MATH70035 ALGEBRA 3

ALESSIO CORTI, NOTES TYPED BY JIEWEI XIONG

Contents

Week 1, lecture 1, 4 October 2024

Broadly there are two parts to this module: I. rings and modules, II. matrix groups.

1. Rings

Example 1.0.1. (1) $\mathbb{Z}$, integers.
(2) $R[x] \coloneqq \{a_0x^n + a_1x^{n-1} + \cdots + a_n : a_0, \ldots, a_n \in R\}$, e.g. $\mathbb{Q}[x]$.
(3) $\mathbb{Z}/m\mathbb{Z}$ where $m \in \mathbb{N}$.
(4) $\mathbb{Z}[i] \coloneqq \{a + bi : a, b \in \mathbb{Z}, i^2 = -1\}$, Gaussian integers.
(5) $R[x_1, \ldots, x_n] = R[x_1, \ldots, x_{n-1}][x_n] = \cdots$.
(6) Monoid rings $R[P]$ where P is a monoid (group without inverses, e.g. $\mathbb{N}$).
(7) $M_{n\times n}(R)$, a noncommutative ring.
(8) $\mathbb{Q}[i, j, k] = \{a + bi + cj + dk : a, b, c, d \in \mathbb{Q}, i^2 = j^2 = k^2 = -1, ij = k, jk = i, ki = j\}$, another noncommutative one.
(9) Weyl algebra $K[x, \delta]$, generated by differentiation operator δ and multiplication by x, with identity (by product rule) $\delta x - x\delta = 1$.
(10) Heisenberg algebra $\mathbb{C}[x, p] = xp - px = i\hbar$.

In ring theory we abstract from these examples and study phenomena that apply to all.

Week 2, lecture 1, 7 October 2024

Definition 1.0.2. A *monoid* is a triple $(P, \cdot, 1)$ where P is a set and $\cdot$ a binary operation $P \times P \to P$ such that $\cdot$ is associative and $a \cdot 1 = 1 \cdot a = a \ \forall a \in P$.

Definition 1.0.3. A *ring* is a quintuple $(R, +, \cdot, 0, 1)$ where R is a set, $+, \cdot$ binary operations $R \times R \to R$ such that $(R, +, 0)$ is an abelian group and $(R, \cdot, 1)$ is a monoid with distributivity: $(a + b)c = ac + bc$ and $c(a + b) = ca + cb$.

Definition 1.0.4. Let R, S be rings. A function $f : R \to S$ is a *ring homomorphism* if $f : (R, +, 0) \to (S, +, 0)$ is a group homomorphism and $f : (R, \cdot, 1) \to (S, \cdot, 1)$ is a monoid homomorphism, i.e.

$$\forall a, b \in R, \ f(ab) = f(a)f(b) \text{ and } f(1) = 1.$$

Example 1.0.5. Let R be a ring and P a monoid, then one has a *monoid ring*, denoted $R[P]$, with elements formal linear combinations of (a, p) (or simply ap) where $a \in R, \ p \in P$. For instance, with $P = G$ and $R = \mathbb{Z}$ one has the group ring / group algebra $\mathbb{Z}[G]$. If there is a group homomorphism

$\varphi : G \to H$ then one can write a map $\widetilde{\varphi} : \mathbb{Z}[G] \to \mathbb{Z}[H]$ defined by $\sum_{g\in G} a_g g \mapsto \sum_{g\in G} a_g \varphi(g)$, which we claim is a ring homomorphism.

Week 2, lecture 2, 8 October 2024

Definition 1.0.6. If $f : R \to S$ is a ring homomorphism, then $\ker f \coloneqq \{r \in R : f(r) = 0\}$ is the *kernel* of f.

Definition 1.0.7. A *left-sided ideal* of a ring R is a subset $I \subseteq R$ such that I is a subgroup with respect to addition and $\forall r \in R, a \in I : ra \in I$.

The definitions for a *right-sided ideal* and a *two-sided ideal* are as expected.

Lemma 1.0.8. $\ker f$ is not a subring of R but is a two-sided ideal.

Proof. Indeed,

(1) $a, b \in I \implies f(a-b) = f(a) - f(b) = 0 - 0 = 0$.
(2) $r \in R, a \in I \implies f(ra) = f(r)f(a) = f(r) \cdot 0 = 0$ and similarly for ar.

□

Example 1.0.9. Let $R = \mathbb{Z}$, then all ideals of R are principal, i.e.

$$\forall I \lhd R,\ \exists m \in R : I = (m) \coloneqq \{rm : r \in R\}.$$

We call these rings *principal ideal domains*, or PID. One can see this by the division algorithm. Let $0 \neq I \lhd R$, then pick the smallest $0 < m \in I$ by the well-ordering principle. We claim $I = (m)$. Indeed, suppose $a \in I$, then $\exists q, r : a = qm + r$ where $0 \leq r < m$, but then $r = a - qm \in I$, hence r must be 0, i.e. $a = qm \in (m)$. This proves that every ring with the division algorithm has this property, i.e.

Proposition 1.0.10. Every euclidean domain is a PID.

Definition 1.0.11. A ring R is an *integral domain* or simply *domain* if $\forall u, v \in R,\ uv = 0 \implies$ either $u = 0$ or $v = 0$.

Definition 1.0.12. A *euclidean domain* is a domain R with a function $\deg : R\backslash\{0\} \to \mathbb{N}$ such that $\forall a, b \in R$ where $b \neq 0,\ \exists q, r : a = qb + r$ where either $\deg r < \deg b$ or $r = 0$.

Example 1.0.13.
(1) If k is a field, then the ring $k[X]$ is a euclidean domain where the degree function is the usual degree of a polynomial.
(2) The gaussian integers form a euclidean domain with $\deg : a + bi \mapsto a^2 + b^2$, usually called the norm.

Week 2, lecture 3, 11 October 2024

(3) The ring of *Hurwitz quaternions* $R = \mathbb{Z}[i, j, k] + \frac{1+i+j+k}{2}\mathbb{Z}$ with $\deg : \alpha \mapsto N(\alpha)$ where for $\alpha = x + iy + jz + kt$ one defines $N(\alpha) \coloneqq \alpha\overline{\alpha} = x^2 + y^2 + z^2 + t^2$ (where $\overline{\alpha} = x - iy - jz - kt$; the definition helps to see this norm is multiplicative) is a noncommutative euclidean domain (for all nonzero $\alpha, \beta \in R,\ \exists \gamma, \rho \in R : \alpha = \beta\gamma + \rho$ with $N(\rho) < N(\beta)$ and $\exists \gamma', \rho' \in R : \alpha = \gamma'\beta + \rho'$ with $N(\rho') < N(\beta)$). The key point to prove this is $\forall \zeta \in \mathbb{R}[i, j, k],\ \exists \alpha \in R : N(\zeta - \alpha) < 1$. This would not be true if $R = \mathbb{Z}[i, j, k]$; consider $\zeta = \frac{1+i+j+k}{2}$.

1.1. **Homomorphism theorems.**

Lemma 1.1.1. A ring homomorphism $\varphi : R \to S$ is injective $\iff \ker\varphi = (0)$.

Proof.

$$\begin{aligned}(\varphi(a) = \varphi(b) \implies a = b) &\iff (\varphi(a-b) = 0 \implies a = b)\\ &\iff ((a-b) \in \ker\varphi \implies a - b = 0)\\ &\iff \ker\varphi = (0).\end{aligned}$$

□

Example 1.1.2. Let G, H be groups and $\varphi : G \to H$ a group homomorphism. Consider $\widetilde{\varphi} : \mathbb{Z}[G] \to \mathbb{Z}[H]$, the *induced* ring homomorphism. Note that if $g \in \ker\varphi$ then $1 - g \in \ker\widetilde{\varphi}$ since then $\widetilde{\varphi}(1-g) = 1 - 1 = 0$. We claim $S = \{1 - g : g \in \ker\varphi\}$ generates $\ker\widetilde{\varphi}$.

Definition 1.1.3. If R is a ring and $I \subset R$ a two-sided ideal, then I defines an equivalence relation on R: $r_1 \sim r_2$ if $r_2 - r_1 \in I$. The equivalence classes are denoted by $[r] = r + I = \{r + a : a \in I\}$. The quotient set of these classes R/I is a ring, called the *quotient ring*, with the natural quotient map $R \to R/I$ a surjective ring homomorphism.

Theorem 1.1.4 (1st homomorphism)**.** Let $\varphi : R \to S$ be a ring homomorphism. Then there is a unique factorisation of φ:

$$\begin{array}{ccc} R & \xrightarrow{\varphi} & S \\ & {\scriptstyle\pi}\searrow \quad \nearrow{\scriptstyle\psi} & \\ & R/\ker\varphi & \end{array}$$

where ψ is injective.

The quotient construction allows us to make a lot of rings. An important class is finitely generated k-algebras (quotients of a polynomial ring over k, $k[x_1, \dots, x_n]$ by a finitely generated ideal $I = (f_1, \dots, f_n)$). In fact, for all ideals $I \subset k[x_1, \dots, x_n]$, $\exists$ finitely many $f_1, \dots, f_r \in I : I = (f_1, \dots, f_r)$: the Hilbert basis theorem. We now define the language we just used more rigorously.

Definition 1.1.5. Let R be a commutative ring. R is a *k-algebra* if $(R, +)$ is a vector space over k.

Definition 1.1.6. A k-algebra R is *finitely generated* if $\exists a_1, \dots, a_n \in R$: the obvious ring homomorphism

$$k[x_1, \dots, x_n] \to R : x_i \mapsto a_i$$

is surjective.

Week 3, lecture 1, 14 October 2024

Theorem 1.1.7 (2nd homomorphism)**.** For rings R, S, if $R \subset S$ and $I \subset S$ is a two-sided ideal, then $R + I$ is a ring and $(R + I)/I = R/(R \cap I)$.

Theorem 1.1.8 (3rd homomorphism)**.** For a ring R, if $I \subset J \subset R$ are two-sided ideals, then

$$\frac{R}{J} = \frac{R/I}{J/I}.$$

1.2. **Commutative rings.** All rings are commutative domains today from now on.

Definition 1.2.1. $a \in R$ is a *unit* if $\exists r \in R : ar = 1$. Write $R^\times \coloneqq \{u \in R : u \text{ is a unit}\}$, which is a group called the *unit group*.

Definition 1.2.2. $m \in R$ is *irreducible* if one writes for $m_1, m_2 \in R$ that $m = m_1 m_2$ then either m_1 or $m_2 \in R^\times$.

$\pi \in R$ is *prime* if $\pi \mid ab \implies$ either $\pi \mid a$ or $\pi \mid b$ $\forall a, b \in R$.

For example, R is a domain $\iff 0$ is prime.

