Math 6310 Lecture Notes

Lectures by Prof. Marcelo Aguiar, notes by Vivian Kuperberg

Calendar

1 August 25th

1.1 Basic Notions of Group Theory

Notions to know well: group, subgroup, normal subgroup, quotient groups.

Group homomorphisms, its kernel (a normal subgroup of the domain), its image (a subgroup of the codomain), injectivity (when the kernel is trivial), surjectivity (when the image is everything), isomorphisms (bijective homomorphisms). The inverse of an isomorphism is also a homomorphism.

1.2 Isomorphism laws

Proposition 1.2.1 (1st Isomorphism Law)**.** *1. Let $N \trianglelefteq G$ and $\pi : G \to G/N$ be the canonical projection, i.e. $\pi(g) = gN = Ng = \overline{g}$. Then π is a surjective homomorphism and the kernel of π is N.*

2. Let $\phi : G \twoheadrightarrow Q$ be another surjective homomorphism with $\ker\phi = N$. Then the map $\hat{\phi} : G/N \to Q$ defined by $\hat{\phi}(gN) = \phi(g)$ is a well-defined isomorphism.

$$\begin{tikzcd}
N \arrow[r, hook] & G \arrow[r, two heads, "\pi"] \arrow[dr, "\varphi"'] & G/N \arrow[d, dashed, "\cong"] \\
& & Q
\end{tikzcd}$$

Informally, quotient groups correspond to surjective homomorphisms.

Example. $SL(n,\mathbb{F}) \hookrightarrow GL(n,\mathbb{F}) \twoheadrightarrow \mathbb{F}^\times$, where the second map is given by taking the determinant. We conclude from the theorem that $GL(n,\mathbb{F})/SL(n,\mathbb{F}) \cong \mathbb{F}^\times$.

Proposition 1.2.2 (Universal Property of the Quotient)**.** *Let $N \trianglelefteq G$ and $\phi : G \to H$ be a homomorphism, with $N \subseteq \ker\phi$. Then there exists a unique homomorphism $\hat{\phi} : G/N \to H$ such that $\hat{\phi} \circ \pi = \phi$.*

$$\begin{tikzcd}
G \arrow[r, "\varphi"] \arrow[d, two heads, "\pi"'] & H \\
G/N \arrow[ur, dashed, "\hat{\varphi}"'] &
\end{tikzcd}$$

Moreover, $\ker\hat{\phi} = \ker\phi/N$, *and* $\operatorname{im}\hat{\phi} = \operatorname{im}\phi$.

Clearly, this proposition includes the previous proposition as a special case.

Remark.

- Given subsets $X, Y \subseteq G$, let $XY = \{xy \mid x \in X, y \in Y\}$ by definition.

- Even if $X \leq G$ and $Y \leq G$, we are *not guaranteed* that XY is a subgroup of G.
- $N_G(X) = \{g \in G \mid gXg^{-1} = X\}$ is the *normalizer* of X in G. For any $X \subseteq G$, $N_G(X) \leq G$.
- We say Y *normalizes* X if $Y \subseteq N_G(X)$, i.e. if $yXy^{-1} = X \; \forall y \in Y$.
- Let $Y \leq G$. Then Y normalizes $X \iff yXy^{-1} \subseteq X \; \forall y \in Y$. Note that containment in X is sufficient because applying the hypothesis to $y^{-1} \in Y$, we get containment in the other direction for y, i.e. $X \subseteq yXy^{-1}$ as well.

Proposition 1.2.3 (2nd Isomorphism Law)**.** *Let, $N, H \leq G$ be such that H normalizes N. Then*

1. *$NH = HN$ and this is a subgroup of G.*
2. *$N \trianglelefteq NH$.*
3. *$NH/N \cong H/N \cap H$.*

Refer to the NH-diagram below.

$$\begin{tikzcd}
 & NH \arrow[dl, dash] \arrow[dr, dash] & \\
N \arrow[dr, dash] & & H \arrow[dl, dash] \\
 & N \cap H &
\end{tikzcd}$$

Proof. 1. This directly follows from the fact that H normalizes N. This should be easy for you.

2. H normalizes N by hypothesis. N normalizes N always. Thus NH normalizes N.

3. Refer to the diagram below. Define $\phi : H \to NH/N$, the restriction of π to H. Then $\ker \phi = H \cap \ker \pi$. but $\ker \pi = N$, so $\ker \phi = H \cap N$. Now what about the image? We claim ϕ is surjective, but this is nontrivial. An element of NH/N is of the form $\overline{nh} = \overline{n}\overline{h} = \overline{1}\overline{h} = \pi(h) = \phi(h)$. Thus ϕ is surjective. By the 1st Isomorphism Law, we then have $H/N \cap H \cong NH/N$.

$$\begin{tikzcd}
N \cap H \arrow[r, hook] & H \arrow[r, two heads] \arrow[d, hook] \arrow[dr, dashed, "\varphi"] & H/(N \cap H) \\
N \arrow[r, hook] & NH \arrow[r, two heads] & NH/N
\end{tikzcd}$$

□

Proposition 1.2.4 (3rd Isomorphism Law)**.** *Let $N, K \trianglelefteq G$, with $N \subseteq K$. Then*

1. *$K/N \trianglelefteq G/N$.*
2. *$\frac{G/N}{K/N} \cong G/K$.*

Proof. Consider the canonical projection $\pi : G \twoheadrightarrow G/K$. By hypothesis, $N \subseteq K = \ker \pi$. The universal property tells us that there exists a unique homomorphism $\hat{\pi} : G/N \to G/K$. Moreover, $\hat{\pi}$ is surjective because π is, and its kernel is $\ker \pi/N = K/N$, which gives normality. Apply the first isomorphism law to $\hat{\pi}$ to get (2).

$$\begin{tikzcd}
G \arrow[r, two heads, "\pi"] \arrow[d, two heads] & G/K \\
G/N \arrow[ur, dashed, "\hat{\pi}"'] &
\end{tikzcd}$$

□

Proposition 1.2.5 (4th Isomorphism Law)**.** *Let $N \trianglelefteq G$.*

1. *If $N \leq H \leq G$ then $H/N \leq G/N$.*
2. *If $Q \leq G/N$, then $\exists! H$ with $N \leq H \leq G$ and $H/N \cong Q$. In words, there is a bijective correspondence between subgroups of the quotient G/N and intermediate subgroups of G containing N.*
3. *This correspondence preserves inclusions and normality. $H_1 \leq H_2 \iff H_1/N \leq H_2/N$, and $H_1 \trianglelefteq H_2 \iff H_1/N \trianglelefteq H_2/N$.*

1.3 Modularity

Let $X, Y, Z \leq G$. Does this hold?

$$X \cap YZ = (X \cap Y)(X \cap Z)$$

No. What's a counterexample? $G = \mathbb{Z}^2$ under addition, and X, Y, Z where Y is one axis, i.e. pairs of the form $(n, 0)$; Z the other axis, i.e. pairs of the form $(0, n)$; and X the line of pairs of the form (n, n). Then the question is whether $X \cap (Y + Z) = X$, but $(X \cap Y) + (X \cap Z) = 0$.

However, a weaker form of that identity holds, and is called the modular law.

Proposition 1.3.1 (Dedekind's modular law)**.** *Let $X, Y, Z \leq G$, and assume $X \supseteq Z$. Then $X \cap YZ = (X \cap Y)Z = (X \cap Y)(X \cap Z)$.*

Proof. $\supseteq$: $X \cap Y \leq X$, $Z \leq X$, so $(X \cap Y)Z \leq X$. Also $(X \cap Y)Z \leq YZ$, so the left side contains the right. $\subseteq$: An element of $X \cap YZ$ is of the form $x = yz$, with $x \in X, y \in Y, z \in Z$. But then $y = x^{-1}z$, so $y \in X$, and thus in $X \cap Y$, and thus in $(X \cap Y)Z$. $\square$

Remark. On posets. A poset (partially ordered set) is a *lattice* when the *join* (least upper bound) and the *meet* (greatest lower bound) of any two elements exists. Explicitly, a join of x, y is an element z such that $x, y \leq z$, and whenever $x, y \leq z'$, then $z \leq z'$. The meet is defined similarly. These don't necessarily exist; when they do, they are unique, following from the definition. We write $x \vee y$ as the join of x, y and $x \wedge y$ as the meet of x, y. The poset of *normal* subgroups of a group under inclusion is a lattice, where $X \wedge Y = X \cap Y$ and $X \vee Y = XY$; note that this DOES NOT HOLD for the poset of all subgroups. This lattice is *not* distributive, i.e. as explored above, $X \wedge (Y \vee Z) \neq (X \wedge Y) \vee (X \wedge Z)$. But it is modular, where a modular lattice is one in which the identity in the above proposition holds.

The poset of all subgroups is in fact still a lattice but under a different join and meet. We must change $X \vee Y$ to $\langle X, Y \rangle$, or the subgroup generated by X, Y by definition. This lattice is non-distributive and non-modular. A counterexample lies, for example, in the dihedral group D_4.

2 August 27th

2.1 Last Time

We refer to the NH diamond; it is easy to remember and very important.

$$\begin{tikzcd}
& NH \arrow[dl, dash] \arrow[dr, dash] & \\
N \arrow[dr, dash] & & H \arrow[dl, dash] \\
& N \cap H &
\end{tikzcd}$$

Recall as well Dedekind's modular law, that $X \cap YZ = (X \cap Y)(X \cap Z) = (X \cap Y)Z$, provided $X \supseteq Z$.

So now we will move on to the butterfly lemma.

2.2 The Butterfly Lemma

Given subgroups of a group G, A, A_1, B, B_1 as follows:

$$\begin{tikzcd}
& G \arrow[dl, dash] \arrow[dr, dash] & \\
A \arrow[d, dash] & & B \arrow[d, dash] \\
A_1 & & B_1
\end{tikzcd}$$

We want to use each of the following chains to refine the other, as follows.

$$\begin{tikzcd}
A \arrow[d, dash] & & B \arrow[d, dash] \\
A_1(A \cap B) \arrow[d, dash] & & (A \cap B)B_1 \arrow[d, dash] \\
A_1(A \cap B_1) \arrow[d, dash] & & (A_1 \cap B)B_1 \arrow[d, dash] \\
A_1 & & B_1
\end{tikzcd}$$

Remark. We first intersected, then multiplied. If instead we first multiply, then intersect, what would happen? Well, $A \cap A_1 B = A_1(A \cap B)$ by the Dedekind modularity, since $A \geq A_1$.

Proposition 2.2.1 (Zassenhaus Butterfly Lemma)**.** *Let $A_1 \trianglelefteq A \trianglelefteq G, B_1 \trianglelefteq B \trianglelefteq G$. Then*

1. $A_1(A \cap B_1) \trianglelefteq A_1(A \cap B) \leq G$ *and* $(A_1 \cap B)B_1 \trianglelefteq (A \cap B)B_1 \leq G$.
2. $\frac{A_1(A \cap B)}{A_1(A \cap B_1)} \cong \frac{(A \cap B)B_1}{(A_1 \cap B)B_1}$.

Proof. 1. $A_1 \trianglelefteq A$, so A normalizes A_1, so $A \cap B$ normalizes A_1, so $A_1(A \cap B) \leq G$. For normality: check each factor normalizes the smaller subgroup. A_1 normalizes $A_1(A \cap B_1)$ because it is contained in it. $A \cap B$ normalizes $A_1(A \cap B_1)$ because A normalizes A_1, so the first factor is good, and the second factor is good because B normalizes B_1.

Thus $A_1(A \cap B_1) \trianglelefteq A_1(A \cap B)$.

2. For this step we build the butterfly diagram.

$$\begin{tikzcd}
A \arrow[d, dash] & & & & B \arrow[d, dash] \\
A_1(A \cap B) \arrow[dd, dash, "="] \arrow[drr, dash] & & & & (A \cap B)B_1 \arrow[dd, dash, "="] \arrow[dll, dash] \\
& & A \cap B \arrow[dd, dash, "="] & & \\
A_1(A \cap B_1) \arrow[dd, dash] \arrow[drr, dash] & & & & (A_1 \cap B)B_1 \arrow[dd, dash] \arrow[dll, dash] \\
& & (A_1 \cap B)(A \cap B_1) \arrow[ddl, dash] \arrow[ddr, dash] & & \\
A_1 \arrow[dr, dash] & & & & B_1 \arrow[dl, dash] \\
& A_1 \cap B & & A \cap B_1 &
\end{tikzcd}$$

We then claim that the top pieces are NH diamonds. Then the second law will yield the isomorphism that we want. To see that the left piece is an NH diamond, we check:

$$\begin{aligned} A_1(A\cap B_1)(A\cap B) &= A_1(A\cap B), \text{ and} \\ \underbrace{A_1}_{Y}\underbrace{(A\cap B_1)}_{Z}\cap\underbrace{(A\cap B)}_{X} &= (A_1\cap A\cap B)(A\cap B_1\cap A\cap B), \text{ using Dedekind modularity} \\ &= (A_1\cap B)(A\cap B_1). \end{aligned}$$

The right hand piece works similarly.

□

2.3 Series

Let G be a group.

Definition 2.3.1. A *series* is a finite sequence of subgroups, each one contained in the preceding and ranging from G to $\{1\}$, i.e.

$$G = G_0 \geq G_1 \geq \cdots G_n = \{1\}.$$

The *length* of such a series is n. It is *proper* if $G_i \neq G_{i+1}$ for all $i = 0 \ldots n-1$. It is *subnormal* if $G_i \trianglelefteq G_{i-1}$ for all i, and it is *normal* if $G_i \trianglelefteq G$ for all i.

Definition 2.3.2. A second series of G is a *refinement* of the first if it consists of the G_i's, plus possibly some intermediate cases.

Let $G = G_0 \trianglerighteq G_1 \cdots \trianglerighteq G_n = \{1\}$ be a subnormal series. The groups G_{i-1}/G_i, $i = 1, \ldots, n$, are the *slices*. Note that the series being proper is equivalent to all slices being nontrivial.

Two subnormal series are *equivalent* if their nontrivial slices are isomorphic, possibly appearing in different orders.

Examples.

1. $D_n = \langle \rho, \sigma : \rho^n = \sigma^2 = 1, \sigma\rho = \rho^{-1}\sigma \rangle$, the dihedral group of order $2n$. There is a series $D_n \triangleright \langle \rho \rangle \triangleright \{1\}$. The slices are $\mathbb{Z}_2$ and $\mathbb{Z}_n$.

2. $\mathbb{Z}_6 \triangleright \langle \overline{2} \rangle \triangleright \{\overline{0}\}$. The slices are $\mathbb{Z}_6/\langle \overline{2} \rangle = \frac{\mathbb{Z}/6\mathbb{Z}}{2\mathbb{Z}/6\mathbb{Z}} \cong \mathbb{Z}/2\mathbb{Z} = \mathbb{Z}_2$.

   Also, there is the series $\mathbb{Z}_6 \triangleright \langle \overline{3} \rangle \triangleright \{\overline{0}\}$. The slices are, by similar arguments, $\mathbb{Z}_6/\langle \overline{3} \rangle \cong \mathbb{Z}_3$, and $\langle \overline{3} \rangle \cong \mathbb{Z}_2$. The two series are equivalent.

3. $S_n \triangleright A_n \triangleright \{1\}$. A_n is simple if $n > 4$, but we do have $S_4 \triangleright A_4 \triangleright V_4 \triangleright \mathbb{Z}_2 \triangleright \{1\}$, where V_4 is the Klein four group, or as it lies in A_4, $\{id, (12)(34), (13)(24), (14)(23)\}$. $\mathbb{Z}_2$ in this series is $\{id, (12)(34)\}$.

4. $GL(n, \mathbb{F}) \trianglerighteq SL(n, \mathbb{F}) \trianglerighteq \mu_n(\mathbb{F}) \trianglerighteq \{1\}$. The first slice is $\mathbb{F}^\times$, as proven last time. The second has a special name; it is $PSL(n, \mathbb{F})$, the projective special linear group.

$$\begin{tikzcd}
\mu_n(\mathbb{F}) \arrow[r, hook] \arrow[d, hook] & \mathbb{F}^\times \arrow[d, hook] \\
\mathrm{SL}(n, \mathbb{F}) \arrow[r, hook] & \mathrm{GL}(n, \mathbb{F})
\end{tikzcd}$$

Remark. Any finite subgroup of $\mathbb{F}^\times$ is cyclic. If $|\mathbb{F}| = q < \infty$, then $\mu_n(\mathbb{F})$ is cyclic. Moreover, $PSL(n, \mathbb{F})$ is simple as well, unless $n = 2$ and $q = 2$ or 3, or if $n = 1$.

Theorem 2.3.3 (Schreier's Refinement Theorem)**.** *Let $\{G_i\}_{0 \leq i \leq n}$ and $\{H_j\}_{0 \leq j \leq m}$ be two subnormal series of a group G. There exist subnormal refinements $\{G'_i\}_{0 \leq i \leq n'}$ and $\{H'_j\}_{0 \leq j \leq m'}$ of the two series, respectively, such that $\{G'_i\} \sim \{H'_j\}$.*

Proof. For each i and j, insert H_j in between G_i and G_{i+1}, as we did for the butterfly lemma. Let $G_{i,j}$ be defined as $G_{i+1}(G_i \cap H_j)$, and do this for every i and for every j. The H_j's decrease, so fix i and let j vary. We get the following: $G_i \geq G_{i,0} \geq \cdots \geq G_{i,m} \geq G_{i+1}$. Note that the first and last are equalities, because $H_0 = G$ and $H_m = \{1\}$.

By the butterfly lemma, in fact, we get a subnormal chain. Piecing these chains together over $i = 0, \ldots, n$, we obtain a subnormal series of G, which is a refinement of the first series $\{G_i\}$. We can do the same process for the other series, inserting G_i between H_j and H_{j+1}, we obtain a subnormal series which is a refinement of $\{H_j\}$. Here the subgroups are $H_{j,i} = (G_i \cap H_j)H_{j+1}$.

Again by the butterfly lemma, $G_{i,j}/G_{i,j+1} \sim H_{j,i}/H_{j,i+1}$ by the butterfly lemma applied to $G_i, G_{i+1}, H_j, H_{j+1}$. □

Remark. The following is an ongoing analogy:

group	positive integer
subnormal series	factorization
simple groups	prime numbers
composition series	prime factorization

2.4 Composition Series

Definition 2.4.1. A group G is *simple* if

- $G \neq \{1\}$
- The only normal subgroups of G are G and $\{1\}$.

Example.

- A group is simple and abelian if and only if it is cyclic of prime order.
- A_n is simple if $n \neq 1, 2, 4$.
- $PSL(n, \mathbb{F}_q)$ is simple unless $n = 2$ and $q = 2$ or 3, and unless $n = 1$.

3 September 1st

3.1 Composition Series, cont.

Definition 3.1.1. A *composition series* of a group is a subnormal series that is proper (no repeated subgroups), and admits no proper refinements other than itself.

Equivalently, all slices are simple. The slices of a composition series have a special name, they are called *composition factors.*

Along with the analogy mentioned just above, composition factors play the role of primes.

Examples. 1. $\mathbb{Z}$ doesn't have a composition series; any such series would begin like: $\mathbb{Z} \trianglerighteq p_1\mathbb{Z}$, so that the quotient is $\mathbb{Z}_{p_1}$. In the second step, we must then have $p_1\mathbb{Z} \trianglerighteq p_1p_2\mathbb{Z}$, so that the second slice is $\mathbb{Z}_{p_2}$. But this process will never terminate at $\{0\}$, so $\mathbb{Z}$ has no composition series.

   Remark. Some infinite groups have composition series. For example, there exist infinite groups G that are simple, which then have the composition series $1 \trianglelefteq G$.

2. $\mathbb{Z}_6 \triangleright \langle\overline{2}\rangle \triangleright \{\overline{0}\}$ and $\mathbb{Z}_6 \triangleright \langle\overline{3}\rangle \triangleright \{\overline{0}\}$ are two composition series of $\mathbb{Z}_6$.

3. Any finite group has a composition series, by induction on the order of the group.

Theorem 3.1.2 (Jordan-Hölder)**.** *Let G be a group with a composition series. Then any two composition series of G are equivalent.*

Proof. Let $\{G_i\}$ and $\{H_j\}$ be two composition series for G.

By Schreier's refinement theorem there are respective refinements $\{G_i'\}$ and $\{H_j'\}$ such that $\{G_i'\} \sim \{H_j'\}$. Since $\{G_i\}$ has no proper refinements, $\{G_i'\}$ has to have the same nontrivial slices as $\{G_i\}$. Hence $\{G_i'\} \sim \{G_i\}$. Similarly, $\{H_j'\} \sim \{H_j\}$. Thus $\{G_i\} \sim \{H_j\}$. □

Remark. It follows that the composition factors are only dependent on the group G and not on the composition series.

On the other hand, nonisomorphic groups may have the same composition factors. For example, with $\mathbb{Z}_6$, the group S_3 has the composition series $S_3 \triangleright \langle(123)\rangle \triangleright \{\text{id}\}$, which has slices $\mathbb{Z}_2$, $\mathbb{Z}_3$.

The slices cannot always be permuted; for example, there's no composition series of S_3 with the slices in the other order.

Example. For $n > 4$, the composition factors of S_n are $\mathbb{Z}_2$ and A_n, $S_n \triangleright A_n \triangleright \{1\}$. For $n = 4$, this can be refined with the Klein four group and a cyclic subgroup that it contains.

3.2 Solvable groups

Definition 3.2.1. A group is *solvable* if admits a subnormal series with all slices abelian.

Examples.
- All abelian groups are solvable.

- D_n is solvable: $D_n \rhd \langle p \rangle \rhd \{1\}$.

What are the basic properties of solvable groups?

Proposition 3.2.2. *Let G be solvable and $H \leq G$. Then*

1. *H is solvable.*
2. *If $H \trianglelefteq G$, then G/H is solvable.*

Proof. Start from a series with abelian slices. $G = G_0 \trianglerighteq G_1 \trianglerighteq \cdots \trianglerighteq G_n = \{1\}$. Then $H = H \cap G_0 \trianglerighteq H \cap G_1 \trianglerighteq \cdots \trianglerighteq H \cap G_n = \{1\}$. When H is normal, we use the canonical projection $\pi : G \to G/H$ to get $G/H = \pi(G_0) \trianglerighteq \cdots \trianglerighteq \pi(G_n) = \{1\}$; the quotients are abelian as well, so G/H is still solvable. □

Proposition 3.2.3. *Let $N \trianglelefteq G$. Then G is solvable if and only if N and G/N are solvable.*

Proof. ($\Rightarrow$) Previous prop.

($\Leftarrow$). Stick together a series for N with abelian slices with the lift to G of a series for G/N, using the fourth isomorphism law. □

Proposition 3.2.4. *Let G be solvable. Then every subnormal series has a refinement with abelian slices.*

Proof. Apply Schreier to the given series and the series with abelian slices that must exist because G is solvable. Note that the slices of a refinement remain abelian. Given $G_i \trianglerighteq N \trianglerighteq G_{i+1}$ with G_i/G_{i+1} abelian, N/G_{i+1} is the subgroup of an abelian group and thus abelian. Quotients of abelian groups are abelian as well, so $G_i/N \cong \frac{G_i/G_{i+1}}{N/G_{i+1}}$ is abelian. Thus the slices remain abelian, so we have found our refinement. □

Proposition 3.2.5. *Let G be a group with a composition series. Then the following are equivalent:*

i G is solvable.

ii All composition factors are abelian.

iii All composition factors are cyclic of prime order.

Proof. $\underline{(i) \Rightarrow (ii)}$ Apply the previous proposition to a composition series.

$\underline{(ii) \Rightarrow (iii)}$ Simple abelian groups are cyclic of prime order.

$\underline{(iii) \Rightarrow (i)}$ Take the composition series as the subnormal series with abelian factors. □

Example. For $n > 4$, S_n is *not* solvable. A_n is not abelian.

Note then that solvable groups with a composition series are finite.

3.3 The derived series

Definition 3.3.1. Let G be a group. The *commutator* of $g, h \in G$ is $[g, h] = ghg^{-1}h^{-1} \in G$. The *commutator* of $H, K \leq G$ is $[H, K]$, the subgroup generated by $\{[h, k] \mid h \in H, k \in K\}$.

Remark. The product of two commutators is not another commutator.

Lemma 3.3.2. *Let $N \leq G$. Then $[G, G] \subseteq N \iff N \trianglelefteq G$ and G/N is abelian.*

Proof. In Homework 2, exercise 11. □

In particular, $[G, G] = \{1\}$ if and only if G is abelian.

Definition 3.3.3. The *derived subgroup* of G is $G^{(1)} = [G, G]$. We also set $G^{(0)} = G$ and $G^{(i)} = [G^{(i-1)}, G^{(i-1)}] = (G^{(i-1)})^{(1)}$ for all $i \geq 1$. We write $G' = G^{(1)}$, $G'' = G^{(2)}$, and so on.

Note that $G = G^{(0)} \geq G^{(1)} \geq \cdots$. By the lemma, we know something about the slices. We know that $G^{(i)} \trianglelefteq G^{(i-1)}$, and the quotient is abelian. BUT our saving grace is that this series may not terminate. We have not, in fact, proven that every group is solvable!

Proposition 3.3.4. *Each $G^{(i)}$ is a* characteristic *subgroup of G. In particular $G^{(i)} \trianglelefteq G$.*

Note that a subgroup is *characteristic* if it is invariant under all automorphisms of G.

Proof. Being characteristic is transitive (unlike normality). This is because given $H \leq K \leq G$ with all inclusions characteristic, we apply an automorphism to G; it keeps K invariant, so it is an automorphism of K, so it keeps H invariant. Thus H is characteristic in G.

Because of this transitivity, it suffices to show that $G^{(1)}$ is characteristic in G. Take $\sigma \in \mathrm{Aut}(G)$. Then showing that $\sigma([g, h])$ is a commutator is sufficient. But $\sigma(ghg^{-1}h^{-1}) = \sigma(g)\sigma(h)\sigma(g^{-1})\sigma(h^{-1}) = [\sigma(g), \sigma(h)] \in G^{(1)}$. Thus we are done. □

Proposition 3.3.5. *Let G be a group. The following are equivalent.*

i G is solvable.

ii $\exists n \geq 0$ such that $G^{(n)} = \{1\}$.

iii G has a normal series with abelian slices.

Proof. $\underline{(i) \Rightarrow (ii)}$ Let $G = G_0 \trianglerighteq G_1 \trianglerighteq \cdots \trianglerighteq G_n = \{1\}$ be a subnormal series with abelian slices. G/G_1 is abelian, so by our lemma $G^{(1)} \leq G_1$. Similarly, $G_1^{(1)} \leq G_2$, so $G^{(2)} \leq G_1^{(1)} \leq G_2$. By induction, $G^{(i)} \leq G_i$, so eventually $G^{(n)}$ must be trivial.

$\underline{(ii) \Rightarrow (iii)}$ The derived series works.

$\underline{(iii) \Rightarrow (i)}$ This is clear simply because normal implies subnormal. □

3.4 Nilpotent Groups

Definition 3.4.1. A group is *nilpotent* if it admits a normal series $G = G_0 \trianglerighteq G_1 \trianglerighteq \cdots \trianglerighteq G_n = \{1\}$ such that $G_{i-1}/G_i \subseteq Z(G/G_i) \forall i$.

Remark.

1. Abelian $\Rightarrow$ nilpotent $\Rightarrow$ solvable.
2. G nilpotent and $G \neq \{1\}$ implies that $Z(G) \neq \{1\}$, since $G_{n-1} \subseteq Z(G)$.

Proposition 3.4.2.

1. *G nilpotent and $H \leq G \Rightarrow H$ nilpotent.*
2. *G nilpotent and $N \trianglelefteq G \Rightarrow G/N$ nilpotent.*
3. *$N \leq Z(G)$ and G/N is nilpotent $\Rightarrow G$ is nilpotent.*

Proof. HW2, Exercise 9. □

Example. HW2, Exercise 10. D_n is nilpotent if and only if n is a power of 2.

3.5 The Lower Central Series

Definition 3.5.1. Define subgroups of G as follows.

$$G^{[0]} = G, \ldots, G^{[i]} = \left[G, G^{[i-1]}\right].$$

Note that $G^{[1]} = G^{(1)}$, $G^{[2]} = [G, [G, G]]$, and so on.
Then $G = G^{[0]} \trianglerighteq G^{[1]} \trianglerighteq \cdots$ is the *lower central series* of G.

Proposition 3.5.2.

1. *Each $G^{[i]}$ is characteristic in G.*
2. *$G = G^{[0]} \trianglerighteq G^{[1]} \trianglerighteq \cdots$ is well-defined.*
3. *$G^{[i-1]}/G^{[i]} \subseteq Z(G/G^{[i]})$*

Proposition 3.5.3. *G is nilpotent if and only if $\exists n \geq 0$ such that $G^{[n]} = \{1\}$.*

4 September 3rd

4.1 Group Actions

Definition 4.1.1. Let G be a group and let Ω be a set. Then a *left action* of G on the set Ω is a function $G \times \Omega \to \Omega$ with $(g, \alpha) \mapsto g \cdot \alpha$ such that $1 \cdot \alpha = \alpha$ and $g \cdot (h \cdot \alpha) = gh \cdot \alpha$. A *right action* is defined analogously, but the function is from $\Omega \times G \to \Omega$.

Given $\alpha, \beta \in \Omega$, we write $\alpha \sim \beta$ if $\exists g \in G$, $g \cdot \alpha = \beta$. Then $\sim$ is an equivalence relation on the set Ω. The equivalence classes are called the *orbits* of the action. The orbit of $\alpha \in \Omega$ is denoted by

$$\begin{aligned} O_G(\alpha) &= \{\beta \in \Omega \mid \beta \sim \alpha\} \\ &= \{\beta \in \Omega \mid \exists g \in G, g \cdot \alpha = \beta\} \\ &= \{g \cdot \alpha \mid g \in G\} \end{aligned}$$

A companion notion is the *stabilizer* of an element $\alpha \in \Omega$. The stabilizer is denoted by

$$S_G(\alpha) = \{g \in G \mid g \cdot \alpha = \alpha\}$$

Note that $O_G(\alpha) \subseteq \Omega$ and $S_G(\alpha) \leq G$. Two basic facts are that distinct orbits are disjoint, and that the union of all orbits is Ω; in other words, as stated above, the orbits are equivalence classes.

Definition 4.1.2. The action is *transitive* when there is only one orbit. Equivalently, $\forall \alpha, \beta \in \Omega, \exists g \in G$ with $g \cdot \alpha = \beta$.

Examples.

1. $G = (\mathbb{R}, +), \Omega = \mathbb{C}$. Let $x \cdot z := e^{ix}z, x \in \mathbb{R}, z \in \mathbb{C}$. This is an action. The orbit of a complex number z is $O_G(z) = \{\omega \in \mathbb{C} \mid \exists x \in \mathbb{R}, e^{ix}z = \omega\} = \{\omega \in \mathbb{C} \mid |\omega| = |z|\}$, or the circle through the origin with radius equal to $|z|$. Moreover, $S_G(Z) = \mathbb{R}$ if $z = 0$, and $2\pi\mathbb{Z}$ if $z \neq 0$.

2. Let G be any group and $\Omega = G$, with $g \cdot h = ghg^{-1}$, or the action of conjugation. $S_G(h) = \{g \in G \mid ghg^{-1} = h\} =$ centralizer of h in G, and $O_G(h) = \{x \in G \mid \exists g \in G, x = ghg^{-1}\} =$ conjugacy class of h in G.

3. S_n acts on $\{1, 2, \ldots, n\}$ by $\sigma \cdot i = \sigma(i)$. This action is transitive because we can always find a permutation sending i to j.

Proposition 4.1.3. *Let G act on Ω, $\alpha \in \Omega$. Then*

1. $S_G(\alpha) \leq G$
2. $|G/S_G(\alpha)| = |O_G(\alpha)|$ *(orbit-stabilizer reciprocity)*
3. *If $\alpha \sim \beta$, then $\exists g \in G$ such that $gS_G(\beta)g^{-1} = S_G(\alpha)$.*

The proof of this is relatively simple and is omitted (at least for now).

Example. Let $k \leq n$ be nonnegative integers. Let $\mathbb{P}_k(n) = \{A \subseteq \{1, 2, \ldots, n\} \mid |A| = k\}$. S_n acts on $\mathbb{P}_k(n)$ by $\sigma \cdot A = \sigma(A)$. Let $A_0 = \{1, 2, \ldots, k\} \in \mathbb{P}_k(n)$. Then $O_{S_n}(A_0) = \mathbb{P}_k(n)$, so the action is transitive. And $S_{S_n}(A_0) = \{\sigma \in S_n \mid \sigma(A_0) \subseteq A_0\} \cong S_k \times S_{n-k}$ under the isomorphism $\sigma \mapsto (\sigma|_{A_0}, \sigma|_{A_0^c})$. Thus,

$$|\mathbb{P}_k(n)| = |S_n/S_{S_n}(A_0)| = \frac{n!}{k!(n-k)!} = \binom{n}{k}.$$

Proposition 4.1.4. *Let G act on a finite set Ω, and suppose all stabilizers are trivial. Then the number of orbits is $|\Omega|/|G|$. In particular, G is finite and $|G|$ divides $|\Omega|$.*

Proof. $S_G(\alpha) = \{1\} \forall \alpha$. Thus for a given α, $|O_G(\alpha)| = |G/S_G(\alpha)| = |G|$, so we know already that G is finite. Let the distinct orbits be $O_G(\alpha_1), \ldots, O_G(\alpha_k)$. Then

$$\Omega = \bigsqcup_{i=1}^{k} O_G(\alpha_i)$$
$$\Rightarrow |\Omega| = \sum_{i=1}^{k} |O_G(\alpha_i)| = k|G|.$$

This completes the proof. □

Example. Let $H \leq G$, with G finite. Consider $G \times H \to G$, $(g, h) \mapsto gh$. This is a right action of H on G.

$O_H(g) = gH$, so the orbits are the same as the H-cosets. In particular, the number of orbits is $[G : H] = |G/H|$.

$S_H(g) = \{h \in H \mid gh = g\} = \{1\}$, so the stabilizers are trivial.

Then by the proposition, $|G/H| = |G|/|H|$, the familiar Lagrange's Theorem.

Example. Let $\mathbb{P}^n(\mathbb{F})$ be the n-dimensional projective space over a field $\mathbb{F}$. This is the set of lines through the origin in $\mathbb{F}^{n+1}$. For $n = 1$, we examine lines in $\mathbb{F}^2$, the plane. It then looks like a horizontal projective line, along with a point at infinity. HW3, Exercise 3 will include a derivation of the result that

$$|\mathbb{P}^n(\mathbb{F}_q)| = 1 + q + q^2 + \cdots + q^n.$$

Definition 4.1.5. For a group G acting on a set Ω, Ω^G is the set of fixed points, $\{\alpha \in \Omega \mid g \cdot \alpha = \alpha \forall g \in G\}$.

