mathbf{X}$ is any subset of $\mathbf{X}$, then the **closure** of $\mathbf{U}$ (denoted $\overline{\mathbf{U}}$) is the *smallest closed subset* of $\mathbf{X}$ containing $\mathbf{U}$. Formally:

$$\overline{\mathbf{U}} \quad = \quad \bigcap_{\substack{\mathbf{U}\subset\mathbf{C}\subset\mathbf{X}\\ \mathbf{C}\in\mathfrak{C}}} \mathbf{C}.$$

Exercise 191 1. Check that $\overline{\mathbf{U}}$ is closed. (**Hint:** *Use property* **(d)**.)

2. Show that, if $\mathbf{C}$ is any closed set and $\mathbf{U} \subseteq \mathbf{C}$, then $\overline{\mathbf{U}} \subseteq \mathbf{C}$.

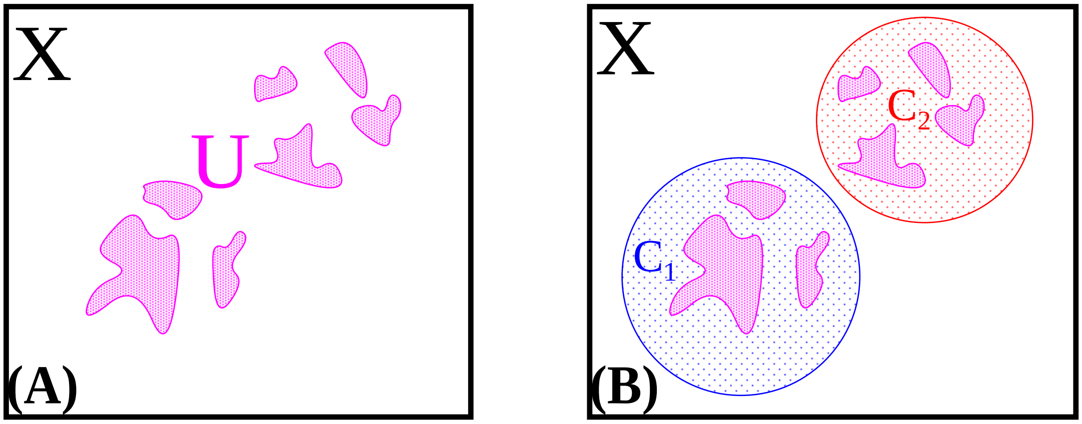


Figure A.4: $\mathbf{U}$ is **disconnected** if there are two disjoint closed subsets $\mathbf{C}_1$ and $\mathbf{C}_2$ in $\mathbf{X}$ so that $\mathbf{U} \subset \mathbf{C}_1 \sqcup \mathbf{C}_2$.

Convergence: Let $\{x_1, x_2, x_3 \ldots\}$ be a sequence of points in $\mathbf{X}$, and let $x_\infty \in \mathbf{X}$ be some other point. We say that the sequence $\{x_1, x_2, x_3 \ldots\}$ **converges** to x_∞ if the *closure* of the set $\{x_1, x_2, x_3 \ldots\}$ is exactly the set $\{x_1, x_2, x_3 \ldots, x_\infty\}$. Formally:

$$\left(\lim_{n\to\infty} x_n \;=\; x_\infty \right) \quad \Longleftrightarrow \quad \left(\overline{\{x_1, x_2, x_3 \ldots\}} \;=\; \{x_1, x_2, x_3 \ldots, x_\infty\} \right).$$

Connectedness: Let $\mathbf{U} \subset \mathbf{X}$ be any subset (Figure A.4A). We say $\mathbf{U}$ is **disconnected** if there are two disjoint closed subsets $\mathbf{C}_1$ and $\mathbf{C}_2$ in $\mathbf{X}$ so that $\mathbf{U} \subset \mathbf{C}_1 \sqcup \mathbf{C}_2$ (Figure A.4B). We say $\mathbf{U}$ is **connected** iff it is not disconnected.