Definition 1.2.3. A ring R is a *unique factorisation domain* if every $m \in R$ can be written as $m = \pi_1 \cdots \pi_r$ where $\pi_i \in R$ are prime and this decomposition is unique up to reordering and units.

Remark 1.2.4. An element is irreducible if it's prime. Suppose π is prime and write $\pi = ab$, then WLOG $\pi \mid a$, so $a = \pi u$, hence $\pi = \pi ub$, so $ub = 1$, i.e. b is a unit.

In a euclidean domain, an element is irreducible only if it's prime. Suppose m is irreducible and $m \mid ab$ and let $c = \gcd(a, m)$, so $m = cu$ for some u. Either c or $u \in R^\times$. If $u \in R^\times$ then $m \mid a$. If $c \in R^\times$ then $\exists p, q \in R : 1 = pa + qm$, so $b = pab + qmb$, hence $m \mid b$.

Remark 1.2.5. A euclidean domain is a UFD. By above it suffices to prove every element has a unique factorisation into irreducibles. For existence, one proves by induction on the norm (if m is not irreducible, write $m = m_1 m_2$ and by inductive hypothesis m_1, m_2 have unique factorisations). For uniqueness, now note that irreducibles are primes, and by definition if a prime appears in a factorisation and divides other, it must appear in the other one as well.

Lemma 1.2.6. If R is a domain then $R[x]$ is a domain.

Proof. It suffices to show for $f, g \in R[x]$, $f, g \neq 0 \implies fg \neq 0$, but mutliplication in $R[x]$ is multiplication of coefficients which are in R, which is a domain. □

Definition 1.2.7. For a commutative ring R (not necessarily a domain), $S \subset R$ is a *multiplicative subset* if $1 \in S$ and $s_1, s_2 \in S \implies s_1 s_2 \in S$.

The *localisation* of R in S is the ring $S^{-1}R \coloneqq \{(a, s) : a \in R, s \in S\}$ with the expected operations and modulo the equivalence relation $(a_1, s_1) \sim (a_2, s_2)$ if $\exists t \in S : t(a_1 s_2 - a_2 s_1) = 0$.

If R is a domain and one takes $S = R \backslash \{0\}$ then $S^{-1}R$ is a field, the *fraction field* of R, denoted by $\mathrm{Frac}(R)$.

Definition 1.2.8. For $f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n \in R[x]$, define the *content* of f to be

$$c(f) \coloneqq \gcd(a_0, \ldots, a_n).$$

f is *content-free* if $c(f) = 1$. Clearly $\frac{1}{c(f)}f$ is content-free $\forall f \in R[x]$.

Definition 1.2.9. An ideal P of a commutative ring R is *prime* if $\forall a, b \in R,\ ab \in P \implies a$ or $b \in P$.

By definition, a principal ideal is prime iff it's generated by a prime.

Remark 1.2.10. R/I is a domain iff I is prime which follows almost immediately from definitions.

Lemma 1.2.11 (Gauss)**.** Let R be a UFD and $K = \mathrm{Frac}(R)$. Suppose for $f \in R[x]$ $\exists h_1, h_2 \in K[x]$ nonconstant such that $f = h_1h_2$, then $\exists \lambda_1, \lambda_2 \in K : \lambda_1\lambda_2 = 1$ and $\widetilde{h_1} = \lambda_1 h_1, \widetilde{h_2} = \lambda_2 h_2 \in R[x]$, so in particular $f = \widetilde{h_1}\widetilde{h_2}$, i.e. a polynomial in a UFD is irreducible iff it's irreducible in the UFD's fraction field.

Week 3, lecture 2, 15 October 2024

Proof. WLOG suppose $c(f) = 1$. Note that $\exists c_1, c_2 \in R : c_1h_1 =: h_1', c_2h_2 =: h_2' \in R[x]$ (clear denominators). One then has the factorisation in $R[x]$: $c_1c_2f = h_1'h_2'$. It remains to get rid of c_1c_2.

The following claim is useful: $\forall h_1, h_2 \in R[x],\ c(h_1h_2) = c(h_1)c(h_2)$. In paticular, if h_1, h_2 are content-free then h_1h_2 is content-free, which is the special case we are now going to prove. Suppose for a contradiction that $c(h_1h_2)$ is not a unit and let $\pi \in R$ be prime such that $\pi \mid$ every coefficient of h_1h_2. Then $[h_1h_2] = 0 \in R/\pi[x]$. But since (π) is prime, R/π is a domain, hence WLOG $[h_1] = 0$, i.e. $\pi \mid$ every coefficient of h_1.

But then $c_1c_2 = c(c_1c_2f) = c(h_1')c(h_2')$, so one can write

$$f = \frac{h_1'}{c(h_1')}\frac{h_2'}{c(h_2')} =: \widetilde{h_1}\widetilde{h_2} \in R[x].$$

□

Again all rings are commutative domains today.

Theorem 1.2.12 (Eisenstein's criterion)**.** Let R be a UFD, $f(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n \in R[x]$ and $\pi \in R$ be prime. If $\pi \nmid a_0,\ \forall i \geq 1,\ \pi \mid a_i$ and $\pi^2 \nmid a_n$, then f is irreducible.

Proof. Consider $\overline{R} = R/\pi$, which is a domain since π is prime, so $\overline{R}[x]$ is a domain as well. One has $\overline{f(x)} = \overline{a_0}x^n$. Suppose for a contradiction $f = hg \in R[x]$, then $\overline{a_0}x^n = \overline{h(x)} \cdot \overline{g(x)}$. It must be that $\overline{h(x)} = a_1x^{k_1}$ and $\overline{g(x)} = a_2x^{k_2}$ since otherwise $\overline{a_0}x^n$ contains a nonzero product of monomials of smallest degree in $\overline{h}$ and $\overline{g}$. But then π divides constant terms of h and g, so $\pi^2 \mid a_n$, a contradiction. □

Week 3, lecture 3, 18 October 2024

Definition 1.2.13. For $a_1, \ldots, a_r \in R,\ (a_1, \ldots, a_r) \coloneqq \{\sum \lambda_i a_i : \lambda_i \in R\}$, which is the left ideal (finitely) *generated* by $a_1, \ldots, a_r$.

A ring R is *left-Noetherian* if every left ideal of R is finitely generateted. R is *Noetherian* if it's both left- and right-Noetherian.

Lemma 1.2.14. A ring is left-Noetherian $\iff$ it satisfies the ascending chain condition, i.e. $\forall$ infinite chain of left ideals $I_1 \subset I_2 \subset \cdots I_r \subset I_{r+1} \subset \cdots,\ \exists N : j \geq N \implies I_N = I_j$.

Proof. $\implies$ Let $I_1 \subset I_2 \subset \cdots$ be an ascending chain of left ideals, then $\bigcup_j I_j = I$ is a left ideal. By assumption, one can write $I = (a_1, \ldots, a_r)$. But then $\exists N : a_j \in I_N\ \forall j = 1, \ldots, r$, so $\forall j \geq N,\ I_j = I_N = I$.

$\impliedby$ Let I be a left-ideal and $a_1 \in I$. If $I = (a_1)$ then we are done, so let $I_1 = (a_1)$ and $a_2 \in I \backslash I_1$. Again if $I = I_2 = (a_1, a_2)$ then we are done, so let $a_3 \in I \backslash I_2$. One continues inductively and obtains an infinite ascending chain of left ideals $I_1 \subset I_2 \subset \cdots$ which stabilises by assumption, i.e. at some point $\nexists a_{r+1} \in I \backslash I_r$ so $I = I_r = (a_1, \ldots, a_r)$.

□

Example 1.2.15. The ring of polynomials with infinite number of variables $K[x_1, x_2, \ldots]$ is clearly not Noetherian since one has the chain $(x_1) \subset (x_1, x_2) \subset \cdots$ not satisfying ACC.

Theorem 1.2.16 (Hilbert basis)**.** If R is Noetherian then $R[x]$ is Noetherian.

Proof. Let $J \subset R[x]$ be an ideal and define

$$I_k \coloneqq \{a : a \text{ is the leading coefficient of } f \in J \text{ where } \deg f = k\}.$$

Note that I_k is an ideal in R (since J is an ideal of $R[x]$) and $I_k \subset I_{k+1}\ \forall k$ (since

$$a \in I_k \implies f = ax^k + \cdots \in J \implies xf = ax^{k+1} + \cdots \in J \implies a \in I_{k+1}).$$

By assumption, the chain of ideals $I_k \subset I_{k+1} \subset I_{k+2} \subset \cdots$ stablises at $I_N = (a_1, \ldots, a_m)$. Then

$$\forall i = 1, \ldots, m,\ \exists f_i = a_i x^N + \cdots \in J.$$

We claim $J = (f_1, \ldots, f_m)$. Indeed, let $f = c_\nu x^\nu + c_{\nu-1}x^{\nu-1} + \cdots + c_0 \in J$ and suppose $\nu \geq N$. Then $c_\nu \in I_\nu = I_N$, i.e. $c_\nu = \lambda_1 a_1 + \cdots + \lambda_m a_m$. Then

$$f - \left(x^{\nu-N}\lambda_1 f_1 + \cdots + x^{\nu-N}\lambda_m f_m\right) \in J$$

has degree strictly less than f. We have therefore shown that J is generated by $f_1, \ldots, f_m$ and polynomials in J with degree $\leq N$. □

Week 4, lecture 1, 21 October 2024

Theorem 1.2.17. If R is a UFD then $R[x]$ is a UFD.

1.3. **Rings of invariants.**

1.3.1. *Motivation and examples.*

Definition 1.3.1. Let a finite group $G \subset \mathrm{GL}_n(\mathbb{C})$ act on $\mathbb{C}[x_1, \ldots, x_n]$. Define the *ring of invariant* as

$$\mathbb{C}[x_1, \ldots, x_n]^G = \{f \in \mathbb{C}[x_1, \ldots, x_n] : \forall \gamma \in G,\ \gamma f = f\}.$$

Example 1.3.2. Consider the symmetric group $S_n \ni \sigma$ with the action

$$\sigma \cdot f(x_1, \ldots, x_n) = f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}).$$

For example, write $S_2 = \{\mathrm{id}, \sigma\}$ then $\sigma \cdot f(x, y) = f(y, x)$ (i.e. $x \mapsto y,\ y \mapsto x$).

Note that polynomials of the form $x_1 + \cdots + x_n,\ \sum_{i<j} x_i x_j,\ x_1 \cdots x_n$ (in general $\sum_{i_1 < \cdots i_k} x_{i_1} \cdots x_{i_k}$) are invariant under these actions. Denote them by $\sigma_1, \sigma_2, \sigma_n$ and σ_k respectively and call them *elementary symmetric functions*. We claim they generate all other such invariants.

Theorem 1.3.3. (1) $\mathbb{C}[x_1, \ldots, x_n]^{S_n}$ is generated by $1, \sigma_1, \ldots, \sigma_n$.