Theorem 4.1.6 (Fixed point lemma)**.** *Let G act on a finite set Ω. Suppose there exists a prime p that divides the index of every proper subgroup of G. Then $|\Omega^G| \equiv |\Omega| \pmod{p}$.*

Proof. List the orbits $O_G(\alpha_1), \ldots, O_G(\alpha_i), O_G(\alpha_{i+1}), \ldots, O_G(\alpha_k)$, where the orbits up to $O_G(\alpha_i)$ are trivial and the ones following it are not. Then $|\Omega| = |\Omega^G| + |G/S_G(\alpha_{i+1})| + \cdots + |G/S_G(\alpha_k)|$. But all of these are divisible by p, because the orbits are nontrivial and thus the stabilizers are proper, except for the fixed points, so $|\Omega| \equiv |\Omega^G| \pmod{p}$. □

This mode of thought will return when we study p-groups, and later on as well.

4.2 Actions and groups of permutations

Suppose G acts on Ω from the left. Given $g \in G$, let $\varphi_g : \Omega \to \Omega$ defined by $\varphi_g(\alpha) = g \cdot \alpha$. Then φ_g is bijective, with $(\varphi_g)^{-1} = \varphi_{g^{-1}}$.

Let $S(\Omega)$ be the group of permutations of Ω, or the group of bijections of Ω under composition. Let $\varphi : G \to S(\Omega)$ be defined by $g \mapsto \varphi_g$. Then φ is a (homo)morphism of groups. In other words, an action gives us a homomorphism from an action to a permutation group. Conversely, starting from a group morphism $\varphi : G \to S(\Omega)$, and definining $g \cdot \alpha = \varphi(g)(\alpha)$, we obtain a left action of G on Ω.

The two constructions above are inverses of each other.

Remark. Right actions would give an anti-homomorphism.

Definition 4.2.1. Let G act on Ω and let $\varphi : G \to S(\Omega)$ be the discussed associated group morphism. We say that $\ker \varphi$ is the *kernel of the action.* The action is *faithful* if $\ker \varphi$ is trivial, or if the associated morphism is injective.

Remark. $g \in \ker \varphi \iff \varphi_g = \mathrm{id}_\Omega \iff g \cdot \alpha = \alpha \forall \alpha$, or if $g \in S_G(\alpha) \forall \alpha$. In other words,

$$\ker \varphi = \bigcap_{\alpha \in \Omega} S_G(\alpha).$$

Examples.

1. Let G act on itself by conjugation. Then $S_G(h)$ is the centralizer of h in G, and $\ker \varphi = \bigcap_{h \in G} S_G(h) = Z(G)$.

2. S_n acts on $\{1, 2, \cdots, n\}$. $S_G(i) = \{\sigma \in S_n \mid \sigma(i) = i\} \cong S_{n-1}$. However, $\ker \varphi = \{\sigma \in S_n \mid \sigma(i) = i \forall i\} = \{\mathrm{id}\}$. The associated morphism is $\varphi : S_n \to S_n$, the identity.

3. G acts on itself by left translations, i.e. $g \cdot h = gh$. $S_G(h) = \{g \in G \mid gh = h\} = \{1\}$. All stabilizers are trivial, so the action must be faithful. The associated morphism $\varphi : G \to S(G)$, injective; this gives Cayley's Theorem, that any group is isomorphic to a subgroup of a permutation group.

4.3 Applications to the existence of normal subgroups

The idea is that actions correspond to this $\ker \varphi$; kernels are normal subgroups, so we should have a correspondence.

Proposition 4.3.1. *Let $H \leq G$ with $|G/H| = n$. Then $\exists N \trianglelefteq G$ such that $N \leq H$ and $|G/N|$ divides $n!$.*

Intuitively, subgroups can't get too large before they start containing normal subgroups.

Proof. Let $\Omega = G/H = \{xH \mid x \in G\}$. Define $g \cdot xH = gxH$. This is an action of G on Ω; let N be the kernel of this action. Then $N \trianglelefteq G$. Since N is the kernel, by definition $N \subseteq S_G(H)$, as $H = 1H \in \Omega$. But $S_G(H) = H$. Thus $N \subseteq H$, so $N \leq H$. By the first isomorphism law, $G/N \hookrightarrow S(\Omega)$, which has order $n!$, so $|G/N|$ divides $n!$. □

Definition 4.3.2. N, as defined in the proof of the previous proposition, is the *core* of H in G. (See HW 3, Exercise 1).

Corollary 4.3.3. *Let G be a finite group and let p be the smallest prime divisor of $|G|$. If $\exists H \leq G$ with $|G/H| = p$, then $H \trianglelefteq G$.*

Proof. Let N be the core of H in G. We'd want to show that $H = N$. By hypothesis, $[G : H] = p$; if $k = [H : N]$, it suffices to show that $k = 1$. We know that $[G : N] = pk$. Then by the above proposition, pk divides $p!$. Cancelling p, we know that k divides $(p-1)!$, so all prime divisors of k are $< p$. But k divides $|G|$, because it is an index, so all prime divisors of k are $\geq p$, because the smallest prime divisor of $|G|$ is p. Then k has no prime divisors, so $k = 1$, and $H = N$, and we are done. □

Corollary 4.3.4. *Let G be finite. If $\exists H \leq G$ with $|G/H| = 2$, then $H \trianglelefteq G$.*

4.4 p-groups

Definition 4.4.1. Let p be a prime. A *finite p-group* is a finite group of order p^k, for some $k \geq 0$.

Theorem 4.4.2 (Fixed point lemma for p-groups)**.** *Let G be a p-group, and Ω a finite set. Suppose G acts on Ω. Then $|\Omega| \equiv |\Omega^G| \pmod{p}$.*

Proof. All proper subgroups have index divisible by p. The index must be a prime power; if the subgroup is proper, that power cannot be 1. We then apply the fixed point lemma. □

5 September 8th

5.1 p- groups

Recall the definition from last time of a p-group. Note that we will assume for now that all of our p-groups are finite. We also discussed the fixed-point lemma for p-groups.

Corollary 5.1.1. *Let G be a nontrivial p-group. Then $Z(G) \neq \{1\}$.*

Proof. Let G act on itself by conjugation. The fixed points of this action are

$$G^G = \{h \in G \mid ghg^{-1} = h, \forall g \in G\} = Z(G).$$

By the fixed point lemma, $|Z(G)| \equiv |G| \pmod{p} \equiv 0 \pmod{p}$. Thus certainly $|Z(G)| \neq 1$, and we are done. □

Corollary 5.1.2. *Every p-group is nilpotent.*

Proof. $G/Z(G)$ is a p-group, and $|G/Z(G)| < |G|$ by the previous corollary, so we can argue by induction on $|G|$, we may assume $G/Z(G)$ is nilpotent by a previous proposition that G/N is nilpotent if and only if G is nilpotent for $N \trianglelefteq G, N \subseteq Z(G)$. □

The following is an important lemma that appeared in the homework.

Lemma 5.1.3. *Let G be a finite abelian group with p a prime divisor of $|G|$. Then G contains an element of order p.*

Proof. HW2 Exercise 18. □

Theorem 5.1.4. *Let G be a nontrivial p-group.*

1. *(big centers). If $N \trianglelefteq G$, $N \neq \{1\}$, then $N \cap Z(G) \neq \{1\}$.*
2. *(subgroups of all possible orders) If we have $N \trianglelefteq G$ and d a divisor of $|N|$, then N has a subgroup of order d which is normal in G.*
3. *(normalizers grow) If $H < G$ then $H < N_G(H)$.*
4. *If $K < G$ is a maximal subgroup of G then $K \triangleleft G$ and $[G : K] = p$.*

Proof.

1. Apply the fixed point lemma to the action of G on N by conjugation. The same proof that tells us that $Z(G) \neq \{1\}$ tells us that $N \cap Z(G) \neq \{1\}$.
2. Write $d = p^\alpha$, and induct on α. If $\alpha = 0$, nothing to do. Assume $\alpha \geq 1$. Then $N \neq \{1\}$. By 1, p divides $|N \cap Z(G)|$. By the lemma, $N \cap Z(G)$ has a subgroup N_1 of order p. Consider G/N_1. We have $N/N_1 \trianglelefteq G/N_1$ and $p^{\alpha-1}$ divides $|N/N_1|$. By induction hypothesis, N/N_1 has a subgroup of order $p^{\alpha-1}$ that is normal in G. By the fourth isomorphism law, this subgroup is of the form N_2/N_1 with $N_1 \leq N_2 \leq N$ and $N_2 \trianglelefteq G$.

   Then $|N_2| = |N_1||N_2/N_1| = p \cdot p^{\alpha-1} = p^\alpha$, so we are done.
3. Let $\Omega = G/H$. Consdier the action of H on Ω by translation: $h \cdot xH = (hx)H$. $xH \in \Omega^H \iff hxH = xH \, \forall h \in H$, which is true if and only if $x^{-1}hx \in H$ for every $h \in H$, which happens if and only if $x^{-1}Hx \subseteq H$, or $x \in N_G(H)$. (Note the subtle-but-unimportant-here-because-finite-groups difference between $x^{-1}Hx \subseteq H$ and $xHx = H$.)

   But note that in each case we care about the coset x, not the element, so $\Omega^H = N_G(H)/H$. The fixed point lemma then tells us that $|N_G(H)/H| \equiv |G/H| \pmod{p} \equiv 0 \pmod{p}$, because H is proper. Then $|N_G(H)/H| \neq \{1\}$, as desired.
4. By 3, $K < N_G(K)$; K is maximal, so $N_G(K) = G$, and $K \triangleleft G$. Now G/K is a p-group. So it must have normal subgroups of all possible orders; but K is maximal, so G/K is simple. This can only happen if there are no possible orders, so $|G/K| = p$, as desired. □

5.2 Sylow Theorems

Definition 5.2.1. Let G be a finite group and p a prime. Write $|G| = p^\alpha m$ with $\alpha \geq 0$ and p does not divide m. A *p-Sylow group* is a subgroup $S \subseteq G$ with $|S| = p^\alpha$. Let $\mathrm{Syl}_p(G)$ be the set of all p-Sylow subgroups of G.

Theorem 5.2.2 (1st Sylow Theorem)**.** *G finite, p prime* $\Rightarrow \mathrm{Syl}_p(G) \neq \varnothing$.

Proof. Induction on $|G|$.

(a) If G has a proper subgroup H of index coprime to p, then $H = p^\alpha m'$, so $\mathrm{Syl}_p(H) \neq \varnothing$ by the inductive hypothesis, and $\mathrm{Syl}_p(H) \subseteq \mathrm{Syl}_p(G)$.

(b) If no such H exists, then all proper subgroups have index divisible by p. By the fixed point lemma, $|\Omega^G| \equiv |\Omega| \pmod{p}$ whenever G acts on a finite Ω. If G has a nontrivial normal p-subgroup N, then we apply the induction hypothesis to the quotient G/N. Then there exists $S/N \leq G/N$ of order $p^{\alpha-\beta}$, then there exists $S \leq G$ and $N \leq S \leq G$ with $|S| = p^\alpha$.

(c) We now must show that either (a) or (b) happens. Consider the conjugation action of G on itself; $|Z(G)| \equiv |G| \pmod{p} \equiv 0 \pmod{p}$ (We can assume $\alpha \geq 1$). By the lemma, $Z(G)$ has a subgroup N of order p, so we're done.

□

Corollary 5.2.3 (Cauchy)**.** *G finite, p divides $|G|$. Then $\exists x \in G$ of order p.*

Proof. Let $S \in \mathrm{Syl}_p(G) \Rightarrow |S| = p^\alpha$. S has some element of order p, so G does as well. □

Theorem 5.2.4 (2nd Sylow Theorem)**.** *Let G be finite, $S \in \mathrm{Syl}_p(G)$. Then every other p-Sylow is conjugate to S. Conversely, although this is easy, any subgroup conjugate to S is a p-Sylow. In fact, let $P \leq G$ be a p-subgroup. Then there exists an element $x \in G$ with $P \leq xSx^{-1}$.*

Proof. Let $\Omega = G/S$. Let P act on Ω by translations, $g \cdot \overline{x} = \overline{gx}$. (Note the bar notation for cosets.) By the fixed point lemma, $|\Omega^P| = |\Omega| \pmod{p}$. Since $S \in \mathrm{Syl}_p(G) \Rightarrow p$ does not divide $|\Omega|$. Then $|\Omega^P| \not\equiv 0 \pmod{p}$, so $\Omega^p \neq \varnothing$. Let $\overline{x} \in \Omega^P$. Then $\overline{gx} = \overline{x}$ for all $g \in P$, which happens if and only if $x^{-1}gx \in S$ for all $g \in P$. This is true iff $x^{-1}Px \subseteq S$, which is what we wanted, because $P \subseteq xSx^{-1}$. □

Corollary 5.2.5. *Let $S \in \mathrm{Syl}_p(G)$. Then $S \trianglelefteq G \iff \mathrm{Syl}_p(G) = \{S\}$.*

Proof. The 2nd Sylow Theorem says the conjugates of S are the p-Sylows. S is normal if and only if it coincides with all of its conjugates. □

Lemma 5.2.6. *G finite, $T, S \in \mathrm{Syl}_p(G)$. If S normalizes T, then $S = T$.*

Proof. $S, T \in \mathrm{Syl}_p(N_G(T))$. But $T \trianglelefteq N_G(T)$, so $S = T$ by the corollary.

$$\begin{tikzcd}
 & G \arrow[d, dash] & \\
 & N_G(T) \arrow[dl, dash] \arrow[dr, dash] & \\
S & & T
\end{tikzcd}$$

□

Notation: $n_p(G) = |\mathrm{Syl}_p(G)|$.

Theorem 5.2.7 (3rd Sylow Theorem)**.** *$|G| = p^\alpha m$, $p \not| m$. Then*

(i) $n_p(G) = |G/N_G(S)|$, where S is any p-Sylow.

(ii) $n_p(G)|m$.

(iii) $n_p(G) \equiv 1 \pmod p$.

Proof. (i) Let $\Omega = \{X \subseteq G \mid |X| = p^\alpha\}$. Let G act on Ω by conjugation and $S \in \Omega$. Then $O_G(S) = \{gSg^{-1} \mid g \in G\} = \mathrm{Syl}_p(G)$, and $S_G(S) = \{g \in G \mid gSg^{-1} = S\} = N_G(S)$. Then orbit-stabilizer reciprocity gives us exactly what we want.

(ii) We have a tower of indices of subgroups, and the smaller one divides the bigger one, so $n_p(G)|m$.

$$\begin{tikzcd}
G \arrow[d, dash] \\
N_G(S) \arrow[d, dash] \\
S
\end{tikzcd}$$

(iii) Let $\Omega = \mathrm{Syl}_p(G)$. S acts on Ω by conjugation. Using the fixed point lemma, we get that $|\Omega^S| \equiv |\Omega| \pmod p$. It suffices to prove that $|\Omega^S| \equiv 1 \pmod p$. But we'll show there's only one fixed point. $T \in \Omega^S \iff gTg^{-1} = T\ \forall g \in S$, which is true if and only if S normalizes T, which by the lemma is only true if $S = T$. Thus $\Omega^S = \{S\}$, so $|\Omega^S| = 1 \equiv 1 \pmod p$, and we're done.

□

6 September 10th

6.1 Sylow Theorems, continued

Proposition 6.1.1 (Frattini's Argument)**.** *Let G be an arbitrary group with $N \trianglelefteq G$ finite, $S \in \mathrm{Syl}_p(N)$ for some prime p, and $H = N_G(S)$. Then $G = NH$.*

Proof. Take $g \in G$. Then $gSg^{-1} \leq gNg^{-1} = N$. Thus $gSg^{-1} \in \mathrm{Syl}_p(N)$. By the second Sylow theorem (applied to N), $\exists n \in N$ with $gSg^{-1} = nSn^{-1}$. Then $n^{-1}gSg^{-1}n = S$, so $n^{-1}g \in N_G(S) = H$, so $g = (n)(n^{-1}g) \in NH$. □

6.2 Direct products: A New Hope

Definition 6.2.1. Given groups A and B, their *direct product* is $A \times B = \{(a, b) \mid a \in A, b \in B\}$ with product $(a_1, b_1) \cdot (a_2, b_2) = (a_1 a_2, b_1 b_2)$. The identity is $(1, 1)$.

If $A_1 \leq A$ and $B_1 \leq B$ then $A_1 \times B_1 \leq A \times B$. Moreover, $A_1 \times B_1 \trianglelefteq A \times B$ if and only if $A_1 \trianglelefteq A$, $B_1 \trianglelefteq B$. However, a word of caution: a subgroup of $A \times B$ need not be of the form $A_1 \times B_1$. For example, letting $A = B = \mathbb{Z}$, and taking the diagonal $D = \{(n, n) \mid n \in \mathbb{Z}\}$, D is not a product of two subgroups.

Okay, so given a group G, how can we tell whether or not $G \cong A \times B$ for two groups A, B?

In fact, this is always possible, with $A = G$ and B trivial, but that's lame. We want a **real** decomposition. Like, a nontrivial one.

Remark. Let $H, K \leq G$.

1. If $H, K \trianglelefteq G \Rightarrow HK \leq G$
2. If $H, K \trianglelefteq G$ and $H \cap K = \{1\}$, then $hk = kh \, \forall h \in H, k \in K$. You can see this by looking at a commutator $hkh^{-1}k^{-1}$, which must be in both H and K.
3. Assume H and K are finite. Then $|HK| = \frac{|H||K|}{|H \cap K|}$.
4. Again with H, K finite. If $\gcd(|H|, |K|) = 1$, then $H \cap K = \{1\}$.

Proposition 6.2.2. *Let $H, K \trianglelefteq G$ be such that*

(i) $G = HK$,

(ii) $H, K \trianglelefteq G$,

(iii) $H \cap K = \{1\}$.

Then $G \cong H \times K$. In this case we say that G is the internal direct product *of H and K.*

Proof. Let $\varphi : H \cap K \to G$ be defined by $\varphi(h, k) = hk$. To check that φ is a homomorphism, we have $\varphi(h_1, k_1)\varphi(h_2, k_2) = h_1 k_1 h_2 k_2 = h_1 h_2 k_1 k_2 = \varphi(h_1 h_2, k_1 k_2)$, because H and K commute. φ is surjective by (i) and injective by (iii), so this is an isomorphism. □

Remark. Suppose $G = A \times B$ for some groups A, B. Let $H = \{(a, 1) \mid a \in A\}$ and let $K = \{(1, b) \mid b \in B\}$. Then $H, K \leq G$ and G is the internal direct product of H and K.

OK, now back to nilpotent groups.

6.3 The Nilpotent Groups Strike Back!

Theorem 6.3.1. *Let G be a finite group. TFAE*

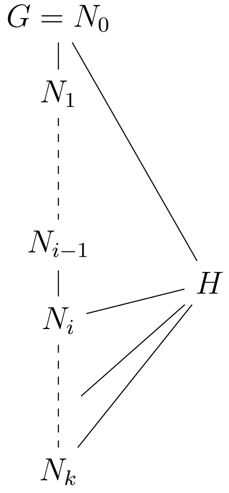
(i) G is nilpotent

(ii) Normalizers grow (as we already know for p-groups, a special case)

(iii) All Sylow subgroups are normal, i.e. there is a unique p-Sylow for each prime p.

(iv) G is a direct product of p-groups for various primes p.

Proof. (i) $\Rightarrow$ (ii): $G = N_0 \trianglerighteq N_1 \trianglerighteq \cdots \trianglerighteq N_k = \{1\}$ with the property that $N_i \trianglelefteq G$ and $N_{i-1}/N_i \subseteq Z(G/N_i)$. Equivalently, $[G, N_{i-1}] \leq N_i$. Let $H < G$. Then there exists i such that $N_i \leq H$ but $N_{i-1} \not\leq H$. Then $[H, N_{i-1}] \leq [G, N_{i-1}] \leq N_i \leq H$. So then $N_{i-1} \leq N_G(H)$. But $N_{i-1} \not\subseteq H$, so normalizers grow, i.e. $H < N_G(H)$.



(ii) $\Rightarrow$ (iii): Let $S \in \mathrm{Syl}_p(G)$. By HW3 Exercise 8, $N_G(N_G(S)) = N_G(S)$, so $N_G(S) = G$, because otherwise its normalizer would be strictly larger. So S is normal.

(iii) $\Rightarrow$ (iv): We show by induction on $|G|$ that G is the direct product of its nontrivial Sylow subgroups. Let $|G| = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$ with p_i distinct. Let S_i be the unique p_i-Sylow subgroup. Let $H = S_1 \cdots S_{r-1}$ and let $K = S_r$. $S_i \trianglelefteq G$, so $H \trianglelefteq G, K \trianglelefteq G$. $|H|$ divides $|S_1||S_2| \cdots |S_{r-1}| = p_1^{\alpha_1} \cdots p_{r-1}^{\alpha_{r-1}}$, so $|H|$ and $|K|$ are coprime, so $H \cap K = \{1\}$. We need only check that G is the product of H and K. Each $S_i \leq H$, or $S_i \in \mathrm{Syl}_{p_i}(H)$, so each $p_i^{\alpha_i}$ divides $|H|$. These are the nontrivial Sylows of H and they are all normal; by the induction hypothesis, $H \cong S_1 \times \cdots \times S_{r-1}$. Moreover, we know in fact that $|H| = p_1^{\alpha_1} \cdots p_{r-1}^{\alpha_{r-1}}$, so $|HK| = \frac{|H||K|}{|H \cap K|} = |G|$, so $HK = G$. Thus $G \cong H \times K \cong S_1 \times \cdots \times S_r$, so we're done.

(iv) $\Rightarrow$ (i): HW4 Exercise 1. A direct product of nilpotent groups is nilpotent. □

Corollary 6.3.2 (Lagrange converse)**.** *Let G be a finite nilpotent group. For each divisor d of $|G|$, there exists $N \trianglelefteq G$ with $|N| = d$.*

Proof. $G \cong P_1 \times \cdots \times P_r$, $|G| = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$. Then $d = p_1^{\beta_1} \cdots p_r^{\beta_r}$ with $\beta_i \leq \alpha_i$. $\exists N_i \trianglelefteq P_i$ with $|N_i| = p_i^{\beta_i}$; take $N = N_1 \times \cdots \times N_r$. □

Theorem 6.3.3. *Let G be a finite group. Then G is nilpotent if and only if all maximal subgroups are normal.*

Proof. $\underline{(\Rightarrow)}$: Let $K < G$ be maximal. Thus $N_G(K) = G$ since normalizers grow, $\underline{(\Leftarrow)}$: Let S be a Sylow subgroup of G. Check $S \trianglelefteq G$. Suppose $N_G(S) < G$. Let K be a maximal subgroup of G containing $N_G(S)$. By hypothesis $K \trianglelefteq G$. By Frattini's argument, $G = KN_G(S) = K$, a contradiction.

$$\begin{tikzcd}
G \arrow[d, dash] \\
K \arrow[d, dash] \\
N_G(S) \arrow[d, dash] \\
S
\end{tikzcd}$$

□

Note that this could prove the structure theorem for finite abelian groups. We won't do that with it, but we could.

6.4 Return of the Semidirect Products

Definition 6.4.1. Let G and A be groups. Suppose G acts on A, with $G \times A \to A$. We say that the action is by *automorphisms* if $g \cdot (ab) = (g \cdot a)(g \cdot b)$ for all $g \in G$, $a, b \in A$.

Remark. 1. In this case $g \cdot 1_A = 1_A$.

2. Let $\varphi : G \to S(A)$ be the associated morphism of groups. Then $\mathrm{Aut}(A) = \{\sigma \in S(A) \mid \sigma \text{ is an isomorphism of groups}\}$; $\mathrm{Aut}(A) \leq S(A)$ and the action is by automorphism if and only if $\operatorname{im} \varphi \subseteq \mathrm{Aut}(A)$. $\varphi_g(ab) = g \cdot ab = (g \cdot a)(g \cdot b) = \varphi_g(a)\varphi_g(b)$, so this is equivalent to saying φ_g is an automorphism of A.

$$\begin{tikzcd}
G \arrow[r, "\varphi"] \arrow[dr, dashed] & S(A) \\
& \mathrm{Aut}(A) \arrow[u, hook]
\end{tikzcd}$$

Example. $G \times G \to G$, $g \cdot h = ghg^{-1}$ is by automorphisms. $G \times G \to G$, $g \cdot h = gh$ is not by automorphisms.

Definition 6.4.2. Suppose G acts on A by automorphisms. Then we have the *semi-direct product* $A \rtimes G$ is defined such that the underlying set is $A \times G$, and the product is $(a, g)(b, h) = (a(g \cdot b), gh)$. The unit is $(1_A, 1_G)$.

Proposition 6.4.3. *$A \rtimes G$ is a group. Proof an optional homework problem.*

Remark. $A \rtimes G$ depends on the action. If the action changes despite A and G remaining the same, we can get a very different semidirect product.

If the action is trivial, i.e. $g \cdot a = a$ for every g, a, then this is a direct product.

But we have the same question as for direct products. Given G, how do we tell if $G \cong A \rtimes B$ for some A, B and some action? That's an exercise that isn't optional, in the homework.

7 September 15th

7.1 Hall subgroups

Definition 7.1.1. Let π be a set of primes and n a positive integer. The *π-part* of n is the largest divisor of n involving only primes from π. The *π'-part* of n is the largest divisor of n *not* involving any of the primes in π.

Example. If $n = 60 = 2^2 \cdot 3 \cdot 5$, and $\pi = \{2, 3\}$, then the π-part is 12 and the π'-part is 5.

Definition 7.1.2. Let G be a finite group and $H \leq G$. Let π be a set of primes. We say H is a *Hall π-subgroup* of G if $|H|$ is the π-part of $|G|$.

Remark.

1. If $\pi = \{p\}$ then π-Hall is the same concept as p-Sylow.
2. $H \leq G$ is Hall for some π if and only if $\gcd(|H|, |G/H|) = 1$.

Our goal is to prove that if G is finite and solvable, then Hall π-subgroups exist for every π.

Lemma 7.1.3. *If G is finite and solvable, and M is a minimal normal subgroup of G, then M is elementary abelian.*

In particular, M is a p-group for some prime p.

Proof. HW2, Exercise 4. □

Lemma 7.1.4. *If G is finite and solvable, and $N \lhd G$, then there exists a p-subgroup $P \leq G$ for some prime p such that*

$$N < NP \trianglelefteq G.$$

Proof. $N \lhd G$, so $G/N \neq \{1\}$, and G/N has a minimal normal subgroup. G is solvable, so any quotient is solvable as well. By Lemma 7.1.3, that subgroup is elementary abelian. It has to be of the form M/N for some $N < M \trianglelefteq G$. Since $|M/N| = p^r$ for some prime p and $r > 0$, so p divides $|M|$. We then choose a p-Sylow P of M.

$$\begin{tikzcd}
 & M \arrow[d, dash] & \\
 & NP \arrow[dl, dash] \arrow[dr, dash] & \\
P & & N
\end{tikzcd}$$

But $[M : P]$ is prime to p, and $[M : N]$ is p^r, so the last index in the diagram, $[M : NP]$, must divide the gcd of p^r and something prime to p, so $[M : NP] = 1$ and $M = NP$. Then we are done. □

Theorem 7.1.5 (Schur-Zassenhaus)**.** *Let G be finite and N be a normal Hall subgroup of G. Then N has a* complement *in G:*

$\exists H \leq G$ with $G = NH$, $N \cap H = \{1\}$.

Proof. This is a hard theorem that we won't prove in this course, but we will prove the case when G is solvable.

It suffices to find $H \leq G$ with $|H| = |G/N|$. For then $|N \cap H|$ divides both $|N|$ and $|H|$, so $|N \cap H| = \{1\}$ since N is Hall. And then $|NH/N| = |H/H \cap N| = |H| = |G/N|$, so $|NH| = |G|$.

We proceed by induction on $|G|$. If $|G| = 1$, we have nothing to do; all relevant groups are trivial and we can go home for lunch. If $N = G$, we again have nothing to do, because we can take $H = \{1\}$. If $N < G$, then by Lemma 7.1.4, there exists a p-subgroup P such that $N < NP \trianglelefteq G$.

We then analyze the following diagram.

$$\begin{tikzcd}
 & G \arrow[d, dash] & \\
 & NP \arrow[dl, dash] \arrow[dr, dash] & \\
P \arrow[dr, dash] & & N \arrow[dl, dash] \\
 & N \cap P \arrow[d, dash] & \\
 & \{1\} &
\end{tikzcd}$$

We then have:

1. $|P|$ is a power of P
2. $|P/N \cap P|$ is a power of p
3. $|NP/N|$ is a power of p (nontrivial, since $N < NP$)
4. $|G/N|$ divisible by p
5. $|N|$ prime to p, since N is Hall

6. $|N/N \cap P|$ prime to p

7. $|NP/P|$ prime to p

Thus P is a p-Sylow of NP.

Observe that $N \cap P = \{1\}$, because its order divides both p and a number prime to p. We now need only a complement of NP. We will apply Frattini's argument. We let $K = N_G(P)$. Then we have the following picture:

$$\begin{tikzcd}
 & G \arrow[dl, dash] \arrow[dr, dash] & \\
NP \arrow[dr, dash] & & K \arrow[dl, dash] \\
 & P &
\end{tikzcd}$$

Then $G = NPK = NK$. But we have to worry about the intersection between N and K; otherwise, K would be a complement. K might be too big, though. So we'd like to find a complement for $N \cap K$ inside K; that complement may do the job. By induction, we can find the complement provided that $N \cap K \leq K$ is a normal Hall subgroup and that $K < G$. Normality is easy; $N \trianglelefteq G$, so $N \cap K \trianglelefteq K$. What is its index? $|K/N \cap K| = |NK/N| = |G/N|$, which is prime to $|N|$ and thus prime to all divisors of $|N|$, including $|N \cap K|$. Thus $N \cap K$ is a normal Hall subgroup of K, so we have two cases.

First, if $K < G$, we can proceed by induction to find a complement H of $N \cap K$ in K, by induction. This isn't appealing to the more general Schur-Zassenhaus, because $K \leq G$ is solvable whenever G is solvable. So then $|H| = |K/N \cap K| = |G/N|$, so we are done by our first claim.

Now we only need consider the case when $K = G$, so we can't use our induction hypothesis. If $K = G$, we have $P \trianglelefteq G$. So we consider G/P, and we claim that NP/P is a normal Hall subgroup of G/P, to which we will finally be able to apply the induction hypothesis, because P is a p-Sylow and thus nontrivial. But first, we must verify our claim. $|NP/P| = |N/N \cap P| = |N|$, which is prime to $|G/N|$. Then $\frac{|G/P|}{|NP/P|} = \frac{|G|}{|NP|} = \frac{|G/N|}{|P|}$, so $|NP/P|$ is prime to $\frac{|G/N|}{|P|}$ because it was prime to the numerator. We now (finally, blessedly!) apply the induction hypothesis to find a complement H/P of NP/P in G/P. Hopefully this H will do the trick. $|H/P| = \frac{|G/P|}{|NP/P|} = \frac{|G/N|}{|P|}$, so $|H| = |G/N|$ and we are (finally, blessedly!) done by our first claim.

□

Theorem 7.1.6 (Hall). *Let G be finite, solvable. For any set of primes π, there exists a Hall π-subgroup H of G.*

Example. $|G| = 60 = 2^2 \cdot 3 \cdot 5$. Then the Sylow theorems find subgroups of orders $4, 3, 5$. Hall's theorem proves that there exist subgroups of orders $12, 15, 20$ as well.

The homework has an exercise proving that A_5 does not have subgroups of order 20 or 15, showing that Hall's theorem does not necessarily hold when G is not solvable.

Proof. Once more we induct on $|G|$. If $G \neq \{1\}$, let M be a minimal normal subgroup of G. G being solvable implies that M is a p-group. G is solvable, so M is a p-group. By the induction hypothesis, there exists a Hall π-subgroup of G/M. Let it be K/M for some $M \leq K \leq G$. Is K a Hall π-subgroup of G?

$|K| = |K/M||M|$, which involves only primes in $\pi \cup \{p\}$, and $|G/K| = \frac{|G/M|}{|K/M|}$ involves only primes in π'. In the good case, $p \in \pi$, and we are done. Assume $p \notin \pi$. Then $|K/M|$ does not involve p, so $\gcd(|M|, |K/M|) = 1$, and M is a Hall subgroup of K with $M \trianglelefteq K$. By Schur-Zassenhaus, there is a complement H of M inside K. Then $|H| = |K/M|$, so $|H|$ involves only primes in π. Meanwhile, $|G/H| = \frac{|G|}{|K/M|} = |G/K| \cdot |M|$, which only involves primes in $\pi' \cap \{p\} = \pi'$. H is the desired subgroup, and we are done. □

7.2 Additional facts about solvable groups

- Complements of Hall's Theorem:
  - Any two π-Hall subgroups are conjugate.
  - If $K \leq G$ with $|K|$ involving only primes in a set π, then there exists a π-Hall H such that $K \leq H$.
- Hall converse: If π-Hall subgroups exist for all π, then G is solvable.
- If $|G| = p^a q^b$, then G is solvable. (Burnside's $p^a q^b$-theorem).
- Feit-Thompson. All groups of odd order are solvable. (outside the scope of this course by far).
- There are about 50 billion groups of order ≤ 2000. Of these, more than 99 percent have order 1024.

8 September 17th

8.1 Simple groups

Let G act on Ω and let $\varphi : G \to S(\Omega)$ be the associated homomorphism. Recall that the following are intuitively equivalent:

- The action is *faithful*
- No nontrivial element of G fixes all elements of Ω
- $$\bigcap_{\alpha \in \Omega} S_G(\alpha) = \{1\}$$
- $\ker \varphi = \{1\}$

- G injects into $S(\Omega)$

Recall further the definition that the action is *transitive* if and only if given any $\alpha, \beta \in \Omega, \exists g \in G$ with $g \cdot \alpha = \beta$. In this case, all stabilizers are conjugate.

Now consider $\Omega^2 \setminus \Delta = \{(\alpha, \alpha') \in \Omega^2 \mid \alpha \neq \alpha'\}$. Suppose that $g \cdot \alpha = g \cdot \alpha'$. Then by acting on both sides with g^{-1}, we see that $\alpha = \alpha'$. Thus G acts on $\Omega^2 \setminus \Delta$ coordinate-wise, and this action is well-defined.

Definition 8.1.1. The action of G on Ω is 2-*transitive* if its action on $\Omega^2 \setminus \Delta$ is transitive.