Continuity: Let $\mathbf{X}$ and $\mathbf{Y}$ be topological spaces, and let $f : \mathbf{X} \longrightarrow \mathbf{Y}$. We say that f is **continuous** if, for any $\mathbf{C} \subset \mathbf{Y}$, $\Big(\mathbf{C} \text{ is closed in } \mathbf{Y} \Big) \Longrightarrow \Big(f^{-1}(\mathbf{C}) \text{ is closed in } \mathbf{X} \Big)$.

We say that f is a **homeomorphism** if:

1. f is continuous.
2. If f is a *bijection* from $\mathbf{X}$ to $\mathbf{Y}$. (Thus, f has an inverse map $f^{-1} : \mathbf{Y} \longrightarrow \mathbf{X}$).
3. f^{-1} is also continuous.

We then say that $\mathbf{X}$ and $\mathbf{Y}$ are **homeomorphic**.

Continuous maps are the *homomorphisms* of topological spaces, and homeomorphisms are the *isomorphisms*. If $\mathbf{X}$ and $\mathbf{Y}$ are homeomorphic, then, for topological purposes, they are identical.

Example 268: Let $\mathbf{X} = \left(\frac{-\pi}{2}, \frac{\pi}{2}\right)$, and let $\mathbf{Y} = \mathbb{R}$. As shown in Figure A.5, let $f : \mathbf{X} \longrightarrow \mathbf{Y}$ be the tangent function: $f(x) = \tan(x)$. Then f is continous and bijective, and $f^{-1} = \arctan$ is also continuous. (**Exercise 192**)

Hence, the interval $\left(\frac{-\pi}{2}, \frac{\pi}{2}\right)$ is homeomorphic to $\mathbb{R}$.

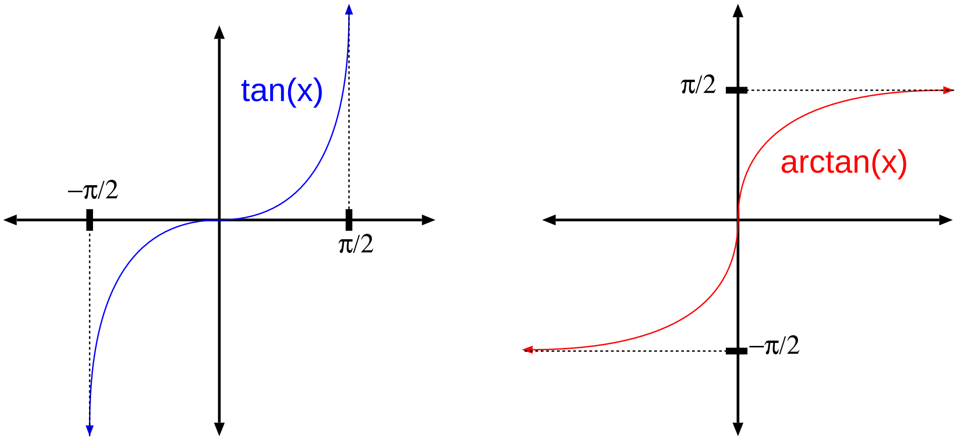


Figure A.5: tan : $\left(\frac{-\pi}{2}, \frac{\pi}{2}\right) \longrightarrow \mathbb{R}$ is a homeomorphism, with inverse homeomorphism arctan : $\mathbb{R} \longrightarrow : \left(\frac{-\pi}{2}, \frac{\pi}{2}\right)$.

A.3 Open sets

An **open set** is the complement of a closed set. In other words, if $\mathbf{X}$ is a topological space, and $\mathbf{U} \subset \mathbf{X}$, then

$$\Big(\ \mathbf{U} \text{ is open}\ \Big) \quad \Longleftrightarrow \quad \Big(\ \mathbf{X} \setminus \mathbf{U} \text{ is closed}\ \Big).$$

Example 269:

(a) The set $\mathbf{U} = (-\infty, 0)$ is open in $\mathbb{R}$, because $\mathbb{R} \setminus \mathbf{U} = [0, \infty)$ is closed.

(b) The interval $\mathbf{U} = (0, 1)$ is open in $\mathbb{R}$, because $\mathbb{R} \setminus \mathbf{U} = (-\infty, 0] \sqcup [1, \infty)$ is closed.