(2) The map

$$\phi : \mathbb{C}[y_1, \ldots, y_n] \to \mathbb{C}[x_1, \ldots, x_n]^{S_n}$$
$$y_i \mapsto \sigma_i$$

is a ring isomorphism, i.e. there are no algebraic relations between the σ_i's and $\mathbb{C}[x_1, \ldots, x_n]^{S_n}$ is itself a polynomial ring.

We will later see a proof.

Example 1.3.4. Consider $\mu_n = \{\zeta : \zeta^n = 1\} \subset \mathbb{C}^\times$ (generated by $\zeta = e^{\frac{2\pi i}{n}}$) with the action

$$\mu_n \curvearrowright \mathbb{C}[x, y] : \zeta \cdot x = \zeta x,\ \zeta \cdot y = \zeta^{-1} y.$$

Note that $w = xy, u = x^n$ and $v = y^n$ are invariant. It turns out that

$$\phi : \mathbb{C}[u, v, w] \to \mathbb{C}[x, y]^{\mu_n} : u \mapsto x^n,\ v \mapsto y^n,\ w \mapsto xy$$

is surjective, i.e. u, v, w generate all invariants, and ϕ has kernel $(uv - w^n)$ as probably expected.

Now consider the same group with a different action: $\zeta \cdot x = \zeta x$ and $\zeta \cdot y = \zeta y$. Then $\mathbb{C}[x, y]^{\mu_n}$ is generated by monomials of degree n, i.e. $x^n, x^{n-1}y, \ldots, y^n$. This time one can write down a similar surjective homomorphism as above, but the kernel is huge: it's the 2×2 minors of

$$\begin{pmatrix} u_0 & u_1 & \cdots & u_{n-1} \\ u_1 & u_2 & \cdots & u_n \end{pmatrix}$$

where $u_i = x^{n-i}y^i$.

Week 4, lecture 2, 22 October 2024

Example 1.3.5. The group $G = \mathrm{BD}_{4n} = \left\langle \mathbb{\mu}_{2n} = \left\{ \begin{pmatrix} \zeta & 0 \\ 0 & \zeta^{-1} \end{pmatrix} : \zeta^{2n} = 1 \right\}, B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right\rangle \subset \mathrm{SL}_2(\mathbb{C})$ acts on $\mathbb{C}[x,y]$. Note that $\mathbb{C}[x,y]^G \subset \mathbb{C}[x,y]^{\mathbb{\mu}_{2n}}$, so it remains to see how B acts on $\mathbb{C}[u,v,w]$ where $u = x^{2n}, v = y^{2n}$ and $w = xy$. One has $B : u \mapsto v, v \mapsto u$ and $w \mapsto -w$. Note that one has B-invariants $\xi = w^2, \eta = u+v, z = w(u-v)$. It turns out ξ, η, z generate $\mathbb{C}[x,y]^G$, but clearly they have some relations. One calculates

$$\xi^2 = w^2(u^2+v^2-2uv) = w^2((u+v)^2 - 4(w^2)^n) = z(\eta^2 - 4z^n).$$

It turns out this relation generates the ideal of relations, i.e. $\mathbb{C}[x,y]^G \cong \mathbb{C}[x,y,z]/(x^2 - y^2z + 4z^{n+1})$.

1.3.2. *Noether's theorem.*

Definition 1.3.6. A ring R is *graded* if $R = \bigoplus_{n \geq 0} R_n$ as additive groups where $R_n R_m \subset R_{m+n}$.
$a \in R$ is *homogeneous* of degree n if $a \in R_n$.

Example 1.3.7. $R = \mathbb{C}[x_1, \ldots, x_n]$ is a graded ring.

Theorem 1.3.8 (Noether)**.** Let finite $G \leq \mathrm{GL}_n(\mathbb{C}) \curvearrowright R = \mathbb{C}[x_1, \ldots, x_n]$. Then R^G is a finitely generated $\mathbb{C}$-algebra.

Observe that

(1) R is graded and G preserves the grading, i.e. $\forall n \geq 0, \gamma \in G, f \in R_n,\ \gamma f \in R_n$.
(2) There is an operation called "averaging over G":

$$\rho : R \to R^G : f \mapsto \frac{1}{|G|} \sum_{\gamma \in G} \gamma f$$

which also preserves the grading: $\operatorname{im} \rho(R_n) \subset R_n^G$.

Note that ρ is additive: $\rho(f+g) = \rho(f) + \rho(g)$, which is clear from the definition, but it's in fact R^G-linear, i.e. if $f \in R^G$ and $g \in R$, then $\rho(fg) = f\rho(g)$.

Week 4, lecture 3, 25 October 2024

Proof 1. Let I be the ideal of R generated by R_+^G (all invariances except the constants). By Hilbert's basis theorem, $\exists a_1, \ldots, a_m \in R : I = (a_1, \ldots, a_m)$. Note that a_i's are not necessarily invariant, but one can write $a_i = \lambda_{i1} r_{i1} + \cdots + \lambda_{ik(i)} r_{rk(i)}$ where $\lambda_{ij} \in \mathbb{C}, r_{ij} \in R^G$ by our first observation, so $I = (r_{11}, r_{12}, \ldots, r_{mk(m)}) = r_N$. We claim $r_1, \ldots, r_N$ generate R^G as a $\mathbb{C}$-algebra. Let S be the ring generated by the r_i's, i.e. let S be the image of the ring homomorphism $\mathbb{C}[y_1, \ldots, y_n] \to R : y_i \mapsto r_i$. Clearly by construction $S \subset R^G$. We prove by induction that $R_n^G \subset S$. Clearly $R_0^G = \mathbb{C} \subset S$, so assume $R_0^G \oplus \cdots \oplus R_{n-1}^G \in S$ and let $f \in R_n^G$. Then $f \in I$, so $\exists a_i \in R : f = a_1 r_1 + \cdots + a_N r_N$. If a_i's are invariant then we are done since $\deg a_i < \deg f = n$ and by inductive hypothesis a_i are $\mathbb{C}$-combinations of r_i's as well. If a_i's are not invariant, recall our second observation and take $f = \rho(f) = \rho(a_1) r_1 + \cdots + \rho(a_N) r_N$ where each $\rho(a_i)$ is invariant. □

Example 1.3.9. Not all subrings $S \subset R = \mathbb{C}[x_1, \ldots, x_n]$ are finitely generated. Take the monoid P to be all monomials in $\mathbb{C}[x,y]$ except $\{x, x^2, \ldots\}$ and the monoid ring $\mathbb{C}[P]$.

Week 5, lecture 1, 28 October 2024

Proof 2 of 1.3.8: the original proof by Emmy. The advantage of this proof is that it's constructive: it gives a way to actually find the ring of invariance.

Consider $R[T]$. $\forall f \in R$, define

$$P^f(T) \coloneqq \prod_{\gamma \in G} (T - \gamma f),$$

a polynomial of degree $|G|$ with its coefficients in R^G. In particular, write

$$P^{x_i}(T) = \prod_{\gamma \in G} (T - \gamma x_i) = T^N + a_{N-1}^i T^{N-1} + \cdots + a_0^i$$

where $a_j^i \in R^G$. Let $S = \mathbb{C}\left[a_j^i : i = 1, \ldots, n, j = 1, \ldots, |G|\right] \subset R^G$ be a subring.

For a monomial $x^r = x_1^{r_1} \cdots x_n^{r_n}$ define $|r| = \max\{r_1, \ldots, r_n\}$. We claim

$$\widehat{S} = \left\{ \sum_r f_r x^r : f_r \in S, |r| < N \right\} = R.$$

It suffices to show $x^r \in \widehat{S}\ \forall r$. If $|r| < N$ then we are done, so suppose $|r| \geq N$. Note that x_i is a root of $P^{x_i}(T)$ since $\mathrm{id} \in G$ and $x_i - \mathrm{id}\, x_i = 0$. Hence

$$P^{x_i}(x_i) = 0 = x_i^N + a_{N_1}^i x_i^{N-1} + \cdots + a_0^i,$$

hence we have a rewriting machine grinding down the degree of any monomial.

We now claim R^G is generated as a ring by S (hence ultimately by the a_j^i) and $\{\rho(x^r) : |r| < N\}$. Let $f \in R^G$. By above, $f = \sum_r s_r x^r$ for some $s_r \in S$ with $|r| < N$. Then $f = \rho(f) = \sum_r s_r \rho(x^r)$. □

Theorem 1.3.10. For $G = S_m \curvearrowright R = \mathbb{C}[x_1, \ldots, x_n]$, one has $R^G = \mathbb{C}[e_1, \ldots, e_m]$ where $e_j(x_1, \ldots, x_n)$ is the jth elementary symmetric polynomial $\sum_{1 \leq k_1 \leq \cdots \leq k_j \leq n} x_{k_1} \cdots x_{k_j}$. Moreover, the natural map $\mathbb{C}[y_1, \ldots, y_m] \to \mathbb{C}^G : y_j \mapsto e_j$ is a ring isomorphism.

Proof. Order monomials lexicographically:

$$x^r = x_1^{r_1} \cdots x_n^{r_n} > x^s = x_1^{s_1} \cdots x_n^{s_n} \quad \text{if for the first } i : r_i \neq s_i, \text{ one has } r_i > s_i.$$

Hence for $f \in R$ one can define $\mathrm{hm}(f)$ to be the highest monomial that appears in f. Now clearly for $f \in R^G$, $\mathrm{hm}(f) = x_1^{m_1} \cdots x_n^{m_n}$ where $m_1 \geq \cdots \geq m_n$ since f is invariant. But then

$$\exists k_1, \ldots, k_n : \mathrm{hm}(f) = \mathrm{hm}\left(e_1^{k_1} \cdots e_n^{k_n}\right).$$

□

Week 5, lecture 2, 29 October 2024

2. Modules

As much as for today's discussion, a ring R is not necessarily commutative.

Definition 2.0.1. A left *R-module* is a triple $(M, +, \cdot, 0)$ where $(M, +, 0)$ is an abelian group and $\cdot : R \times M \to M$ satisfies $r_1(r_2 m) = (r_1 r_2)m\ \forall r_1, r_2 \in R, m \in M$ and $(r_1 + r_2)m = r_1 m + r_2 m$.

Example 2.0.2. Any abelian group A is a $\mathbb{Z}$-module: define ma by $\underbrace{a + \cdots + a}_{m \text{ times}}$ if $m \geq 0$ and $\underbrace{-a - \cdots - a}_{-m \text{ times}}$ if $m < 0$.

$R^{\oplus n} = \{(r_1, \ldots, r_n) : r_i \in R\}$ is an R-module. In fact, any direct sums of modules is a module: $M_1 \oplus M_2 = \{(m_1, m_2) : m_i \in M_i\}$.