Explicitly, this means that given $\alpha \neq \alpha'$ and $\beta \neq \beta'$ in Ω, there exists $g \in G$ with $g \cdot \alpha = \beta$ ***and also*** $g \cdot \alpha' = \beta'$. Note that we can have $\alpha = \beta$ or $\alpha' = \beta'$.

Example. The action of S_4 on $[4] = \{1, 2, 3, 4\}$ is 2-transitive. Given $a \neq a'$ and $b \neq b'$ in $[4]$, we need a permutation that sends a to b and a' to b'. Let $\sigma = (a, b)(a', b')$. There are a couple of adjustments to be done if $b' = a$, or $a' = b$, or whatever, but it's very possible.

Proposition 8.1.2. *Suppose the action of G on Ω is 2-transitive and $|\Omega| \geq 2$. Then*

(a) It is transitive.

(b) All stabilizers are maximal subgroups of G.

Proof. (a) Given $\alpha, \beta \in \Omega$, we need $g \in G$ such that $g \cdot \alpha = \beta$. Pick any $\alpha' \neq \alpha$ in Ω; this is possible because $\Omega \neq \{\alpha\}$. Further pick any $\beta' \neq \beta$, again possible because $\Omega \neq \{\beta\}$. By hypothesis there exists g with $g \cdot \alpha = \beta$.

(b) Let $\alpha \in \Omega$, $H = S_G(\alpha)$. If $H = G$ then $\Omega = \{\alpha\}$, a contradiction because $|\Omega| \geq 2$. So $H < G$. Suppose that there exists K with $H < K < G$, and we will derive a contradiction. With such a K, we know that there exists $k \in K \setminus H$ and $g \in G \setminus K$. Thus $k \cdot \alpha \neq \alpha$ because $k \notin H$, and $g \notin K$ so $g \cdot \alpha \neq \alpha$ as well. By 2-transitivity, there exists $f \in G$ with $f \cdot \alpha = \alpha$ and $f \cdot k \cdot \alpha = g \cdot \alpha$. Thus $f \in H$ and $k^{-1} f^{-1} g \in H$. So $g \in fkH \subseteq K$, but $g \notin K$, a contradiction.

□

Definition 8.1.3. G is *perfect* if $G^{(1)} = G$, where we recall that $G^{(1)} = [G, G]$.

Remark.

1. G solvable and nontrivial means that G is not perfect, because the derived series must terminate.

2. If G is simple, it is always either perfect or abelian, and never both. Being nonabelian makes $G^{(1)} \neq \{1\}$, and $G^{(1)} \trianglelefteq G$.

3. Not every perfect group is simple. For example, let S be simple and nonabelian. Take $G = S \times S$. G is not simple, because $S \times \{1\}$ and $\{1\} \times S$ is normal. However, G is perfect, because $G^{(1)} = S^{(1)} \times S^{(1)} = S \times S = G$.

Theorem 8.1.4 (Iwasawa's Lemma)**.** *Let G be a nontrivial perfect group. Suppose G acts on a set Ω such that*

(a) The action is faithful and 2-transitive.

(b) There exists a point with stabilizer H containing a subgroup A such that

(i) $A \trianglelefteq H$

(ii) A is abelian

(iii) The set

$$\bigcup_{g \in G} gAg^{-1}$$

generates G.

Then G is simple.

Remark. Under (a), all stabilizers are conjugate. Thus for hypothesis (b), any stabilizer H should work.

Proof. First we note that G is nontrivial and the action is faithful, so Ω has at least two points. We will then happily appeal to the previous proposition, which tells us that stabilizers are maximal given 2-transitivity.

Suppose $\exists N$ with $1 < N \lhd G$. Then there exists a stabilizer H with $N \not\leq H$; if not, N would be in the intersection of all stabilizers, but this intersection is trivial by the fidelity of the action. Thus $H < NH$, so $NH = G$ because H is maximal. As explained in the remark, we can assume that H satisfies the hypothesis by containing the appropriate subgroup A. We pick any $g \in G$ with $g = nh$, $n \in N$, $h \in H$. Then

$$\begin{aligned} gAg^{-1} &= nhAh^{-1}n^{-1} \\ &= nAn^{-1}, \text{ because } A \trianglelefteq H \\ &\subseteq NAN = NNA = NA, \text{ by normality of } N. \end{aligned}$$

So $G = NA$, because gAg^{-1} generates when ranging over $g \in G$. Then $G/N = NA/N \cong A/N \cap A$, which is abelian because A is abelian. So $[G, G] \leq N$, so $G^{(1)} = G \leq N$ and N is trivial, a contradiction. □

Note that the converse like definitely doesn't hold, but this is a nice sufficient criterion for simplicity. The simplest (ha, ha) application is to the simplicity of A_5, but first we will have a couple of basic facts about A_n.

Fact 8.1.5 (Facts About A_n)**.**

- *The $(2, 2)$-cycles form a conjugacy class of A_n for $n \geq 4$.*

- *The 3-cycles generate A_n for $n \geq 3$.*

- *The $(2,2)$-cycles generate A_n for $n \geq 5$.*
- *A_n is perfect for $n \geq 5$. Since $[(a,b,c),(a,b,d)] = (a,b)(c,d)$, the $(2,2)$-cycles are in the derived subgroup; for $n \geq 5$, this means that the derived subgroup generates.*

Corollary 8.1.6. *A_5 is simple.*

Proof. A_5 acts on $[5] = \{1,2,3,4,5\}$. This action is faithful, because the map to the group of permutations is just an inclusion. It's also 2-transitive; take $a \neq a'$ and $b \neq b'$, and let $\sigma = (a,b)(a',b')$. If $a \neq a'$ and $b \neq b'$, then $\sigma \in A_5$, and this is fine. If $a = b$, choose $\sigma = (c,d)(a',b')$ where c and d are the other two elements in [5]. Similarly if $a' = b'$. If $a = b$ and $a' = b'$, we pick the identity.

Now let $H = S_{A_5}(5)$, so $H \cong A_4$. Let $A = \{1, (12)(34), (13)(24), (14)(23)\}$, the Klein group. Then $A \trianglelefteq H$ and $A \cong V_4$ and thus A is abelian. The conjugates of A are all $(2,2)$-cycles in A_5, which generate.

Thus by Iwasawa's Lemma, A_5 is simple. □

We will now prove one last result, that A_n is simple for $n \geq 5$.

Corollary 8.1.7. *A_n is simple for all $n \geq 5$.*

Note that we can't use Iwasawa, because we won't find a normal abelian subgroup of any stabilizer.

Proof. By induction on n. The base case is done above, by Iwasawa. Take $n \geq 6$. Suppose there exists N with $\{1\} < N \triangleleft A_n$. Take the action of A_n on $[n]$ and take $H = S_{A_n}(n)$. Then $H \cong A_{n-1}$, so by hypothesis H is simple. But $N \cap H \trianglelefteq H$. Either $N \cap H = \{1\}$ or $N \cap H = H$.

If $N \cap H = H$, then $H \leq N$, so all conjugates of H are in N as well. Stabilizers are conjugate, so all stabilizers are in N; among many, many other elements, this includes all $(2,2)$-cycles, which generate A_n. Thus $A_n \subseteq N$, so we are done.

If $N \cap H = \{1\}$, we have further that $N \cap S_{A_n}(i) = \{1\}$. If a permutation in N fixes any point, it fixes all of them. Let $\sigma \in N$ be nontrivial, because $\{1\} < N$. σ has no fixed points, so either σ is entirely transpositions, i.e. $\sigma = (ab)(cd)\cdots$, or $\sigma = (abc\cdots)\cdots$. Choose $x \neq y \in [n]$ with $x, y \neq a, b, c, d$, possible because $n \geq 6$. Conjugate σ by $\gamma = (cxy)$, and define $\tau = \gamma\sigma\gamma$.

There are two cases for τ. If $\sigma = (ab)(cd)\cdots$, then $\tau = (\gamma(a)\gamma(b))(\gamma(c)\gamma(d))\cdots = (ab)(xd)\cdots$. Otherwise, $\tau = (\gamma(a)\gamma(b)\gamma(c)\cdots)\cdots = (abx\cdots)\cdots$. In both cases, $\tau \neq \sigma$ by choice of x, so $\sigma\tau^{-1} \neq 1$. But, $\sigma\tau^{-1}$ fixes b.

Thus N contains a permutation that is nontrivial but has a fixed point, so we have a contradiction.

Thus no such N exists, and A_n is simple. □

9 September 22nd

9.1 The projective linear groups

Definition 9.1.1. Let $\mathbb{F}$ be a field. Then we define the *projective general linear group* as $\mathrm{PGL}(n,\mathbb{F}) = \mathrm{GL}(n,\mathbb{F})/Z(\mathrm{GL}(n,\mathbb{F}))$, and the *projective special linear group* as $\mathrm{PSL}(n,\mathbb{F}) = \mathrm{SL}(n,\mathbb{F})/Z(\mathrm{SL}(n,\mathbb{F}))$.

Remark. Refer to the Homework problem stating that $Z(\mathrm{GL}(n,\mathbb{F})) = \{a \cdot I_n \mid a \in \mathbb{F}^\times\}$ and that $Z(\mathrm{SL}(n,\mathbb{F})) = \{a \cdot I_n \mid a \in \mu_n(\mathbb{F})\}$. We use the notation that $\mathrm{GL}(n,q) = \mathrm{GL}(n,\mathbb{F}_q)$, and so on.

Remark.

1. $|\mathrm{GL}(n,q)| = (q^n-1)(q^n-q)\cdots(q^n-q^{n-1})$
2. $\mathbb{F}_2^\times = \{1\}$, so $\mathrm{GL}(n,2) = \mathrm{SL}(n,2) = \mathrm{PSL}(n,2) = \mathrm{PGL}(n,2)$. In general, we have a diagram

$$\begin{tikzcd}
 & \mathrm{GL}(n,\mathbb{F}) \arrow[dr, "s"] & \\
\mathrm{SL}(n,\mathbb{F}) \arrow[ur, "i"] \arrow[dr, "s"'] & & \mathrm{PGL}(n,\mathbb{F}) \\
 & \mathrm{PSL}(n,\mathbb{F}) \arrow[ur, "i"'] &
\end{tikzcd}$$

   where arrows labeled by i are injective and those labeled by s are surjective.

3. $\mathrm{PGL}(1,\mathbb{F}) = \{1\}$; $\mathrm{PSL}(2,2) \cong S_3$; $\mathrm{PSL}(2,3) \cong A_4$ (these last two are on the homework). Our goal is to show that in all other cases, i.e. with $n > 2$ and any q or $n = 2$ and $q > 3$, $\mathrm{PSL}(n,q)$ is simple.

We will use Iwasawa's Lemma, as discussed last time, which is basically our only tool. But for now we have other lemmas!

Lemma 9.1.2. *Let $\mathbb{F}$ be any field.* $\mathrm{SL}(2,\mathbb{F})$ *is generated by matrices of the form* $\begin{pmatrix}1 & *\\ 0 & 1\end{pmatrix}$ *and* $\begin{pmatrix}1 & 0\\ * & 1\end{pmatrix}$. *These matrices are called* transvections *or* shear mappings*; although those words are mainly used to apply to the transformations represented by those matrices.*

Proof. Take $\begin{pmatrix}a & b\\ c & d\end{pmatrix} \in \mathrm{SL}(2,\mathbb{F})$. There are three cases to consider.

<u>$b \neq 0$:</u>

$$\begin{pmatrix}a & b\\ c & d\end{pmatrix} = \begin{pmatrix}1 & 0\\ \frac{d-1}{b} & 1\end{pmatrix}\begin{pmatrix}1 & b\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ \frac{a-1}{b} & 1\end{pmatrix}.$$

$\underline{c \neq 0}$:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & \frac{a-1}{c} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} \begin{pmatrix} 1 & \frac{d-1}{c} \\ 0 & 1 \end{pmatrix}.$$

$\underline{b = c = 0}$:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ d-1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ a-1 & 1 \end{pmatrix} \begin{pmatrix} 1 & -d \\ 0 & 1 \end{pmatrix}.$$

Note that all these equations hold in the context that $ad - bc = 1$. They're found via row-reduction. Yay.... □

For the purposes of today, let $U = \{\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \mid b \in \mathbb{F}\}$ and $B = \{\begin{pmatrix} a & b \\ 0 & \frac{1}{a} \end{pmatrix} \mid a \in \mathbb{F}^\times, b \in \mathbb{F}\}$.

Lemma 9.1.3. *(a)* $U \trianglelefteq B \trianglelefteq \mathrm{SL}(2, \mathbb{F})$, *and* $B \cong U \rtimes \mathbb{F}^\times$.

(b) U *is abelian*

(c) U *and its conjugates generate* $\mathrm{SL}(2, \mathbb{F})$.

Proof. (a) Exercise

(b) $\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \beta \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & b+\beta \\ 0 & 1 \end{pmatrix}$ so $U \cong (\mathbb{F}, +)$.

(c) $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -b & 1 \end{pmatrix}$

□

Lemma 9.1.4. *If* $|\mathbb{F}| \geq 4$ *(where* $\mathbb{F}$ *isn't necessarily finite), then* $\mathrm{SL}(2, \mathbb{F})$ *is perfect.*

Proof. It suffices to show that any $\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$ is a commutator; in that case the derived subgroup contains U, and thus all its conjugates because the derived subgroup is normal, but U and its conjugates generate $\mathrm{SL}(2, \mathbb{F})$, so $\mathrm{SL}(2, \mathbb{F})' = \mathrm{SL}(2, \mathbb{F})$. Perfection!

Consider

$$\left[\begin{pmatrix} a & 0 \\ 0 & \frac{1}{a} \end{pmatrix}, \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}\right] = \begin{pmatrix} 1 & (a^2-1)b \\ 0 & 1 \end{pmatrix}.$$

Then it suffices to show that $\forall c \in \mathbb{F}, \exists a \in \mathbb{F}^\times, b \in \mathbb{F}$ with $c = (a^2 - 1)b$.

For this, it suffices to find $a \in \mathbb{F}^\times$ with $a^2 - 1 \neq 0$; then $b = \frac{c}{a^2-1}$. So we need $a \in \mathbb{F}$ with $a^3 - a \neq 0$. The polynomial $x^3 - x$ has at most 3 roots in $\mathbb{F}$, one of which is 0, so since $|\mathbb{F}| \geq 4$ we can always find a scalar that is not a root. □

Recall $\mathrm{SL}(2, \mathbb{F})$ acts on $\mathbb{F}^2$ by $A \cdot v = Av$ and hence also on $\mathbb{P}^1(\mathbb{F})$ = set of 1-dim subspaces of $\mathbb{F}^2$.

Lemma 9.1.5. *(a) The action is* 2*-transitive.*

(b) The stabilizer of the x-axis is B.

(c) The kernel of the action is the center of the group.

Proof. (a) Let (l_1, l_2) and (r_1, r_2) be pairs of lines through the origin, with $l_1 \neq l_2$ and $r_1 \neq r_2$. We need $A \in \mathrm{SL}(2, \mathbb{F})$ such that $Al_1 = r_1$ and $Al_2 = r_2$. Choose $v_1 \in l_1, v_2 \in l_2, w_1 \in r_1, w_2 \in r_2$, all non-zero. Then $\{v_1, v_2\}$ and $\{w_1, w_2\}$ are bases of $\mathbb{F}^2$. So there is $T \in \mathrm{GL}(\mathbb{F}^2)$ with $T(v_1) = w_1$ and $T(v_2) = w_2$. Let A be the matrix of T in the canonical basis, and $D = \det(A) \neq 0$. Let $S \in \mathrm{GL}(\mathbb{F}^2)$ be such that $S(w_1) = \frac{1}{D} w_1$, $S(w_2) = w_2$. Then $BAv_1 = Bw_1 = \frac{1}{D} w_1$, so $BAl_1 = r_1$, and $BAv_2 = Bw_2 = w_2$, so $BAl_2 = r_2$. Moreover, $\det(BA) = \det(B)\det(A) = 1$.

(b) By verification.

(c) By verification.

□

Theorem 9.1.6. *If* $|\mathbb{F}| \geq 4$*,* $\mathrm{PSL}(2, \mathbb{F})$ *is simple.*

Proof. Apply Iwasawa's Lemma to the action of $\mathrm{PSL}(2, \mathbb{F})$ on $\mathbb{P}^1(\mathbb{F})$, noting that quotients will inherit the 2-transitivity and the perfection. □

Remark. Why PSL? GL also acts on $\mathbb{P}^1(\mathbb{F})$. Well, recall the diagram as before.

$$\begin{tikzcd}
& \mathrm{GL}(n, \mathbb{F}) \arrow[dr, "s"] & \\
\mathrm{SL}(n, \mathbb{F}) \arrow[ur, hook, "i"] \arrow[dr, "s"'] & & \mathrm{PGL}(n, \mathbb{F}) \\
& \mathrm{PSL}(n, \mathbb{F}) \arrow[ur, hook, "i"'] &
\end{tikzcd}$$

GL is not perfect; $\mathrm{GL}' \leq \mathrm{SL}$. The action is also not faithful. SL is perfect, but the action isn't faithful; PGL isn't perfect and has a faithful action.

Compare with

$$\begin{tikzcd}
& S_n \\
A_n \arrow[ur, "i"] &
\end{tikzcd}$$

and the canonical action on $[n]$. The action of S_n is faithful but S_n is not perfect; the action of A_n is faithful and A_n is perfect. (nb: $n > 5$). This is the limiting nonexistent case, of a field of 1 element.

Theorem 9.1.7. *If* $n \geq 3$, $\mathrm{PSL}(n, \mathbb{F})$ *is simple.*

Proof.

- $\mathrm{PSL}(n, \mathbb{F})$ acts on $\mathbb{P}^{n-1}(\mathbb{F})$.
- The action is 2-transitive and the kernel is $Z(\mathrm{SL}(n, \mathbb{F}))$.
- The stabilizer of the 1-dimensional subspace spanned by $(1, 0, \ldots, 0) \in \mathbb{F}^n$, or

  $$\left\{ \begin{pmatrix} a & v \\ 0 & A \end{pmatrix} \mid a \in \mathbb{F}^\times, A \in \mathrm{GL}(n-1, \mathbb{F}), v \in \mathbb{F}^{n-1}, a \det(A) = 1 \right\}.$$

  It contains a normal abelian subgroup $U = \left\{ \begin{pmatrix} 1 & v \\ 0 & I_{n-1} \end{pmatrix} \mid v \in \mathbb{F}^{n-1} \right\}$, called transvections. $U \cong (\mathbb{F}, +)$, so it is abelian.
- U and its conjugates generate $\mathrm{SL}(n, \mathbb{F})$ because any matrix of determinate 1 is a product of matrices of the form $E_{ij}(\lambda) = I_n + \lambda e_{ij}$ with $i \neq j$, and any $E_{ij}(\lambda)$ is a conjugate of an element of U. $E_{1j}(\lambda) \in U$, and any two matrices $E_{ij}(\lambda)$, varying i and j, are conjugate for $n \geq 3$.
- For $n \geq 3$, $\mathrm{SL}(n, \mathbb{F})$ is perfect.

  $$\left[\begin{pmatrix} 1 & \lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \right] = \begin{pmatrix} 1 & 0 & \lambda \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = E_{13}(\lambda),$$

  so the commutator subgroup contains the set of generators.

More comments will follow next time. □

10 September 24th

10.1 Projective Linear Groups, continued

Recall from last time that we were mired in the proof that $\mathrm{PSL}(n, q)$ is simple when $n > 2$ or when $n = 2$ and $q > 3$.

Note that $|\mathrm{GL}(n, q)| = (q^n - 1)(q^n - q) \cdots (q^n - q^{n-1})$, so $|\mathrm{PSL}(n, q)| = |\mathrm{GL}(n, q)| / \gcd(n, q-1) \cdot (q-1)$. We have the following table of orders.

$n \backslash q$	2	3	4	5	7
2	6	12	60	60	168
3	168	5616	20160	372000	
4	20160				

The entries in red represent the groups that aren't simple. Note the following facts:

Fact 10.1.1. *1.* $\mathrm{PSL}(2, 2) \equiv S_3$, $\mathrm{PSL}(2, 3) \equiv A_4$ *(HW)*

2. There is a unique simple group of order 60*, so* $\mathrm{PSL}(2,4) \cong \mathrm{PSL}(2,5) \cong A_5$*.*

3. There is a unique simple group of order 168*, so* $\mathrm{PSL}(2,7) \cong \mathrm{PSL}(3,2)$*.*

4. $\mathrm{PSL}(4,2) \cong A_8$ *but* $\mathrm{PSL}(3,4)$ *is not isomorphic to them.*

Our short-term goal is to understand number 3. But first, we will comment on number 2. Given a simple group G of $|G| = 60$, find $H \leq G$ of $|G/H| = 5$, which can be counted by Sylows. Then G acts on G/H, which is a map from $G \to S_5$, and so on.

10.2 Projective Geometries

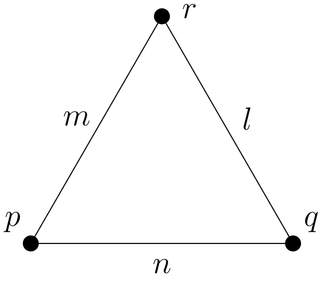
Definition 10.2.1. An *incidence geometry* of rank 2 (briefly, a *plane*) is $\mathcal{G} = (\mathcal{G}_0, \mathcal{G}_1, R)$ where $\mathcal{G}_0$ and $\mathcal{G}_1$ are sets and R is a relation between $\mathcal{G}_0$ and $\mathcal{G}_1$. The elements of $\mathcal{G}_0$ are called *points*, and the elements of $\mathcal{G}_1$ are called *lines*. Given $p \in \mathcal{G}_0$ and $l \in \mathcal{G}_1$, if pRl, we say one of "p lies in l," "l goes through p," or "p and l are incident."

Definition 10.2.2. A plane is *projective* if

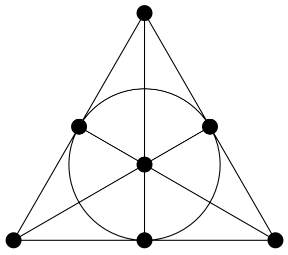
- Given two distinct points, ∃! line going through them.
- Given two distinct lines, ∃! point lying in both.
- There are at least 3 noncollinear points.

Example.

- The smallest projective plane is a triangle with points $\mathcal{G}_0 = \{r, p, q\}$, lines $\mathcal{G}_1 = \{l, m, n\}$, and R the triangle relation as shown in figure below.



- The Fano plane, $\mathcal{F}$. Assume that there are at least 4 points, no three of which are collinear; by adding points and lines until we get a minimal plane, we end up with the picture below.



It's often referred to as the smallest projective geometry, becomes sometimes people exclude by definition the triangle. But we don't.

- Let $\mathbb{F}$ be a field. Let $\mathrm{PG}(2,\mathbb{F})$ be the plane with points the one-dimensional subspaces in $\mathbb{F}^3 = \mathbb{P}^2(\mathbb{F})$, lines the two-dimensional subspaces in $\mathbb{F}^3$, and incidence defined by containment.

Proposition 10.2.3. $\mathrm{PG}(2,\mathbb{F}_2) \cong \mathcal{F}$, *the Fano plane.*

Proof. We may think of $\mathbb{F}_2^3$ as $\{(000),(001),(010),(100),(110),(101),(011),(111)\}$. The number of points in $\mathrm{PG}(2,\mathbb{F}_2)$ is the number of lines through the origin in $\mathbb{F}_2^3$. But each nonzero element spans its own line, consisting of itself and the origin, so there are 7. We claim there are also 7 lines in this geometry. The lines are 2-dimensional subspaces; there are the three coordinate planes. There are then three more consisting of one coordinate axis and the cube's opposite edge. The last plane is $\{(000),(110),(101),(011)\}$, with equation $x+y+z=0$.

By direct inspection, containment follows the Fano plane picture. □

Definition 10.2.4. A *symmetry* of a plane $\mathcal{G}$ is a $\sigma = (\sigma_0,\sigma_1)$ where $\sigma_0 : \mathcal{G}_0 \to \mathcal{G}_0$ and $\sigma_1 : \mathcal{G}_1 \to \mathcal{G}_1$ such that p is incident to l if and only if $\sigma_0(p)$ is incident to $\sigma_1(l)$.

Let $\mathrm{Aut}(\mathcal{G})$ be the group of symmetries of $\mathcal{G}$.

Lemma 10.2.5. $|\mathrm{Aut}(\mathcal{F})| \leq 168$.

Proof. Choose 3 noncollinear points p,q,r. A symmetry σ is determined by the values $\sigma_0(p)$, $\sigma_0(q)$, $\sigma_0(r)$, because all other points in the plane can be filled in eventually as the last point in some line that must already be defined.

Thus the map $\mathrm{Aut}(\mathcal{F}) \to \{(x,y,z) \in \mathcal{F}^3 \mid x \neq y \neq z, x \neq z\}$, with $\sigma \mapsto (\sigma_0(p),\sigma_0(q),\sigma_0(r))$, is injective. Thus $|\mathrm{Aut}(\mathcal{F})| \leq 7*6*5 = 210$. But that's not quite what we wanted. Recall further that p,q,r weren't collinear. Thus their images must also be noncollinear, so r can't be mapped to the last point on the line containing p and q; this allows only 4 remaining options for the image of r. Thus $|\mathrm{Aut}(\mathcal{F})| \leq 7*6*4 = 168$. □

Proposition 10.2.6. $\mathrm{Aut}(\mathcal{F}) \cong \mathrm{PSL}(3,2) = \mathrm{GL}(3,2)$.

Proof. $\mathrm{GL}(3,2)$ acts on $\mathbb{F}_2^3$ linearly. It sends i-subspaces to i-subspaces and it preserves containment, so $\mathrm{GL}(3,2)$ acts on $\mathrm{PG}(2,\mathbb{F}_2)$ by symmetries.

So we get a map $\mathrm{GL}(3,2) \to \mathrm{Aut}(\mathrm{PG}(2,\mathbb{F}_2))$.

Generally the action of $\mathrm{GL}(n+1,\mathbb{F})$ on $\mathbb{P}^n(\mathbb{F})$ is not faithful. The kernel is the space of scalar matrices. But we are in the case where the field has two elements, so the action is faithful in our case.

Thus $\mathrm{GL}(3,2) \hookrightarrow \mathrm{Aut}(\mathrm{PG}(2,\mathbb{F}_2))$. But $|\mathrm{GL}(3,2)| = 168$ and $|\mathrm{Aut}(\mathrm{PG}(2,\mathbb{F}_2))| \leq 168$, so the sizes must be equal and the map is an isomorphism. □

Remark. One can define incidence (and projective) geometries of higher rank. $\mathrm{PG}(n,\mathbb{F})$ is a projective geometry of rank n for which:

- points = 1-subspaces of $\mathbb{F}^{n+1}$

- lines = 2-subspaces of $\mathbb{F}^{n+1}$
- planes = 3-subspaces of $\mathbb{F}^{n+1}$

and so on.

This gives us

Fact 10.2.7 (The Fundamental Theorem of Projective Geometry)**.** $\mathrm{Aut}(\mathrm{PG}(n, \mathbb{F})) \cong \mathrm{PGL}(n+1, \mathbb{F}) \rtimes \mathrm{Aut}(\mathbb{F})$.

Fact 10.2.8. *Let G be a simple group of order 168. Then $G \cong \mathrm{PSL}(3, 2)$.*

Proof. Not really a full proof, but here's the key idea. We have to come up with an action on the Fano plane. Let $\mathcal{C}$ be the collection of subgroups of G isomorphic to V_4.

There are 14 such subgroups. You can see this via looking at Sylows, or something.

They come in 2 conjugacy classes of size 7. Call these two classes $\mathcal{H}$ and $\mathcal{K}$ and let $\mathcal{G}$ be the plane for which $\mathcal{H}$ is the set of points, $\mathcal{K}$ is the set of lines, and H is incident to K if there is a 2-Sylow D containing both H and K.

Then one can claim that $\mathcal{G} \cong \mathcal{F}$. We have then an action on the plane by conjugation, which leads to a map from G to $\mathrm{Aut}(\mathcal{G})$, which ends up being an isomorphism. □

10.3 Other simple groups

This will be just a rough outline.

- Cyclic groups of prime order, $\mathbb{Z}_p$
- Alternating groups, A_n
- Linear $\mathrm{PSL}(n, q)$ (type A)
- Orthogonal groups $\mathrm{PSO}(2n+1, q)$ (type B) and $\mathrm{PSO}^+(2n, q)$ (type D, with oriflamme geometry)
- Symplectic $\mathrm{PS}_p(2n, q)$ (type C, involving polar geometries)
- Unitary $\mathrm{PSU}(n, q)$

and so on and so forth. These are Chevalley groups, and Steinberg groups. There are 19 families in total, plus 26 sporadic cases including the monster and the baby monster.

11 September 29th

11.1 Words and Free Monoids...Free Groups to Come!

Definition 11.1.1. A *monoid* is a set M with a binary operation that is associative (just like a group!) and unital, i.e. it has an identity (just like a group!). A group is simply a monoid with inverses.

Definition 11.1.2. Let S be a set and let

$$\begin{aligned} S^* &= \{(s_1, \dots, s_n) \mid s_i \in S, n \geq 0\} \\ &= \bigcup_{n \geq 0} S^n, \end{aligned}$$

or the set of all finite sequences including the empty sequence. Then concatenation is defined as follows:

$$(s_1, \dots, s_i) \cdot (t_1, \dots, t_j) = (s_1, \dots, s_i, t_1, \dots, t_j).$$

Proposition 11.1.3.

1. *S^* is a monoid under concatenation. The unit is the empty sequence.*
2. *Let M be a monoid and $m : S \to M$ be an arbitrary map. Then there exists a unique morphism of words $\hat{m} : S^* \to M$ such that the following diagram commutes, where $i(s) = s$.*

$$\begin{tikzcd} S \arrow[r, hook, "i"] \arrow[dr, "m"'] & S^* \arrow[d, dashed, "\hat{m}"] \\ & M \end{tikzcd}$$

Proof. Define $\hat{m}(s_1, \dots, s_n) = m(s_1) \cdots m(s_n)$; the rest of the proof of both parts is just checking axioms. □

A bit of terminology that we use is that an element of S^* is a *word*. An element of S is a *letter* an S is the *alphabet*. Given elements $\{m_s\}_{s \in S}$ and a word $w \in S^*$, the *evaluation* of w at $\{m_s\}_{s \in S}$ in M is $w(m_s)_{s \in S} = \hat{m}(w) \in M$, where $m : S \to M$ is $m(s) = m_s$.

For instance, if $S = \{a, b\}$, given $S \to M$ sending $a \mapsto \alpha$ and $b \mapsto \beta$, and $w = (a, b, b, a, a)$, then the evaluation of w at m is $\alpha\beta^2\alpha^2$.

We want to move from monoids to groups, so we'll start moving in that direction. Consider a group G and $g : S \to G$ a map. Let $S^{-1} = \{s^{-1} \mid s \in S\}$, or another copy of S. We then extend g to $S \cup S^{-1}$ by $g(s^{-1}) = g(s)^{-1}$. By the proposition, there exists a unique morphism of monoids $\hat{g} : (S \cup S^{-1})^* \to G$ that extends g, in the sense that the diagram commutes.

Proposition 11.1.4. *The image of $\hat{g}$ is a subgroup of G. Moreover, it is the smallest subgroup of G containing $g(S)$, so it is the subgroup of G generated by $g(S)$.*

Proof. Since $\hat{g}$ is a morphism of monoids, we are guaranteed that its image $\operatorname{im}\hat{g}$ is a submonoid. So we need only check that it is closed under inverses. $\hat{g}(t_1,\dots,t_n)^{-1} = (g(t_1)\cdots g(t_n))^{-1} = g(t_n)^{-1}\cdots g(t_1)^{-1} = g(t_n^{-1})\cdots g(t_1^{-1}) = \hat{g}(t_n^{-1},\dots,t_1^{-1}) \in \operatorname{im}\hat{g}$, and is thus a subgroup.

The rest can be checked pretty easily. □

So now we have to look very quickly at monoid quotients before we get our group. We're almost there. We have the right type of universal mapping property.

11.2 Monoid quotients

Definition 11.2.1. Let M be a monoid and $\sim$ an equivalence relation. We say the relation is *left compatible* if $a \sim b$ implies $xa \sim xb$ for all $x, a, b \in M$. It is *right compatible* if $a \sim b$ implies $ax \sim bx$ for all $x, a, b \in M$. It is *two-sided compatible* if $a \sim b$ and $x \sim y$ implies $ax \sim by$ for all $x, y, a, b \in M$.

Remark. Left and right compatible is the same as two-sided.

Proof. Assume $\sim$ is left and right compatible, and let $a \sim b$ and $x \sim y$. Then $ax \sim bx$ and $bx \sim by$, so $ax \sim by$ and $\sim$ is two-sided compatible. The converse implication is even simpler and easier! □

Let $M/\sim$ be the set of equivalence class $\overline{a}$, for $a \in M$. Consider defining an operation on $M/\sim$ by $\overline{a}\cdot\overline{b} = \overline{ab}$.

Proposition 11.2.2. *This operation is well-defined if and only if $\sim$ is two-sided compatible. In this case, the quotient $M/\sim$ is in fact a monoid.*

Proof. Left as an exercise. □

Note that we don't have to care about any of this for groups. That's because two-sided relations are exactly those given by cosets of a normal subgroup, so that's special. Precisely, we have the following.

Proposition 11.2.3. *Let G be a group and $\sim$ an equivalence relation on G. Then*

(i) $\sim$ is left compatible if and only if there exists $H \le G$ such that $a \sim b$ if and only if $a^{-1}b \in H$.

(ii) $\sim$ is right compatible if and only if there exists $H \le G$ such that $a \sim b$ if and only if $ab^{-1} \in H$.

(iii) $\sim$ is two-sided compatible if and only if there is an $N \trianglelefteq G$ such that either of the above holds. In this case, both hold, and the quotient is a group.

We're not going to prove this, because it's kind of beside the point, but it's a fun exercise.

Okay okay okay, now we can finally get to what we all know we've wanted to be talking about all along.

11.3 Free Groups

Our goal, informally, is that given a set S, we will construct a group $F(S)$ such that

- $F(S)$ is generated by S
- There are no relations among the elements of S other than those forced by the group axioms.

Example. Let $a, b \in S$. $abb^{-1}a^{-1} = 1$ is what we would call a forced relation. It holds in any group. But $aba^{-1} = b^2$ is not forced.

Definition 11.3.1. Let S be a set, G a group and $g : S \to G$ a map. (Equivalently, let $\{g_s\}_{s \in S}$ be a collection of elements of G). Given two words $w_1, w_2 \in (S \cup S^{-1})^*$, we say $\{g_s\}$ *satisfies the relation* $w_1 = w_2$ if $w_1(g_s)_{s \in S} = w_2(g_s)_{s \in S}$ in G. We say that the relation $w_1 = w_2$ is *forced* if it is satisfied by all families $g : S \to G$ in all groups G.