An equivalent definition of openness is:

$$\Big(\ \mathbf{U} \text{ is open}\ \Big) \quad \Longleftrightarrow \quad \Big(\ \text{No point in } \mathbf{U} \text{ is a limit of a sequence in } \mathbf{X} \setminus \mathbf{U}\ \Big). \quad (\textbf{Exercise 193})$$

Note: It is *not* true that **U** is open if **U** is not closed. *Most* subsets are neither open nor closed.

Example 270:

(a) $[0, 1)$ is neither open nor closed in $\mathbb{R}$.

(b) $\mathbb{Q}$ is neither open nor closed in $\mathbb{R}$.

Open subsets satisfy properties which are 'dual' to properties **(C1)** to **(C4)** of closed sets:

(O1) The empty set $\emptyset$ is open.

(O2) The set **X** is open.

(O3) If subsets $\mathbf{O}_1$ and $\mathbf{O}_2$ are open, then their union $\mathbf{O}_1 \cap \mathbf{O}_2$ is also open.

(O4) If $\mathfrak{O}$ is any collection of closed sets, then the union $\displaystyle\bigcup_{\mathbf{O}\in\mathfrak{O}} \mathbf{O}$ is also open

<u>**Exercise 194**</u> Prove **(O1)** to **(O4)**, by applying `de Morgan's laws` to the `Closed Set Axioms` **(C1)** to **(C4)**.

Remark: The complement of every open set is closed, and vice versa. Thus the collection of *all open sets* in a space completely determines the collection of *all closed sets*, and vice versa. For this reason, we can define a topological space just as easily by specifying which sets are *open*, rather than by specifying which sets are *closed* (as we've done above). Indeed, in most texts, a 'topological space' is defined to be a set $\mathbf{X}$ and a collection of *open* sets obeying axioms **(O1)** to **(O4)**.

A.4 Compactness

Consider the subsets $[0,\infty)$ and $[0,1]$ in $\mathbb{R}$. Both sets are closed, but there is a difference between them: $[0,1]$ is bounded, and this endows it with fundamentally different topological properties. We say that $[0,1]$ is *compact.*

Formally, a subset $\mathbf{K} \subset \mathbb{R}$ is **compact** if:

1. $\mathbf{K}$ is *closed.*

2. $\mathbf{K}$ is *bounded* (meaning that there is some $M > 0$ so that $|k| < M$ for all $k \in \mathbf{K}$).

Example 271:

(a) If $-\infty < a < b < \infty$, then the closed interval $[a,b]$ is compact.

(b) Any finite union of finite closed intervals is compact. In other words, if $-\infty < a_1 < b_1 < a_2 < b_2 < \ldots < a_n < b_n < \infty$, then the set $[a_1,b_1] \sqcup [a_2,b_2] \sqcup \ldots \sqcup [a_n,b_n]$ is compact. _

The concept of compactness generalizes to abstract topological spaces. A topological space $\mathbf{X}$ is **compact** if it has any of the following three logically equivalent properties:

1. The **Finite Subcover Property**.

2. The **Chinese Box Property**.

3. The **Cluster Point Property**.

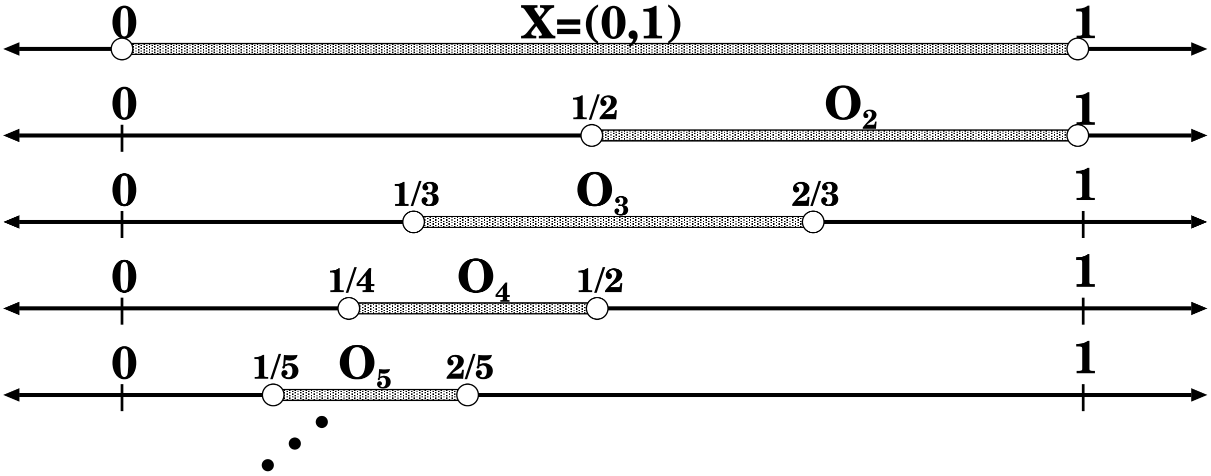


Figure A.6: An open cover of (0, 1) which has no finite subcover.

Finite Subcover Property: *Let $\mathfrak{O}$ be a collection of open sets which* **covers X**, *in the sense that* $\bigcup_{\mathbf{O}\in\mathfrak{O}} \mathbf{O} = \mathbf{X}$. *Then there is a finite subcollection* $\{\mathbf{O}_1, \mathbf{O}_2, \ldots, \mathbf{O}_N\} \subset \mathfrak{O}$ *which* also *covers* **X** *—that is,* $\mathbf{O}_1 \cup \mathbf{O}_2 \cup \ldots \cup \mathbf{O}_N = \mathbf{X}$.

This is usually expressed by the slogan, *Every open cover has a finite subcover.*

Example 272: Let $\mathbf{X} = (0,1)$. For all $n \in \{2,3,4,\ldots\}$, let $\mathbf{O}_n = \left(\frac{1}{n}, \frac{2}{n}\right)$, as in Figure A.6. Let $\mathfrak{O} = \{\mathbf{O}_2, \mathbf{O}_3, \mathbf{O}_4, \ldots\}$. Then $\bigcup_{\mathbf{O}\in\mathfrak{O}} \mathbf{O} = \bigcup_{n=2}^{\infty} \mathbf{O}_n = (0,1)$.

Hence, $\mathfrak{O}$ is an open cover of $(0,1)$. However, no finite subcollection of $\mathfrak{O}$ covers $(0,1)$. Thus, $(0,1)$ *is not compact.*

(Weak) Chinese Box Property: *Suppose* $\mathbf{C}_1 \supset \mathbf{C}_2 \supset \mathbf{C}_3 \supset \cdots$ *is a descending sequence of nonempty closed subsets of* **X**, *as in Figure A.7. Then* $\bigcap_{n=1}^{\infty} \mathbf{C}_n$ *is* also *nonempty.*

Example 273:

(a) Let $\mathbf{X} = [-1,1]$, and for all $n \in \mathbb{N}$, let $\mathbf{C}_n = \left[\frac{-1}{n}, \frac{1}{n}\right]$, as in Figure A.8. Let $\mathfrak{C} = \{\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3, \ldots\}$. If $\{\mathbf{C}_{m_1}, \mathbf{C}_{m_2}, \ldots, \mathbf{C}_{m_k}\} \subset \mathfrak{C}$ is any finite subset, then

$$\mathbf{C}_{m_1} \cap \mathbf{C}_{m_2} \cap \ldots \cap \mathbf{C}_{m_k} = \left[\frac{-1}{M}, \frac{1}{M}\right], \quad \text{where } M = \max\{m_1, m_2, \ldots, m_K\}.$$

Thus, all finite intersections are nonempty, and thus, since $[0,1]$ is compact, we expect that the infinite intersection $\bigcap_{n=1}^{\infty} \mathbf{C}_n$ is also nonempty. Indeed, $\bigcap_{n=1}^{\infty} \mathbf{C}_n = \{0\}$.