If $I \subset R$ is a left ideal, then I and $R/I = \{r + I : r \in R\}$ are R-modules.

2.1. Construction of R-modules.

Definition 2.1.1. $N \subset M$ is a *submodule* if it is an additive subgroup of M and $rn \in N\ \forall r \in R, n \in N$.

If $N \subset M$ is a submodule, one can make the *quotient module* M/N, which is the usual quotient group with an induced multiplication.

Definition 2.1.2. For modules M_1, M_2, a map $\phi : M_1 \to M_2$ is a *module homomorphism* if

$$\phi : (M_1, +) \to (M_2, +)$$

is a group homomorphism and

$$\phi(rm) = r\phi(m)\ \forall r \in R, m \in M.$$

(Analogous to linear maps.)

For a homomorphism $\varphi : M_1 \to M_2$, $\varphi(M_1) \subset M_2$ is a submodule, called the *image* of ϕ and denoted by $\mathrm{im}\,\varphi$. The set $\{m \in M_1 : \varphi(m) = 0\} \subset M_1$ is a submodule, called the *kernel* of ϕ and denoted by $\ker \varphi$.

The quotient module $M_2/\mathrm{im}\,\varphi$ is called the *cokernel* of φ and denoted by $\mathrm{coker}\,\varphi$.

Definition 2.1.3. An R-module M is *finitely generated* if $\exists \varphi : R^n \to M$, a surjective homomorphism of R-modules. Equivalently, $\exists m_1, \ldots, m_n \in M : \forall m \in M,\ \exists r_1, \ldots, r_n \in R : r_1 m_1 + \cdots + r_n m_n = m$.

For the rest of our discussion of modules, we focus on classification of finitely generated modules over commutative PIDs.

Theorem 2.1.4 (Homomorphism theorems)**.**

(1) $\varphi : M \to N$ homomorphism $\implies M/\ker\varphi = \mathrm{im}\,\varphi$.
(2) $(A + B)/A = B/(A \cap B)$.
(3) $L \subset N \subset M \implies M/N = (M/L)/(N/L)$.

Week 5, lecture 3, 1 November 2024: example class

Week 6, lecture 1, 4 November 2024

Definition 2.1.5. A left R-module M is *finitely generated* if $\exists$ a surjective R-module homomoprhism $R^n \twoheadrightarrow M$, i.e. $\exists m_1, \dots, m_n \in M : \forall m \in M,\ m = r_1 m_1 + \cdots + r_n m_n$ for some $r_1, \dots, r_n \in R$.

Definition 2.1.6. For a set S, write

$$R^S = \bigoplus_{s \in S} R = \{a : S \to R : a_s = 0 \text{ for all but finitely many } s \in S\}.$$

An element of R^S is a tuple, with addition and multiplication defined pointwise. We call this type of modules *free*.

Week 6, lecture 2, 5 November 2024: in-class test

Week 6, lecture 3, 8 November 2024

Definition 2.1.7. Let M be an R-module and $m : S \hookrightarrow M$ where $S \subset M$.
We say S *generates* M if $\forall x \in M,\ \exists$ a finite expression $x = \sum r_s m_s$ where $r_s \in R, m_s \in S$.
We say S is *linearly independent* if $0 = \sum_{\text{finite}} r_s m_s \implies$ all $r_s = 0$.

Proposition 2.1.8. $M \cong R^S \iff S$ generates M and S is linearly independent.

Definition 2.1.9. An ideal $I \subset R$ is *maximal* if any ideal $I' \supset I$ is either I or R.

Proposition 2.1.10. $I \subset R$ is maximal $\iff R/I$ is a field.

Proof. Recall the correspondence between ideals $\overline{R} \subset R/I$ and $J : I \subset J \subset R$.
If R/I is a field then, then the only ideals $\overline{J}$ of R is 0 or R/I, so J is either I or R.
If I is maximal, let $\overline{x} \in R/I$ where $\overline{x}$ is the image of $x \in R$, and suppose $x \notin I$ (so $\overline{x} \neq 0$). Consider $(\overline{x})$, an ideal of R/I. Then its preimage is either I or R, but $x \notin I$ so it's R. But then $(\overline{x}) = R/I$, i.e. $\exists \overline{y} : \overline{yx} = 1$. □

Theorem 2.1.11. Every commutative ring R has a maximal ideal.

Proof sketch. Let I_1 be an ideal of R. If it's maximal then we're done. If it's not maximal, then $\exists$ ideal $I_2 \supset I_1$. One continue inductively and obtains a chain of ideals. If R is Noetherian we can stop here. If not, we have to believe in the axoim of choice. □

Theorem 2.1.12. If R is a commutative ring, then $R^n \cong R^m$ (as R-modules) $\iff n = m$.

Remark 2.1.13. This is false for modules over noncommutative rings.
The theorem is surprisingly not a trivial result. The proof is a brutal reduction to the case of vector spaces over fields.

Proof sketch. Suppose $R^n \cong R^m$ and by 2.1.11 and 2.1.10, let I be a maximal ideal of R and

$$(R/I)^n \cong (R/I)^m$$

as vector spaces, then by linear algebra $n = m$. □

2.2. **Classification of finitely generated modules over a PID.**

Theorem 2.2.1. Let R be a PID and M a finitely generated R-module. Then $M \cong R^n \oplus \bigoplus_{i=1}^{k} R/d_i$ where $d_i \mid d_{i+1}\ \forall i = 1, \dots, k-1$. Moreover, m, called the *rank* of M, k and all d_i's are uniquely (up to multiplication by units in R) determined.

The proof is slightly easier for Euclidean domains (but only slightly). But it's the worth the effort, since in practice there's virtually no way to know if something is a Euclidean domain or not. In particular, to prove it is, you have to construct a norm function a priori from scratch, and even worse, to prove it's not, you have to prove that every possible function doesn't work.

Week 7, lecture 1, 11 November 2024

Proposition 2.2.2. If R is a Noetherian ring, N is a finitely generated R-module and $M \subset N$ is a submodule, then M is also finitely generated.

Proof sketch. Consider the short exact sequence of R-modules $0 \to M' \xrightarrow{\varphi} M \xrightarrow{\pi} M'' \to 0$ where φ is injective, π is surjective and $\operatorname{im}\varphi = \ker\pi$. In this situation, clearly M', M'' are finitely generated $\implies M$ is finitely generated.

Now since N is assumed to be finitely generated, there is a surjective homomorphism $\psi : R^n \twoheadrightarrow N$ for some n. We prove by induction on n. The base case $n = 1$ follows from that R is Noetherian. Take M' to be the $\psi(R^{n-1}) \cap M$ where R^{n-1} is the first $n-1$ copies of R in R^n, and M'' to be $\psi(R)$ where R is the last copy. By inductive hypothesis, M' and M'' are finitely generated as submodules of $\psi(R^{n-1})$, so by our observation above, M is finitely generated. □

The result 2.2.1 is a consequence of the existence of the Smith normal form for matrices with coefficients in R:

Theorem 2.2.3. Let R be a PID. For $A \in M_{n\times m}(R)$,

$$\exists P \in \mathrm{GL}_n(R), Q \in \mathrm{GL}_m(R) : PAQ = \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_r \end{pmatrix}, \quad 0 \text{ elsewhere}$$

where $d_i \mid d_{i+1}\ \forall i = 1, \dots, r-1$ are uniquely determined by A. This is called the *Smith normal form*.

Proof of 2.2.1. Since M is finitely generated, $\exists$ a surjective homomorphism $\pi : R^m \twoheadrightarrow M$. Let $N = \ker\pi$, which as a submodule of R^m is finitely generated by 2.2.2, so there is a surjective homomorphism $\varphi : R^n \twoheadrightarrow N \subset R^m$. Then φ correspond to a matrix $A \in M_{n\times m}$ with coefficients in R, which we can assume is is SNF by 2.2.3 by change of bases for R^n and R^m. □

Proof of 2.2.3. We first prove first for a Euclidean domain R and then explain how to change the proof so that it works for PIDs.

Note that for a general matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, the multipliation on the left by $\begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix}$ gives $\begin{pmatrix} a+rc & b+rd \\ c & d \end{pmatrix}$, so it corresponds to the row operation of adding r times row 2 to row 1. Similarly, $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ multiplies row 1 by -1, and $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ swaps the two rows. In general, multiplications on the left correspond to all row operations, and on the right correspond to all column operations. Hence it suffices to show that after row and column operations, a matrix can be brought to the desired form. We show this in steps.

(1) We first achieve that a_{11} divides all a_{1i} and all a_{j1}, i.e. elements in the same row and column. To do this, we first make sure $|a_{11}| = \min\{|a_{ij}|\}$ by swapping rows and columns. Now if $a_{11} \nmid a_{l1}$, write $a_{l1} = qa_{11} + r$ with $|r| < |a_{11}|$, then subtract row l by q times row 1 so that now $a_{l1} = r$, and bring r to a_{11} again by swapping rows and columns. Similarly if $a_{11} \nmid a_{1l}$. This can't go on indefinitely since $|\cdot|$ is bounded below, hence the desired is achieved.

(2) Now by row/columns operations, our matrix is of the form $\left(\begin{array}{c|ccc} a_{11} & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array}\right)$.

(3) We now achieve that a_{11} divides each b_{ij}. If $a_{11} \nmid b_{ij}$, we add row i to row 1 and repeat step 1, which reduces $|a_{11}|$ and again this can't go on indefinitely.

(4) Repeat steps 1–3 for $B, B', \dots$

Now if R is a PID, we are not allowed to do induction on the norm (because there isn't one).

Define $N(a) \coloneqq \#\{\text{prime factors in the prime decomposition of } a, \text{ counting multiplicities}\}$, i.e.

$$\text{if } a = \prod_i p_i^{r_i}, \text{ then } N(a) = \sum_i r_i.$$

We now prove by induction on $N(A) \coloneqq \min\{N(a_{ij}) : 1 \le i \le n, 1 \le j \le m\}$. If $a = 0$ then $N(a) = \infty$. This is our "norm".

To have something similar to the Euclidean algorithm, recall that for any $a, b \in R$ we have $(a, b) = (c)$ for some $c \in R$: denote this by $c = \gcd(a, b)$. Tautologically, in this case, $\exists x, y \in R : ax + by = c$, and as naively expected, $d \mid a, b \implies d \mid c$, and $c \mid a, b$. Also, in the form $ax + by = c$, one has $\gcd(x, y) = 1$, since if we write $a = ca', b = cb'$, then $c(a'x + b'y) = c$, so $a'x + b'y = 1$ (since R is a domain).

The key mechanism of the proof is then the following. Let $a, b \in R$ and $c = \gcd(a, b)$ with x, y, a', b' as above. Consider the calculation

$$\begin{pmatrix} x & y \\ -b' & a' \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} ax + by \\ -b'a + a'b \end{pmatrix} = \begin{pmatrix} c \\ 0 \end{pmatrix}.$$

We are now prepared to construct the Smith normal form following the construction for Euclidean domains.