Okay so formally, our goal is:
Given a set S, construct a group $F(S)$ and a map $i : S \to F(S)$ such that

(i) $F(S)$ is generated by $i(S)$

(ii) Given two words $w_1, w_2 \in (S \cup S^{-1})^*$, if the elements $\{i_s\}_{s \in S}$ satisfy the relation $w_1 = w_2$, then the relation $w_1 = w_2$ is forced, just like our smiles!

Note that we can't set $F(S) = (S \cup S^{-1})^*$, because this monoid is not a group. For example, $(s) \cdot (s^{-1}) = (s, s^{-1}) \neq ()$. But that's all that fails, so all we have to do is fix it.

Define an equivalence relation on $(S \cup S^{-1})^*$ by $w \sim w'$ if and only if we can obtain one word from the other by finitely many insertions or deletions of subwords of the form (s, s^{-1}), $s \in S \cup S^{-1}$.

Example.

$$(a, b^{-1}, a, a^{-1}, b) \sim (a, b^{-1}, b) \sim (a) \sim (b, b^{-1}, a).$$

Fact 11.3.2. *Did you know? Free groups are called free because you get them for free! Unlike other groups, which are on average sold at a price of* \$2.28 *an element, the establishment will just* <u>*give*</u> *you free groups!*

Proposition 11.3.3. *The relation* $\sim$ *on the monoid* $(S \cup S^{-1})^*$ *is two-sided compatible.*

Proof. Let's be real it's pretty clear. □

Definition 11.3.4. Let $F(S) = (S \cup S^{-1})^* / \sim$. It is a monoid. Let $[s_1, \ldots, s_n]$ denote the equivalence class of $(s_1, \ldots, s_n)$ with $s_i \in S \cup S^{-1}$.

Proposition 11.3.5. *$F(S)$ is a group. Moreover, if we let $i : S \to F(S)$ be $i(s) = [s]$, then the two formal properties are satisfied.*

Proof. Let's find some inverses. We claim that $[s_1, \dots, s_n]^{-1} = [s_n^{-1}, \dots s_1^{-1}]$.

$$\begin{aligned}[s_1, \dots, s_n][s_n^{-1}, \dots, s_1^{-1}] &= [s_1, \dots, s_n, s_n^{-1}, \dots, s_1^{-1}] \\ &= [s_1, \dots, s_{n-1}, s_{n-1}^{-1}, \dots, s_1^{-1}] \\ &\vdots \\ &= [s_1, s_1^{-1}] \\ &= [].\end{aligned}$$

So it's a group.

Now we claim that given $w \in (S \cup S^{-1})^*$, $\hat{i}(w) = [w] \in F(S)$. We have the diagram

$$\begin{tikzcd} S \arrow[r, hook] \arrow[dr, "i"'] & S \cup S^{-1} \arrow[r] \arrow[d, "i"] & (S \cup S^{-1})^* \arrow[dl, "\hat{i}"] \\ & F(S) & \end{tikzcd}$$

If $w = (s_1, \dots, s_n)$, then $\hat{i}(w) = i(s_1) \cdots i(s_n) = [s_1] \cdots [s_n] = [s_1 \cdots s_n] = [w]$. Thus $\hat{i}$ is surjective. But we saw before that the image of $\hat{i}$ is the subgroup generated by $i(S)$. So $F(S)$ is generated by $i(S)$, so we have the first thing we wanted to have.

Now let $w_1, w_2 \in (S \cup S^{-1})^*$ and suppose that the elements $\{i_s\}_{s \in S}$ satisfy $w_1 = w_2$. Then $\hat{i}(w_1) = \hat{i}(w_2)$, so $[w_1] = [w_2]$, so $w_1 \sim w_2$. We can assume that we can pass between w_1 and w_2 by a single insertion or deletion, and proceed inductively. More precisely, we're assuming that $w_1 = (s_1, \dots, s_n)$ and $w_2 = (s_1, \dots, s, s^{-1}, \dots, s_n)$, for some $s \in S \cup S^{-1}$. Now if G is any group and $\{g_s\}_{s \in S}$ are arbitrary elements of G, we have $w_1(g_s)_{s \in S} = g_{s_1} \cdots g_{s_n}$ and $w_2(g_s)_{s \in S} = g_{s_1} \cdots g(s) g(s^{-1}) \cdots g_{s_n} = g_{s_1} \cdots g(s) g(s)^{-1} \cdots g_{s_n} = g_{s_1} \cdots g_{s_n} = w_1(g_s)_{s \in S}$, so the relation $w_1 = w_2$ is satisfied in all groups and it's forced and we're done. □

12 October 1st

12.1 Free Groups, continued

Recall that given a set S:

1. S^* is a monoid and given a map $m : S \to M$ to another monoid, there is a unique morphism of monoids $\hat{m} : S^* \to M$ extending m, so that the following diagram commutes.

$$\begin{tikzcd} S \arrow[r, hook, "i"] \arrow[dr, "m"'] & S^* \arrow[d, dashed, "\hat{m}"] \\ & M \end{tikzcd}$$

2. $F(S) = (S \cup S^{-1})^*/\sim$, where $\sim$ is the relation generated by insertion or deletion of pairs (s, s^{-1}) for $s \in S \cup S^{-1}$. We also have $i : S \to F(S)$ given by $i(s) = [s]$. The appropriate diagram is as follows.

$$\begin{tikzcd}
S \arrow[r, hook] \arrow[drr, "i"'] & S \cup S^{-1} \arrow[r, hook] \arrow[dr, "i"'] & (S \cup S^{-1})^* \arrow[d, "\hat{i}"] \\
 & & F(S)
\end{tikzcd}$$

We know:

- $F(S)$ is a group.
- It is generated by $i(S)$.
- $\hat{i}(w) = [w]$ for all $w \in (S \cup S^{-1})^*$.
- Let $g : S \to G$ be a map to another group. If $\hat{i}(w_1) = \hat{i}(w_2)$, then $\hat{g}(w_1) = \hat{g}(w_2)$, with $\hat{g}$ the equivalent map in the appropriate diagram. (i.e. take the diagram above, just simply replacing all the i's with g's.

Proposition 12.1.1. *Let G be a group and $g : S \to G$ a map. Then there exists a unique morphism of groups $\widetilde{g} : F(S) \to G$ such that the diagram*

$$\begin{tikzcd}
S \arrow[r, hook, "i"] \arrow[dr, "g"'] & F(S) \arrow[d, dashed, "\widetilde{g}"] \\
 & G
\end{tikzcd}$$

commutes.

Proof. Define $\widetilde{g}([w]) = \hat{g}(w)$. Is this well-defined? Suppose $[w_1] = [w_2]$. Then $\hat{i}(w_1) = \hat{i}(w_2)$, so $\hat{g}(w_1) = \hat{g}(w_2)$, so then $\widetilde{g}([w_1]) = \widetilde{g}([w_2])$, so $\widetilde{g}$ is well-defined.

Now we show commutativity of the diagram. Take $s \in S$. $(\widetilde{g} \circ i)(s) = \widetilde{g}([s]) = \hat{g}((s)) = g(s)$, so the diagram commutes.

To check that $\widetilde{g}$ is a group morphism, take $[w], [w'] \in F(S)$. $\widetilde{g}([w] \cdot [w']) = \widetilde{g}([w \cdot w']) = \hat{g}(w \cdot w') = \hat{g}(w)\hat{g}(w') = \widetilde{g}([w])\widetilde{g}([w'])$.

All that remains is uniqueness. For uniqueness, we use uniqueness of $\hat{g}$. □

Proposition 12.1.2. *Let F be a group and $j : S \to F$ a map. Suppose that (j, F) satisfies the universal property in place of $(i, F(S))$. Then there exists a unique isomorphism of groups $F(S) \to F$ such that the following diagram commutes.*

$$\begin{tikzcd}
F(S) \arrow[rr, "\cong"] & & F \\
& S \arrow[ul, "i"] \arrow[ur, "j"'] &
\end{tikzcd}$$

Proof. Consider $\widetilde{j} : F(S) \to F$, the universal property for $F(S)$ applied to j, and $\widetilde{i} : F \to F(S)$, the universal property for F. Use the uniqueness for both to deduce that $\widetilde{j} \circ \widetilde{i} = \mathrm{id}_F$, $\widetilde{i} \circ \widetilde{j} = \mathrm{id}_{F(S)}$. □

So any group satisfying the property is isomorphic to $F(S)$, which we then can morally call *the* free group on S.

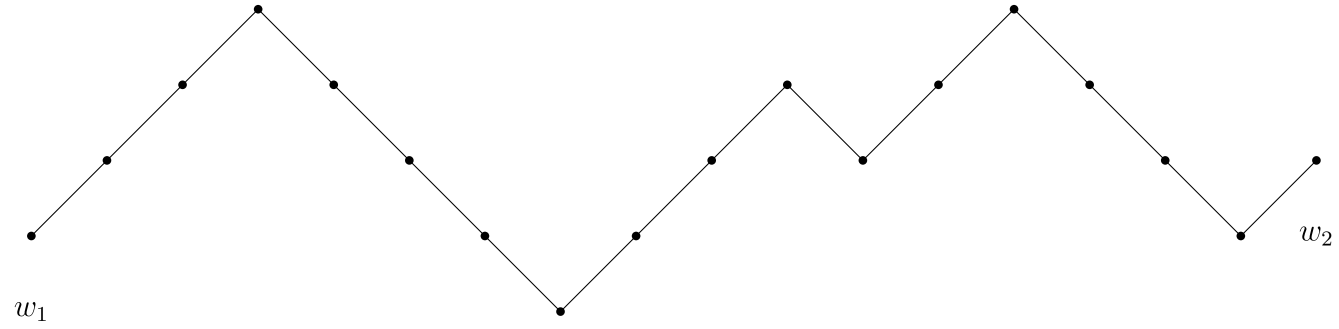
Definition 12.1.3. A pair of consecutive letters in a word $w \in (S \cup S^{-1})^*$ is *cancellable* if it is of the form (s, s^{-1}) for some $s \in S \cup S^{-1}$.

Definition 12.1.4. A word w is *reduced* if it contains no cancellable pairs.

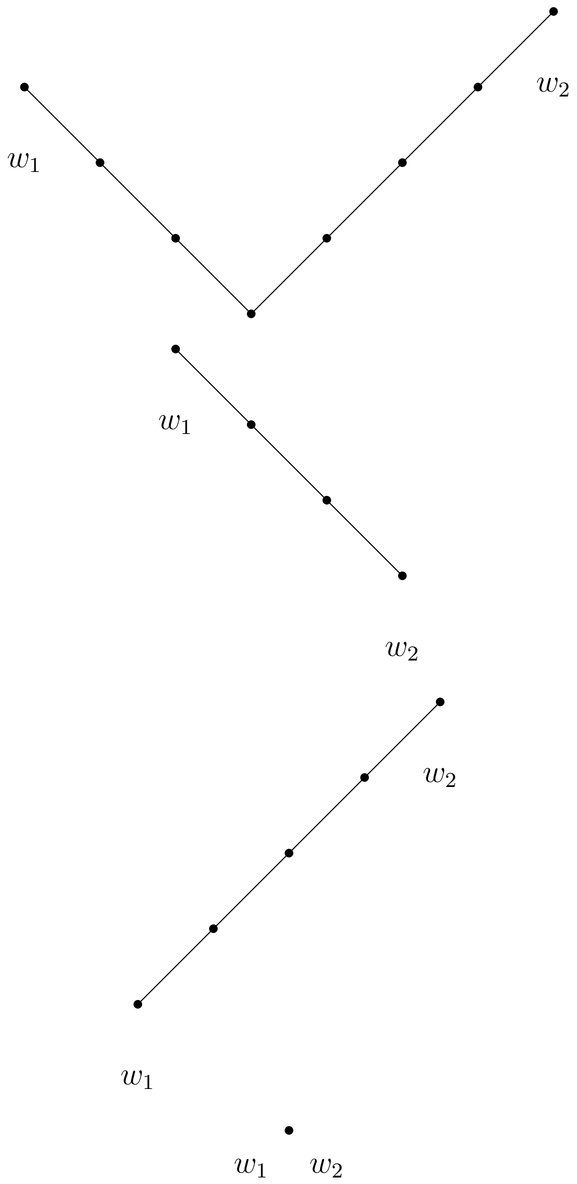
Proposition 12.1.5. *Each equivalence class of words contains a unique reduced word. Thus, the set $F(S)$ is in bijection with the set of reduced words. Note that Dummit and Foote defines the free group this way. We don't.*

Proof. Existence: Given a class, choose a representative w. If it is reduced, we are done. If not; there is some cancellable pair. Delete that pair; we then get a smaller element of the same class as our new word. But this process only ever shortens the word, so eventually it must terminate. When it terminates, we have a reduced word.

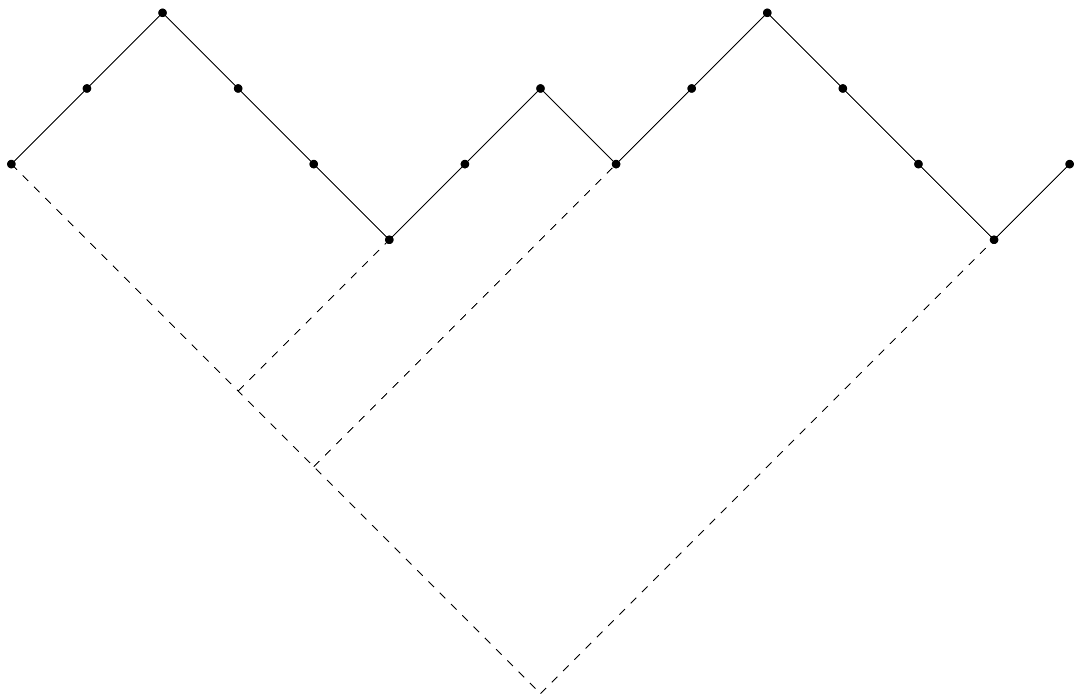
Uniqueness: Suppose w_1 and w_2 are two words in the same class. They are connected by a sequence of insertions and deletions of cancellable pairs. (This proof is like making mountains out of molehills.)



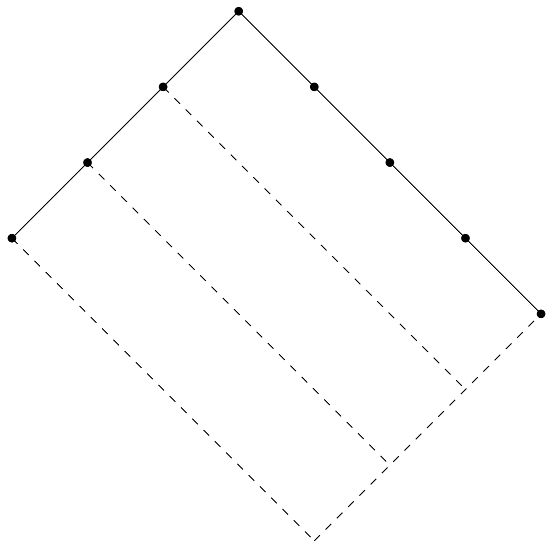
An increasing segment in the above imaginary picture corresponds to an insertion, and a decreasing segment corresponds to a deletion. For uniqueness, it suffices to show that there is a path of only deletions and then only insertions (i.e. a valley), where either of the two sides, or both, may have length 0 (all cases are pictured below). Why? Well, if both w_1 and w_2 are reduced, the only possible case is $w_1 = w_2$; all others imply the existence of a cancellable pair in at least one of the two words; going from w_1 or w_2 involves going "down" on both sides of the following picture.



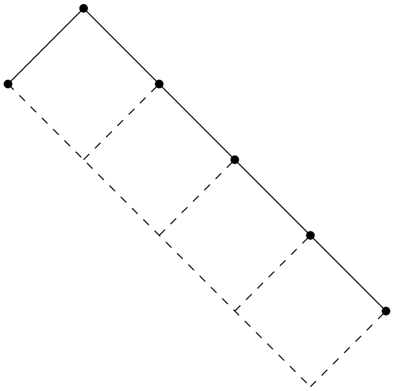
Okay, but to prove *that* it suffices to show that any peak can be resolved into a valley. Combinatorially, then, we can kill every peak until there are no peaks left (see below). This property is called *confluence*.



Let's prove us some confluence! To show confluence, it is enough to show the special case when either side of the peak has length 1, as below.



But for *that* it is sufficient to show the special case in which both sides have length 1, *provided* we can resolve in at most one step, as below.



Okay, finally, let's prove that! This case may arise in two ways. First, $w_1 = (u, v, t, t^{-1}, w)$, with u, v, w words; an insertion gives $(u, s, s^{-1}, v, t, t^{-1}, w)$ and $w_2 = (u, s, s^{-1}, v, w)$. In this disjoint case, the insertions and deletions pretty clearly commute, so we can make it a valley in one step. If however they're nondisjoint, we have something like $w_1 = (u, s, v)$, adding to get (u, s, s^{-1}, s, v), and deleting to get $(u, s, v) = w_2$, in which case we can simply collapse the molehill to a point. There are a couple of other cases, but they all boil down to the same thing.

So we're done. □

Corollary 12.1.6. *$i : S \to F(S)$, $i(s) = [s]$ is injective.*

Proof. If $i(s) = i(t)$, we have two reduced representatives for the same word, so we must have $s = t$. □

12.2 Group presentations

Let S be a set and take words $w_1, \ldots, w_n$ and $w'_1, \ldots, w'_n$ in $(S \cup S^{-1})^*$. Let N be the smallest normal subgroup of $F(S)$ containing $[w_i][w'_i]^{-1}$ for all i. Then we define a new group $\langle S \mid w_i = w'_i \forall i \rangle = F(S)/N$. This is called the "group generated by S and subject to the relations $w_i = w'_i$ for all i."

Proposition 12.2.1. *Let $g : S \to G$ be a map to a group G such that the $\widetilde{g}$ coming from the universal mapping property satisfies $\widetilde{g}([w_i]) = \widetilde{g}([w'_i])$. Then there is a unique morphism of groups $\overline{g} : \langle S \mid w_i = w'_i \rangle \to G$ such that the following diagram commutes.*

$$\begin{tikzcd}
S \arrow[r, "j"] \arrow[dr, "g"'] & \langle S \mid w_i = w'_i \rangle \arrow[d, dashed, "\overline{g}"] \\
 & G
\end{tikzcd}$$

Proof. Use the universal properties of free groups and of quotients. □

Example. 1. $G = \langle a, b \mid a^7 = 1, b^3 = 1, bab^{-1} = a^2 \rangle$. Let $A = \langle a \rangle$, and $B = \langle b \rangle$. $A \trianglelefteq G$, and A has order 1 or 7, and B has order 1 or 3. $A \cap B = \{1\}$ because of the orders, and $AB = G$. So immediately $G = A \rtimes B$. We claim $|G| = 21$. Consider $\mathbb{Z}_7 \rtimes \mathbb{Z}_3$; we will construct a map from this to G. We need an action by automorphisms of $\mathbb{Z}_3$ on $\mathbb{Z}_7$. Send $\overline{1}$ to the map consisting of multiplication by $\overline{2}$. $\overline{2}$ has order 3 in $\mathbb{Z}_7$, so this is well-defined. We then map $G \to \mathbb{Z}_7 \rtimes \mathbb{Z}_3$, with $a \mapsto (\overline{1}, \overline{0})$, $b \mapsto (\overline{0}, \overline{1})$. Generators are mapped to generators; it's easy to see this is an isomorphism.

2. $G = \langle a, b \mid a^7 = 1, b^3 = 1, bab^{-1} = a^3 \rangle$. We want to do the same analysis, but it doesn't work. In this group, $a^2 = 1$.

13 October 8th

Today marks a bridge of a transition in the course between groups and other algebraic objects, via Zorn's Lemma. But there is one more thing to say about groups, in particular group presentations, that will be covered in an extra class after the break.

13.1 Zorn's Lemma

Definition 13.1.1. A *poset* is a set along with a partial ordering on it. Let X be a poset. A subset C of X is a *chain* if it is totally ordered by the ordering of X: given $x, y \in C$, either $x \preceq y$ or $y \preceq x$. Given a subset S of X, an *upper bound* for S is an element $u \in X$ such that $x \preceq u$ for all for all $x \in S$.

Remark. 1. A chain may be uncountable.

2. $\varnothing$ is a chain in X.

3. $\varnothing$ has an upper bound in X if and only if $X \neq \varnothing$.

Definition 13.1.2. For a poset X, an element $m \in X$ is said to be *maximal* if there is no element $x \in X$ with $m \prec x$. It is a *maximum* if for all $x \in X$, $x \preceq m$.

Remark. 1. A maximum is maximal.

2. Maximal elements need not exist or be unique.

3. If a maximum element exists, it is unique.

Proposition 13.1.3 (Baby Zorn's Lemma)**.** *Let X be a finite non-empty poset. Then X has a maximal element.*

Proof. Choose $x_0 \in X$, with $X \neq \varnothing$. If x_0 is not maximal, choose x_1 with $x_0 \preceq x_1$; proceed by induction. □

Theorem 13.1.4 (Zorn's Lemma)**.** *Let X be a poset such that every chain in X has an upper bound. Then X has a maximal element.*

Remark. $X \neq \varnothing$ is implied by the hypothesis.

Before we prove this, examine the following propositions, all of which are helpful applications.

Proposition 13.1.5. *Any finitely generated nontrivial group has a maximal (proper) subgroup.*

Proof. Let X be the poset of proper subgroups of G, ordered by inclusion. $X \neq \varnothing$ because $\{1\} < G$ because G is nontrivial. Let $C = \{H_\alpha\}_{\alpha \in I}$ be a chain in X, with I totally ordered. Then let

$$H = \bigcup_{\alpha \in I} H_\alpha.$$

The key is to check that $H \leq G$. This uses that I is totally ordered, so any multiplication lies in some H_α. Moreover, we have to check crucially that H is proper, so that it's an element of our poset and hence an upper bound for our chain.

Suppose $H = G$. G is finitely generated, so choose generators $\{g_1, \dots, g_r\}$ for G. If $H = G$, then $g_i \in H$ for all i, so $g_i \in H_{\alpha_i}$ for some α_i. Let α be the maximum of the α_i's. Then H_α is a maximum, so $g_i \in H_\alpha$ for all i, so $H_\alpha = G$, contradicting $H_\alpha \in X$.

Thus every chain has an upper bound, so we apply Zorn's Lemma to get our result. □

Definition 13.1.6. Let R be a ring and M a left R-module (definition presumably seen in previous courses). A *basis* of M is a subset that is both linearly independent, i.e. the only zero linear combination is the trivial one, and generating.

Proposition 13.1.7. *Let R be a* division ring*, i.e. a ring in which every nonzero element is invertible. Then any non-trivial R-module has a basis.*

Proof. Let X be the poset of linearly independent subsets of M. If $C = \bigcap_{s \in C} s$ is an upper bound in X. It is linearly independent because a linear combination only involves finitely many elements at a time. Now check that a maximal linearly independent subset B of M is generating. If not, there exists $m \in M$ that is not generated by B. We claim that $B \cup \{m\}$ is still linearly independent, but is bigger than B (a contradiction). $B \cup \{m\}$ is still linearly independent because any linear combination is of the form

$$\lambda_1 b_1 + \cdots + \lambda_n b_n + \lambda m = 0, \ \lambda_i, \lambda \in R.$$

$\lambda \neq 0$, because B is linearly independent. Dividing by λ, which is possible because we are in a division ring, we know that m is generated by $b_1, \dots, b_n$, a contradiction.

Thus every chain has an upper bound, so by Zorn's Lemma we have the desired result. □

Remark. If every left module over a ring R has a basis, then R is a division ring. This is left as an exercise.

Fact 13.1.8. *Fun fact: Division rings arise alongside projective geometries, by adding the Desargues axiom. Look it up!*

Proposition 13.1.9. *Let R be a non-trivial ring with identity* 1*, i.e.* $0 \neq 1 \in R$*. Then R has a maximal ideal (left, right, or two-sided).*

Proof. Let X be the poset of proper (left) ideals of R. If $C = \{I_\alpha\}_{\alpha \in A}$ is a chain in X, then $\bigcup_{\alpha \in A} I_\alpha$ is a (left) ideal. We need to check that I is proper; if $I = R$, then $1 \in I$ so $1 \in I_\alpha$, so $R \subseteq I_\alpha$, a contradiction.

We then apply Zorn's Lemma. □

Clearly this lemma is the bees' knees of lemmas. It's so important, maybe we'd even want to choose it as an axiom! Now we'll go through the proof. To do so, we'll need a couple more definitions.

Definition 13.1.10. A poset is *well-ordered* if it is totally ordered and every nonempty subset S has a minimum $m \in S$.

Proposition 13.1.11 (Transfinite induction)**.** *Let A be a well-ordered poset, and let P be a property on A, where a* property *is a function $P : A \to \{T, F\}$ (think 'True,' 'False.') Suppose for any $b \in A$, that if P holds for all $a \prec b$, then it holds for b, or that P is* inductive*. Then P holds for all elements of A.*

Proof. Otherwise, the nonempty set $\{x \in A \mid P(x) = F\}$ has a minimum b. But P is inductive, so P holds for b, a contradiction. □

Why is it that we don't seem to need a base case? Well, that's sort of inherent in our definition of inductive. If $b = \min(A)$, then vacuously for all $a \prec b$, P holds; there is no such element a.

Interlude: A brief ode to induction. Sung to the tune of "Do You Hear the People Sing," from "Les Miserables." Songwriting creds to Luke Sciarappa and Susan Durst.

You can prove a theorem's true,
Prove it for arbitrary n
If you can show that every instance means that it occurs again!
First you prove it for the base,
now assume n, and you'll be done
if you can show that the result holds for $n + 1$.
Will you rise among the ranks of the mathematical elite?
Definitions by recursion can be magical and neat!
You'll find that transfinite induction is totally sweeeeet!

Lastly, we introduce the very important Axiom of Choice.

Axiom of Choice. Let X be a set and $\{A_\alpha\}_{\alpha\in X}$ a family of nonempty sets. Then

$$\Pi_{\alpha\in X} A_\alpha \neq \varnothing.$$

In other words, there exists $f : X \to \bigcup_{\alpha\in X} A_\alpha$ such that $f(\alpha) \in A_\alpha$ for all α. Such an f is called a *choice function.*

Now we dive in to the proof of Zorn's Lemma.

Proof. Suppose X has no maximal element. Choose for each $x \in X$ an element x^+ such that $x \prec x^+$. This is possible by the assumption that X has no maximal element and by the Axiom of Choice. Now for each chain C in X, we Choose an upper bound $u(C) \in X$, again possible by the hypothesis and by the Axiom of Choice.

Let A be a well-ordered set. We define a sequence $\{x_a\}_{a\in A}$ in X such that if $a < b$ in A, then $x_a \prec x_b$ in X. We do this by transfinite induction. Suppose we have defined x_a for all $a < b$. We have an increasing sequence $\{x_a\}_{a<b}$ in X. This is a chain in X, so we can define $x_b = u(\{x_a\}_{a<b})^+$. Then $x_b > u(\{x_a\}_{a<b}) \geq x_a$ for all $a < b$. By transfinite induction, we have $\{x_a\}_{a\in A}$ strictly increasing in X. This contradicts Hartog's Lemma, a result from set theory on high to which we will appeal.

Lemma 13.1.12 (Hartog's Lemma)**.** *Given a set X, there exists a well-ordered set A such that there is no injection $A \hookrightarrow X$.*

This lemma isn't too hard to prove with the Axiom of Choice, or with the well-ordering principle, but (fun fact!) it's independent of the Axiom of Choice. Our sequence is in particular an injective map $A \hookrightarrow X$, so we can't do this for every well-ordered A. We have a contradiction, so X must have a maximal element. □

14 October 15th

14.1 RINGS

Definition 14.1.1. A *ring* $(R, +, 0, \cdot, 1)$ consists of an abelian group $(R, +, 0)$ and a monoid $(R, \cdot, 1)$ such that for all $a, b, c \in R$,

- $a \cdot (b + c) = a \cdot b + a \cdot c$ and
- $(a + b) \cdot c = a \cdot c + b \cdot c$,

referred to as *distributivity.*

Examples.

- $\mathbb{Z}$ and $\mathbb{Z}_n$
- $R[x]$, or polynomials with coefficients in a ring R.
- $M_n(R)$, or $n \times n$ matrices with entries in a ring R.

- R^X, or functions from a set X to a ring R with pointwise operations (see problem 15 on HW 7).

Remark. This will be mainly notation.

1. The inverse of an element a in $(R, +, 0)$ is denoted $-a$ and called the *opposite* of a.

   The inverse of a in $(R, \cdot, 1)$ is denoted a^{-1}. The set of invertible elements in $(R, \cdot, 1)$ is denoted $R^\times$; $(R^\times, \cdot, 1)$ is then a group.

2. For any $a \in R$, $a \cdot 0 = 0 = 0 \cdot a$, a property known as *absorption.* Sometimes a *near-ring* is defined, where instead of having an abelian group with addition, we just take a monoid; in this case absorption must be an axiom. In fact, in the standard definition of a ring, one can prove that the ring under addition must be abelian. The proof of absorption is one line and is left as an exercise.

3. Can $0 = 1$ in a ring R? Suppose $0 = 1$. Take $a \in R$. Then $a = a \cdot 1 = a \cdot 0 = 0$, so every element in R is trivial. So there's exactly one ring R with $0 = 1$; namely, $R = \{0\} = \{1\}$, the trivial ring. I have heard this referred to not as the trivial ring but as the "stupid" ring, in an only-slightly-less-formal setting.

4. An element $z \in R$ is a *zero-divisor* if there exists $w \in R$, $w \neq 0$, with $zw = 0$ or $wz = 0$. Note that while sometimes 0 is not defined as a zero-divisor, this definition includes 0 as a zero-divisor except in the stupid ring. We denote by R^z the set of zero-divisors in R.

5. Assume $R \neq \{0\}$. Then $0 \in R^z$; also, $1 \in R^\times$. We think of zero-divisors as "like zero" and invertible elements as "like one." In fact, $R^\times \cap R^z = \varnothing$.

Examples.

1. $R = \mathbb{Z}$. $R^z = \{0\}$ and $R^\times = \{\pm 1\}$.

2. $R = M_n(\mathbb{F})$, for $\mathbb{F}$ a field. $R^\times = \mathrm{GL}(n, \mathbb{F})$. $R^z = M_n(\mathbb{F}) \backslash \mathrm{GL}(n, \mathbb{F})$. Take $A \notin \mathrm{GL}(n, \mathbb{F})$, so there exists $x \in \ker A$, $x \in \mathbb{F}^n$, $x \neq 0$. Let B be n copies of x; then $AB = 0$.

Definition 14.1.2. A nontrivial ring R is a *domain* if $R^z = \{0\}$. It is a *division ring* if $R^\times = R \setminus \{0\}$. Division rings are also called *skew fields.* It is an *integral domain* if it is a commutative domain. Finally, a commutative division ring is a *field.*

So we have a schematic

$$\begin{array}{ccccc} \textit{division rings} & \subset & \textit{domains} & \subset & \textit{rings} \\ \cup & & \cup & & \cup \\ \textit{fields} & \subset & \textit{integral domains} & \subset & \textit{commutative rings} \end{array}$$

Definition 14.1.3. Let R be a ring. A subset $S \subseteq R$ is a *subring* if it is a subgroup of $(R, +, 0)$ and a *submonoid* of $(R, \cdot, 1)$. In this case, S is a ring.

Definition 14.1.4. Let R_1, R_2 be rings. A function $\varphi : R_1 \to R_2$ is a *morphism of rings* if it is a morphism of groups and monoids. Explicitly,

$$\begin{aligned} \varphi(a+b) &= \varphi(a) + \varphi(b) \\ \varphi(ab) &= \varphi(a)\varphi(b) \\ \varphi(1) &= 1. \end{aligned}$$

As a consequence, $\varphi(0) = 0$, $\varphi(-a) = -\varphi(a)$, and $\varphi(a^{-1}) = \varphi(a)^{-1}$, when such an element exists.

Definition 14.1.5. Let R be a ring and $I \subseteq R$ a subgroup of $(R, +, 0)$. I is a *left ideal* if for any $a \in R$, $x \in I$, we have $ax \in I$. *Right ideals* and *two-sided ideals* are defined similarly.

Let I be a two-sided ideal of R. Write $a \equiv b \pmod{I}$ if $a - b \in I$. This is an equivalence relation, and it is compatible with both $+$ and $\cdot$.

Why is the above equivalence relation compatible with $+$? Well, I is a subgroup of an abelian group, so it is a normal subgroup of $(R, +, 0)$. Why's it compatible with $\cdot$? By distributivity. This is a tiny exercise. It's worth checking.

Hence, R/I, the set of equivalence classes, is a ring, with $\overline{a} + \overline{b} = \overline{a+b}$, and $\overline{a} \cdot \overline{b} = \overline{a \cdot b}$. This is the *quotient* of R by I.

Example. $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$, where $n\mathbb{Z}$ is notably an ideal of $\mathbb{Z}$.

Now for the ring isomorphism theorems. Proofs are omitted given that they're all easy and all really similar to the group proofs.

Proposition 14.1.6 (1st isomorphism theorem). *Let $\varphi : R \to R'$ be a morphism of rings. Then*

1. $\ker \varphi = \{x \in R : \varphi(x) = 0\}$ *is a two-sided ideal of R.*

2. $\operatorname{im} \varphi$ *is a subring of R'.*

3. $R/\ker \varphi \cong \operatorname{im} \varphi$*, with* $\overline{a} \mapsto \varphi(a)$.