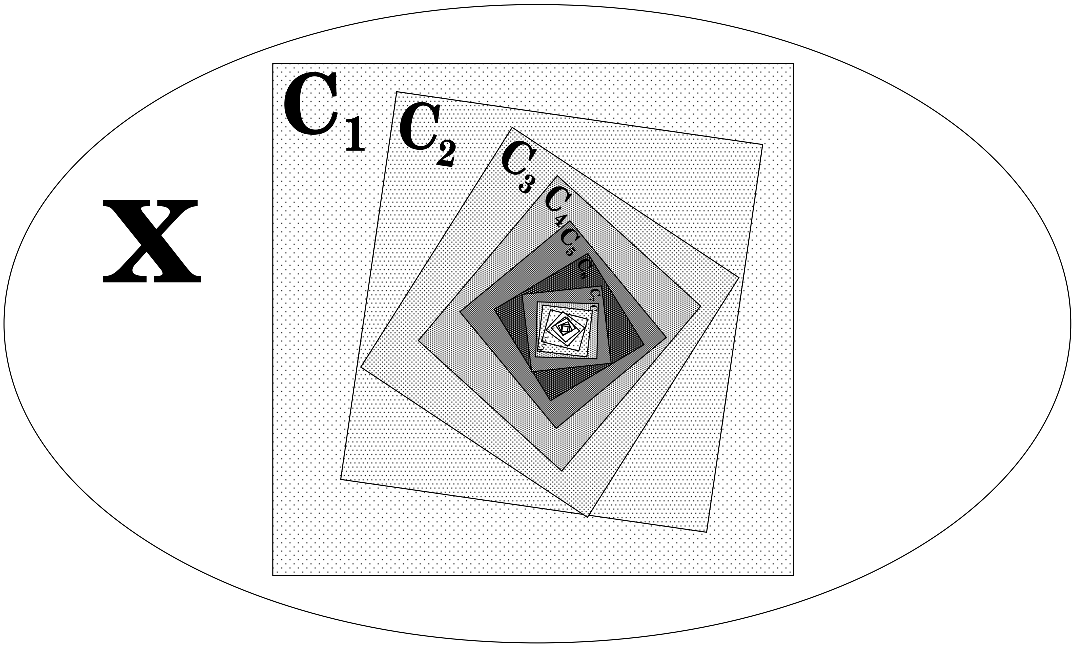


Figure A.7: The Chinese Box property.

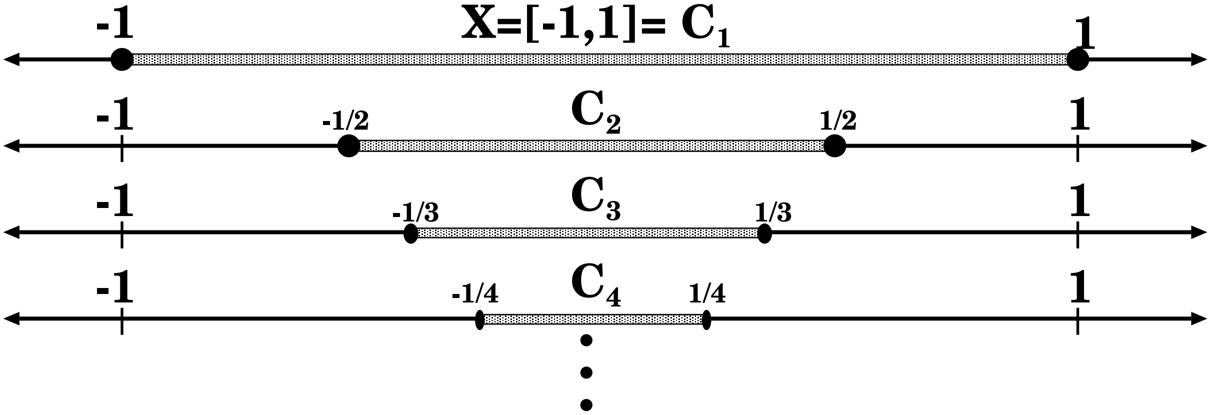


Figure A.8: The Chinese Box Property in the unit interval.