(1) Make sure $N(A) = N(a_{11})$ and that $a_{11} \mid a_{k1}, a_{1l}$ for all $k = 1, \dots, n, l = 1, \dots, m$. If $a_{11} \nmid a_{k1}$, WLOG assume $k = 2$ and let $c = \gcd(a_{11}, a_{21})$, then $N(c) < N(a_{11})$. Write $c = xa_{11} + ya_{21}$, then the multiplying the matrix

$$\left(\begin{array}{cc|c} x & y & \\ -a'_{21} & a'_{11} & 0 \\ \hline & 0 & \end{array}\right)$$

on the left gives us $\begin{pmatrix} c & \cdots \\ 0 & \\ \vdots & \end{pmatrix}$, so we have a matrix with smaller "norm".

(2) Again by row/column operations we get the form same as in step 2 of the Euclidean domain process.

(3) Again make sure a_{11} divides each a_{kl} for $k = 2, \dots, n, l = 2, \dots, m$. Suppose $a_{11} \nmid a_{kl}$, then add row k to row 1, so a_{kl} is in row 1 now as a_{k1}, and we repeat step 1 and again have a matrix of strictly smaller "norm".

(4) Repeat these steps to inner submatrices...

□

Week 7, lecture 2, 12 November 2024

Remark 2.2.4. In fact, the torsion part can be further decomposed:

$$M \cong R^n \oplus R/(p_1^{n_1}) \oplus \cdots \oplus R/(p_k^{n_k}) \qquad \text{where } n_i \geq 0 \text{ and } p_1, \dots, p_k \text{ are irreducibles of } R.$$

Example 2.2.5. Take $R = \mathbb{Z}$, then R-modules are exactly abelian groups, so 2.2.1 are simply the fundamental theorem of finitely generated abelian groups.

Suppose A is an abelian group generated by a, b, c subject to $2a + 3b + c = 0, a + 2b = 0$ and $5a + 6b + 7c = 0$. Then $A = \mathbb{Z}^3 / \left\langle \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 5 \\ 6 \\ 7 \end{pmatrix} \right\rangle$. Let $X = \begin{pmatrix} 2 & 1 & 5 \\ 3 & 2 & 6 \\ 1 & 0 & 7 \end{pmatrix}$. Let's reduce X to SNF:

$$\begin{pmatrix} 2 & 1 & 5 \\ 3 & 2 & 6 \\ 1 & 0 & 7 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 5 \\ 2 & 3 & 6 \\ 0 & 1 & 7 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & 5 \\ 0 & -1 & -4 \\ 0 & 1 & 7 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & 1 & 7 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & 0 & 3 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix},$$

hence $A \cong \mathbb{Z}^3 / \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} \right\rangle \cong \mathbb{Z}/\mathbb{Z} \oplus \mathbb{Z}/\mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z} \cong C_3$.

2.3. **Jordan normal form.** Let F be a field, V a F-vector space and $\alpha : V \to V$ a linear map. The key idea is to endow V with the structure of a module over $F[x]$ by letting x acts as α. Thus a polynomial f acts on V as $f(\alpha)$: this means if $f(x) = a_n x^n + \cdots + a_0$ where $a_i \in F$, then $f(x) \cdot v = a_n \alpha^n(v) + \cdots + a_1 \alpha(v) + a_0 v$. So V becomes an $F[x]$-module; call it V_α.

Lemma 2.3.1. If V is a finite-dimensional F-vector space, then V_α is a finitely generated $F[x]$-module.

Proof. Take a finite basis $v_1, \dots, v_n$ of V, then they generate V_α as an $F[x]$-module.
(Trivially $F \subset F[x]$). □

Consider $F[x]/d$ where $0 \neq d \in F[x]$.

Example 2.3.2. Take $d(x) = x^r$. Then as usual, images of $1, x, \dots, x^{r-1}$ form a basis of $F[x]/d$ as a F-vector space. The action of left mutliplication by $x \in F[x]$ can be written as the matrix

$$\begin{pmatrix} 0 & & & & \\ 1 & & & & \\ & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & 0 \end{pmatrix} \quad \text{everywhere else is 0,}$$

now if $\alpha : F[x]/d \to F[x]/d$ is represented by this matrix, then V_α is $F[x]/d$.

Example 2.3.3. Let $V_\alpha \cong F[x]/((x-\lambda)^r)$ for some $\lambda \in F$ and $\beta = \alpha - \lambda I$. Then $\beta : V \to V$ is a linear map with matrix as in above, so α is given by

$$\begin{pmatrix} \lambda & & & & \\ 1 & \lambda & & & \\ & 1 & \lambda & & \\ & & \ddots & \ddots & \\ & & & 1 & \lambda \end{pmatrix}$$

which is the Jordan normal form (sometimes the 1's are above the diagonal but it actually doesn't matter).

Example 2.3.4. What do we get in general? Suppose $V_\alpha \cong F[x]/(f)$ where $f = x^r + a_{r-1}x^{r-1} + \cdots + a_0$ where $a_i \in F$. As before, $1, x, \ldots, x^{r-1}$ form a basis of V as an F-vector space. Then α is given by the matrix (same as 2.3.2 except the last column)

$$C(f) = \begin{pmatrix} 0 & & & -a_0 \\ 1 & & & -a_1 \\ & 1 & & -a_2 \\ & & \ddots & \vdots \\ & & 1 & -a_{r-1} \end{pmatrix}, \quad \text{the companion matrix of } f$$

Theorem 2.3.5 (Rational canonical form)**.** Let F be any field, V a finite dimensional F-vector space, and $\alpha : V \to V$ a linear map. Then as $F[x]$-modules,

$$V_\alpha \cong \frac{F[x]}{(f_1)} \oplus \cdots \oplus \frac{F[x]}{(f_r)} \text{ where } f_i \neq 0 \qquad \text{with } f_i \mid f_{i+1}\ \forall i = 1, \ldots, r-1.$$

There is a basis for V in which α is given by the block diagnoal matrix

$$\begin{pmatrix} C(f_1) & & & \\ & C(f_2) & & \\ & & \ddots & \\ & & & C(f_r) \end{pmatrix}.$$

The word "rational" means you don't have to leave your ground field to its algebraic closure, so for example you can work over $\mathbb{Q}$, the "rationals", as desired.

Proof. By 2.3.1 and 2.2.1, one can write

$$V_\alpha \cong \frac{F[x]}{(f_1)} \oplus \cdots \oplus \frac{F[x]}{(f_r)} \oplus F[x]^n$$

with the desired properties for f_i's, so it suffices to show that $n = 0$, but this is trivial: V is finite dimensional over F, but $F[x]$ is not. The desired basis is

$$x_1^0, x_1, x_1^2, \ldots, x_1^{\deg f_1 - 1}, x_2^0, x_2, \ldots, x_2^{\deg f_2 - 1}, \ldots$$

where x_i is the image of x in $F[x]/(f_i)$. □

Example 2.3.6. Consider the field $\mathbb{Q}$ and $\mathbb{Q}[x]/(x^2-1)$, which is already in rational canonical form. It's clearly isomorphic to $\mathbb{Q}[x]/(x-1) \oplus \mathbb{Q}[x]/(x+1)$, but divisibility doesn't hold, so it's not a rational canonical form.

Week 7, lecture 3, 15 November 2024

The lecture started by continuing the proof of 2.2.3 for principal ideal domains, so see Week 7, lecture 1, 11 November 2024.

Corollary 2.3.7 (Jordan normal form)**.** Let V be a $\mathbb{C}$-vector space and $T : V \to V$ a $\mathbb{C}$-linear map. Then $\exists$ a basis of V such that the matrix of T has the form

$$\begin{pmatrix} J_{n_1}(\lambda_1) & & \\ & \ddots & \\ & & J_{n_r}(\lambda_r) \end{pmatrix}$$

where $J_{n_i}(\lambda_i)$, a *Jordan block*, is a $n_i \times n_i$ matrix with λ on the diagonal and 1 below.

Remark 2.3.8 (Alternative form of classification of finitely generated modules over a PID)**.** We have

$$M \cong R^{\oplus r} \oplus \bigoplus_i R/(p_i^{r_i}),$$

indeed, consider a $\mathbb{C}[x]$-module V with the action $x \cdot v = Tv$. Then $J_n(\lambda) \leftrightarrow \mathbb{C}[x]/((x-\lambda)^n)$...

Week 8, lecture 1, 18 November 2024

3. Matrix Lie groups

3.1. **Subgroups of the general linear group.** For a commutative ring R, define

$$\mathrm{GL}_n(R) \coloneqq \{A \in M_{n\times n}(R) : \exists B \in M_{n\times n}(R) : AB = BA = I_{n\times n}\} = \{A \in M_{n\times n}(R) : \det A \in R^\times\}.$$

We will work with closed subgroups $G \leq \mathrm{GL}_n(F)$ where $F = \mathbb{R}$ or $\mathbb{C}$. The point of this is that $\mathrm{GL}_n(F)$ is a topological group:

Definition 3.1.1. A *topological group* is a group G that is also a topological space such that the multiplication map $m : G \times G \to G : (g_1, g_2) \mapsto g_1 g_2$ and the inverse map $i : G \to G : g \mapsto g^{-1}$ are both continuous.

Lemma 3.1.2. $\mathrm{GL}_n(F)$ is a topological group in a natural way.