Example. Let $\varphi : \mathbb{R}[x] \to \mathbb{C}$ be $\varphi(p(x)) = p(i)$. Then φ is a surjective homomorphism of rings with kernel $\ker \varphi$ generated by $x^2 + 1$, so $\mathbb{R}[x]/\langle x^2 + 1 \rangle \cong \mathbb{C}$.

Proposition 14.1.7 (The other isomorphism theorems). *Let R be a ring.*

1. *Let S be a subring and I a two-sided ideal of R. Then*

   (a) *$S + I$ is a subring of R.*

   (b) *$S \cap I$ is a two-sided ideal.*

   (c) *$(S + I)/I \cong S/(S \cap I)$.*

2. *Let I and J be two-sided ideals of R, with $I \subseteq J$. Then*

   (a) *J/I is a two-sided ideal of R/I.*

   (b) $\frac{R/I}{J/I} \cong R/J$.

3. *Let I Be a fixed two-sided ideal of R. Then there is a bijective correspondence between the set of two-sided ideals J of R containing I and the set of two-sided ideals of the quotient R/I, with $J \mapsto J/I$.*

There are variants for left and right ideals or subrings.

Definition 14.1.8. For two ideals I, J, we have

$$I + J = \{a + b \mid a \in I, b \in J\}$$
$$I \cdot J = \left\{ \sum_{i=1}^{n} a_i b_i \mid a_i \in I, b_i \in J, i = 1, \ldots, n, n \geq 0 \right\}.$$

Note that $IJ, JI \subseteq I \cap J$.

Definition 14.1.9. Let R be a nontrivial ring. A proper ideal I of R (left, right, two-sided...) is *maximal* if it is a maximal element of the poset of proper ideals of R under inclusion.

Proposition 14.1.10. *A nontrivial ring has at least one maximal ideal (left, right, two-sided...). More generally, any proper ideal is contained in a maximal one.*

Proof. The first statement is an application of Zorn's Lemma. A similar proof gives a second statement, using the bijective correspondence from that one isomorphism law. □

Proposition 14.1.11. *Let R be nontrivial and assume that R is commutative (multiplicatively; for addition we don't need to specify, so we never do). The following are equivalent:*

1. *R is a field*

2. *$\{0\}$ is the only proper ideal*

3. *$\{0\}$ is a maximal ideal*

Proof. Left as an exercise. It's very simple, but it's important to be able to do it eyes closed. □

Corollary 14.1.12. *Let I be an ideal of a commutative ring R. I is maximal if and only if the quotient R/I is a field.*

Proposition 14.1.13. *Let R be nontrivial. Then TFAE:*

1. *R is a division ring*

2. *0 is the only proper left ideal*

3. *0 is the only proper right ideal*

Definition 14.1.14. A ring is *simple* if it is nontrivial and 0 is the only proper two-sided ideal.

Remark. Division rings are simple, but the converse does not necessarily hold. For example, $M_n(\mathbb{F})$ is simple but is not a division ring. See homework 8, problem 15. If $n > 1$, this is not a division ring.

Definition 14.1.15. Let I and J be two two-sided ideals of a ring R. They are called *comaximal* if $I + J = R$. Equivalently, no proper ideal contains both I and J.

Theorem 14.1.16 (Chinese Remainder Theorem)**.** *Let I and J be comaximal two-sided ideals of a ring R. Let $\varphi : R \to R/I \times R/J$ be $\varphi(a) = (\overline{a}, \overline{a})$.*

1. *φ is surjective and* $\ker \varphi = I \cap J$.
2. *$R/(I \cap J) \cong R/I \times R/J$, an isomorphism of rings.*
3. *$I \cap J = IJ + JI$.*

Proof.

1. Given $a, b \in R$, we need $x \in R$ with $\overline{x} = \overline{a}$ in R/I and $\overline{x} = \overline{b}$ in R/J. In other words, we need $x \equiv a \pmod{I}$ and $x \equiv b \pmod{J}$. $I + J = R$, so $1 = e + f$ for some $e \in I$, $f \in J$. Thus $e \equiv 0 \pmod{I}$ and $e \equiv 1 \pmod{J}$; meanwhile $f \equiv 1 \pmod{I}$ and $f \equiv 0 \pmod{J}$. Let $x = be + af$; this x works. Thus the mapping is surjective. The kernel being the intersection is immediate from the definition.
2. A consequence of 1, and the first isomorphism law.
3. Needs a sentence; think about it on your own, or maybe it will be covered next time.

□

15 October 20th

15.1 And God liked the integers, so He put a ring on them: Rings Part 2

For future reference, if we say "ideals" without specifying left or right, we mean two-sided.

Theorem 15.1.1 (Chinese Remainder Theorem)**.** *Let I, j be comaximal ideals of a ring R, and consider $\varphi : R \to R/I \times R/J$ with $a \mapsto (\overline{a}, \overline{a})$. Then*

1. *φ is surjective and* $\ker \varphi = I \cap J$.
2. *$R/(I \cap J) \cong R/I \times R/J$*
3. *$I \cap J = JI + IJ$*

Proof. Last time, we did all but number 3, so we'll do that now. $1 = e + f$, for $e \in I, f \in J$. If $x \in I \cap J$, then $x = xe + xf$. $xe \in JI$ and $xf \in IJ$, so $x \in JI + IJ$. Now, $JI + IJ \subseteq I \cap J$ holds always, regardless of comaximality, because $IJ \subseteq I$, and $IJ \subseteq J$, and $JI \subseteq I$, etc., etc. □

Examples. 1. $R = \mathbb{Z}$. We know $n\mathbb{Z}$ is an ideal, and every ideal is of this form for a unique $n \in \mathbb{N}$. Also, $n\mathbb{Z} \subseteq m\mathbb{Z}$ if and only if $m|n$. Thus, the poset of ideals (under inclusion, as always) of $\mathbb{Z}$ is *anti-isomorphic* to the poset of $\mathbb{Z}$ under divisibility. So the Chinese Remainder Theorem is as follows. Let $I = n\mathbb{Z}$ and let $J = m\mathbb{Z}$. The following picture applies.

$$\begin{tikzcd}
 & I+J \arrow[dl, dash] \arrow[dr, dash] & & & & \gcd(n,m) \arrow[dl, dash] \arrow[dr, dash] & \\
I \arrow[dr, dash] & & J \arrow[dl, dash] & \longleftrightarrow & n \arrow[dr, dash] & & m \arrow[dl, dash] \\
 & I \cap J \arrow[d, dash] & & & & \operatorname{lcm}(n,m) \arrow[d, dash] & \\
 & IJ & & & & nm &
\end{tikzcd}$$

Explicitly, if $\gcd(n, m) = 1$, then $x \equiv a \pmod{n}$ and $x \equiv b \pmod{m}$ has a unique solution modulo nm.

2. $R = \mathbb{F}^X$, with X a finite set and $\mathbb{F}$ a field. We know from the homework that every ideal of $\mathbb{F}^X$ is of the form $\mathcal{I}(S)$ for a unique $S \subseteq X$, where $\mathcal{I}(S) = \{f \in \mathbb{F}^X : f|_S = 0\}$. Also, $\mathbb{F}^X/\mathcal{I}(S) \cong \mathbb{F}^S$. Moreover, the poset of ideals of $\mathbb{F}^X$ is anti-isomorphic to the poset of subsets of X. Let $I = \mathcal{I}(S)$, $J = \mathcal{I}(T)$. The picture below applies.

$$\begin{tikzcd}
 & I+J \arrow[dl, dash] \arrow[dr, dash] & & & & S \cap T \arrow[dl, dash] \arrow[dr, dash] & \\
I \arrow[dr, dash] & & J \arrow[dl, dash] & \longleftrightarrow & S \arrow[dr, dash] & & T \arrow[dl, dash] \\
 & I \cap J \arrow[d, equal] & & & & S \cup T \arrow[d, equal] & \\
 & IJ & & & & S \cup T &
\end{tikzcd}$$

The Chinese Remainder Theorem tells us that if $S \cap T = \varnothing$, then $\mathbb{F}^{S \cup T} \cong \mathbb{F}^S \times \mathbb{F}^T$. Explicitly, if $S \cap T = \varnothing$, then a function of $S \cup T$ is the same as two functions, one on S and the other on T.

15.2 Noetherian Rings

Remark. I've heard it said that Emmy Noether was quite the BAMPh – Brilliant Algebraist, Mathematician, and Physicist. Citation to Ruthi Hortsch for this cool acronym.

Definition 15.2.1. A poset satisfies the *ascending chain condition* or ACC if every countable ascending chain $x_0 \leq x_1 \leq x_2 \leq \cdots$ stabilizes, i.e. if $\exists n \geq 0$ with $x_N = x_{N+1} = \cdots$.

Proposition 15.2.2. *A poset X satisfies the ascending chain condition if and only if every nonempty subset $S \subseteq X$ has a maximal element (by which we mean an element $m \in S$ such that if $m < x$, then $x \notin S$).*

Proof. First suppose that X satisfies the ACC and that $\exists S \neq \varnothing$ without a maximal element. Choose $x_0 \in S$; by assumption there exists $x_1 \in S$, $x_0 < x_1$. Inductively, we define a strictly increasing chain $x_0 < x_1 < x_2 < \cdots$, which violates the hypothesis. Note that for this proof we need at least the axiom of countable choice.

Now we show the converse. Given an ascending chain $x_0 \leq x_1 \leq x_2 \leq \cdots$, let $S = \{x_i : i \in \mathbb{N}\}$. S has a maximal element x_N by hypothesis, so $x_M = x_N$ for all $M \geq N$. Thus we have the ascending chain condition. □

Definition 15.2.3. A ring is *left-noetherian* if the poset of left ideals satisfies the ACC. *Right-noetherian* is defined in exactly the same way.

Fun fact! Rings can be left-noetherian but not right-noetherian. These definitions are far too zany for that drivel.

Definition 15.2.4. Given a subset A of a ring R, the *left ideal generated by A* is the set

$$RA = \left\{ \sum_{i \in F} r_i a_i \mid F \text{ is any finite set}, r_i \in R, a_i \in A, \forall i \in F \right\}.$$

It is the smallest left ideal containing A. A left ideal I of R is *finitely generated* if there exists a finite $A \subseteq R$ such that $I = RA$. It is *principal* if $\exists a \in R$ such that $I = R\{a\}$.

Proposition 15.2.5. *Let R be left noetherian, and let I be a two-sided ideal. Then R/I is left noetherian.*

Proof. A chain of left ideals in R/I is $\overline{I_0} \subseteq \overline{I_1} \subseteq \cdots$, where $I_0 \subseteq I_1 \subseteq I_2 \subseteq \cdots$ is a chain of left ideals in R containing I. R is noetherian so this stabilizes, and thus the given chain in R/I stabilizes as well. □

Proposition 15.2.6. *A ring is left noetherian if and only if every left ideal is finitely generated.*

Proof. $\underline{(\Rightarrow)}$: Let I be a left ideal of R. Let $\mathcal{F} = \{RA \mid A \text{ a finite subset of } I\}$. $\mathcal{F} \neq \varnothing$, because $\overline{\{0\}} \in \mathcal{F}$, so $\mathcal{F}$ has a maximal element RA. We claim that $RA = I$. If not, $\exists x \in I$,

$x \notin RA$. Let $A' = A \cup \{x\}$. A' is a finite subset of I, and $RA \subsetneq RA'$, because one but not the other contains x. This is a contradiction.

$\underline{(\Leftarrow)}$: Let $I_0 \subseteq I_1 \subseteq \cdots$ be a chain of left ideals. Then let $I = \bigcup_{n\geq 0} I_n$, which is also a left ideal. I is an ideal, so it is finitely generated by some generating set $A = \{a_1, \ldots, a_k\}$. Then $a_i \in I_{n_i}$ for some n_i; let $N = \max(n_i)$. All generators are in I_N, so $I \subseteq I_N$, and thus $I_M \subseteq I_N$ for all $M \geq N$, and we have stabilization. □

Examples. 1. Any division ring is (left and right) noetherian. The only ideals are $\{0\}$ and R.

2. Any PID (principal ideal domain, an integral domain for which every ideal is principal) is noetherian.

   The next two are non-examples:

3. $R[x_1, x_2, \ldots]$, the ring of polynomials in countably many variables is <u>not</u> an example. The chain $Rx_1 \subset R\{x_1, x_2\} \subset R\{x_1, x_2, x_3\} \subset \cdots$ does not stabilize.

4. Let X be an infinite set. R^X is not (left) noetherian. Let $X_0 \supset X_1 \supset X_2 \supset \cdots$ be an infinite strictly decreasing chain of subsets of X. Then $\mathcal{I}(X_0) \subset \mathcal{I}(X_1) \subset \cdots$ is the chain of ideals that does not stabilize.

Depending on which mathematical road you travel by, these concepts will be either really really relevant all the time, or not.

Theorem 15.2.7 (Hilbert's Basis Theorem)**.** *Let R be a left noetherian ring. Then $R[x]$ is also left noetherian.*

Proof. Let I be a left ideal of $R[x]$. We show it is finitely generated. Suppose not. In particular, $I \neq \{0\}$, so we can choose a polynomial $f_0 \in I$ of minimal degree. Now $I \neq R[x]\{f_0\}$, so we can choose $f_1 \in I \setminus R[x]\{f_0\}$, again of minimal degree. f_0 was of minimal dgree of all polynomials in I, and $f_1 \in I$, so $\deg(f_0) \leq \deg(f_1)$. Iterate the process to get a bunch of f_i's. Given $f_0, f_1, \ldots, f_{n-1}$, $I \neq R[x]\{f_0, f_1, \ldots, f_{n-1}\}$, so we can choose $f_n \in I \setminus R[x]\{f_0, f_1, \ldots, f_{n-1}\}$ and of minimal degree. By induction, we get f_n for all $n \in \mathbb{N}$. Let $d_n = \deg(f_n)$. Note that $d_n \leq d_{n+1}$ for all n. Write $f_n = a_n x^{d_n} +$ lower terms, with $a_n \in R$ nonzero. We'll use the notation $\mathrm{LT}(f_n) = a_n x^{d_n}$, where LT stands for "leading term".

Consider the chain of left ideals in R, $R\{a_0\} \subseteq R\{a_0, a_1\} \subseteq R\{a_0, a_1, a_2\} \subseteq \cdots$. R is left noetherian, so it stabilizes, say at $N \in \mathbb{N}$. Thus $a_n \in R\{a_0, a_1, \ldots, a_N\}$ for all $n \geq N$. In particular, $a_{N+1} = \sum_{i=0}^{N} r_i a_i$. Consider the polynomial

$$g = \sum_{i=0}^{N} r_i f_i x^{d_{N+1} - d_i},$$

which is well-defined as a polynomial because the d_i's are increasing.

Now

$$\begin{aligned}
\mathrm{LT}(g) &= \sum_{i=0}^{N} r_i \mathrm{LT}(f_i) x^{d_{N+1}-d_i} \\
&= \sum_{i=0}^{N} r_i a_i x^{d_{N+1}} \\
&= a_{N+1} x^{d_{N+1}} \\
&= \mathrm{LT}(f_{N+1}).
\end{aligned}$$

So $\deg(f_{N+1}-g) < \deg(f_{N+1})$. But $g \in R[x]\{f_0, \dots, f_N\}$, and $f_{N+1} \in I \backslash R[x]\{f_0, \dots, f_N\}$, so $f_{N+1} - g \in I \setminus R[x]\{f_0, \dots, f_N\}$, which contradicts the minimality of f_{N+1}. □

16 October 22nd

16.1 Modules

Definition 16.1.1. Let R be a ring. A *left R-module* consists of an abelian group $(M, +, 0)$ with $R \times M \to M$, $(a, m) \mapsto a \cdot m$, such that for all $a, b \in R$ and for all $m \in M$,

(i) $a \cdot (b \cdot m) = (ab) \cdot m, 1 \cdot m = m$

(ii) $a \cdot (m + n) = a \cdot m + a \cdot n$

(iii) $(a + b) \cdot m = a \cdot m + b \cdot m$

Remark. 1. If R is a field, then an R-module is a vector space over R.

2. If we replace $m \in M$ in the definition by $c \in R$, we obtain the ring axioms. In particular, $M = R$ is a left R-module.

3. (i) holds if and only if the monoid $(R, \cdot, 1)$ acts on the set M. (ii) holds if and only if the action is by endomorphisms of $(M, +, 0)$. (ii) and (iii) hold if and only if the map $R \times M \to M$ is *biadditive*, preserving addition in both variables.

As seen in homework 7. If $(M, +, 0)$ is an abelian group, then $\mathrm{End}_{\mathbb{Z}}(M) = \{f : M \to M \mid f \text{ is a group homomorphism}\}$ is a ring under $(f + g)(m) = f(m) + g(m)$, and $(f \circ g)(m) = f(g(m))$. A map $R \times M \xrightarrow{\cdot} M$ gives rise to a map $R \xrightarrow{l} M^M = \{f : M \to M \mid f \text{ is a function}\}$. Under this, $a \mapsto l_a$, where $l_a(m) = a \cdot m$. (ii) holds if and only if $l_a \in \mathrm{End}_{\mathbb{Z}}(M)$, and (i) and (iii) hold if and only if $l : R \to \mathrm{End}_{\mathbb{Z}}(M)$ is a morphism of rings. To conclude, given an abelian group M, a left R-module structure on it is equivalent to a ring homomorphism $R \to \mathrm{End}_{\mathbb{Z}}(M)$.

Proposition 16.1.2. *1. Let R be a ring. There exists a unique ring homomorphism $\mathbb{Z} \to R$ (we say as a result that $\mathbb{Z}$ is the* initial ring*, the one ring to bring them all, and in the darkness bind them).*

2. *Let M be an abelian group. $\exists!$ (left) $\mathbb{Z}$-module structure on M. (Hence $\mathbb{Z}$-modules $\equiv$ abelian groups).*

Proof. 1. Define $\varphi : \mathbb{Z} \to R$ by

$$\varphi(n) = \begin{cases} 0 \text{ if n} = 0, \\ 1 + \cdots + 1 \ n \text{ times, if } n \geq 1, \\ -\varphi(-n) \text{ if } n \leq -1. \end{cases}$$

2. By 1, $\exists!$ ring homomorphism $\mathbb{Z} \to \mathrm{End}_{\mathbb{Z}}(M)$.

□

Definition 16.1.3. Let M be a left R-module. A subset $N \subseteq M$ is a *submodule* if it is a subgroup of $(M, +, 0)$ and $a \cdot n \in N$ for all $a \in R, n \in N$. In this case, N is a left R-module $(N \leq M)$. Also, M/N is a left R-module with $a \cdot \overline{m} = \overline{a \cdot m}$.

Example. 1. Let $M = R$ be the canonical left R-module. Then the submodules of M are the left ideals of R.

2. Let $R = \mathbb{Z}$ and M an abelian group. Then the submodules of M are precisely the subgroups of M.

Definition 16.1.4. Let M be a left R-module and $A \subseteq M$ a subset. The R-submodule *generated* by A is

$$RA = \{\sum_{i \in F} r_i a_i \mid F \text{ finite}, r_i \in R, a_i \in A\}.$$

A module M is *finitely generated* if there exists a finite $A \subseteq M$ such that $M = RA$. It is *cyclic* if $\exists a \in M$, $M = R\{a\}$.

Definition 16.1.5. Let M, N be left R-modules. A *homomorphism* is a function $f : M \to N$ such that $f(m_1 + m_2) = f(m_1) + f(m_2)$ and $f(a \cdot m) = a \cdot f(m)$.

This leads us to products and sums, a new subsection in our lives.

16.2 Products and sums

Definition 16.2.1. Let I be a set and $\{M_i\}_{i \in I}$ a collection of left R-modules. On the cartesian product $\prod_{i \in I} M_i$, define operations

$$(m_i)_{i \in I} + (m_i')_{i \in I} := (m_i + m_i')_{i \in I},$$

and

$$a \cdot (m_i)_{i \in I} = (a \cdot m_i)_{i \in I}.$$

Then $\prod_{i \in I} M_i$ is a left R-module, called the *direct product* of $\{M_i\}_{i \in I}$.

Definition 16.2.2. Let $\bigoplus_{i\in I} M_i$ be the subset of $\prod_{i\in I} M_i$ consisting of sequences $(m_i)_{i\in I}$ with *finite support*, i.e. with $\{i \in I \mid m_i \neq 0\}$ finite. Then $\bigoplus_{i\in I} M_i \leq \prod_{i\in I} M_i$. The left R-module $\bigoplus_{i\in I} M_i$ is called the *direct sum* of the $\{M_i\}_{i\in I}$. If I is finite, then clearly the direct sum and direct product are the same.

For each $j \in I$, there are morphisms of R-modules $\pi_j : \prod M_i \to M_j$, with $(m_i)_{i\in I} \mapsto m_j$, and morphisms $\sigma_j : M_j \to \bigoplus M_i$, with $m \mapsto (m_i)_{i\in I}$, where $m_i = m$ if $i = j$ and 0 otherwise.

Proposition 16.2.3. *Let $\{M_i\}_{i\in I}$ be a collection as before. Let M be another R-module.*

1. *For each $j \in I$, let $\varphi_j : M \to M_j$ be a morphism of R-modules. Then there exists a unique morphism of R-modules $\varphi : M \to \prod_{i\in I} M_i$ such that $\pi_j \circ \varphi = \varphi_j$.*

$$\begin{tikzcd}
M \arrow[r, dashed, "\varphi"] \arrow[dr, "\varphi_j"] & \prod M_i \arrow[d, "\pi_j"] \\
 & M_j
\end{tikzcd}$$

2. *For each $j \in I$, let $\psi_j : M_j \to M$ be a morphism of R-modules. Then there exists a unique morphism of R-modules $\psi : \bigoplus M_i \to M$ such that all diagrams such as below commute.*

$$\begin{tikzcd}
\bigoplus M_i \arrow[r, dashed, "\psi"] & M \\
M_j \arrow[u, "\sigma_j"] \arrow[ur, "\psi_j"] &
\end{tikzcd}$$

Proof.

1. Define $\varphi(m) = (\varphi_i(m))_{i\in I}$. This works.
2. Define $\psi((m_i)_{i\in I}) = \sum_{i\in I} \psi_i(m_i)$. Note that this is finite because the I-tuple has finite support.

□

Remark. The following is a general category-theoretic organization of thoughts.

General Categories	Category of R-modules	Category of Groups	Category of Sets
product	direct product	direct product	cartesian product
coproduct	direct sum	free product	disjoint union

Moreover the universal property tells us for products that a map to the product is a bunch of maps to each factor, and for the coproduct that a map from the coproduct is a bunch of maps from each factor.

Note also for modules that $M \oplus N \cong M \times N$, but for groups $G * H \ncong G \times H$, and for sets $X \sqcup Y \ncong X \times Y$.

Proposition 16.2.4. *Let M be a left R-module and $\{M_i\}_{i\in I}$ a collection of submodules of M. Suppose*

1. *The $\{M_i\}$ are* independent*: given $(m_i)_{i\in I}$ of finite support with $m_i \in M_i$, if $\sum_{i\in I} m_i = 0$, then $m_i = 0$ for all i.*
2. *The $\{M_i\}$* generate *M: given $m \in M$, $\exists (m_i)_{i\in I}$ of finite support with $m_i \in M_i$ such that $m = \sum_{i\in I} m_i$.*

Then $M \cong \bigoplus_{i\in I} M_i$.

Proof. Let $\psi_i : M_i \hookrightarrow M$ be the inclusion . The UMP gives a morphism of R-modules

$$\begin{tikzcd}
\bigoplus M_i \arrow[r, dashed, "\psi"] & M \\
M_j \arrow[u, "\sigma_j"] \arrow[ur, "\psi_j"] &
\end{tikzcd}$$

But $(m_i)_{i\in I} = \sum \sigma_i(m_i)$, so $\psi((m_i)_{i\in I}) = \sum_{i\in I} \psi(\sigma_i(m_i)) = \sum_{i\in I} m_i$. (1) and (2) say precisely that ψ is injective and surjective, and thus an isomorphism. □

Conversely, for arbitrary modules M_i, you can find independent generating sets (basically) inside the direct sum.

17 October 27th

17.1 Free modules

Definition 17.1.1. Let I be a set. For each $i \in I$, let $M_i = R$ be the canonical left R-module. The *free R-module on I* is by definition

$$R^{(I)} = \bigoplus_{i\in I} M_i.$$

For each $j \in I$, let $e_j \in R^{(I)}$ be the tuple $e_j = (a_i)_{i\in I}$ with $a_i = 1$ if $i = j$, and 0 if not.

Proposition 17.1.2. *Let I be a set and M a left R-module. Given a collection $\{m_i\}_{i\in I}$ with $m_i \in M$ for all i, there exists a unique morphism of R-modules $\varphi : R^{(I)} \to M$ such that $\varphi(e_j) = m_j$ for all $j \in I$.*

$$\begin{tikzcd}
I \arrow[r] \arrow[dr] & R^{(I)} \arrow[d, dashed, "\varphi"] \\
 & M
\end{tikzcd}$$

Proof. Let $\varphi_j : M_j \to M$ be $\varphi_j(a) = a \cdot m_j$. This is a morphism of R-modules; apply the universal property of the direct sum to obtain φ. Then $\varphi(e_j) = \varphi\sigma_j(1) = \varphi_j(1) = 1{\cdot}m_j = m_j$. Uniqueness is easy. □

Definition 17.1.3. Let M be a left R-module and $S \subseteq M$ a subset. For each $s \in S$, let $M_s = R\{s\}$. We say that S is *linearly independent* if $\{M_s\}_{s\in S}$ is independent. Moreover, S *generates* M if $\{M_s\}_{s\in S}$ generates M, and S is a *basis* if it is linearly independent and generates M.

Proposition 17.1.4. *Let I be a set.*

1. *The set $\{e_i\}_{i\in I}$ is a basis of the left R-module $R^{(I)}$*

2. *Suppose M is a left R-module with a basis I. Then $M \cong R^{(I)}$.*

Proof. Both statements reduce to earlier propositions, proven for direct sums. □

Note that the free R-module on I is strictly contained in the set $R^I = \prod_{i\in I} M_i$.

17.2 Noetherian Modules

Definition 17.2.1. A left R-module is *neotherian* if the poset of submodules satisfies the ascending chain condition. Equivalently, if every nonempty set of submodules has a maximal element.

Remark. R is left noetherian (as a ring) if and only if it is noetherian as a left R-module.

Proposition 17.2.2. *A left R-module M is noetherian if and only if every R-submodule of M is finitely generated.*

In particular, M is noetherian $\Rightarrow$ M is finitely generated.

Proof. Same as for ideals, so we omit. □

Proposition 17.2.3. *Let M be a left R-module and $N \leq M$. Then M is noetherian if and only if N and M/N are noetherian.*

Proof. $\underline{(\Rightarrow)}$: Submodules of N are submodules of M, so they must be finitely generated. Submodules of M/N are of the form L/N for $N \leq L \leq M$. L is also finitely generated, so basically everything under the sun must be finitely generated and nothing goes wrong. $\underline{(\Leftarrow)}$: Let $L \leq M$, so $L \cap N \leq N$, so $L \cap N = R\{x_1, \ldots, x_r\}$. Let $\pi : M \to M/N$. Then $\pi(L) \leq M/N$, so $\pi(L) = R\{\overline{y_1}, \ldots, \overline{y_s}\}$. It is an easy check that $L = R\{x_1, \ldots, x_r, y_1, \ldots, y_r\}$. □

Proposition 17.2.4. *Let $\{M_i\}_{i\in I}$ be noetherian left R-modules. Suppose I is finite. Then $\bigoplus_{i\in I} M_i$ is noetherian.*

Proof. By induction, we reduce to the case when $I = \{1, 2\}$. Let $M = M_1 \oplus M_2$. We have $M_1 \leq M$ and $M/M_1 \cong M_2$, so by the previous proposition the module M is noetherian, and we are done. □

Proposition 17.2.5. *Let R be left noetherian and M a left R-module. If M is finitely generated, then M is noetherian.*

Proof. Let $M = R\{x_1, \dots, x_k\}$. Then $R^k \to M$ with $e_j \mapsto x_j$ is a surjection. So M is a quotient of R^k and R^k is notehrian because it is a finite direct sum of noetherian modules. □

Remark. Let R be arbitrary and let M be a finitely generated left R-module. Is every submodule of M also finitely generated?

NO. Take any non-noetherian ring R and $M = R$.

17.3 Tensor Products of Modules

Or Kronecker products, Outer products, etc., etc., etc.

Let R be a commutative ring. Let M and N be R-modules. A homomorphism of R-modules $M \to N$ is also called a *linear map*. Their set is $\mathrm{Hom}_R(M, N)$. Let M, N, L be R-modules.

Definition 17.3.1. A function $\beta : M \times N \to L$ is *bilinear* if $\beta(m + m', n) = \beta(m, n) + \beta(m', n)$, and $\beta(m, n + n') = \beta(m, n) + \beta(m, n')$, and $\beta(am, n) = a\beta(m, n) = \beta(m, an)$, for all $m, m' \in M$, $n, n' \in N$, $a \in R$. We denote by $\mathrm{Hom}_R(M, N; L)$ the set of these bilinear maps.

We can similarly define *multilinear* maps $\mathrm{Hom}_R(M_1, \dots, M_n; L)$ as maps that are linear in each variable.

Modules and linear maps constitute a category, the study of which is linear algebra. Modules and multilinear maps constitute a multicategory, the study of which is sometimes called multilinear algebra. So that's fun.

Tensor products reduce bilinear maps to linear maps, i.e.

$$\mathrm{Hom}_R(M, N; L) \cong \mathrm{Hom}_R(M \otimes N, L).$$

But linear maps and bilinear maps really aren't the same thing, so tensor products need to do something. Let's construct it.

Proposition 17.3.2. *Given R-modules M, N, there exists an R-module X with a bilinear map $\theta : M \times N \to X$ such that given any other R-module L with a bilinear map $\beta : M \times N \to L$, there is a unique linear map $\hat{\beta} X \to L$ making the following diagram commute.*

$$\begin{tikzcd}
M \times N \arrow[r, "\beta"] \arrow[d, "\theta"'] & L \\
X \arrow[ur, dashed, "\hat{\beta}"'] &
\end{tikzcd}$$

Moreover, (X, θ) are unique up to isomorphism.

Proof. First we cover existence. (In the beginning we proved the existence of the heavens and the earth...)

Let F be the free R-module on the set $M \times N$. Then F has a basis $\{e_{(m,n)} \mid m \in M, n \in N\}$. Let F' be the submodule generated by all elements of the forms

- $e_{(m+m',n)} - e_{(m,n)} - e_{(m',n)}$,
- $e_{(m,n+n')} - e_{(m,n)} - e_{(m,n')}$,
- $e_{(am,n)} - ae_{(m,n)}$, and
- $e_{(m,an)} - ae_{(m,n)}$.

Let $X = F/F'$, and let $\theta : M \times N \hookrightarrow F \twoheadrightarrow F/F' = X$, with $(m,n) \mapsto e_{(m,n)} \mapsto \overline{e_{(m,n)}}$. We claim that θ is bilinear, and that (X, θ) satisfy the desired universal property. We won't write out the details, but you can extend any map β to a map $F \to L$, and then bilinearity tells you you can keep going to the quotient.

This reduces to the universal property for free modules plus the universal property for quotients.

□

Definition 17.3.3. X, as in the proof above, is called the *tensor product* of M and N, and is denoted $M \otimes N$. It's unique up to isomorphism, so this definition makes sense. When convenient, if there are like FIVE DISTINCT GOLDEN RIIIIINGS around getting us confused, we can write $M \otimes_R N$. We write $m \otimes_R N$ for the element $\theta(m,n) \in M \otimes N$.

Note that a general element of $M \otimes N$ is of the form $\sum_{i \in A} m_i \otimes n_i$, where A is a finite set and $m_i \in M$, $n_i \in N$. This is the same as $\sum_{i \in A} \overline{e_{(m_i,n_i)}}$.

Example. $\mathbb{Z}_m \otimes_{\mathbb{Z}} \mathbb{Z}_n \cong \mathbb{Z}_d$, with $d = \gcd(m,n)$. Define $\varphi : \mathbb{Z}_d \to \mathbb{Z}_m \otimes_{\mathbb{Z}} \mathbb{Z}_n$, with $\overline{a} \mapsto a \cdot \overline{1} \otimes \overline{1} = \overline{a} \otimes \overline{1} = \overline{1} \otimes \overline{a}$.

Is this well-defined? We need that $d \cdot \overline{1} \otimes \overline{1} = \overline{0} \otimes \overline{0}$. Well, $d = \gcd(m,n)$, so $d = rm + sm$ for $r, s \in \mathbb{Z}$. Then

$$\begin{aligned} d \cdot \overline{1} \otimes \overline{1} &= (rm + sn) \cdot \overline{1} \otimes \overline{1} \\ &= r \cdot \overline{m} \otimes \overline{1} + s \cdot \overline{1} \otimes \overline{n} \\ &= r \cdot \overline{0} \otimes \overline{1} + s \cdot \overline{1} \otimes \overline{0} \\ &= \overline{0} \otimes \overline{0} = 0. \end{aligned}$$

Now we define $\psi : \mathbb{Z}_m \otimes_{\mathbb{Z}} \mathbb{Z}_n \to \mathbb{Z}_d$. We want to say that $\overline{a} \otimes \overline{b} \mapsto \overline{ab}$, but that's not actually defining it on all elements of the tensor product. So we first define a map $\mathbb{Z}_m \times \mathbb{Z}_n \to \mathbb{Z}_d$, with $(\overline{a}, \overline{b}) \mapsto \overline{ab}$, and we check that it is well-defined and bilinear. This will show that a linear ψ exists by the universal property. This can be checked.

Then the two maps φ and ψ compose to the identity, which can also be checked. So yay! Isomorphism.

18 October 29th

18.1 Divisibility

We assume that R is an integral domain.

Definition 18.1.1. Let $a, b \in R$. We say that

- b *divides* a if $a = bc$ for some $c \in R$, and we write $b \mid a$.
- a and b are *associates* if $b \mid a$ and $a \mid b$, and we write $a \sim b$.

Let $u \in R$. We say u is a *unit* if $u \mid 1$.

Proposition 18.1.2. *1. u is a unit $\iff$ u is invertible.*

2. $b \mid a \iff (a) \subseteq (b)$, where (a) is the ideal generated by a and (b) is the ideal generated by (b).

3. $a \sim b \iff (a) = (b) \iff \exists$ a unit $u \in R, a = ub$.

Proof. Left as a simple exercise. □

Definition 18.1.3. Let $p \in R$. Assume that $p \neq 0$, and p is not a unit. Then

- p is *irreducible* if whenever $p = ab$, either a or b is a unit.
- p is *prime* if whenever $p \mid ab$, either $p \mid a$ or $p \mid b$.