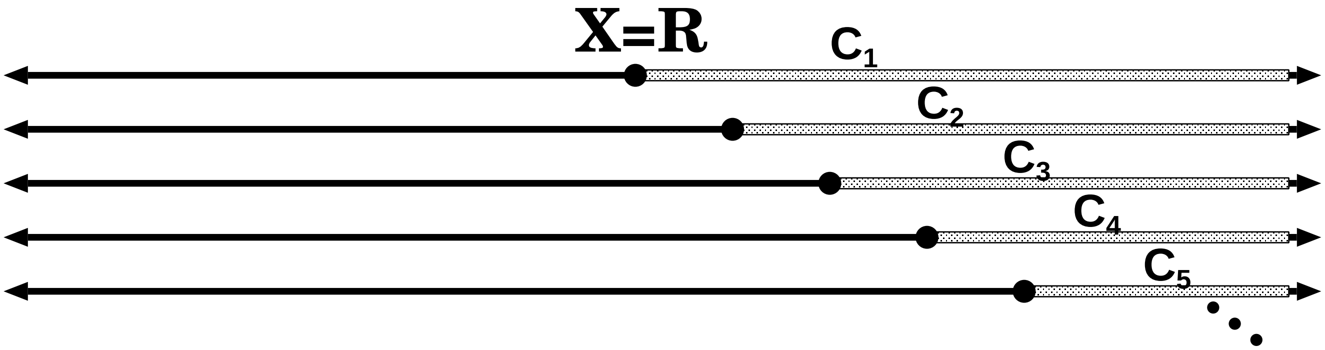


Figure A.9: No Chinese Box Property for the real line.

(b) Let $\mathbf{X} = \mathbb{R}$, and for all $n \in \mathbb{N}$, let $\mathbf{C}_n = [n, \infty)$, as in Figure A.9. Let $\mathfrak{C} = \{\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3, \ldots\}$. If $\{\mathbf{C}_{m_1}, \mathbf{C}_{m_2}, \ldots, \mathbf{C}_{m_k}\} \subset \mathfrak{C}$ is any finite subset, then

$$\mathbf{C}_{m_1} \cap \mathbf{C}_{m_2} \cap \ldots \cap \mathbf{C}_{m_k} \quad = \quad [M, \infty), \quad \text{where } M = \max\{m_1, m_2, \ldots, m_K\}.$$

Thus, all finite intersections are nonempty. But $\mathbb{R}$ is not compact: $\bigcap_{n=1}^{\infty} \mathbf{C}_n \;=\; \emptyset$. ______

The 'weak' Chinese Box Property is actually equivalent to the

(Strong) Chinese Box Property: *Let $\mathfrak{C}$ be a collection of closed sets. Suppose that, for any finite subset $\{\mathbf{C}_1, \mathbf{C}_2, \ldots, \mathbf{C}_N\} \subset \mathfrak{C}$, the intersection $\mathbf{C}_1 \cap \mathbf{C}_2 \cap \ldots \cap \mathbf{C}_N$ is nonempty. Then the full intersection $\bigcap_{\mathbf{C} \in \mathfrak{C}} \mathbf{C}$ is also nonempty.*

Lemma 274 *Let $\mathbf{X}$ be a topological space. Then*

$$\Big(\mathbf{X} \textit{ has the } \texttt{Weak Chinese Box Property}\Big) \iff \Big(\mathbf{X} \textit{ has the } \texttt{Strong Chinese Box Property}\Big).$$

Proof: '$\Longleftarrow$' **Exercise 195** Hint: Let $\mathfrak{C} \;=\; \{\mathbf{C}_1 \supset \mathbf{C}_2 \supset \mathbf{C}_3 \supset \cdots\}$.

'$\Longrightarrow$' **Exercise 196** Hint: Use `Zorn's Lemma`. If you don't know `Zorn's Lemma`, forget about it. □

When we speak of 'the' `Chinese Box Property`, we mean the 'strong' one.

Cluster Point Property: *Let $\mathbf{C} \subset \mathbf{X}$ be any infinite closed set. Then $\mathbf{C}$ contains a* **cluster point** *—that is, a point $c \in \mathbf{C}$ so that $\mathbf{C} \setminus \{c\}$ is not closed.*

This is usually expressed by the slogan, *Every infinite set has a cluster point.*

Example 275:

(a) Let $\mathbf{X} = [0,1]$, and let $\mathbf{C} = \{\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \ldots, 0\}$. Then $\mathbf{C} \subset \mathbf{X}$ is a closed subset, and has a cluster point —namely, 0. If we remove 0, then the set $\mathbf{C} \setminus \{0\} \;=\; \{\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \ldots\}$ is no longer closed.