Proof. To give it a topology, note that $\mathrm{GL}_n(F)$ is a subset of $M_{n\times n}(F)$, which, as a set, can be considered as F^{n^2}, which has the usual Euclidean topology. Since $\det : F^{n^2} \to F$ is continuous (it's a polynomial), $\mathrm{GL}_n(F)$ as the preimage of $F\backslash\{0\}$, an open set, is open. We then endow $\mathrm{GL}_n(F)$ with induced topology: $v \in \mathrm{GL}_n(F)$ is open $\iff$ $v \in F^{n^2}$ is open. Then m and i are continuous since one is a polynomial and the other is a rational function. □

Example 3.1.3. Special linear group

$$\mathrm{SL}_n(F) = \{g \in \mathrm{GL}_n(F) : \det g = 1\}.$$

In particular, $\mathrm{SL}_n(\mathbb{C})$ acts on $\mathbb{C} \cap \{\infty\}$: consider the Möbius transformation $z \mapsto \frac{az+b}{cz+d}$.
$\mathrm{SL}_n(\mathbb{R})$ acts on $\mathfrak{h} = \{z \in \mathbb{C} : \Im z > 0\}$ again by $z \mapsto \frac{az+b}{cz+d}$.
Unitary group $\mathrm{U}_n = \{g \in \mathrm{GL}_n(\mathbb{C}) : \overline{{}^t A} A = I_{n\times n}\}$ acts on $S^{2n-1} = \{z \in \mathbb{C}^n : ||z|| = 1\}$.
$\mathrm{SU}_n = \{g \in \mathrm{U}_n : \det g = 1\}$ also acts on this sphere.
Euclidean group

$$\begin{aligned}\mathrm{Eucl}_n(\mathbb{R}) &= \{g \in \mathrm{GL}_n(\mathbb{R}) : d(gv - gu) = d(v-u)\}\\ &= \{\text{rigid motions with respect to its natural metric}\}.\end{aligned}$$

Consider the subgroup $\mathbb{R}^n$ of translations, which is a normal subgroup of $\mathrm{Eucl}_n(\mathbb{R})$, with

$$\mathrm{Eucl}_n(\mathbb{R})/\mathbb{R}^n = \mathrm{O}_n(\mathbb{R}) = \{g \in \mathrm{GL}_n(\mathbb{R})\} = \{{}^t gg = \mathrm{id}\} = \{gx \cdot gy = x \cdot y\},$$

the orthogonal group. Choosing an "origin" $O \in \mathbb{R}^n$, one can consider $\mathrm{O}_n(\mathbb{R})$ as a group of "rotations", and $\mathrm{Eucl}_n(\mathbb{R}) = \mathbb{R}^n \rtimes \mathrm{O}_n(\mathbb{R})$.

Affine group $\mathrm{Aff}_n(F)$, on the other hand, can be similarly constructed as $F^n \rtimes \mathrm{GL}_n(F)$. But how is this a subgroup of $\mathrm{GL}_n(F)$? Well, it's not, it's a subgroup of $\mathrm{GL}_{n+1}(F)$, and one can think of an element as

$$\left(\begin{array}{ccc|c} & n\times n & & b_1\\ & & & b_2\\ & & & \vdots\\ & & & b_n \\ \hline 0 & \cdots & 0 & 1 \end{array}\right),$$

and think of $\mathrm{Aff}_n(F)$ as

$$\{g \in \mathrm{GL}_{n+1}(F) : g\pi \subset \pi\} \quad \text{where } \pi = \{x \in F^{n+1} : x_{n+1} = 1\}, \text{ the hyperplane}$$

Projective general linear group

$$\mathrm{PGL}_n(F) = \mathrm{PSL}_n(F) = \mathrm{GL}_n(F)/\{\lambda I_{n\times n} : \lambda \in F^\times\}$$

acts on $\mathbb{P}^n(F) = F^{n+1}\backslash\{0\}/\sim$ where $v \sim \lambda v$ $\forall \lambda \in F^\times$. It's not entirely obvious that this is a subgroup of some $\mathrm{GL}_m(F)$.

Week 8, lecture 2, 19 November 2024

Special orthogonal group $\mathrm{SO}_n(F) = \{A \in \mathrm{O}_n : \det A = 1\}$.

Indefinite orthogonal group $\mathrm{O}_{p,q} = \{A \in \mathrm{GL}_n(\mathbb{R}) : {}^tAB_{p,q}A = B_{p,q}\}$ where $B_{p,q}$ is the diagonal matrix with first p 1's and q -1's.

Any finite subgroup G of any of the above.

Example 3.1.4 (A very special case)**.** Let $\mathrm{SL}_2(\mathbb{C})$ act on $\mathbb{R}^4$: identify $x = \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}$ as the matrix $x = \begin{pmatrix} x_0 + x_3 & x_1 - ix_2 \\ x_1 + ix_2 & x_0 - x_3 \end{pmatrix}$ (which satisfies ${}^t\overline{x} = x$; in fact all such matrices can be written in this way and corresponds uniquely to an element of $\mathbb{R}^4$, so we are identifying $\mathbb{R}^4$ as $\{x \in M_{2\times 2}(\mathbb{C}) : {}^t\overline{x} = x\}$, a vector space isomorphism), and define $g \cdot x$ as $gx{}^t\overline{g}$. This action preserves determinants, and since

$$\begin{vmatrix} x_0 + x_3 & x_1 - ix_2 \\ x_1 + ix_2 & x_0 - x_3 \end{vmatrix} = x_0^2 - x_1^2 - x_2^2 - x_3^2,$$

so we have a group homomorphism $\mathrm{SL}_2(\mathbb{C}) \to \mathrm{O}_{1,3}$. Now

$$\begin{array}{ccc} \mathrm{SL}_2(\mathbb{C}) & \longrightarrow & \mathrm{O}_{1,3}(\mathbb{C}) \\ \uparrow & & \uparrow \\ \mathrm{SU}_2(\mathbb{C}) & \longrightarrow & \mathrm{SO}_3(\mathbb{C}) \end{array}$$

where $\mathrm{SU}_2(\mathbb{C}) = \left\{ \begin{pmatrix} a & b \\ -\overline{b} & \overline{a} \end{pmatrix} : a, b \in \mathbb{C}, |a|^2 + |b|^2 = 1 \right\}$, which is homeomorphic to the three dimensional sphere S^3, which is also homeomorphic to the unit quaternions

$$\{y_0 + y_1 I + y_2 J + y_3 K : y_0^2 + y_1^2 + y_2^2 + y_3^2 = 1\}.$$

Now recall that elements of $\mathrm{SO}_3(\mathbb{C})$ have the canonical form $\begin{pmatrix} a & b & 0 \\ -b & a & 0 \\ 0 & 0 & \pm 1 \end{pmatrix}$. Hence $\mathrm{SU}_2(\mathbb{C}) \to \mathrm{SO}_3(\mathbb{C})$ is not an isomorphism, but a 2-to-1 map, with the kernel $\{\pm I_2\}$. In fact, SO_3 is topologically $\mathbb{P}^3(\mathbb{R})$ (identify antipodal points as equivalent).

Lemma 3.1.5. $\mathrm{GL}_n(\mathbb{C})$ is path-connected, i.e.

$$\forall x, y \in \mathrm{GL}_n(\mathbb{C}),\ \exists \gamma : [0,1] \to \mathrm{GL}_n(\mathbb{C}) \text{ continuous} : \gamma(0) = x, \gamma(1) = y.$$

Proof. Let $A \in \mathrm{GL}_n(\mathbb{C})$. It suffices to find a $\gamma : \gamma(0) = I_n, \gamma(1) = A$. If A is a Jordan block, then pick $z \in \mathbb{C} : e^z = \lambda$, and define $\alpha : t \mapsto \begin{pmatrix} e^{tz} & & & \\ t & \ddots & & \\ & \ddots & \ddots & \\ & & t & e^{tz} \end{pmatrix}$, which is the desired γ. If A is not a Jordan block but for some P one has $P^{-1}AP$ is a Jordan block, then define γ as $P\alpha P^{-1}$. If the Jordan normal form of A is several Jordan blocks, do the necessary modifications. □

Week 8, lecture 3, 22 November 2024

3.2. **Matrix exponentials.** The plan for the rest of the module is a tiny bit of theory on Lie algebras $\mathfrak{g} = \mathrm{Lie}\, G$, which really come from groups, something that a whole course on Lie algebras may not even mention, which is a little perverse, but the reason algebraists just go for Lie algebras and forget about the group is that technically one needs to learn about manifolds first. Moreover, as expected, a group homomorphism $f : G \to H$ induces a Lie algebra homomorphism $f_* : \mathfrak{g} \to \mathfrak{h}$. That's the theory. The rest is going to be examples of finite matrix groups (e.g. finite reflection groups and Coxeter groups).

Today we cannot quite get started with the theory, but will be a preliminary discussion of exponentials of matrices so that we are able to talk about Lie algebras.

Definition 3.2.1. For a matrix $A \in M_{n\times n}(F)$ where F is either $\mathbb{R}$ or $\mathbb{C}$, define

$$e^A := \sum_{k \geq 0} \frac{A^k}{k!}.$$

Remark 3.2.2 (Issues of convergence)**.** We are in the $n \times n$-dimensional vector space of matrices over $\mathbb{R}$ or $\mathbb{C}$. How do we define the notion of convergence? With the norm

$$||A|| = \max\{|a_{ij}| : i = 1, \dots, n, j = 1, \dots, n\},$$

the series does converge. Note that $||AB|| = \max\{|c_{ij}|\}$ where for any i, j one has

$$|c_{ij}| = \left|\sum_{k=1}^{n} a_{ik}b_{kj}\right| \leq \sum_{k=1}^{n} |a_{ik}||b_{kj}| \leq n||A||||B||,$$

so $||AB|| \leq n||A||||B||$ and

$$\frac{||A^{k+1}||}{(k+1)!}\frac{k!}{||A^k||} \leq \frac{n}{k+1}||A||,$$

so the (absolute) convergence follows from the ratio test.

Hence, the thing defined above is a C^∞ function $M_{n\times n}(F) \to M_{n\times n}(F)$. Also, if $AB = BA$, then $e^{A+B} = e^A e^B$ (which implies that every e^A is invertible with inverse e^{-A}). If A is invertible, then $\forall B,\ e^{A^{-1}BA} = A^{-1}(e^B)A$.

Suppose A is diagonalisable, i.e. $\exists P : P^{-1}AP = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$, then

$$e^A = Pe^{P^{-1}AP}P^{-1} = P\begin{pmatrix} e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_n} \end{pmatrix}P^{-1}.$$

If A is not diagonalisable, recall the part on functions of matrices in MA251 Linear algebra.

Finally, we claim $\det(e^A) = e^{\operatorname{tr} A}$. B can be put in upper triangular form, i.e.

$$\exists P \in \mathrm{GL}_n(\mathbb{C}) : P^{-1}AP = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

so $e^{P^{-1}AP} = \begin{pmatrix} e^{\lambda_1} & & * \\ & \ddots & \\ & & e^{\lambda_n} \end{pmatrix}$, hence

$$\det e^A = \det\left(P^{-1}e^B P\right) = \det\left(e^{P^{-1}AP}\right) = e^{\lambda_1}\cdots e^{\lambda_n} = e^{\lambda_1+\cdots+\lambda_n} = e^{\operatorname{tr} A}.$$

Week 9, lecture 1, 25 November 2024

Proposition 3.2.3. Fix $B \in M_{n\times n}(F)$ and consider $\gamma_B : F \to \mathrm{GL}_n(F)$ defined by $t \mapsto e^{tB}$. Then

(1) $\gamma_B : (F, +) \to (\mathrm{GL}_n(F), \cdot)$ is a continuous group homomorphism.
(2) In fact it's differentiable with derivative

$$\left.\frac{\mathrm{d}\gamma_B}{\mathrm{d}t}\right|_{t=t_0} = Be^{t_0 B} \in M_{n\times n}(F).$$

Note that B and $e^{t_0 B}$ commute.