Proposition 18.1.4. *Every prime element is irreducible.*

Proof. Let p be a prime. Suppose $p = ab$; then $p \mid ab$, so $p \mid a$ or $p \mid b$. Assume without loss of generality that $p \mid a$. But $a \mid p$, because $p = ab$, so $a \sim p$, so $p = ua$ for some unit $u \in R$. Then $p = ua = ba$, so $u = b$ because we're working in a domain. Thus if $p = ab$, either a or b is a unit, so p is irreducible. □

There are a lot of similar little arguments in this section. We're going to tackle some more interesting ones.

Proposition 18.1.5. *Suppose $p \sim q$.*

1. p prime $\Rightarrow$ q prime.

2. p irreducible $\Rightarrow$ q irreducible.

The proof, neither particularly difficult nor particularly interesting, is omitted.

Proposition 18.1.6 (Uniqueness of Prime Factorization)**.** *Let I and J be finite sets. Let $\{p_i\}_{i\in I}$ and $\{q_j\}_{j\in J}$ be primes. In general, including right now, when writing elements in this form we're not necessarily assuming that the p_i or q_j are distinct.*

Suppose

$$\prod_{i\in I} p_i \sim \prod_{j\in J} q_j.$$

Then there exists a bijection $\sigma : I \to J$ such that $p_i \sim q_{\sigma(i)}$.

Proof. By induction on $|I|$. If $|I| = 0$ we have $1 \sim \prod_{j\in J} q_j$. If we have that $|J| \neq 0$, then there exists $j \in J$. Then $q_j \mid 1$, so q_j is a unit, q contradiction. Thus $|J| = 0$, completing the base case.

Suppose $|I| \geq 1$. Then $\exists i_0 \in I$, so $p_{i_0} \mid \prod_{i\in I} p_i = u\prod_{j\in J} q_j$. p_{i_0} is prime, so $p_{i_0} \mid q_{j_0}$ for some $j_0 \in J$, since a prime can't divide a unit. Thus $q_{j_0} = p_{i_0}c$ for some $c \in R$. But q_{j_0} is irreducible and p_{i_0} is not a unit, so c is a unit and $p_{i_0} \sim q_{j_0}$.

We then have that

$$\prod_{i\in I} p_i = u\prod_{j\in J} q_j \Rightarrow \prod_{i\neq i_0} p_i = uc\prod_{j\neq j_0} q_j \sim \prod_{j\neq j_0} q_j.$$

By our inductive hypothesis there exists a bijection $\sigma : I\setminus\{i_0\} \to J\setminus\{j_0\}$ such that $p_i \sim q_{\sigma(i)}$ for all $i \in I \setminus \{i_0\}$. Extend σ by $\sigma(i_0) = j_0$, to get the desired bijection. □

Remark. 1. The above result is begging for the result that factorization into irreducibles is unique, but this is not necessarily the case.

2. Suppose $\{p_i\}$ are prime and $\{q_j\}$ are irreducible; then the above proof still works, so there is a bit of a partial result.

Proposition 18.1.7 (Existence of Irreducible Factorizations)**.** *Let R be noetherian. Let $a \in R$, $a \neq 0$, a not a unit. There is a finite set I and irreducibles $\{p_i\}_{i\in I}$ such that $a = \prod_{i\in I} p_i$.*

Proof. Suppose not. Let $\mathcal{F} = \{(x) \mid x \neq 0, x \neq \text{ unit}, x \neq \text{ finite product of irreducibles}\}$. Then $(a) \in \mathcal{F}$, so $\mathcal{F} \neq \varnothing$. Since R is noetherian, there exists a maximal element $(x) \in \mathcal{F}$. x is not a finite product of irreducibles, so x is not irreducible. Also $x \neq 0$ and is not a unit. Thus there exist $y, z \in R$ with $x = yz$, y, z not units. Since $x \neq 0$, $y, z \neq 0$. If both y and z are a product of irreducibles, so is x. Thus one of y and z is not a finite product of irreducibles; without loss of generality, say y. Then $(y) \in \mathcal{F}$. But $(x) \subseteq (y)$. By maximality, $(x) = (y)$. Thus $x = uy$ for a unit u, a contradiction because z was not a unit!

This completes the proof. ☺ □

Remark. This can also be done using *König's Lemma*, which says that an infinite tree with finite branching has to have an infinite branch. One looks at the tree of divisibility.

18.2 Unique Factorization Domains

Consider, as before, integral domains R.

Fact 18.2.1. *UFD's, while a fairly natural thing to look at, are surprisingly thought of as particularly alien. In fact, a misreading of the acronym UFD as "UFO" led to many science fiction tales regarding misconceptions about these strange objects.*

Proposition 18.2.2. *Consider the following statements about R.*

(1) Every $a \in R$, a nonzero and nonunit, admits a factorization into irreducibles.

(2) Any factorization into irreducibles is unique up to units and reordering.

(3) Every irreducible is prime.

Then (1) and (2) hold if and only if (1) and (3) hold.

Proof. $\underline{(\Leftarrow):}$ Uniqueness holds for prime factorizations; but (3) tells us that prime and irreducible are equivalent concepts. Thus (2) holds.

$\underline{(\Rightarrow):}$ Let $p \in R$ be irreducible. Suppose $p \mid ab$. We want either $p \mid a$ or $p \mid b$. If $a = 0$, done. If a is a unit, then $p \mid ab \sim b$, so $p \mid b$. We can then assume that both a and b are nonzero nonunits.

$p \mid ab$, so $ab = pc$. If $c = 0$, then $ab = 0$ so $a = 0$ or $b = 0$, a contradiction. If c is a unit, then $ab \sim p$, so either a is a unit or b is a unit, another contradiction. Thus c is another nonzero nonunit. By (1), we can factor. Let's say $a = \prod_{i \in I} p_i$, $b = \prod_{j \in J} q_j$, and $c = \prod_{h \in H} r_h$. But $ab = pc$, so

$$\prod_{i \in I} p_i \prod_{j \in J} q_j = p \cdot \prod_{h \in H} r_h.$$

By (2), $p \sim p_i$ or $p \sim q_j$. Thus $p \mid a$ or $p \mid b$, as desired. □

This leads to the definition of a unique factorization domain.

Definition 18.2.3. An integral domain is a *unique factorization domain*, or UFD, if (1) and (2) hold, as in the proposition. Equivalently, an integral domain is a UFD if (1) and (3) hold.

Thus in a UFD, every nonzero nonunit has a unique factorization into irreducibles/primes.

Remark.

- There are UFDs that are not noetherian (HW 11): $\mathbb{F}[x_1, x_2, x_3, \ldots]$.
- There are noetherian domains that are not UFDs (HW1 10): $\mathbb{Z}[\sqrt{-3}]$.

Proposition 18.2.4. *An integral domain is a UFD if and only if (3) and (4) hold, where (3) is as in Proposition 18.2.2 and (4) is that the principal ideals satisfy the ascending chain condition.*

Proof. Seen in the Homework. □

18.3 Principal Ideal Domains

Definition 18.3.1. An integral domain is a *principal ideal domain* or PID if every ideal is principal.

Proposition 18.3.2. *Let R be an integral domain and $p \in R$.*

1. *p is prime if and only if (p) is a prime ideal.*
2. *p is irreducible if and only if (p) is maximal among the proper principal ideals.*

Proofs are left as homework problems.

Corollary 18.3.3. *PID $\Rightarrow$ UFD.*

(Hahaha! Take that for brevity! You never knew it, but the soul of wit was in a particular statement of a well-known and somewhat rote theorem in commutative algebra all along!)

Proof. PID $\Rightarrow$ noetherian $\Rightarrow$ (4), from the previous section. Let p be irreducible; then (p) is maximal among principal ideals, which is all ideals. Thus (p) is maximal, so (p) is prime, and so p is prime, so (3) holds, and we have a UFD. □

18.4 Euclidean domains

Definition 18.4.1. Let R be an integral domain. A *Euclidean norm* on R is a function $\delta : R \setminus \{0\} \to \mathbb{N}$ such that for all $a \in R$ and for all $b \in R \setminus \{0\}$, there exists $q, r \in R$ with

(i) $a = bq + r$

(ii) if $r \neq 0$, then $\delta(r) < \delta(b)$.

Definition 18.4.2. A *Euclidean domain* is an integral domain with a Euclidean norm.

Proposition 18.4.3. *ED $\Rightarrow$ PID.*

Our friend brevity is here once again, especially in the proof of this proposition, which we omit.

Examples.

1. $\mathbb{Z}$ is a ED with $\delta(a) = |a|$.
2. If F is a field, then $F[x]$ is a ED with $\delta(a) = \deg(p(x)) \in \mathbb{N}$.
3. $\mathbb{Z}[i]$ is a ED with $\delta(z) = |z|^2 \in \mathbb{N}$.

19 November 3rd

19.1 Integrality

Definition 19.1.1. Let $R \subseteq S$ be integral domains and $\alpha \in S$. We say that α is *integral* over R if there is a monic polynomial $p(x) \in R[x]$ such that $p(\alpha) = 0$. Recall that a *monic* polynomial is one with leading coefficient 1.

Remark. If $\alpha \in R$, then α is integral over R; it is the root of $x - \alpha \in R[x]$, for example.

Example. $\sqrt{2} \in \mathbb{R}$ is integral over $\mathbb{Z}$, because it is the root of $x^2 - 2 \in \mathbb{Z}[x]$.

Definition 19.1.2. Let R be an integral domain and let F be its field of fractions. We say that R is *integrally closed* if the only elements of F that are integral over R are elements of R.

Example. $\mathbb{Z}$ is integrally closed. Indeed, let $\alpha \in Q$ be integral over $\mathbb{Z}$, and let $p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \in \mathbb{Z}[x]$ be such that $p(\alpha) = 0$. Write $\alpha = \frac{a}{b}$ with $a, b \in \mathbb{Z}$ and $\gcd(a, b) = 1$. Then multiplying out by b, we have

$$\begin{aligned} a^n + a_{n-1}a^{n-1}b + \cdots + a_1 a b^{n-1} + a_0 b^n &= 0 \\ &\Rightarrow a^n \equiv 0 \pmod{b}, \end{aligned}$$

so $b | a^n$. Since $\gcd(a, b) = 1$, we then know that $b = \pm 1$. What is the reasoning here? Well, if $b \neq \pm 1$, there exists a prime $p | b$. Thus $p | a^n \Rightarrow p | a \Rightarrow p | \gcd(a, b) = 1$, a contradiction.

Since $b = \pm 1$, $a \in \mathbb{Z}$.

As a corollary, $\sqrt{2} \notin \mathbb{Q}$, because $\sqrt{2} \notin \mathbb{Z}$.

Proposition 19.1.3. *Any UFD is integrally closed. The same argument works.*

Example. $\mathbb{Z}[\sqrt{-3}] = \{a + b\sqrt{-3} \mid a, b \in \mathbb{Z}\}$, and $\mathbb{Q}[\sqrt{-3}] = \{a + b\sqrt{-3} \mid a, b \in \mathbb{Q}\}$. $\mathbb{Q}[\sqrt{-3}]$ is the field of fractions of $\mathbb{Z}[\sqrt{-3}]$. Let $\omega = \frac{-1+\sqrt{-3}}{2}$. Then $\omega^2 + \omega + 1 = 0$, so ω is integral over $\mathbb{Z}[\sqrt{-3}]$. But $w \in \mathbb{Q}[\sqrt{-3}]/\mathbb{Z}[\sqrt{-3}]$, because $\frac{1}{2} \in \mathbb{Q}/\mathbb{Z}$. Thus $\mathbb{Z}[\sqrt{-3}]$ is not closed, so it is not a UFD.

In fact, $\mathbb{Z}[\sqrt{-3}]$ is a nice small example of an integral domain that is not a UFD.

19.2 Quadratic integers

Given $\alpha \in \mathbb{C}$, let $\mathbb{Q}(\alpha)$ be the smallest subfield of $\mathbb{C}$ containing α, and $\mathbb{Z}[\alpha]$ be the smallest subring of $\mathbb{C}$ containing α.

Definition 19.2.1. $\alpha \in \mathbb{C}$ is a *quadratic integer* if there exists a monic polynomial $p(x) \in \mathbb{Z}[x]$ of degree 2 such that $p(\alpha) = 0$: $\alpha^2 + m\alpha + n = 0$, for some $m, n \in \mathbb{Z}$.

If α is a quadratic integer, then $\mathbb{Q}(\alpha) = \{a+b\alpha \mid a, b \in \mathbb{Q}\}$ and $\mathbb{Z}[\alpha] = \{a+b\alpha \mid a, b \in \mathbb{Z}\}$.

Note that $\mathbb{Q}(\alpha) = \mathbb{Q}(\sqrt{d})$ where $d = m^2 - 4n \in \mathbb{Z}$. We can write $d = K^2 D$ where $K \in \mathbb{Z}$ and D is *square-free* (a product of distinct primes with no repetition). Then $\mathbb{Q}(\sqrt{d}) = \mathbb{Q}(\sqrt{D})$. From now on, assume that D is squarefree and look at $\mathbb{Q}(\sqrt{D})$. Let $\beta = \frac{-1+\sqrt{D}}{2} \in \mathbb{Q}(\sqrt{D})$. Then $\beta^2 + \beta + \frac{D-1}{4} = 0$.

If $D \equiv 1 \pmod 4$, then β is integral over $\mathbb{Z}$, but $\beta \notin \mathbb{Z}[\sqrt{D}]$. Hence in this case $\mathbb{Z}[\sqrt{D}]$ is not integrally closed, and thus not a UFD.

Definition 19.2.2. D is assumed to be squarefree. The ring of *quadratic integers* of *discriminant D* is

$$\mathcal{O}(D) = \begin{cases} \mathbb{Z}[\sqrt{D}] \text{ if } D \equiv 2, 3 \pmod 4 \\ \mathbb{Z}\left[\frac{-1+\sqrt{D}}{2}\right] \text{ if } D \equiv 1 \pmod 4 \end{cases}.$$

$\mathcal{O}(-1) = \mathbb{Z}[i]$ is the ring of Gaussian integers; $\mathcal{O}(-3) = \mathbb{Z}\left[\frac{-1+\sqrt{-3}}{2}\right]$ is the ring of Eisenstein integers, with $\omega = \frac{-1+\sqrt{-3}}{2}$. $\mathcal{O}(-2)$ still definitely exists, it just doesn't have a fancy name.

Fact 19.2.3.

- *$\mathcal{O}(D)$ is integrally closed.*
- *$\mathcal{O}(D)$ is a UFD $\iff$ it is a PID.*
- *The Stark-Heegner Theorem tells us that for $D < 0$, $\mathcal{O}(D)$ is a UFD if and only if $D \in \{-1, -2, -3, -7, -11, -19, -46, -67, -163\}$. (side note: what the heck?!). Moreover, $\mathcal{O}(D)$ is a euclidean domain if and only if $D \in \{-1, -2, -3, -7, -11\}$. This is hard to prove. Also, what the heck?!*
- *The* class number *$h(D)$ is an invariant that measures how far $\mathcal{O}(D)$ is from being a UFD. $h(D) \to +\infty$ as $D \to -\infty$.*
- *$\mathcal{O}(D)$ is always a Dedekind domain, and any domain that is both a UFD and Dedekind is a PID.*

So we have the following concept diagram, with inclusion going from right to left.

$$\begin{tikzcd}
 & & & \text{UFD} \arrow[dr, dash] & \\
\text{Fields} \arrow[r, dash] & \text{ED} \arrow[r, dash] & \text{PID} \arrow[ur, dash] \arrow[dr, dash] & & \text{ID} \\
 & & & \text{DD} \arrow[ur, dash] &
\end{tikzcd}$$

Where given $\mathbb{F}$ a field, we have:

- $\mathbb{Z}$, $\mathbb{F}[x]$, $\mathbb{Z}[i]$, $\mathbb{Z}[\omega]$ are ED's
- $\mathcal{O}(-19)$ is a PID

- $\mathcal{O}(-5)$ and $\mathcal{O}(10)$ are DD's
- $\mathbb{Z}[x_1, x_2, \dots]$ and $\mathbb{Z}[x_1, \dots, x_n]$ are UFD's.

Conjecture 19.2.4 (Gauss)**.** *$\mathcal{O}(D)$ is a UFD for infinitely many values $D > 0$.*

There is also a theorem, depending on the generalized Riemann Hypothesis, saying that if it holds, then for $D > 0$, $\mathcal{O}(D)$ is Euclidean exactly when it is a UFD.

19.3 Polynomial rings over integral domains

- If R is an integral domain, then $R[x]$ is an integral domain as well. Moreover, $\deg(f \cdot g) = \deg(f) + \deg(g)$. Also, $f(x) \in R[x]^\times \iff f(x) = u$, with u a unit in R.
- If F is a field, then $F[x]$ is a Euclidean domain, with $\delta(f(x)) = \deg(f)$. In particular, it is a PID.
- R is a PID $\not\Rightarrow$ $R[x]$ is a PID. As a counterexample, $\mathbb{Z}[x]$ is not a PID. The ideal generated by 2 and x is not principal.

Our goal is to show that if R is a UFD, then $R[x]$ is a UFD.

A brief preview of that which is to come: Given a UFD R, let F be the field of fractions of R. We will relate factorizations and irreducibles in $F[x]$ to factorizations and irreducibles in $R[x]$. The following are the key facts to keep in mind:

- If $f(x) \in R[x]$ factors in $F[x]$, then you can juggle constants to get a factorization in $R[x]$.
- Irreducibles in $R[x]$ are the same as in $F[x]$, except for “obvious” differences involving constants.

Example. 1. Consider $f(x) = 8x^2 + 2x - 15 \in \mathbb{Z}[x]$. Roots are $\frac{5}{4}$ and $\frac{-3}{2}$, so $f(x) = 8(x - \frac{5}{4})(x + \frac{3}{2})$ in $\mathbb{Q}[x]$. So the point is, we can break the 8 into two pieces and bring it inside, to make both factors integral. In this case, we have $f(x) = 4(x - \frac{5}{4}) \cdot 2(x + \frac{3}{2}) = (4x - 5)(2x + 3)$. We want to prove that this is always the case. The fact that this can be done is called Gauss's Lemma, which we will prove next time.

2. $f(x) = 2 \in \mathbb{Z}[x]$ and $g(x) = 2x + 4 \in \mathbb{Z}[x]$.
   (a) $f(x)$ is irreducible in $\mathbb{Z}[x]$ but not in $\mathbb{Q}[x]$, because it's a unit in $\mathbb{Q}[x]$.
   (b) $g(x)$ is irreducible in $\mathbb{Q}[x]$ but not in $\mathbb{Z}[x]$. In $\mathbb{Z}[x]$, we can factor out the 2; in $\mathbb{Q}[x]$, factoring out the 2 doesn't make g not irreducible, because 2 is a unit in $\mathbb{Q}[x]$.

Next time, we will prove the one, the only, the ultimate...

Theorem 19.3.1 (Gauss's Lemma)**.** *Let R be a UFD, and let F be the field of fractions of R. Let $h(x) \in R[x]$. Suppose that there exists $f(x), g(x) \in F[x]$ such that $h(x) = f(x)g(x)$. Then there exists $A, B \in F^\times$ such that*

(i) $\widetilde{f}(x) = Af(x) \in R[x]$, *and* $\widetilde{g}(x) = Bg(x) \in R[x]$

(ii) $h(x) = \widetilde{f}(x)\widetilde{g}(x)$

20 November 5th

20.1 Gauss's Lemma: A Continuation Of The Study Of UFD's

Theorem 20.1.1 (Gauss's Lemma)**.** *Let R be a UFD, and let F be the field of fractions of R. Let $h(x) \in R[x]$. Suppose that there exists $f(x), g(x) \in F[x]$ such that $h(x) = f(x)g(x)$. Then there exists $A, B \in F^\times$ such that*

(i) $\widetilde{f}(x) = Af(x) \in R[x]$, *and* $\widetilde{g}(x) = Bg(x) \in R[x]$

(ii) $h(x) = \widetilde{f}(x)\widetilde{g}(x)$

Proof. $\exists d \neq 0$ in R such that $dh(x) = f_1(x)g_1(x)$, where

- $f_1(x), g_1(x) \in R[x]$
- $f_1(x)$ and $g_1(x)$ are multiples of $f(x)$ and $g(x)$.

If $d \in R^\times$, multiply by $\frac{1}{d} \in R^\times$, and we're done. Otherwise, d factors into irreducibles (= primes) of R. We claim that if p is a prime in R that divides d, then either p divides all coefficients of $f_i(x)$ or all coefficients of $g_1(x)$. We can then cancel p from both sides.

To prove this claim, set $\overline{R} = R/(p)$, which is an integral domain because p is a prime. Consider the homomorphism $R[x] \to \overline{R}[x]$ given by reducing all coefficients modulo p. We have $dh(x) = f_1(x)g_1(x)$, so $\overline{0} = \overline{f_1}(x)\overline{g_1}(x)$. □

Recall (from homework 10), that given a, b in a UFD R, $\gcd(a, b)$ exists, and it is unique up to units.

Fact 20.1.2. *If* $d \in \gcd(a, b)$, *then* $a = da'$, $b = db'$, *with* $\gcd(a', b') = 1$.

We can also use $\gcd(a_1, \ldots, a_n)$.

Definition 20.1.3. A polynomial $f(x) = \sum_{i=0}^{n} a_i x^i \in R[x] \backslash \{0\}$ is *primitive* if $\gcd(a_0, a_1, \ldots, a_n) \in R^\times$.

Given any $f(x) \in R[x] \backslash \{0\}$, we can write $f(x) = c(f) \cdot f_1(x)$ where $c(f) \in \gcd(a_0, \ldots, a_n)$ and $f_1(x)$ is primitive. $c(f)$ is called the *constant* of $f(x)$.

Corollary 20.1.4. *Let R be a UFD, and F be its field of fractions. Let $h(x) \in R[x] \subseteq F[x]$.*

1. *Suppose* $\deg(h) = 0$. *Write* $h(x) = p \in R$. *Then* $h(x)$ *is irreducible in* $R[x]$ *if and only if* p *is irreducible in* R.

2. *Suppose* $\deg(h) \geq 1$. *Then* $h(x)$ *is irreducible in* $R[x]$ *if and only if* $h(x)$ *is irreducible in* $F[x]$ *and* $h(x)$ *is primitive, i.e.* $h(x) = 2x + 4$.

Proof. 1. Use $R[x]^\times = R^\times$. This is left as an exercise.

2. ($\Rightarrow$): We have $h(x) = c(h) \cdot h_1(x)$. h is irreducible, so $c(h)$ or $h_1(x)$ must be in $R[x]^\times$. But $\deg(h_1) = \deg(h) \geq 1$, so $h_1(x) \notin R[x]^\times$. Thus $c(h) \in R[x]^\times = R^\times$, so $h(x)$ is primitive.

   Suppose that $h(x) = f(x)g(x)$ with $f(x), g(x) \in F[x]$. Gauss tells us that $h(x) = \widetilde{f}(x)\widetilde{g}(x)$, with $\widetilde{f}(x), \widetilde{g}(x) \in R[x]$. $\deg(\widetilde{f}) = \deg(f)$ and $\deg(\widetilde{g}) = \deg(g)$, so $\widetilde{f}(x)$ or $\widetilde{g}(x) \in R[x]^\times = R^\times$, so $f(x)$ or $g(x) \in F^\times$.

   ($\Leftarrow$): Suppose $h(x) = f(x)g(x)$ with $f(x), g(x) \in R[x] \subseteq F[x]$. Then $f(x)$ or $g(x) \in F[x]^\times = F^\times$. Suppose the former; then $f(x) = a \in R$, so $a | c(h) \in R^\times$, so $a \in R^\times$. □

Examples. 1. Let $a, b \in R$, with $a \neq 0$. The linear polynomial $ax + b$ is irreducible in $R[x]$ if and only if $\gcd(a, b) \ni 1$.

2. For any $m \geq 0$, $x^m + y$ is irreducible in $\mathbb{K}[x, y]$, with $\mathbb{K}$ a field. Indeed, $y + x^m$ is a linear polynomial in $\mathbb{K}[x][y]$ and $1 \in \gcd(1, x^m)$.

But what was our motivation behind this? Well, we wanted the following theorem.

Theorem 20.1.5. *Let* R *be a UFD. Then* $R[x]$ *is a UFD.*

Proof. We will show existence of factorizations into irreducibles, and we know that every irreducible is prime from the homework. This is sufficient.

Let $f(x) \in R[x]$, $f(x) \neq 0$, $f(x) \neq R[x]^\times$. We write $f(x) = c(f) \cdot f_1(x)$ with f_1 primitive. Either $c(f) \in R^\times$ or we can factor $c(f) = \prod_{i \in I} p_i$ into irreducible elements in R. By the corollary, each p_i is also irreducible in $R[x]$. So, it suffices to factor $f_1(x)$ into irreducibles in $R[x]$. Let F be the field of fractions. Since $F[x]$ is a UFD (in fact, a ED), we can factor, so $f_1(x) = p_1(x) \cdots p_n(x)$ into irreducibles in $F[x]$. Applying Gauss's Lemma several times, we obtain that $f_1(x) = \widetilde{p_1}(x) \cdots \widetilde{p_n}(x)$ with $\widetilde{p_i}(x) \in R[x]$ a nonzero multiple of $p_i(x)$.

Hence $\widetilde{p_i}(x) \sim p_i(x)$ in $F[x]$, so $\widetilde{p_i}(x)$ is irreducible in $F[x]$. Since $c(\widetilde{p_i}) | c(f_1)$, $\widetilde{p_i}(x)$ must be primitive. By the corollary, each $\widetilde{p_i}(x)$ is irreducible in $R[x]$. □

20.2 Modules over a PID

Definition 20.2.1. Let R be an integral domain and M an R-module. The *rank* of M is

$$rk(M) = \max\{|S| \mid S \text{ a linearly independent subset of } M\}.$$

Proposition 20.2.2. *Suppose $M \cong R^r$ with $r < \infty$ (Note that R is still an integral domain). Then*

(i) $rk(M) = r$

(ii) Any basis of M has r elements.

Proof. (i) Assume that $M = R^r = R \oplus \cdots \oplus R$. Let F be the field of fractions of R. Then $M \subseteq F^r = F \oplus \cdots \oplus F$. We claim that if $S \subseteq M$ is linearly independent over R, it's also linearly independent over F. Take $\sum_{s \in S} a_s s = 0$ with $a_s \in F$ for all $s \in S$. Choose some $d \in F^\times$ such that $da_s \in R$ for all s. Then $\sum_{s \in S} da_s s = 0$, so we have a linear dependence in R, so $da_s = 0$ for all s; since we're in a domain and $d \in F^\times$, $a_s = 0$ for all s.

Thus $rk(M) \leq \dim_F F^r = r$.

But $\{e_1, \ldots, e_r\}$ is a basis of M over R, so the rank $rk(M) \geq r$, so $rk(M) = r$.

(ii) Suppose T is a basis of M, and $|T| = t$. Then $M \cong R^{(T)} = R^t$, so by (i) $rk(M) = rk(R^t) = t$, so $r = t$. One has to be careful to check that t is finite - this follows because $t \leq r$, since a basis is linearly independent and we can check the definition of rank.

□

Remark. If R is a PID, every non-zero ideal is a free R-module of rank 1.

Proof. $I = (a), a \neq 0$. So $\{a\}$ generates I by definition. If $ra = 0$, then $r = 0$, because R is a domain, so clearly $\{a\}$ is linearly independent. Thus it is a basis, so we have a free module of rank 1. □

Lemma 20.2.3. *R a general ring. Let M be a left R-module and $N \leq M$. Suppose M/N is a free R-module. Then $M \cong N \oplus M/N$.*

Proof. Exercise. □

Proposition 20.2.4. *Let R be a PID. Let M be a free R-module of rank $r < \infty$ and $N \leq M$. Then N is free and of rank $s \leq r$.*

Proof. Assume $M = R^r$. We proceed by induction on r.

Let $r = 1$ (if $r = 0$, there's really nothing going on). Then N is a submodule of R, so it is an ideal. Thus N is free of rank 1 if N is nonzero, and of rank 0 if it is the zero ideal. So we are fine.

Let $r \geq 2$. Consider $\varphi : R^r \to R$ be the projection onto the last coordinate, with $(a_1, \ldots, a_r) \mapsto a_r$. Let $I = \varphi(N) \leq R$, an ideal of R. By the base case, I is free of rank 1 or 0.

$$\begin{tikzcd}
\ker\varphi \arrow[r, hook] & R^r \arrow[r, two heads, "f"] & R \\
N\cap\ker\varphi \arrow[u, hook] \arrow[r, hook] & N \arrow[u, hook] \arrow[r, two heads] & I \arrow[u, hook]
\end{tikzcd}$$

□

$\ker\varphi = \{(a_1,\dots,a_r) \mid a_r = 0\} \cong R^{r-1}$, which is free. By the induction hypothesis, $N\cap\ker\varphi$ is free of rank $s \le r-1$. Since I is free, by the lemma,

$$\begin{aligned} N &\cong (N\cap\ker\varphi)\oplus I \\ &\cong R^s \oplus R^\varepsilon \cong R^{s+\varepsilon} \end{aligned}$$

Thus $s+\varepsilon \le r-1+1 = r$.

21 November 10th

21.1 Last time: PIDs

Let M be a free R-module for R a PID and $N \le M$. Then recall that N is free.

Remark. 1. Let $W \le V$ vector spaces. Given a basis of V, does there exist a subset which is a basis of W? Not necessarily. For example, if V is $\mathbb{R}^2$ under any basis, take W diagonal space.

   But there is *some* basis of V containing a basis of W.

2. Is the previous fact true for general free R-modules? Well, no. Take $R=\mathbb{Z}, M=\mathbb{Z}, N=2\mathbb{Z}$. Or, take $R=\mathbb{Z}$, $M=\mathbb{Z}^2$, and $N=\{(a,b)\mid a=b \text{ even}\}$.

   But, there is a basis of M that is "stacked upon" a basis of N. What do we mean by that? Well. This calls for the stacked basis theorem.

Theorem 21.1.1 (Stacked basis theorem)**.** *Let R be a PID, and let M be a free R-module with rank $r<\infty$, and $N \le M$. Then there exist bases $\{m_1,\dots,m_r\}$ of M and $\{n_1,\cdots,n_s\}$ of N (for some s with $0\le s\le r$) and elements $a_1,\dots,a_s \in R\setminus\{0\}$ such that*

(i) $n_i = a_i\cdot m_i$ *for all* $i=1,\dots,s$.

(ii) $a_1|a_2|\cdots|a_s$.

Before we prove this, here is a corollary.

Corollary 21.1.2 (Structure theorem for finitely generated modules over a PID, invariant factor version)**.** *Let R be a PID, M a finitely generated R-module. Then there exist $r, s \geq 0$ and elements $a_1, \ldots, a_s \in R$ such that $a_i \neq 0$, a_i nonunits for all i, and*

(i) $M \cong R^r \oplus R/(a_1) \oplus \cdots \oplus R/(a_s)$

(ii) $a_1 | a_2 | \cdots | a_s$

Proof. Let $\{x_1, \ldots, x_k\}$ be generators of M. There is a projection $\pi : R^k \to M$, with $\pi(e_i) = x_i$. Then $M \cong R^k/N$, where $N = \ker \pi$. Choose stacked bases $\{m_1, \ldots, m_k\}$ of R^k and $\{n_1, \ldots, n_h\}$ of N. $n_i = a_i \cdot m_i$ for some $a_i \in R \setminus \{0\}$; $a_1|a_2|\cdots|a_h$.

Then $R^k = Rm_1 \oplus \cdots \oplus Rm_k$, and $N = Rn_1 \oplus \cdots \oplus Rn_h$, so $R^k/N = Rm_1/Rn_1 \oplus \cdots \oplus Rm_h/Rn_h \oplus \cdots \oplus Rm_k$. In general, if $N_i \leq M_i$, then $\oplus M_i / \oplus N_i \cong \oplus(M_i/N_i)$.

Now, $R \cong Rm_i$ via $1 \mapsto m_i$. This map sends $a \mapsto a \cdot m_i$. Hence it sends $a_i \mapsto a_i \cdot m_i = n_i$, and thus $(a_i) \mapsto Rn_i$, so $R/(a_i) \cong Rm_i/Rn_i$, and $M \cong R^k/N \cong R/(a_1) \oplus \cdots \oplus R/(a_h) \oplus R \oplus \cdots \oplus R$. If some $a_i \in R^\times$, then $(a_i) = R$, so the quotient is 0 and we can simply remove it from the sum. This then gives us the desired result. □

There are a couple of comments about this that lead to another version of the structure theorem.

Given $a \in R$, $a \neq 0$, a nonunit. Write $a \sim p_1^{\alpha_1} \cdots p_k^{\alpha_k}$ with each p_i irreducible, and p_i and p_j nonassociate for $i \neq j$, and $\alpha_i \geq 0$. Then $1 \in \gcd(p_1^{\alpha_1}, p_j^{\alpha_j})$, so $(p_i^{\alpha_i}) + (p_j^{\alpha_j}) = R$. Also, $(a) = (p_1^{\alpha_1}) \cdots (p_j^{\alpha_j})$.

By the Chinese Remainder Theorem, there is an isomorphism of (commutative) rings

$$R/(a) \cong R/(p_1^{\alpha_1}) \times \cdots \times R/(p_k^{\alpha_k}),$$

with $\overline{x} \mapsto (\overline{x}, \ldots, \overline{x})$. This is also an isomorphism of R-modules.

Definition 21.1.3. A *partition* $\lambda = (l_1, \ldots, l_n)$ is a sequence of positive integers such that $l_1 \geq \cdots \geq l_h$. h is the *number of parts.*

Given a partition λ and an irreducible $p \in R$, let $R/p^\lambda = R/(p^{l_1}) \oplus \cdots \oplus R/(p^{l_n})$. We then have

Corollary 21.1.4 (Structure theorem, elementary divisor version)**.** *Let R be a PID, and M a finitely generated R-module. Then there exists $r \geq 0$ and irreducibles $p_1, \ldots, p_k \in R$ and partitions $\lambda_1, \ldots, \lambda_k$ (of various lengths) such that*

$$M \cong R^r \oplus R/p_1^{\lambda_1} \oplus \cdots \oplus R/p_k^{\lambda_k}.$$

Proof. We have

$$M \cong R^r \oplus R/(a_1) \oplus \cdots \oplus R/(a_s),$$

by the previous corollary.