(b) Let $\mathbf{X} = \mathbb{R}$, and let $\mathbf{C} = \mathbb{N} = \{1, 2, 3, 4, \ldots\}$. Then $\mathbb{N}$ is a closed subset, but it has no cluster points, because we can remove any $n \in \mathbb{N}$ to get the set $\mathbb{N} \setminus \{n\}$, which is *still* closed. Hence, $\mathbb{R}$ is *not* compact. ______

Proposition 276 *Let* $\mathbf{X}$ *be a topological space. Then the following are equivalent:*

(a) $\mathbf{X}$ *has the* `Open Cover Property`.

(b) $\mathbf{X}$ *has the* `Chinese Box Property`.

(c) $\mathbf{X}$ *has the* `Cluster Point Property`.

Proof: **(a)** $\Longleftrightarrow$ **(b)** **Exercise 197** Hint: use de Morgan's Laws.

(b) $\Longrightarrow$ **(c)** Suppose $\mathbf{X}$ has the `Chinese Box Property`, and let $\mathbf{D} \subset \mathbf{X}$ be an infinite, closed subset. We want to show that $\mathbf{D}$ has a cluster point.

Suppose not. Then for every $d \in \mathbf{D}$, let $\mathbf{C}_d = \mathbf{D} \setminus \{d\}$. Since d is not a cluster point of $\mathbf{D}$, the set $\mathbf{C}_d$ is also closed. Let $\mathfrak{C} = \{\mathbf{C}_d \; ; \; d \in \mathbf{D}\}$. This is an infinite collection of closed sets. Furthermore, if $\{d_1, d_2, \ldots, d_n\} \subset \mathbf{D}$ is any finite subset, then

$$\mathbf{C}_{d_1} \cap \mathbf{C}_{d_2} \cap \ldots \cap \mathbf{C}_{d_n} \quad = \quad \mathbf{D} \setminus \{d_1, d_2, \ldots, d_n\}$$

is nonempty (because $\mathbf{D}$ is infinite).

Thus, the `(strong) Chinese Box Property` implies that $\bigcap_{d \in \mathbf{D}} \mathbf{C}_d$ is *also* nonempty. But

$$\bigcap_{d \in \mathbf{D}} \mathbf{C}_d \quad = \quad \mathbf{D} \setminus \{d \; ; \; d \in \mathbf{D}\} \quad = \quad \mathbf{D} \setminus \mathbf{D} \quad = \quad \emptyset.$$

Contradiction.

(c) $\Longrightarrow$ **(b)** **Exercise 198** . ______ □

Further Reading: A very friendly introduction to topology is [5]. An excellent and very complete text is [10].

Bibliography

[1] Michael Artin. *Algebra.* Prentice-Hall, Englewood Cliffs, NJ, 1991.

[2] John B. Conway. *A Course in Functional Analysis.* Springer-Verlag, New York, second edition, 1990.

[3] Robin Hartshorne. *Algebraic Geometry*, volume 52 of *GTM.* Springer-Verlag, New York, 1977.

[4] Nathan Jacobson. *Basic Algebra II.* W.H. Freeman and Co., New York, second edition, 1989.

[5] John G. Hocking and Gail S. Young. *Topology.* Dover, New York, 1961.

[6] Serge Lang. *Algebra.* Addison-Wesley, Reading, MA, 2nd edition, 1984.

[7] Oscar Zariski and P. Samuel. *Commutative Algebra*, volume I & II. van Nostrand, Princeton, NJ, 1958,1960.

[8] I. R. Shavarevich. *Basic Algebraic Geometry*, volume 213 of *Grundlehren.* Springer-Verlag, Heidelberg, 1974.

[9] Frank M. Warner. *Foundations of Differentiable Manifolds and Lie Groups.* Springer-Verlag, New York, 1983.

[10] Stephen Willard. *General Topology.* Addison-Wesley, London, 1970.