(3) The exponential map $\exp : M_{n\times n}(F) \to \mathrm{GL}_n(F)$ defined last time is differentiable itself, and

$$\mathrm{D}\exp|_{B=0} = \mathrm{id} : M_{n\times n}(F) \to: M_{n\times n}(F).$$

Proof. 1 and 2 are proved last time; for the derivative let's first look at how $\left.\frac{\mathrm{d}\delta_B}{\mathrm{d}t}\right|_{t=0} = Be^0 = B$:

$$\begin{aligned}\left.\frac{\mathrm{d}\gamma_B}{\mathrm{d}t}\right|_{t=0} &= \lim_{h\to 0}\frac{\gamma_B(h) - \gamma_B(0)}{h} = \lim_{h\to 0}\frac{\sum_{k=0}^{\infty}\frac{(hB)^k}{k!} - I}{h} \\ &= \lim_{h\to 0}\frac{I + hB + O(h^2) - I}{h} = \lim_{h\to 0}(B + O(h)) = B.\end{aligned}$$

and for a general $t = t_0$:

$$\begin{aligned}\left.\frac{\mathrm{d}\gamma_B}{\mathrm{d}t}\right|_{t=t_0} &= \lim_{h\to 0}\frac{\gamma_B(t_0 + h) - \gamma_B(t_0)}{h} = \lim_{h\to 0}\frac{\gamma_B(t_0)\gamma_B(h) - \gamma_B(t_0)}{h} \\ &= \gamma_B(t_0)\lim_{h\to 0}\frac{\gamma_B(h) - I}{h} = e^{t_0 B}B.\end{aligned}$$

Finally, note that γ_B is a composition of the map $\varphi_B : F \to M_{n\times n}(F)$ defined by $t \mapsto tB$ and exp. Then by the chain rule,

$$B = \left.\frac{\mathrm{d}\delta_B}{\mathrm{d}t}\right|_{t=0} = \mathrm{D}\exp|_{B=\varphi(0)=0}\, \mathrm{D}\varphi_B|_{t=0} = \mathrm{D}\exp|_{B=0}\, B,$$

so $\mathrm{D}\exp|_{B=0} = \mathrm{id}$ as desired. □

Theorem 3.2.4. Suppose $\gamma : \mathbb{R} \to \mathrm{GL}_n(F)$ is a continuous group homomorphism where $\mathbb{R}$ is seen as an additive group. Then $\exists B \in M_{n\times n}(F)$ such that $\gamma(t) = \gamma_B(t) = e^{tB}$.

Proof. Recall the inverse function theorem, which implies that for $0 \in M_{n\times n}(F)$ and $\exp 0 = I \in M_{n\times n}(F)$, there are open neighbourhoods U, V which contain 0 and I (so that $V \subset \mathrm{GL}_n(F)$) respectively such that $\exp(U) \subset V$ and $\exp|_U : U \to V$ is invertible. Call its inverse log. By 3 above, $\mathrm{D}\log|_I = I$.

Week 9, lecture 2, 26 November 2024

In particular, $\exists \delta > 0 : \gamma([-\delta, \delta]) \subset V$. Let $\beta : [-\delta, \delta] \to U$ be defined by $\log \circ \gamma$. By possibly choosing a smaller δ, we can also assume that $\forall t_1, t_2 \in [-\delta, \delta]$, by continuity of addition, $\beta(t_1) + \beta(t_2) \in U$.

We claim $\forall t \in \left[-\frac{\delta}{2}, \frac{\delta}{2}\right]$ and $\forall r \in [-1, 1]$, $\beta(rt) = r\beta(t)$, i.e. β behaves additively. Let's first look at how $\beta(t) = 2\beta\left(\frac{t}{2}\right)$. First of all by our choice of δ, both $\beta(t), 2\beta\left(\frac{t}{2}\right) \in U$. To show they are equal, it suffices to show their exponentials are equal, but $\exp(\beta t) = \gamma(t)$ and $\exp\left(2\beta\left(\frac{t}{2}\right)\right) = \left(\exp\beta\left(\frac{t}{2}\right)\right)^2 = \gamma\left(\frac{t}{2}\right)^2 = \gamma(t)$ by 3.2.3.1. Moreover, $\forall t \in \left[-\frac{\delta}{2}, \frac{\delta}{2}\right]$ and $\forall k, p \in \mathbb{Z}$ with $0 \le p \le 2^k$, one has $\beta\left(\frac{p}{2^k}t\right) = \frac{p}{2^k}\beta(t)$. This will imply our claim by continuity (binary decimal expansion). We prove by induction on k. We just looked at the $k = 1$ case. Now if p is even then $\frac{p}{2^k} = \frac{p/2}{2^{k-1}}$ and we are done by induction, so suppose p is odd. Then

$$\frac{p}{2^k} = \frac{1}{2}\left(\frac{p-1}{2^k} + \frac{p+1}{2^k}\right),$$

and by induction,

$$\beta\left(\frac{p-1}{2^k}t\right) = \frac{p-1}{2^k}\beta(t), \quad \beta\left(\frac{p+1}{2^k}t\right) = \frac{p+1}{2^k}\beta(t),$$

so we need to show that

$$\beta\left(\frac{p}{2^{k-1}}\right) = \beta\left(\frac{p-1}{2^k}\right) + \beta\left(\frac{p+1}{2^k}\right).$$

Again, since they are both in U, it suffices to show their exponentials are the same. By induction on

$$\beta(t) = 2\beta\left(\frac{t}{2}\right)$$

and 3.2.3.1, one has that the LHS is $\gamma\left(\frac{t}{2^{k-1}}\right)^p$. Now write $p = 2q + 1$, then $\frac{p-1}{2^k} = \frac{q}{2^{k-1}}$, $\frac{p+1}{2^k} = \frac{q+1}{2^{k-1}}$, so the RHS is

$$\begin{aligned}\exp\left(q\beta\left(\frac{t}{2^{k-1}}\right) + (q+1)\beta\left(\frac{t}{2^{k-1}}\right)\right) &= \exp\left(q\beta\left(\frac{t}{2^{k-1}}\right)\right)\exp\left((q+1)\beta\left(\frac{t}{2^{k-1}}\right)\right)\\ &= \gamma\left(\frac{t}{2^{k-1}}\right)^q \gamma\left(\frac{t}{2^{k-1}}\right)^{q+1} = \gamma\left(\frac{t}{2^{k-1}}\right)^{2q+1} = \gamma\left(\frac{t}{2^{k-1}}\right)^p.\end{aligned}$$

Finally, to prove the theorem, we claim $\forall t \in \mathbb{R}$, one can write $\gamma(t) = e^{t\frac{\beta(\delta)}{\delta}}$. Indeed, $\exists N \in \mathbb{N} : \left|\frac{t}{N}\right| \le \delta$, so $\frac{t}{N} = r\delta$ for some $r \in [-1, 1]$. Then

$$\gamma\left(\frac{t}{N}\right) = \gamma(r\delta) = \exp\beta(r\delta) = \exp(r\beta(\delta)),$$

so

$$\gamma(t) = \gamma\left(\frac{t}{N}\right)^N = \exp(r\beta(\delta))^N = \exp(rN\beta(\delta)) = \exp\left(t\frac{\beta(\delta)}{\delta}\right)$$

□

Week 9, lecture 3, 29 November 2024: problem class

Worksheet 3 Q 1. let R be a commutative ring and M, M' be R-modules with submodules $N \le M, N' \le M'$. Suppose $N \cong N'$ and $M/N \cong M'/N'$. (i) If N is free, does this imply $M \cong M'$? (ii) What about M/N?

Solution. The situation can be summarised as a short exact sequence $0 \to N \to M \to M/N \to 0$, and the natural question is: given a module N and P, we want to study all modules M such that $N \le M$ and $M/N = P$. It's like in the study of finite groups, when we want to understand all groups of a certain

order, we find two smaller groups and see how we can "multiply" them to form the bigger one. Hence M is called an *extension* of P by N. The point is to make new modules out of known modules.

The answer to (i) is false. For example, take $R = N = \mathbb{Z}$ (which is free) and $P = \mathbb{Z}/2\mathbb{Z}$. Then there are at least 2 M that contains N as a submodule and P as the quotient: take $M = \mathbb{Z}$ and inject N into M by multiplication by 2, so $M/N = \mathbb{Z}/2\mathbb{Z} = P$, or take $M = \mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$ and inject N into M trivially.

But (ii) is true, i.e. given N and P, then if P is free, all extensions of P by N are isomorphic to $N \oplus P$.

Indeed, let M be such an extension, then there is an injective homomorphism $i_1 : N \hookrightarrow M$ and a surjective homomorphism (projection) $M \twoheadrightarrow P$. Now the universal property of direct sum $M_1 \oplus M_2$ of two modules M_1, M_2 is:

(1) There are injective (inclusion) maps $i_1 : M_1 \to M_1 \oplus M_2$ and $i_2 : M_2 \to M_1 \oplus M_2$;
(2) For any N and homomorphisms $f_1 : M_1 \to N$ and $f_2 : M_2 \to N$, there is a unique $f : M_1 \oplus M_2 \to N$ such that $fi_1 = f_1, fi_2 = f_2$.

Therefore, in order to prove $M = M_1 \oplus M_2$, we simply need to show M satisfies the two properties.

We already have i_1, so what we need is a homomorphism i_2 from P back to M. The thing about free modules is that they are absolutely tailored for defining morphisms from them to something. Recall that P is free iff $\exists B = \{e_1, \dots, e_r\} \subset P$ is a basis, so defining $P \to M$ is just giving r elements in M. For each i, choose m_i such that $\pi(m_i) = e_i$, and define $i_2 : P \to M$ by $e_i \mapsto m_i$.

Now f is unique since i_2 is defined to be the inverse of π.

Week 10, lecture 1, 2 December 2024

3.3. **Lie algebras.**

Definition 3.3.1. For $A \in G \leq GL_n(F)$, the *tangent space* to G at A is

$$T_A G \coloneqq \left\{ \dot\gamma(0) = \left.\frac{\mathrm{d}\gamma}{\mathrm{d}t}\right|_{t=o} : \gamma : (-\varepsilon, \varepsilon) \to G \text{ is differentiable as a function with values in } M_n(F),\ \gamma(0) = A \right\}$$

Call $T_I G$ the *Lie algebra* of G and denote it by $\mathfrak{g}$.

Remark 3.3.2 (Properties of the tangent space)**.** (1) $T_A G = A(T_I G)$: indeed, take $\gamma : (-\varepsilon, \varepsilon) \to G$ and consider $\gamma' = A\gamma$, then $\gamma(0) = \mathrm{id} \iff \gamma'(0) = A$, and $\dot\gamma'(0) = A\dot\gamma(0)$ by chain rule.