We write $a_s \sim p_1^{\alpha_{s1}} \cdots p_k^{\alpha_{sk}}$, then the same for a_{s-1}, and so on, up to $a_2 \sim p_1^{\alpha_{21}} \cdots p_k^{\alpha_{2k}}$ and $a_1 \sim p_1^{\alpha_{11}} \cdots p_k^{\alpha_{1k}}$. Since we're allowing some of the α_i's to be 0, we can make sure we always have a square matrix. Each p_i is irreducible and $\alpha_{si} \geq \cdots \geq \alpha_{2i} \geq \alpha_{1i} \geq 0$, because of the divisibility condition on the a_i's. Let λ_i be this sequence, with any tailing 0's removed. The remarks above concerning partitions, and rearrangement, give the desired result. Specifically, $R/(a_i) \cong R/(p_1^{\alpha_{i1}}) \oplus \cdots \oplus R/(p_k^{\alpha_{ik}})$, which we can plug in and rearrange. □

Remark. Both theorems concerned existence. But in fact, for both results, uniqueness holds. The integer r in both cases is the rank of the module M, so it is uniquely determined by the module M. The elements a_i are uniquely determined by M, up to associates. They are called the *invariant factors* of M; the irrducibles p_i and the partitions λ_i are also uniquely determined. The $p_i^{\alpha_{ij}}$ are the *elementary divisors.*

So now to turn to the proof of the Stacked Basis Theorem.

21.2 Stacked Basis Theorem

We repeat, for convenience, the statement of the theorem here.

Theorem 21.2.1 (Stacked basis theorem)**.** *Let R be a PID, and let M be a free R-module with rank $r < \infty$, and $N \leq M$. Then there exist bases $\{m_1, \ldots, m_r\}$ of M and $\{n_1, \cdots, n_s\}$ of N (for some s with $0 \leq s \leq r$) and elements $a_1, \ldots, a_s \in R \setminus \{0\}$ such that*

(i) $n_i = a_i \cdot m_i$ *for all* $i = 1, \ldots, s$.

(ii) $a_1 | a_2 | \cdots | a_s$.

We need some preliminary discussion on dual modules. Let R be a ring and let M be a left R-module. Let $M^* = \mathrm{Hom}_R(M, R)$, the set of left R-module homomorphisms $\varphi : M \to R$.

Then

- M^* is a right R-module via $(\varphi \cdot a)(x) = \varphi(x)a$, for all $\varphi \in M^*, a \in R, x \in M$. Note $(\varphi \cdot a)(bx) = \varphi(bx)a = b\varphi(x)a = b(\varphi \cdot a)(x)$, so $\varphi \cdot a$ is a homomorphism of left R-modules.
- If R is commutative (which, let's be honest, it will be), then $(\varphi \cdot a)(x) = \varphi(x)a = a\varphi(x) = \varphi(a \cdot x)$, so we can afford to be careless with where we put our variables.
- Given $x \in M$, let $\varepsilon_x : M^* \to R$ be $\varepsilon_x(\varphi) = \varphi(x)$. Then ε_x is a homomorphism of right R-modules.
- Suppose now that M is free of rank r and $\{x_1, \ldots, x_r\}$ is a basis. For each $i = 1, \ldots, r$, define $\pi_i : M \to R$ by $\pi_i \left(\sum_{j=1}^r a_j \cdot x_j \right) = a_i$. This is well-defined because we had a basis of x_i's. π_i is a homomorphism of left R-modules, so $\pi_i \in M^*$.

With the π_i's above, we have the following useful lemma.

Lemma 21.2.2. $\{\pi_1, \dots, \pi_r\}$ *form a basis of* M^*. *In particular,* M^* *is free. We say that* $\{\pi_1, \dots, \pi_r\}$ *is the* dual basis *of* $\{x_1, \dots, x_r\}$. *Often, we write* x_i^* *instead of* π_i. *This is misleading!* π_i *depends on the whole basis. But, go ahead and use this terribly misleading and downright devilish notation if you feel so inclined. Heathens.*

Proof. Given $x \in M$, write $x = \sum_{j=1}^r a_j x_j = \sum_{j=1}^r \pi_j(x) x_j$. Now, given $\varphi \in M^*$, we have $\varphi(x) = \sum_{j=1}^r \pi_j(x)\varphi(x_j)$, with $\varphi(x_j) \in R$. So $\varphi(x) = \sum_{j=1}^r (\pi_j \cdot \varphi(x_j))(x)$, for all $x \in M$. Thus $\varphi = \sum_{j=1}^r \pi_j \cdot \varphi(x_j)$. So $\{\pi_1, \dots, \pi_r\}$ generate M^*. But we now must show linear independence. Suppose $\sum_{i=1}^r \pi_i \cdot a_i = 0$ for some $a_i \in R$. Then we evaluate on x_j:

$$0 = \sum_{i=1}^r (\pi_i \cdot a_i)(x_j) = \sum_{i=1}^r \pi_i(x_j) a_i = a_j.$$

Thus linear independence holds. □

22 November 12th

22.1 Stacks on stacks on stacks

Let M be a free R-module of rank r, with $\{x_1, \dots, x_r\}$ a basis for M and $\{\pi_1, \dots, \pi_r\}$ the dual basis for M^*. Then for all $x \in M$, $x = \sum_{i=1}^r \pi_i(x) \cdot x_i$. This is known as the Fourier expansion, since that's pretty much what's happening. Now for, finally, the stacked basis theorem.

We repeat, for convenience, the statement of the theorem here.

Theorem 22.1.1 (Stacked basis theorem)**.** *Let* R *be a PID, and let* M *be a free* R*-module with rank* $r < \infty$*, and* $N \leq M$. *Then there exist bases* $\{m_1, \dots, m_r\}$ *of* M *and* $\{n_1, \cdots, n_s\}$ *of* N *(for some* s *with* $0 \leq s \leq r$*) and elements* $a_1, \dots, a_s \in R \setminus \{0\}$ *such that*

(i) $n_i = a_i \cdot m_i$ *for all* $i = 1, \dots, s$.

(ii) $a_1 | a_2 | \cdots | a_s$.

Proof. If $N = \{0\}$, any basis of N works. The statement is vacuously true, like most of my jokes. Now assume that $N \neq \{0\}$. Then $\{\varphi(N) \mid \varphi \in M^*\}$ is a collection of ideals of R; it is nonempty since $M^* \neq \varnothing$ (at least $0 \in M^*$). R is a PID, so it is noetherian and there is an ideal $\varphi_1(N)$ that is maximal in this collection (for the motivation behind choosing this ideal, see the remark following this proof). We claim that $\varphi_1(N) \neq 0$.

Proof of claim that $\varphi_1(N) \neq 0$**.** $N \neq \{0\}$, so there exists $y \in N$, $y \neq 0$. Write $y = \sum_{i=1}^r a_i \cdot x_i$. Then there exists i with $a_i \neq 0$, so $\pi_i(y) = a_i \neq 0$, so $\pi_i(N) \neq \{0\}$. Thus $\varphi_1(N) \neq \{0\}$.

R is a PID, so $\varphi_1(N) = (a_1)$ for some $a_1 \in R$, $a_1 \neq 0$. Choose $n_1 \in N$ such taht $\varphi_1(n_1) = a_1$. We now claim that for every $\varphi \in M^*$, $a_1 | \varphi(n_1)$.

Proof of claim that $a_1|\varphi(n_1)$. Consider $\varepsilon : M^* \to R$, $\varepsilon(\varphi) = \varphi(n_1)$ for fixed n_1. We want to show that a_1 divides every element in $\operatorname{im}\varepsilon$, or equivalently that $\operatorname{im}\varepsilon \subseteq (a_1)$. Let $\operatorname{im}\varepsilon = (b)$; choose $\varphi_0 \in M^*$ such that $\varepsilon(\varphi_0) = b$. We then have $a_1 = \varphi_1(n_1) = \varepsilon(\varphi_1)$, so $a_1 \in \operatorname{im}\varepsilon$, so $(a_1) \subseteq \operatorname{im}\varepsilon$. We wanted the opposite inclusion. But $b = \varepsilon(\varphi_0) = \varphi_0(n_1) \in \varphi_0(N)$, so $(b) \subseteq \varphi_0(N)$. Thus $(a_1) \subseteq \operatorname{im}\varepsilon = (b) \subseteq \varphi_0(N)$, so by maximality $(a_1) = \operatorname{im}\varepsilon = (b) = \varphi_0(N)$.

Now we claim that there exists $m_1 \in M$ with $n_1 = a_1 \cdot m_1$.

Proof of claim that such an m_1 exists. By the previous claim, $a_1|\pi_i(n_1)$ for all i, so $\pi_i(n_1) = a_1 b_i$ for some $b_1 \in R$, for all i. Then

$$n_1 = \sum_{i=1}^{r} \pi_i(n_1) \cdot x_i = \sum_{i=1}^{r} a_1 b_i \cdot x_i = a_1 \left(\sum_{i=1}^{r} b_i \cdot x_i \right).$$

Define m_1 as the last sum, i.e. $m_1 = \sum_{i=1}^{r} b_i \cdot x_i$.

Now $a_1 = \varphi_1(n_1) = \varphi_1(a_1 \cdot m_1) = a_1\varphi_1(m_1)$. But R is an integral domain, so $\varphi_1(m_1) = 1$. So we have

$$\begin{tikzcd}
M \arrow[r, two heads, "\varphi_1"] & R \\
N \arrow[u, hook] \arrow[r, two heads, "\varphi_1|_N"] & (a_1) \arrow[u, hook]
\end{tikzcd}$$

with $m_1 \mapsto 1$, so $M/\ker\varphi_1 \cong R$ (free), and $n_1 \mapsto a_1 \Rightarrow N/\ker\varphi_1|_N \cong (a_1)$ (also free). Since free quotients split (not proven here but it is a fun and true fact) we deduce that $M = \ker\varphi_1 \oplus Rm_1$ and $N = \ker\varphi_1|_N \oplus Rn_1$. This gives us a new pair of modules, and we can proceed by induction. In particular, the direct sum decompostions reduce the problem to finding a basis for $\ker\varphi_1$ and $\ker\varphi_1|_N$, both of which are free and have rank one smaller. Throwing in m_1 and n_1 gives us the desired bases.

More precisely, applying the induction hypothesis to the free module $\ker\varphi_1$ and the submodule $\ker\varphi_1|_N$, we obtain bases $\{m_2, \ldots, m_r\}$ of $\ker\varphi_1$ and $\{n_2, \ldots, n_s\}$ of $\ker\varphi_1|_N$, such that $n_i = a_i \cdot m_i$ for all $i = 2, \ldots, s$. Then $\{m_1, \ldots, m_r\}$ and $\{n_1, \ldots, n_s\}$ of M and N respectively. (i) is satisfied easily, but for (ii) we still need to show that $a_1|a_2$.

Since $\{m_1, \ldots, m_r\}$ is a basis of M, there is $\varphi \in M^*$ with $\varphi(m_1) = \varphi(m_2) = 1$, and $\varphi(m_i)$ has no restrictions for all $i \geq 3$. Then $\varphi(n_1) = \varphi(a_1 \cdot m_1) = a_1$, so $a_1 \in \varphi(N)$, so $(a_1) \subseteq \varphi(N)$, so $(a_1) = \varphi(N)$¿ But also, $\varphi(n_2) = \varphi(a_2 \cdot m_2) = a_2$, so $a_2 \in \varphi(N) = (a_1)$, so $a_1|a_2$. □

Remark. Suppose there exists a pair of stacked bases for $N \leq M$. Take $\varphi \in M^*$ and $n \in N$; write $n = \sum_{i=1}^{s} b_i \cdot n_i = \sum_{i=1}^{s} b_i a_i \cdot m_i$. Then $\varphi(n) = \sum_{i=1}^{s} b_i a_i \cdot \varphi(m_i) \in (a_i) \subseteq (a_1)$, so the whole expression is in (a_1). This tells you that $\varphi(N) \subseteq (a_1)$.

Also, $\pi_1(n_1) = \pi_1(a_1 \cdot m_1) = a_1$, so $(a_1) \subseteq \pi_1(N)$, and in particular $(a_1) = \pi_1(N)$. Thus (a_1) is the largest of all the ideals $\varphi(N)$ as φ ranges in M^*.

Note that the proof is nonconstructive because of that one time we used properties of noetherian rings. There is a constructive proof of Smith Normal form, which is useful because computers, but it's not this proof and it's more ~~of a pain~~ gruntwork.

Now we move on to field extensions.

22.2 Field characteristic

Let F be a field. Recall that $\mathbb{Z}$ is the initial ring, so there is a unique ring homomorphism $\varphi : \mathbb{Z} \to F$. Explicitly, $\varphi(n) = 1 + \cdots + 1$, with n summands, for $n \geq 0$, and with -1 summands for negative n. $\operatorname{im}\varphi$ is a subring of a field, so $\operatorname{im}\varphi$ is an integral domain. Thus $\ker\varphi$ is a prime ideal of $\mathbb{Z}$, so $\ker\varphi = \{0\}$ or (p) for some prime p. Then $\operatorname{im}\varphi \cong \mathbb{Z}$ or $\mathbb{Z}_p = \mathbb{F}_p$. This leads to the definition of field characteristic.

Definition 22.2.1. Using F and φ as defined above, the *characteristic* of F, denoted $\mathrm{char}F$, is 0 if $\operatorname{im}\varphi \cong \mathbb{Z}$, and p if $\operatorname{im}\varphi \cong \mathbb{Z}_p$.

- If $\mathrm{char}F = p$, F contains a subfield isomorphic to $\mathbb{F}_p$.
- If $\mathrm{char}F = 0$, $\varphi : \mathbb{Z} \to F$ is injective. By the universal property of fields of fractions, φ extends to $\mathbb{Q}$, with $\varphi : \mathbb{Q} \to F$. The extension is still injective (any homomorphism from a field to a nontrivial ring is injective). So F contains a subfield isomorphic to $\mathbb{Q}$.

So every field F contains a copy either of $\mathbb{F}_p$ or of $\mathbb{Q}$. This copy is called the *prime subfield* of F. This is the smallest subfield of F.

22.3 Extensions and degree

Definition 22.3.1. A *field extension* is a pair consisting of a field K and a subfield F. We write $F \leq K$ or $K|F$ or F—K

In this case, K is a vector space over F and is in fact an F-algebra.

Definition 22.3.2. The *degree* of the extension is $[K : F] = \dim_F K$ (which could be infinite).

A *ring extension* and its degree are defined similarly. Yay field extensions! What can we say about them?

Proposition 22.3.3. *Let $F \leq L \leq K$ be field extensions. Then*

(a) If $[L : F] < \infty$ and $[K : L] < \infty$, then $[K : F] = [K : L][L : F] < \infty$.

(b) $[K : F] = \infty \iff [L : F] = \infty$ OR $[K : L] = \infty$.

Proof. (a) Let $\{\alpha_1, \ldots, \alpha_n\}$ be an F basis of L. and $\{\beta_1, \ldots, \beta_m\}$ be an L-basis for K. Then check that $\{\alpha_i\beta_j \mid 1 \leq i \leq n, 1 \leq j \leq m\}$ is a basis for K over F.

(b) The ($\Rightarrow$) implication follows from (a); if $[K : F] = \infty$, it can't be true that both of the smaller degrees are finite. Now consider ($\Leftarrow$); this is easy, considering bases of K.

□

Definition 22.3.4. Given a ring extension $R \leq S$ and $\alpha \in S$, let $R[\alpha]$ be the smallest subring of S containing both R and α. Given $F \leq K$ and $\alpha \in K$, $F[\alpha]$ is then defined analogously, as a subring. In addition, let $F(\alpha)$ be the smallest subfield of K containing F and α.

The two definitions for a field may or may not be equal. But we do know that $F \leq F[\alpha] \leq F(\alpha) \leq K$. $F(\alpha)$ is the field of fractions of $F[\alpha]$. Explicitly,

$$F[\alpha] = \{f(\alpha) \in K \mid f(x) \in F[x]\},$$

and

$$F(\alpha) = \{f(\alpha)/g(\alpha) \in K \mid f(x), g(x) \in F[x], g(\alpha) \neq 0\}.$$

23 November 17th

23.1 Field Extensions, extended

We have, from last time, $F \leq F[\alpha] \leq F(\alpha) \leq K$, with $\alpha \in K$, $F[\alpha]$ the smallest subring of K containing F and α, and $F(\alpha)$ the smallest subfield.

Definition 23.1.1. We say that K is a *simple extension* of F if $\exists \alpha \in K$ such that $K = F(\alpha)$. In this case we also say that α is a *primitive element* for the extension $K|F$, and that K is obtained from F by *adjoining* α.

Given $K|F$ and $\alpha \in K$, the universal property of the polynomial algebra $F[x]$ yields a homomorphism of F-algebras, $\varphi : F[x] \to K$ such that $\varphi(x) = \alpha$. We have $\varphi(a) = a$ for all $a \in F$ and $\varphi(f(x)) = f(\alpha)$ for all $f(x) \in F[x]$, so $\operatorname{im} \varphi = F[\alpha]$. $\operatorname{im} \varphi$ is a subring of K, so $\operatorname{im} \varphi$ is an integral domain, so $\ker \varphi$ is a prime ideal in $F[x]$. But then $\ker \varphi = \{0\}$ or $(p(x))$, for some irreducible $p(x) \in F[x]$, because $F[x]$ is a PID. Moreover, $F[x]/\ker \varphi \cong F[\alpha]$.

If $\ker \varphi = \{0\}$, we say α is *transcendental over F*. Equivalently there exists no $g(x) \in F[x]$, $g(x) \neq 0$, such that $g(\alpha) = 0$. In this case, $F[x] \cong F[\alpha]$ and this extends to $F(x) \cong F(\alpha)$. In this setting,

$$F(x) = \{f(x)/g(x) \mid f(x), g(x) \in F[x], g(x) \neq 0\}$$

is the field of *rational* functions.

If $\ker \varphi \neq \{0\}$, we say that α is *algebraic* over F. Equivalently, there is a nonzero polynomial $g(x) \in F[x]$ with $g(\alpha) = 0$. In this case, there exists an irreducible $p(x) \in F[x]$ unique up to scalars such that $\ker \varphi = (p(x))$, and we have $F[x]/(p(x)) \cong F[\alpha]$. $p(x)$ is irreducible, so $(p(\alpha))$ is maximal; so $F[\alpha]$ is a field. Thus $F(\alpha) = F[\alpha] = \{f(\alpha) \mid f(x) \in F[x]\}$. The unique *monic* irreducible $p(x) \in F[x]$ such that $p(\alpha) = 0$ is the *minimum polynomial* of α

over F (or *minimal polynomial*, but minimum is better, because the polynomial is unique). We denote the minimum polynomial by $m_{\alpha,F}(x)$ or $\operatorname{irr}(\alpha, F)(x)$. The degree of $m_{\alpha,F}(x)$ is the *degree of α over F*, $\deg \alpha$.

Proposition 23.1.2. *Given $K|F$, let $\alpha \in K$ be algebraic over F. Then* $\deg \alpha = [F(\alpha) : F]$.

Proof. By the preceding discussion, $F(\alpha) \cong F[x]/(p(x))$, where $p(x) = m_{\alpha,F}(x)$. This is an isomorphism of F-algebras.

We claim that if $\deg \alpha = n$, then $\{1, \overline{x}, \overline{x}^2, \ldots, \overline{x}^{n-1}\}$ is an F-basis for $F[x]/(p(x))$. □

Corollary 23.1.3. *Let $K|F$ and $\alpha \in K$. α is algebraic over $F \Leftrightarrow [F(\alpha) : F] < \infty$.*

Proof. $(\Rightarrow)$: is shown by the proposition.

$(\Leftarrow)$: If α is transcendental, then $F(\alpha) \cong F(x) \geq F[x]$. Then $[F(\alpha) : F] \geq \dim_F F[x] = \infty$. □

23.2 Finite, algebraic, and finitely generated extensions

Definition 23.2.1. We say $K|F$ is *finite* if $[K : F] < \infty$. It is *algebraic* if every $\alpha \in K$ is algebraic over F. It is *transcendental* if it is not algebraic, or if there exists some $\alpha \in K$ that is transcendental.

Proposition 23.2.2. *$K|F$ is finite $\Rightarrow$ $K|F$ is algebraic.*

Proof. Let $\alpha \in K$, so we have $F \leq F(\alpha) \leq K$. Thus $[F(\alpha) : F] \leq [K : F] \leq \infty$. Thus α is algebraic over F.

Thus, F is algebraic. □

Assume we have $F \leq K$ and $\alpha_1, \ldots, \alpha_n \in K$. We define $F[\alpha_1, \ldots, \alpha_n]$ to be the smallest subring of K containing F and $\alpha_1, \ldots, \alpha_n$. Similarly, we have $F(\alpha_1, \ldots, \alpha_n)$, the smallest subfield of K containing F and $\alpha_1, \ldots, \alpha_n$.

We have $F[\alpha, \beta] = F[\alpha][\beta]$. Similarly, $F[\alpha_1, \ldots, \alpha_n] = F[\alpha_1, \ldots, \alpha_{n-1}][\alpha_n]$, and so on.

Definition 23.2.3. If there exists $\alpha_1, \ldots, \alpha_n \in K$ such that $K = F(\alpha_1, \ldots, \alpha_n)$, we say that $K|F$ is finitely generated.

Proposition 23.2.4. *Let $K|F$ and $\alpha_1, \ldots, \alpha_n \in K$ be such that α_i is algebraic over $F(\alpha_1, \ldots, \alpha_{i-1})$ for all $i \geq 1$. Then $F(\alpha_1, \ldots, \alpha_n)|F$ is finite, and $F(\alpha_1, \ldots, \alpha_n) = F[\alpha_1, \ldots, \alpha_n]$. In particular, this holds when each α_i is algebraic over F.*

Proof. By induction on $n \geq 0$. For $n = 0$, this is trivial.

Assume $n \geq 1$. α_n is algebraic over $L = F(\alpha_1, \ldots, \alpha_{n-1})$, so $L(\alpha_n) = L[\alpha_n]$ and $[L(\alpha) : L] < \infty$. By the induction hypothesis, $L|F$ is finite and $L = F[\alpha_1, \ldots, \alpha_{n-1}]$. Thus $[L(\alpha_n) : F] = [L(\alpha_n) : L][L : F] < \infty$ and $F(\alpha_1, \ldots, \alpha_n) = L(\alpha_n) = L[\alpha_n] = F[\alpha_1, \ldots, \alpha_n]$. □

Corollary 23.2.5. *$K|F$ is finite if and only if $K|F$ is algebraic and finitely generated.*

Proof. $(\Leftarrow)$: By the proposition.

$(\Rightarrow)$: If $K|F$ is finite, then it is algebraic. Pick $\alpha \in K \setminus F$; then $[F(\alpha) : F] > 1$. If $K = F(\alpha)$, we're done; if not, pick $\beta \in K \setminus F(\alpha)$. Then $[F(\alpha)(\beta) : F(\alpha)] > 1$, so these dimensions are always increasing. The process must terminate, because $K|F$ is finite. □

Definition 23.2.6. Given $K|F$, the *algebraic closure* of F in K is the set $\Omega_K(F) = \{\alpha \in K \mid \alpha \text{ is algebraic over } F\}$.

Corollary 23.2.7. *$\Omega_K(F)$ is a subfield of K (containing F).*

Proof. Let $\alpha, \beta \in K$ be algebraic over F. We have to show that $\alpha \pm \beta$, $\alpha\beta$, and α/β (with $\beta \neq 0$) are algebraic over F. But all these elements lie in $F(\alpha, \beta)$, which is algebraic over F by the above proposition and corollary, so all elements it contains are algebraic, so we're done. □

Remark (Variant of the polynomial UMP). Let $\varphi : F \to R$ be a homomorphism of rings from a field F to some ring R, and $\alpha \in R$. In that case there is a unique homomorphism of rings $F[x] \to R$ extending φ and sending x to α.

$$\begin{tikzcd}
F[x] \arrow[r, dashed, "\varphi"] & R \\
F \arrow[u, hook] \arrow[ur, "\varphi"'] &
\end{tikzcd}$$

In particular, a homomorphism of fields $\varphi : F \to \widetilde{F}$ extends to a homomorphism of rings sending x to x.

$$\begin{tikzcd}
F[x] \arrow[r, dashed, "\varphi"] & \widetilde{F}[x] \\
F \arrow[u, hook] \arrow[r, "\varphi"'] & \widetilde{F} \arrow[u, hook]
\end{tikzcd}$$

23.3 Root extension

Proposition 23.3.1. *Let F be a field and $p(x) \in F[x]$ irreducible. Let $K = F[x]/(p(x))$ and $\alpha = \overline{x} \in K$. Then*

(1) K is a field, $F \hookrightarrow K$ and $p(\alpha) = 0$.

(2) If K' is another field with $F \hookrightarrow K'$ and $\alpha' \in K'$ such that $p(\alpha') = 0$, then there exists a unique homomorphism of fields $\varphi : K \to K'$ such that $\varphi|_K$ is the identity, $\varphi(\alpha) = \alpha'$ and in general the diagram below commutes.

$$\begin{tikzcd}
\alpha \in K \arrow[rr, dashed, "\varphi"] & & K' \ni \alpha' \\
& F \arrow[ul, hook] \arrow[ur, hook] &
\end{tikzcd}$$

Proof. (1) The map $F \hookrightarrow F[x] \twoheadrightarrow F[x]/(p(x)) = K$, with codomain nontrivial because $p(x)$ is nonconstant, is injective because it is a homomorphism of fields. Moreover $p(\alpha) = p(\overline{x}) = \overline{p(x)} = \overline{0}$.

(2) The UMP of $F[x]$ yields $F[x] \xrightarrow{\psi} K'$ such that $\psi|_F$ is the identity and $\psi(x) = \alpha'$. $\psi(p(x)) = p(\alpha') = 0$, so ψ descends to a homomorphism φ from the quotient, as in the following diagram.

$$\begin{tikzcd}
F[x] \arrow[r, dashed, "\psi"] \arrow[d, two heads] & K' \\
K \arrow[ur, dashed, "\varphi"'] &
\end{tikzcd}$$

□

Corollary 23.3.2. *Let F be a field and $f(x) \in F[x]$ with $\deg f(x) = n \geq 1$. Then there exists a field $K \geq F$ and $\alpha \in K$ such that $f(\alpha) = 0$. Moreover, K can be chosen so that $[K : F] \leq n$.*

Proof. Apply (1) to an irreducible factor $p(x)$ of $f(x)$. Note that $[K : F] = \deg p(x) \leq \deg f(x) = n$. □

24 November 19th

24.1 Splitting fields

Definition 24.1.1. Let F be a field and $f(x) \in F[x]$. A *splitting field* of $f(x)$ over F is a field $K \geq F$ containing elements $\alpha_1, \dots, \alpha_n$ such that

(i) $f(x) \sim (x - \alpha_1) \cdots (x - \alpha_n)$ in $K[x]$

(ii) $K = F(\alpha_1, \dots, \alpha_n)$

Remark. (i) says that $f(x) = u(x - \alpha_1) \cdots (x - \alpha_n)$ where $u \in K^\times$. But then u is the leading coefficient of $f(x)$, so $u \in F^\times$.

Proposition 24.1.2 (Existence of splitting fields)**.** *Let $f(x) \in F[x]$, $n = \deg(f(x))$. There exists a splitting field K of $f(x)$ over F with $[K : F] \leq n!$.*

Proof. Let $L \geq F$ be a field and $\alpha_1 \in L$ be such that $f(\alpha_1) = 0$, and $[L : K] \leq n$, possible by a previous corollary. Let $F_1 = F(\alpha_1) \leq L$. In $F_1[x]$, we have $f(x) = (x - \alpha_1)g(x)$, with $\deg(g(x)) = n - 1$.

By induction, there exists a splitting field K of $g(x)$ over F_1, with $[K : F_1] \leq (n-1)!$, so there exist $\alpha_2, \dots, \alpha_n \in K$ such that $g(x) \sim (x - \alpha_2) \cdots (x - \alpha_n)$ and $K = F_1(\alpha_2, \dots, \alpha_n)$. But then $[K : F] = [K : F_1][F_1 : F] \leq n!$, $f(x) = (x - \alpha_1)g(x) \sim (x - \alpha_1) \cdots (x - \alpha_n)$, and $K = F(\alpha_1, \alpha_2, \dots, \alpha_n)$. □

Lemma 24.1.3. *Let $f(x) \in F[x]$ and K be a splitting field of $f(x)$ over F. Let $\alpha \in K$ be a root of $f(x)$, and write $f(x) = (x - \alpha)f_1(x)$ with $f_1(x) \in K[x]$. Let $F_1 = F(\alpha)$. Then $f_1(x) \in F_1[x]$ and K is a splitting field of $f_1(x)$ over F_1.*

Proof. We have $f(x) \sim (x - \alpha_1) \cdots (x - \alpha_n)$ in $K[x]$, and $K = F(\alpha_1, \dots, \alpha_n)$. We can assume $\alpha = \alpha_1$.

Then $f_1(x) \sim (x - \alpha_2) \cdots (x - \alpha_n)$, and $K = F(\alpha_1, \dots, \alpha_n) = F(\alpha_1)(\alpha_2, \dots, \alpha_n) = F_1(\alpha_2, \dots, \alpha_n)$. □

Now we can address uniqueness of splitting fields.

Proposition 24.1.4 (Uniqueness of splitting fields)**.** *Let $\varphi : F \to \widetilde{F}$ be an isomorphism of fields. Let $f(x) \in F[x]$ and $\widetilde{f}(x) = \varphi(f(x)) \in \widetilde{F}[x]$. Let K be a splitting field of $f(x)$ over F and let $\widetilde{K}$ be a splitting field of $\widetilde{f}(x)$ over $\widetilde{F}$. Then φ can be extended to an isomorphism $K \to \widetilde{K}$.*

$$\begin{tikzcd}
K \arrow[r, dashed, "\varphi"] & \widetilde{K} \\
F \arrow[u, hook] \arrow[r, "\varphi"'] & \widetilde{F} \arrow[u, hook]
\end{tikzcd}$$

Proof. By induction on $n = \deg(f(x))$. The degree 0 case is clear, because there are no roots.

Assume $n \geq 1$. Let $\alpha \in K$ be a root of $f(x)$. Let $F_1 = F(\alpha)$ and $f_1(x) \in F_1[x]$ be as in the lemma. So K is a splitting field of $f_1(x)$ over F_1 and $f(x) = (x - \alpha)f_1(x)$.

Let $p(x) = m_{\alpha,F}(x) \in F[x]$, and $\widetilde{p}(x) = \varphi(p(x)) \in \widetilde{F}[x]$.

$f(\alpha) = 0$, so $p(x) | f(x)$ in $F[x]$, and so $\widetilde{p}(x) | \widetilde{f}(x)$ in $\widetilde{F}[x]$. But $\widetilde{f}(x)$ factors completely in $\widetilde{K}$, so there exists $\widetilde{\alpha} \in \widetilde{K}$ such that $\widetilde{p}(\widetilde{\alpha}) = 0$. Then $\widetilde{p}(x) = m_{\widetilde{\alpha},\widetilde{F}}(x)$. Let $\widetilde{F}_1 = \widetilde{F}(\widetilde{\alpha})$. We then have isomorphisms of F-algebras

$$F_1 \cong F[x]/(p(x)) \xrightarrow{\varphi} \widetilde{F}[x]/(\widetilde{p}(x)) \cong \widetilde{F}_1,$$

sending $\alpha \leftrightarrow \overline{x} \leftrightarrow \overline{x} \leftrightarrow \widetilde{\alpha}$, and extending $\varphi : F \to \widetilde{F}$.

Consider $\widetilde{f_1}(x) = \varphi(f_1(x)) \in \widetilde{F_1}[x]$. The proof follows by induction if we can show that $\widetilde{K}$ is a splitting field for $\widetilde{f}(x)$ over $\widetilde{F_1}$. This follows from the lemma, and we are done. □

Corollary 24.1.5. *Let F be a field and K and K' are two splitting fields for the same $f(x) \in F[x]$. Then there exists an isomorphism $\varphi : K \to K'$ such that $\varphi|_F$ is the identity.*

Proof. Apply the Uniqueness proposition to the identity on F.

$$\begin{tikzcd}
K \arrow[rr, dashed, "\varphi"] & & K' \\
& F \arrow[ul, hook] \arrow[ur, hook] &
\end{tikzcd}$$

□

24.2 Separability

Definition 24.2.1. A polynomial $f(x) \in F[x]$ is *separable* if all its roots in some splitting field are distinct. (This is independent of the choice of splitting field, by uniqueness.)

Lemma 24.2.2. *Let $f(x) \in F[x]$ be separable and K some extension of F. Then all roots of $f(x)$ that are in K, if any, are distinct.*

Proof. Let $\widetilde{K}$ be a splitting field of $f(x)$ over K, the extension, so $\widetilde{K} = K(\alpha_1, \dots, \alpha_n)$. Then $F(\alpha_1, \dots, \alpha_n)$ is a splitting field of $f(x)$ over F.

So the α_i are distinct. □

Definition 24.2.3. Given $f(x) = \sum a_i x^i \in F[x]$, its *derivative* is $f'(x) = \sum i a_i x^{i-1} \in F[x]$.

Recall that if $a \in F$ and $i \in \mathbb{N}$, then $ia = a + \cdots + a$ with i terms, or $(1 + \cdots + 1)a$, where there are i terms in the sum of ones, or $(i \cdot 1)a$. The element $i \cdot 1$ is in the prime field of F, i.e. $\mathbb{Q}$ or $\mathbb{F}_p$, so it may be 0 even if $i \neq 0$.

We have:

- $(f+g)' = f' + g'$
- $(fg)' = f'g + fg'$
- $f(g(x))' = f'(g(x))g'(x)$

Proposition 24.2.4. *Let $f(x) \in F[x]$. Let K be a splitting field. The following are equivalent:*

(i) $f(x)$ is separable.

(ii) $f(x)$ *and* $f'(x)$ *have no common roots in* K*.*

(iii) $1 \in \gcd(f(x), f'(x))$*, with* gcd *taken in the PID* $F[x]$*.*

Proof. (i) ⇒ (ii): Suppose $\exists \alpha \in K$ such that $f(\alpha) = f'(\alpha) = 0$. Then

$$\begin{aligned}
f(x) &= (x-\alpha)g(x) \text{ with } g(x) \in K[x] \\
\Rightarrow f'(x) &= g(x) + (x-\alpha)g'(x) \\
\Rightarrow f'(\alpha) &= g(\alpha) \Rightarrow g(\alpha) = 0 \\
\Rightarrow g(x) &= (x-\alpha)h(x) \\
\Rightarrow f(x) = (x-\alpha)^2 h(x) \text{ in } K[x].
\end{aligned}$$

(ii) ⇒ (i): Suppose $f(x) = (x-\alpha)^2 h(x)$ with $h(x)$ in $k[x]$. Then $f'(x) = 2(x-\alpha)h(x) + (x-\alpha)^2 h'(x)$, so $f'(\alpha) = 0$. Thus inseparability implies common roots.