(2) $T_I G$ is a vector space: the key thing is to prove $X, Y \in T_I G \implies X + Y \in T_I G$. Let $\alpha, \beta : (-\varepsilon, \varepsilon) \to G$ with $\alpha(0) = \beta(0) = I$ and $\dot\alpha(0) = X, \dot\beta(0) = Y$. Consider $\gamma = \alpha\beta$. Then $\dot\gamma(0) = \dot\alpha(0)\beta(0) + \alpha(0)\dot\beta(0) = X + Y$.

Example 3.3.3. Consider $G = \mathrm{O}_n(\mathbb{R}) \subset \mathrm{GL}_n(\mathbb{R})$. What's $\mathfrak{g} = \mathfrak{o}_n$? Write $\gamma : (-\varepsilon, \varepsilon) \to G$ as $\gamma(t) = I + tX + O(t)$ where $X = \dot\gamma(0)$. Now

$$\begin{aligned} \gamma(t) \in \mathrm{O}_n &\iff I = {}^t\gamma\gamma = (I + t\,{}^tX + O(t))(I + tX + O(t)) = I + t({}^tX + X) + O(t) \\ &\iff {}^tX + X = 0, \end{aligned}$$

so $\mathfrak{o}_n$ is precisely $\{x \in M_{n\times n}(\mathbb{R}) : {}^tX = -X\}$, the antisymmetric matrices (which have only 0's on the diagonal).

Example 3.3.4. Consider $\{A \in M_{n\times n}(\mathbb{C}) : {}^t\overline{A}A = I\} = U_n \leq \mathrm{GL}_n(\mathbb{C})$. By the same process as above,

$$\mathfrak{u}_n = \{X \in M_{n\times n}(\mathbb{C}) : {}^t\overline{X} + X = 0\},$$

so in particular $a_{ii} \in \mathbb{R}\ \forall i$ and $\mathfrak{u}_n$ is a real vector space, or U_n is a "real" Lie group.

Example 3.3.5. Consider SL_n. For $\gamma(t) \in \mathrm{SL}_n$,

$$1 = \det\gamma(t) = \det(I + tX + O(t)) = 1 + t\,\mathrm{tr}\,X + O(t) \iff \mathrm{tr}\,X = 0.$$

Remark 3.3.6 (Key properties of $\mathfrak{g}$)**.** (1) Again, $\mathfrak{g}$ is a vector space.

(2) G acts on $\mathfrak{g}$ by $(g, X) = gXg^{-1}$.

(3) For $X, Y \in \mathfrak{g}$, the *Lie bracket* defined by $[X, Y] \coloneqq XY - YX \in \mathfrak{g}$, which is an anti-commutative operation (i.e. $[X, Y] = -[Y, X]$). Indeed, if $X \in \mathfrak{g}$ then suppose $\gamma(-\varepsilon, \varepsilon) \to G$ has $\gamma(0) = I$ and $\dot\gamma(0) = X$, and write $\gamma(t) = I + tX + O(t)$. Then by (2),

$$g \ni \gamma(t)Y\gamma(t)^{-1} = (I + tX + O(t))Y(I - tX + O'(t)) = I + t(XY - YX) + O''(t),$$

so $\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \gamma(t)Y\gamma(t)^{-1} = XY - YX \in \mathfrak{g}$.

(4) All $X, Y, Z \in \mathfrak{g}$ satisfy the Jacobi identity:

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

Week 10, lecture 2, 3 December 2024

Example 3.3.7. $G = \mathrm{O}_n(\mathbb{R}) \subset \mathrm{GL}_n(\mathbb{R})$. Then we said $\mathrm{Lie}(G) = \mathfrak{o}_n(\mathbb{R}) = \{X \in M_{n\times n}(\mathbb{R}) : {}^tX + X = 0\}$ by series expansion. Now suppose ${}^tX + X = 0$. Then in particular ${}^tX$ and X commute. Consider $\gamma(t) = e^{tX}$. We claim $\gamma(t) \in \mathrm{O}_n(\mathbb{R})$. Indeed, ${}^t\gamma(t)\gamma(t) = {}^t\left(e^{tX}\right)e^{tX} = e^{t^tX}e^{tX} = e^{t({}^tX+X)} = e^0 = I$. It's also clear that $\gamma(0) = I$ and $\dot\gamma(0) = X$.

Theorem 3.3.8. We can identify a Lie algebra with

$$\mathfrak{g} = \{\gamma : (\mathbb{R}, +) \to G : \gamma \text{ is a continuous group homomorphism}\} = \{t \mapsto e^{tX}\}.$$

Proof. Clearly $\{t \mapsto e^{tX}\} \subset \mathfrak{g}$, so now suppose $X \in \mathrm{Lie}(G)$, i.e. $\exists\gamma : (-\varepsilon, \varepsilon) \to G : \gamma(0) = I,\ \dot\gamma(0) = X$. We claim that for all $t \in (-\varepsilon, \varepsilon)$, one has

$$e^{tX} = \lim_{n\to\infty} \gamma\left(\frac{t}{n}\right)^n,$$

and since G is closed, if we can prove this, we have $e^{tX} \in G$. Choose, as we have done in the past, open neighbourhoods V of $0 \in M_{n\times n}(\mathbb{F})$ and V of $I \in \mathrm{GL}_n(\mathbb{F})$ such that $\exp|_U : U \xrightarrow{\sim} V$ is differentiable. By shrinking ε, we may assume $\gamma(-\varepsilon, \varepsilon) \subset V$, and $\gamma = e^\beta$ where $\beta : (-\varepsilon, \varepsilon) \to V$ and $\beta(0) = 0$. By the chain rule, $\dot\beta(0) = \dot\gamma(0) = X$. Then

$$\lim_{n\to\infty} \gamma\left(\frac{t}{n}\right)^n = \lim_{n\to\infty} \left(e^{\beta\left(\frac{t}{n}\right)}\right)^n = \lim_{n\to\infty} \left(e^{\frac{t}{n}X + o\left(\frac{t}{n}\right)}\right)^n = \lim_{n\to\infty} \left(e^{tX + no\left(\frac{t}{n}\right)}\right) = e^{tX}.$$

□

Theorem 3.3.9. Let G, H be matrix Lie groups and $f : G \to H$ a continuous group homomorphism. Define $f_* : \mathfrak{g} \to \mathfrak{h}$ using the above identification (if $\gamma : (\mathbb{R}, +) \to G$ is a continuous group homomorphism then $f_*(\gamma) = f \circ \gamma$). Then f_* is a homomorphism of Lie algebras which (i) is a linear map of real vector spaces and (ii) $\forall X, Y \in \mathfrak{g},\ f_*[X, Y] = [f_*X, f_*Y]$.

Remark 3.3.10. We don't know how to prove this theorem without using exponentials (other than heavy manifold theory) and this is the real reason for spending time on exponentials.

Week 10, lecture 3, 6 December 2024

Proof. Let $\gamma_X, \gamma_Y \in \mathfrak{g}$ be given by $t \mapsto e^{tX}$ and $t \mapsto e^{tY}$. Consider $\delta(t) = \gamma_X(t)\gamma_Y(t) = \gamma_{X+Y}(t)$. Note that $\delta(0) = I$ and by the product rule, $\dot\delta(0) = X + Y$. Then by the proof of 3.3.8,

$$\gamma_{X+Y}(t) = \lim_{n\to\infty} \left(\gamma_X\left(\frac{t}{n}\right)\gamma_Y\left(\frac{t}{n}\right)\right)^n,$$

and so

$$\begin{aligned}
\gamma_{f_*X+f_*Y}(t) &= \lim_{n\to\infty} \left(\gamma_{f_*X}\left(\frac{t}{n}\right)\gamma_{f_*Y}\left(\frac{t}{n}\right)\right)^n = \lim_{n\to\infty} \left(f\circ\gamma_X\left(\frac{t}{n}\right) f\circ\gamma_Y\left(\frac{t}{n}\right)\right)^n \\
&= \lim_{n\to\infty} f\left(\left(\gamma_X\left(\frac{t}{n}\right)\gamma_Y\left(\frac{t}{n}\right)\right)^n\right) \qquad \text{since } f \text{ is a group homomorphism} \\
&= f\left(\lim_{n\to\infty} \left(\gamma_X\left(\frac{t}{n}\right)\gamma_Y\left(\frac{t}{n}\right)\right)^n\right) \qquad \text{since } f \text{ is continuous} \\
&= f\left(\gamma_{X+Y}(t)\right) \qquad \text{by above}
\end{aligned}$$

Now let $X, Y \in \mathfrak{g}$ and for $s \in \mathbb{R}$, consider $\delta_s : \mathbb{R} \to G : t \mapsto \gamma_X(s)\gamma_Y(t)\gamma_X(s)^{-1}$, which is a group homomorphism, and $\dot\delta_s(0) = \gamma_X(s)Y\gamma_X(s)^{-1} \in \mathfrak{g}$. Now consider the map $\mathbb{R} \to \mathfrak{g} : s \mapsto \dot\delta_s(0)$, whose derivative at $s = 0$ is precisely $XY - YX = [X, Y]$. Now similarly consider $\beta_s(t) = \gamma_{f_*X}(s)\gamma_{f_*Y}(t)\gamma_{f_*X}(s)^{-1}$. By the same argument, $s \mapsto \dot\beta_s(0)$ again has derivative $[f_*X, f_*Y]$, but $\beta_s(t) = f \circ \delta_s(t)$, so by definition $\dot\beta_s(0) = f_*\dot\delta_s(0)$, and the desired follows precisely. □

Week 11, lecture 1, 9 December 2024: problem class

Problem. Suppose $G \le \mathrm{SL}_3(\mathbb{C})$ is the subgroup generated by

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

and $G \curvearrowright \mathbb{C}[x, y, z]$ in the usual way. Compute a set of generators of ring of invariants $\mathbb{C}[x, y, z]^G$.

Solution. We use Noether's method. Note that $G \cong C_2 \times C_2$. Then for x one has

$$p_x(T) = \prod_{g \in G} (T - gx) = (T - x)(T - Ax)(T - Bx)(T - ABx) = (T - x)^2(T + x)^2 = (T^2 - x^2)^2,$$

and similarly $p_y(T) = (T^2 - y^2)^2$, $p_z(T) = (T^2 - z^2)^2$. From coefficients of these three polynomials we have invariants $x^2, x^4, y^2, y^4, z^2, z^4$. We also need to average all monomials $x^a y^b z^c$ with $\max\{a, b, c\} < 2 = |G|$. Now $x - x = 0$ and $xy - xy = 0$, so we are left with xyz, and hence a set of generators of ring of invariants is x^2, y^2, z^2, xyz.

Jiewei Xiong, Department of Mathematics and Statistics, Mathematics Building, University of Reading, Whiteknights campus, Reading RG6 6AX, United Kingdom

Email address: jiewei.xiong@pgr.reading.ac.uk