(iii) ⇒ (ii): Suppose there is $\alpha \in K$ with $f(\alpha) = f'(\alpha) = 0$. Then $x - \alpha$ divides both $f(x)$ and $f'(x)$ in $K[x]$. But then their gcd cannot be 1; if it were, we'd have $1 = r(x)f(x) + s(x)f'(x)$.

The rest of the proof, that (i) and (ii) imply (iii), is left as an exercise. □

Corollary 24.2.5. *Let* $p(x) \in F[x]$ *be irreducible. Then* $p(x)$ *is separable* $\iff$ $p'(x) \neq 0$*.*

Proof. (⇒): If $p'(x) = 0$, then $p(x)$ and $p'(x)$ have common roots (all the roots of $p(x)$).

(⇐): Let $d(x) \in \gcd(p(x), p'(x))$. Suppose that p is not separable; then $d(x)$ is not constant. But $d(x)|p(x)$, so $d(x) \sim p(x)$. Thus $\deg(d(x)) = \deg(p(x))$, but $d(x)|p'(x)$, which has degree one smaller, so this is a contradiction. □

Corollary 24.2.6. *Suppose* $\operatorname{char} F = 0$*. Then any irreducible polynomial is separable.*

Remark. The homework will show that the same result holds if F is a finite field.

Example. $F = \mathbb{F}_p(y)$, the field of rational functions in y.

Let $f(x) = x^p - y \in F[x]$. We know that $f(x)$ is irreducible in $\mathbb{F}_p[x,y] = \mathbb{F}_p[x][y]$. Since it is primitive, it is irreducible in $\mathbb{F}_p(y)[x] = F[x]$, a consequence of Gauss's Lemma.

But $f'(x) = px^{p-1} = 0$, because the coefficients are in $\mathbb{F}_p$, so $f(x)$ is not separable.

Let K be a splitting field of $f(x)$ over F. There is $\alpha \in K$ such that $f(\alpha) = 0$, so $\alpha^p = y$. Then $f(x) = x^p - y = x^p - \alpha^p = (x-\alpha)^p$, because the characteristic of our field is p. So it's *really* not separable.

25 November 24th

25.1 More Separability

Definition 25.1.1. Given $K|F$, an element $\alpha \in K$ is *separable over* F if it is algebraic over F and if $m_{\alpha,F}(x) \in F[x]$ is separable.

Theorem 25.1.2 (Primitive Element Theorem)**.** *Let $K = F(\alpha_1, \alpha_2, \ldots, \alpha_n)$. with all α_i algebraic over F and $\alpha_2, \ldots, \alpha_n$ separable. Then there exists $\gamma \in K$ such that $K = F(\gamma)$.*

Proof. We will separate the proof into two cases: when F is finite, and when F is not finite.

If F is finite, then K is finite as well. Then $K^\times$ is cyclic. (see HW 4 and HW 12). Let γ be a generator of $K^\times$. Then $K = F(\gamma)$ so we're done.

Assume F is infinite. Induction reduces the proof to the case when $n = 2$. Assume that $K = (\alpha, \beta)$, with α, β algebraic and β separable. We will show that $\gamma = \alpha + \lambda\beta$ is primitive for a suitable $\lambda \in F$. Specifically, suppose γ is not primitive; we'll show then that λ must belong to a certain finite subet $S \subseteq F$; since F is infinite, this will suffice. Let γ be not primitive.

First note $F(\alpha, \beta) = F(\gamma, \beta)$, because $\gamma \in F(\alpha, \beta)$ and $\alpha \in F(\gamma, \beta)$. γ is not primitive, so $F(\gamma) < F(\gamma, \beta)$. Thus $\deg_{F(\gamma)} \beta \geq 2$. Let $f(x) = m_{\alpha,F}(x)$, $g(x) = m_{\beta,F}(x)$ and $d(x) = m_{\beta,F(\gamma)}(x)$.

Then β satisfies $g(\beta) = 0$ and $f(\gamma - \lambda\beta) = 0$, i.e. β is a root of $g(x)$ and of $h(x) = f(\gamma - \lambda x)$, both in $F(\gamma)[x]$. Thus $d(x)$ divides both $g(x)$ and $h(x)$ in $F(\gamma)[x]$. Let $\widetilde{K}$ be an extension of K where $g(x)$ splits completely. Hence, $d(x)$ does as well. $\deg d(x) \geq 2$, so d has at least two distinct roots. Thus $g(x)$ and $h(x)$ have at least two distinct common roots. At least one is different from β; call it β'.

Now $h(\beta') = 0$, so $f(\gamma - \lambda\beta') = 0 \Rightarrow \gamma - \lambda\beta' = \alpha'$ for some root α' of $f(x)$, so $\gamma = \alpha' + \lambda\beta' \Rightarrow \alpha + \lambda\beta = \alpha' + \lambda\beta'$. Thus $\lambda = \frac{\alpha-\alpha'}{\beta-\beta'}$.

Thus λ belongs to the set $S = \{\lambda \in F \mid \lambda = \frac{\alpha-\alpha'}{\beta-\beta'}, \alpha'$ is a root of $f(x), \beta'$ is a root of $g(x), \alpha', \beta' \in \widetilde{K}\}$, which is finite, so we are done. □

25.2 Algebraic independence

Definition 25.2.1. Let $K|F$ be a field extension. Elements $\alpha_1, \ldots, \alpha_n \in K$ are *algebraically independent*, or a.i., over F if there exists no nonzero polynomial $f(x_1, \ldots, x_n) \in F[x_1, \ldots, x_n] \setminus \{0\}$ such that $f(\alpha_1, \ldots, \alpha_n) = 0$. Equivalently, the unique morphism of F-algebras $\varphi : F[x_1, \ldots, x_n] \to K$ such that $\varphi(x_i) = \alpha_i$ is injective.

This is a multivariate version of the definition of transcendental elements. In this case, it is necessarily true that $\dim_F K = \infty$, if $n \geq 1$, and φ extends to an injective homomorphism $F(x_1, \ldots, x_n) \to K$ such that $F(x_1, \ldots, x_n) \cong F(\alpha_1, \ldots, \alpha_n) \leq K$.

Proposition 25.2.2. *Given $K|F$ and $\alpha_1, \ldots, \alpha_n \in K$, the following are equivalent:*

(i) $\alpha_1, \ldots, \alpha_n$ are algebraically independent.

(ii) Each α_i is transcendental over $F(\alpha_1, \ldots, \hat{\alpha}_i, \ldots, \alpha_n)$.

(iii) Each α_i is transcendental over $F(\alpha_1, \ldots, \alpha_{i-1})$.

Proof. (i) ⇒ (ii): Follows easily from the definition, so we omit.

(ii) ⇒ (iii): Clear since $F(\alpha_1, \dots, \alpha_{i-1}) \leq F(\alpha_1, \dots, \hat{\alpha_i}, \dots, \alpha_n)$.

(iii) ⇒ (i): Do it for $n = 2$. Given α transcendental over F and β transcendental over $F(\alpha)$, we want to show that $\{\alpha, \beta\}$ is algebraically independent. Choose $f(x, y) \in F[x, y] \setminus \{0\}$. Write $f(x, y) = \sum_{i,j} a_{ij} x^i y^j$, with some $a_{nm} \neq 0$.

So we have

$$f(x, y) = \sum_j \left(\sum_i a_{ij} x^i \right) y^j,$$

with $a_j(x)$ defined to be $\sum_i a_{ij} x^i$. Now $a_m(x) = \sum_i a_imx^i \in F[x] \setminus \{0\}$, since $a_{nm} \neq 0$. Thus $a_m(\alpha) \neq 0$, since α is transcendental over F. Thus $f(\alpha, y) = \sum_j a_j(\alpha) y^j \in F(\alpha)[y] \setminus \{0\}$ is nonzero because it has at least one nonzero coefficient. Thus $f(\alpha, \beta) \neq 0$. □

Definition 25.2.3. $K|F$ is *purely transcendental* if there are $\alpha_1, \dots, \alpha_n$ algebraically independent over F such that $K = (\alpha_1, \dots, \alpha_n)$. In particular, each α_i is transcendental over F.

Definition 25.2.4. Let $K|F$ be a field extension. A set of elemetns $\alpha_1, \dots, \alpha_d \in K$ form a (finite) *transcendence basis* for $K|F$ if

(i) $\alpha_1, \dots, \alpha_d$ are a.i.

(ii) $K|F(\alpha_1, \dots, \alpha_d)$ is algebraic.

Proposition 25.2.5. *Let $K|F$ be a finitely generated extension: $K = F(S)$ for some finite $S \subseteq K$. Then S contains a transcendence basis for $K|F$.*

Proof. Apply the last criterion from Proposition 25.2.2 to elements of S repeatedly. □

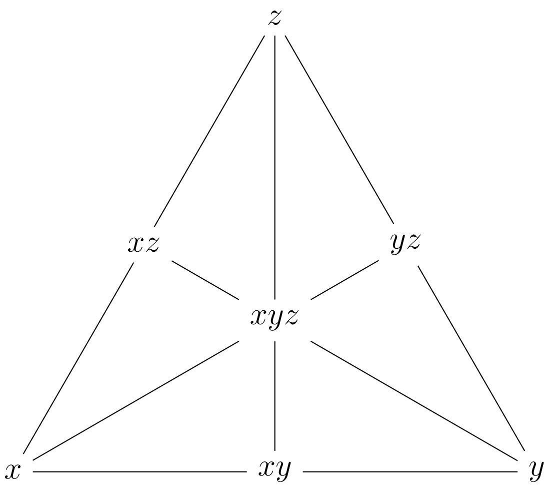
Now we adopt as our goal to show that any two bases have the same size d, or to show that the following definition is well-defined.

Definition 25.2.6. Such a size d is called the *transcendence degree* of $K|F$. Sometimes denoted $d = \mathrm{td}_F K$.

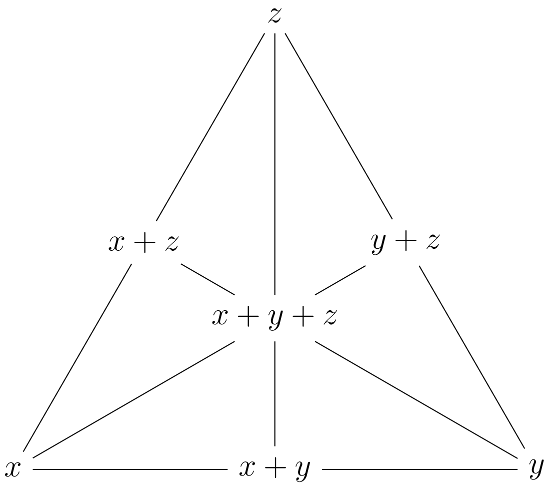
Remark. $d = 0$ is the same thing as saying that the empty set is a transcendence basis, or that the extension is algebraic.

Example. Let $K = F(x, y, z)$, the field of rational functions in three variables. Clearly x, y, z are a.i., since evaluation in this case is simply the identity map, which must have trivial kernel.

We claim that xy, yz, and zx are a.i., which is not easy (see Homework 12). The algebraic dependences among x, y, z, xy, yz, xz, xyz are summarized in the diagram below; any two are a.i., and when three are algebraically dependent they are placed in a line.



This is in contrast with *linear dependencies* among $x, y, z, x+y, x+z, y+z, x+y+z$. Both are one line off from the Fano plane, unless the field has characteristic 2. In characteristic 2 this really just is the Fano plane.



This leads, as will come next time, quite naturally to the concept of a matroid.

Example. $\mathbb{R}[S^1] = \mathbb{R}[x,y]/(x^2+y^2-1)$, the ring of polynomial functions on S^1. We claim first of all that $x^2+y^2-1 \in \mathbb{R}[x,y]$ is irreducible. To prove this, view it as $x^2+y^2-1 \in \mathbb{R}[y][x] \subseteq \mathbb{R}(y)[x]$. Let $F = \mathbb{R}(y)$, and $R = \mathbb{R}[y]$. It is primitive (monic), so it suffices to check that it is irreducible in $F[x]$. If not, being of degree 2, it would have a root α in F. Being monic, α has to be in R (using the rational root theorem, because R is a UFD).

Write $\alpha = \sum a_i y^i$, so $\left(\sum a_i y^i\right)^2 + y^2 - 1 = 0$ in $\mathbb{R}[y]$. But then $a_0^2 - 1 = 0$, $2a_0a_1 = 0$, $a_1^2 + 1 = 0$, which is a contradiction because the characteristic of $\mathbb{R}$ is not 2. Thus x^2+y^2 is irreducible.

Thus $\mathbb{R}[S^1]$ is an integral domain. Let $\mathbb{R}[S^1]$ be its field of fractions; let $\alpha = \overline{x}$ and $\beta = \overline{y} \in \mathbb{R}[S^1] \subset \mathbb{R}(S^1)$. Then $\mathbb{R}[S^1] = \mathbb{R}[\alpha,\beta]$ and $\mathbb{R}(S^1) = \mathbb{R}(\alpha,\beta)$. Thus $\mathbb{R}(S^1)|\mathbb{R}(\alpha)$ is algebraic, because β satisfies $\alpha^2+\beta^2-1=0$. We claim that α is transcendental over $\mathbb{R}$.

Proof. If not, $\exists f(x) \in \mathbb{R}[x] \setminus \{0\}$ such that $f(\alpha) = 0$. Then $\overline{f(x)} = \overline{0}$ in $\mathbb{R}[x,y]/(x^2+y^2-1)$, so $x^2+y^2-1|f(x)$ in $\mathbb{R}[x,y]$. For each $(x_0, y_0) \in S^1$, $f(x_0) = 0$, so $f(x)$ has infinitely many roots and is thus 0, which is a contradiction. □

So we have

$$\begin{tikzcd}
\mathbb{R}(S^1) \arrow[d, dash, "\text{alg}"] \\
\mathbb{R}(\alpha) \arrow[d, dash, "\text{p. trans}"] \\
\mathbb{R}
\end{tikzcd}$$

and in particular, $\{\alpha\}$ is a transcendence basis for $\mathbb{R}(S^1)|\mathbb{R}$, and the transcendence degree of $\mathbb{R}(S^1)$ over $\mathbb{R}$ is 1. This is a special case of Noether's normalization lemma.

26 December 1st

26.1 Matroids

Definition 26.1.1. Let S be a finite set, let 2^S be the set of subsets of S, and let $\sigma : 2^S \to 2^S$ be a function. Then σ is a *closure operator* if for all $A, B, \in 2^S$, we have

(1) $A \subseteq \sigma(A)$

(2) $A \subseteq B \Rightarrow \sigma(A) \subseteq \sigma(B)$

(3) $\sigma\sigma(A) = \sigma(A)$.

σ is a *matroid* if in addition we have

(4) If $b \in \sigma(A \cup \{a\})$, but $b \notin \sigma(A)$, then $a \in \sigma(A \cup \{b\})$. This is known as them *MacLane-Steinitz exchange axiom.*

The following propositions will give a couple of examples of matroids, some more familiar than others.

Proposition 26.1.2. *Let V be a vector space over a field F and S a finite subset of V. Given $A \in 2^S$, define $\sigma(A) = \{b \in S \mid b \in \mathrm{Span}_F(A)\}$. Then σ is a matroid.*

Proof. (1)-(3) are clear. For (4), if b is spanned by $A \cup \{a\}$ but not by A, then we can write b as a linear combination of vectors in $A \cup \{a\}$ with nonzero coefficient of a. Thus we can solve for a and write it in terms of $A \cup \{b\}$, just as the axiom says. □

Proposition 26.1.3. *Let $K|F$ be a field extension and S a finite subset of K. Given $A \in 2^S$, define $\sigma(A) = \{\beta \in S \mid \beta$ is algebraic over $F(A)\}$. Then σ is a matroid.*

Proof. (1)-(2) are clear.

(3) uses transitivity of algebraic extensions, as seen in homework 12.

For (4), let $\widetilde{F} = F(A)$. We are given that β is algebraic over $F(A \cup \{\alpha\}) = \widetilde{F}(\alpha)$ but not over $F(A) = \widetilde{F}$. We need that α is algebraic over $F(A \cup \{\beta\}) = \widetilde{F}(\beta)$. Since β is algebraic, there exists $f(x) \in \widetilde{F}[\alpha][x] \setminus \{0\}$ with $f(\beta) = 0$. But $f(x) \in \widetilde{F}[\alpha][x] = \widetilde{F}[x][\alpha]$, so

$$f(x) = \sum_i g_i(x)\alpha^i,$$

with $g_i(x) \in \widetilde{F}[x]$. $f(x)$ is nonzero, so at least one of the coefficients $g_i(x)$ are nonzero. Then $g_i(\beta) \neq 0$, since β is not algebraic over $\widetilde{F}$.

So consider $g(x) = \sum_i g_i(\beta)x^i \in \widetilde{F}[\beta][x]$. At least one coefficient is nonzero. Then $g(\alpha) = f(\beta) = 0$, so α is aglebraic over $\widetilde{F}(\beta)$, and we are done. □

Definition 26.1.4. Given a matroid σ on a finite set S:

- an element $s \in S$ is *spanned* by $A \in 2^S$ if $s \in \sigma(A)$;
- a subset $A \in 2^S$ is *spanning* if $\sigma(A) = S$;
- a subset $A \in 2^S$ is *independent* if no $a \in A$ is spanned by $A \setminus \{a\}$, i.e. $a \notin \sigma(A \setminus \{a\})$;
- a subset $A \in 2^S$ is a *basis* if it is independent and spanning.

Proposition 26.1.5. *1. A subset is a basis if and only if it is a maximal independent subset, which is true if and only if it is a minimal spanning subset. In particular, bases exist.*

2. All bases have the same cardinality.

Proof. Matroid theory. Omitted, but not difficult. □

26.2 Back to field extensions

Let $K = F(S)$ be a finitely generated extension of F. Let σ_S be the matroid in S of Proposition 26.1.3. Let $A = \{\alpha_1, \dots, \alpha_d\} \in 2^S$. Then

- A is independent if and only if $\alpha_1, \dots, \alpha_d$ are algebraically independent.
- A is spanning if and only if $K|F(\alpha_1, \dots, \alpha_d)$ is algebraic.
- A is a basis if and only if $F(A)|F$ is purely transcendental, and $\{\alpha_1, \dots, \alpha_d\}$ is a transcendence basis.

In particular, transcendence bases exist because matroid bases exist, so there exist transcendence bases contained in S and all such have the same cardinality.

Fact 26.2.1 (Fun unproven extra exercise)**.** *If $K = F(T)$ for some other T then* $\operatorname{rank}(\sigma_S) = \operatorname{rank}(\sigma_T)$.

26.3 Finitely generated algebras

Let $R \subseteq S$ be commutative rings. Given $\alpha_1, \dots, \alpha_n \in S$, $R[\alpha_1, \dots, \alpha_n]$ denotes the smallest subring of S containing R and $\alpha_1, \dots, \alpha_n$. It is the image of the unique homomorphism of R-algebras $\varphi : R[x_1, \dots, x_n] \to S$ such that $\varphi(x_i) = \alpha_i$.

If there are elements $\alpha_1, \dots, \alpha_n \in S$ such that $S = R[\alpha_1, \dots, \alpha_n]$, we say that S is *finitely generated* as an R-algebra.

In this case, φ is onto, so $S \cong R[x_1, \dots, x_n]/I$ for some ideal I. This allows us to prove the following proposition.

Proposition 26.3.1. *Suppose that R is noetherian and S is a finitely generated R-algebra. Then S is noetherian too.*

Proof. Hilbert's Basis Theorem tells us that polynomial rings over noetherian rings are noetherian; then quotients of that polynomial ring are still noetherian. □

Remark. We may also view S as an R-module. We already knew that S was noetherian as a module; this says that S is noetherian as a ring. It is parallel to the familiar result about finitely generated noetherian modules.

So if S is finitely generated as an R-module, then S is certainly finitely generated as an R-algebra; we're allowed all polynomials to generate, instead of just linear ones. But the converse does not hold. $R[x]$ is finitely generated as an R-algebra, but not as an R-module.

Now let $R \le S \le T$ be commutative rings. Here are some more fun results:

- Suppose S is f.g. as an R-algebra and T is f.g. as an S-algebra. Then T is f.g. as an R-algebra. (If $S = R[\alpha_1, \dots, \alpha_m]$, and $T = S[\beta_1, \dots, \beta_n]$, then $T = R[\alpha_1, \dots, \alpha_m, \beta_1, \dots, \beta_n]$.

- Suppose that T is f.g. as an R-algebra. Is T f.g. as an S-algebra? Yes. The same generators will do. Is S finitely generated as an R-algebra? Not necessarily. This is harder to see. As one counterexample, let $R = F$, a field. Let $T = F[x, y]$, which is finitely generated over R. Let $S = F[y, xy, x^2y, x^3y, \dots] = \{f(x, y) \in F[x, y] \mid f(x, 0) = f(0, 0)\}$. It is an exercise to see that S is not f.g. as an F-algebra.

However, we have the following proposition.

Proposition 26.3.2 (Artin-Tate)**.** *Assume we have a chain of commutative rings $R \le S \le T$. Suppose R is noetherian, T is f.g. as an R-algebra, and T is f.g. as an S-module. Then S is f.g. as an R-algebra.*

A way to make the statement more intuitive is to think that "if S is pretty similar to T, then S is finitely generated as well."

Proof. Write $T = R[\alpha_1, \ldots, \alpha_m]$. Moreover, $T = S\beta_1 + \cdots + S\beta_n$. Then $\alpha_i = \sum_j s_{ij}\beta_j$, for some $s_{ij} \in S$, and $\beta_i\beta_j = \sum_k s_{ijk}\beta_k$, for some $s_{ijk} \in S$.

Let $A = \{s_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n\} \cup \{s_{ijk} \mid 1 \leq i, j, k \leq n\}$, a finite set, and let $S_0 = R[A]$, a finitely generated R-subalgebra of S.

R is noetherian, so S_0 is noetherian as well. Note that our expressions with s_{ijk}'s allow us to turn polynomial expressions into linear expressions in the β_k's with only elements of S_0 as coefficients. Thus T is f.g. as a S_0 module, since $T = S_0\beta_1 + \cdots + S_0\beta_n$. So T is noetherian as an S_0-module.

Now $S_0 \leq S \leq T$, so S is an S_0-submodule of T, so S is finitely generated as an S_0-module and thus as an S_0-algebra. But S_0 was also finitely generated as an R-algebra, so we are done, and S is f.g. as an R-algebra. □

Consider now a field extension $F \leq K$. There are in fact three types of finite generation.

(1) K is f.g. as an F-module, which holds if and only if K is finite dimensional as a vector space over F, or $K|F$ is finite.

(2) K is f.g. as an F-algebra, true $\iff$ $K = F[\alpha_1, \ldots, \alpha_n]$, using brackets because the RHS in principal is only a ring.

(3) K is f.g. as as a field extension of F, which holds $\iff$ $K = F(\alpha_1, \ldots, \alpha_n)$.

Clearly (1) ⇒ (2) ⇒ (3). Do the converses hold?

Does (3) ⇒ (2)? No. $F(x)$ is a f.g. field extension of F, but it is not f.g. as an algebra. There is no $\alpha_1, \ldots, \alpha_n \in F(x)$ such that $F(x) = F[\alpha_1, \ldots, \alpha_n]$. We will see the proof of this next time, in the context of a more general theorem.

Does (2) ⇒ (1)? Yes, as it turns out! This is Zariski's Theorem.

Theorem 26.3.3 (Zariski)**.** *If K is f.g. as an F-algebra, then* $\dim_F K < \infty$.

Proof next time.

27 December 3rd

27.1 Zariski's Theorem and Nullstellensätze

Theorem 27.1.1 (Zariski)**.** *If K is f.g. as an F-algebra, then* $\dim_F K < \infty$.

Proof. K is f.g. as an F-algebra, so K is f.g. as a field extension of F. So it suffices to show that $K|F$ is algebraic, in which case it will be finite, so we are done. We know that there exist $\alpha_1, \ldots, \alpha_d \in K$ such that

$$\begin{tikzcd}
K \arrow[d, dash] \\
F(\alpha_1, \dots, \alpha_d) \arrow[d, dash] \\
F
\end{tikzcd}$$

with $K|F$ algebraic and $F(\alpha_1, \dots, \alpha_d)|F$ purely transcendental. We want $d = 0$. We will use Artin-Tate.

Since K is f.g. as an F-algebra, K is f.g. as an $F(\alpha_1, \dots, \alpha_d)$-module because $K|F(\alpha_1, \dots, \alpha_d)$ is algebraic and f.g. as a field extension. By Artin-Tate, $F(\alpha_1, \dots, \alpha_d)$ is f.g. as an F-algebra. But we use the lemma below:

Lemma 27.1.2. *If $d \geq 1$, $F(x_1, \dots, x_d)$ is not f.g. as an F-algebra.*

to get that $d = 0$. We therefore need only prove the lemma.

Proof. Suppose $F(x_1, \dots, x_d)$ is f.g. as an F-algebra. Then $F(x_1, \dots, x_d) = F[\alpha_1, \dots, \alpha_n]$ for some rational functions $\alpha_1, \dots, \alpha_n$ (note that these are different α's from those used above). Write $\alpha_i = f_i/g_i$ with $f_i, g_i \in F[x_1, \dots, x_d]$. If all g_i are units in $F[x_1, \dots, x_d]$, then $\alpha_i \in F[x_1, \dots, x_d] \Rightarrow F(x_1, \dots, x_d) = F[x_1, \dots, x_d]$, which is clearly impossible if $d \geq 1$, because in that case $F[x_1, \dots, x_d]$ is not a field. So not all g_i's can be units. It follows that $1 + g_1 \cdots g_n$ has positive degree. Hence it has an irreducible factor $p \in F[x_1, \dots, x_d]$.

Now $\frac{1}{p} \in F(x_1, \dots, x_d) = F[\alpha_1, \dots, \alpha_n]$, so there exists $N \geq 0$ such that $\frac{(g_1 \cdots g_{)}^N}{p} \in F[x_1, \dots, x_d]$. Thus $p|(g_1 \cdots g_n)^N$ in $F[x_1, \dots, x_d]$, so $p|g_i$ for some i. But $p|(1 + g_1 \cdots g_n)$, so $p|1$, a contradiction. □

□

Corollary 27.1.3 (Weak Nullstellensatz)**.** *Note that the name literally means "zero locus theorem."*

Let F be a field and R a f.g. F-algebra (everything is commutative), and let M be a maximal ideal of R with $K = R/M$. Then K is a finite field extension of F.

Proof. The homomorphism $F \hookrightarrow R \twoheadrightarrow R/M = K$ is injective, because F is a field. R is f.g. as an F-algebra, so $R = F[\alpha_1, \dots, \alpha_n]$, and $K = F[\overline{\alpha_1}, \dots, \overline{\alpha_n}]$. Thus K is f.g. as an F-algebra, so Zariski tells us that $\dim_F K < \infty$, and we are done. □

Given a commutative ring R, its *maximal spectrum* is the set $\mathrm{Spec}_m(R) = \{M \subseteq R \mid M \text{ maximal ideal of } R\}$. If R and S are F-algebras, with F a field, let $\mathrm{Alg}_F(R, S)$ denote the set of F-algebra homomorphisms $\varphi : R \to S$. Then we have a couple of helpful corollaries.

Corollary 27.1.4. *Suppose F is algebraically closed and R is a f.g. F-algebra. Then $\mathrm{Alg}_F(R,S) \to \mathrm{Spec}_m(R)$ given by $\varphi \mapsto \ker\varphi$ is a bijection.*

Proof. We first need show that it is well-defined. Take $\varphi \in \mathrm{Alg}_F(R,F)$. Then $\varphi|_F = \mathrm{id}_F$, so $\varphi : R \to F$ is onto and $\ker\varphi$ is maximal, and thus an element of $\mathrm{Spec}_m(R)$.

Now we check injectivity. Taking $\varphi, \psi \in \mathrm{Alg}_F(R,F)$, suppose that $\ker\varphi = \ker\psi$. Let $a \in R$; then $a - \varphi(a)\cdot 1 \in \ker\varphi$, since $\varphi(1) = 1$. Thus $a - \varphi(a)\cdot 1 \in \ker\psi$, so $\psi(a) = \varphi(a)$, since $\psi(1) = 1$. Thus $\varphi = \psi$.

Lastly, surjectivity. Take $M \in \mathrm{Spec}_m(R)$. Let $K = R/M$. Then K is a field and $K|F$ is a finite extension. $K|F$ is algebraic, and F is algebraically closed, so $K = F$. Hence we have a homomorphism of F-algebras $\varphi : R \to R/M = K = F$, and $\ker\varphi = M$. □

Definition 27.1.5. Given $S \subseteq F[x_1,\dots,x_n]$, let $\mathcal{Z}(S) = \{a \in F^n \mid f(a) = 0\ \forall\, f \in S\}$ be the *zero set* or *zero locus* of S.

Now let I be an ideal of $F[x_1,\dots,x_n]$ and $R = F[x_1,\dots,x_n]/I$. Given $a = (a_1,\dots,a_n) \in F^n$, consider the homomorphism of F-algebras

$$\varphi_a : F[x_1,\dots,x_n] \to F; \varphi_a(x_i) = a_i \Rightarrow \varphi_a(f) = f(a).$$

So when does φ_a factor through R?

$$\begin{tikzcd}
F[x_1,\dots,x_n] \arrow[r] \arrow[d, two heads] & F \\
R \arrow[ur, dashed, "\hat{\varphi}_a"'] &
\end{tikzcd}$$

Why, precisely when $\varphi_a(f) = 0$ for every $f \in I$, of course!

Thus $f(a) = 0$ for all $f \in I$, which is true if and only if $a \in \mathcal{Z}(I)$. In this situation, there is a bijection $\mathcal{Z}(I) \to \mathrm{Alg}_F(R,F)$ with $a \mapsto \hat{\varphi_a}$,

Corollary 27.1.6. *F algebraically closed, I an ideal of $F[x_1,\dots,x_n]$, $R = F[x_1,\dots,x_n]/I$. There is a bijection $\mathcal{Z}(I) \to \mathrm{Spec}_m(R)$, with $a \mapsto \ker\hat{\varphi_a}$.*

Proof. Combine the previous bijections. □

Corollary 27.1.7 (unproven but not forgotten)**.** *Suppose that F is algebraically closed and R is a f.g. F-algebra. Let $f \in R$. If f is not nilpotent, there exists a maximal ideal M of R that does not contain f.*

Proof. See Exercises 3 & 4 in the final. The main tool is that of *localizations,* ~~who knows what those are~~. □

Let F be a field. Define

$$\{\text{subsets of } F^n\} \rightleftharpoons \{\text{subsets of } F[x_1, \dots, x_n]\},$$

with rightward map $\mathcal{I}$ and leftward map $\mathcal{Z}$. In particular, for $S \subseteq F[x_1, \dots, x_n]$, $\mathcal{Z}(S) = \{a \in F^n \mid f(a) = 0 \forall f \in S\}$, and for $A \subseteq F^n$, $\mathcal{I}(A) = \{f \in F[x_1, \dots, x_n] \mid f(a) = 0 \forall a \in S\}$. Both domains are ordered by inclusion.

Proposition 27.1.8. *(1) $\mathcal{I}$ and $\mathcal{Z}$ are order-reversing.*

(2) For any $A \subseteq F^n$, $A \subseteq \mathcal{Z}\mathcal{I}(A)$. For any $S \subseteq F[x_1, \dots, x_n]$, $S \subseteq \mathcal{I}\mathcal{Z}(S)$.

(3) $\mathcal{I}\mathcal{Z}\mathcal{I} = \mathcal{I}$ and $\mathcal{Z}\mathcal{I}\mathcal{Z} = \mathcal{Z}$.

(4) $\mathcal{Z}$ and $\mathcal{I}$ induce inverse bijections

$$\operatorname{im} \mathcal{Z} \rightleftharpoons \operatorname{im} \mathcal{I}.$$

Remark. We say $\mathcal{I}$ and $\mathcal{Z}$ form a *Galois connection,* because (1) and (2) hold. ((3) and (4) follow.)

Definition 27.1.9. A subset $A \subseteq F^n$ is *algebraic* if it is in the image of $\mathcal{Z}$, or if it is the set of solutions of a system of polynomial equations, which may be infinite.

Such sets are in bijection with the subsets of $F[x_1, \dots, x_n]$ in the image of $\mathcal{I}$. What are those subsets?

The answer is provided by the **Strong Nullstellensatz**. But first! A final definition.

Definition 27.1.10. Let R be a commutative ring and I an ideal of R. The *radical* of I is

$$\sqrt{I} = \{a \in R \mid \exists n \in \mathbb{N}, a^n \in I\}.$$

Theorem 27.1.11 (Strong Nullstellensatz)**.** *Let F be algebraically closed. Let I be an ideal of $F[x_1, \dots, x_n]$. Then $\mathcal{I}\mathcal{Z}(I) = \sqrt{I}$.*

Proof. Let $R = F[x_1, \dots, x_n]/I$. It is a f.g. F-algebra. Consider the projection $F[x_1, \dots, x_n] \twoheadrightarrow R$ with $f \mapsto \overline{f}$. Recall that for $a \in F^n$, we have

$$\begin{tikzcd}
F[x_1, \dots, x_n] \arrow[r, "\varphi_a"] \arrow[d, two heads] & F \\
R \arrow[ur, "\hat{\varphi}_a"'] &
\end{tikzcd}$$

with $\varphi_a(f) = f(a)$.
Now:

$$\begin{aligned} f \in \mathcal{I}\mathcal{Z}(I) &\iff f(a) = 0 \forall a \in \mathcal{Z}(I) \\ &\iff \varphi_a(f) = 0 \forall a \in \mathcal{Z}(I) \\ &\iff \hat{\varphi}_a(\overline{f}) = 0 \forall a \in \mathcal{Z}(I) \\ &\overset{\text{Cor27.1.6}}{\iff} \overline{f} \in \text{ every max. ideal of } R \\ &\overset{\text{Cor27.1.7}}{\iff} \overline{f} \text{ nilpotent} \end{aligned}$$

which holds if and only if there exists $n \in \mathbb{N}$ with $f^n \in I \iff f \in \sqrt{I}$. □

Definition 27.1.12. An ideal I is *radical* if $I = \sqrt{I}$.

Corollary 27.1.13. *If F is alg. closed, the image of $\mathcal{I}$ consists precisely of the radical ideals of $F[x_1, \ldots, x_n]$.*

Proof. If I is radical, then $I = \sqrt{I} = \mathcal{I}\mathcal{Z}(I) \in \operatorname{im} \mathcal{I}$. For any $A \subseteq F^n$, $\mathcal{I}(A)$ is an ideal of $F[x_1, \ldots, x_n]$. In addition, $\mathcal{I}(A) = \mathcal{I}\mathcal{Z}\mathcal{I}(A) = \sqrt{\mathcal{I}(A)}$. □