Notes of Zelmanov's Algebraic Lectures

taken by Li Yunsheng based on the contents of the lectures

Lecture Information

Lecturer: Efim Zelmanov
E-mail: efim.zelmanov@gmail.com
Website of the Course: https://sustech-math.github.io/zelmanov.html
Office Hour: Tuesday 10:00 a.m.~12:00 a.m. at office at the Math Center
No midterm
No formal final exam, only an oral presentation in person.
No textbook to follow

A Remark about the Notes

The topics of Lectures 1 and 2 are very basic, hence are not written down in great detail; they are presented as a list of concepts that one needs to know, with a minimal amount of explanation. Comprehensive details are given for contents from and after Lecture 3.

Contents

Lecture 1

§1 Review of Abstract Algebra

- Groups, Subgroups $H < G$, Homomorphisms
- Normal Subgroups $H \lhd G$, Quotient of Groups, the First Isomorphism Theorem
- Natural Homomorphism $G/H_1 \to G/H_2$, $H_1 \lhd H_2 \lhd G$.
- Commutation
- Groups Generated by a Subset: $X \subset G$,

$$\langle X \rangle := \{x_1^{\varepsilon_1} \cdots x_n^{\varepsilon_n} \mid x_1, \cdots, x_n \in X, \varepsilon_i = \pm 1\}$$

- Rings (with identity), Subrings, Homomorphisms
- Ideals $I \trianglelefteq R$, Quotient by Ideals (Factor Ring), the First Isomorphism Theorem
- Natural Homomorphism $R/I_1 \to R/I_2$, $I_1 \subset I_2 \trianglelefteq R$.
- Cartesian Products of Groups $\overline{\prod_{i\in I} G_i}$, Direct Products of Groups $\prod_{i\in I} G_i$
- Direct Sums of Rings $\bigoplus_{i\in I} R_i$
- Fields, (not necessarily commutative) Algebras over a Field A/F: A is an F-algebra if it is a vector space over F equipped with a F-bilinear multiplication operator $A \times A \to A$. Similarly an algebra over a commutative ring is defined, i.e. an algebra over a commutative ring R is an R-module equipped with an R-bilinear multiplication operator.
- Algebras Generated by a Subset: $X \subset A$,

$$\langle X \rangle_F := \left\{ \sum \alpha x_1 \cdots x_n \mid \alpha \in F; x_1, \cdots, x_n \in X \right\}$$

Note that the above formula holds for all kinds of (associative) algebras

- Vector Spaces, Linear Transformations $\mathrm{Lin}_F(V)$
- Modules over an Algebra: A/F, V a vector space over F. A bilinear map $A \times V \to V$ satisfying $a(bv) = ab \cdot v$ and $1_A v = v$ makes V into a (left) module over A. Note that a bilinear map is equivalent to a homomorphism of algebras $\varphi\colon A \to \mathrm{Lin}_F(V)$.
- Semigroup (with identity), Congruence Relation: $x \sim y \Rightarrow xz \sim yz$ and $zx \sim zy$.
- the "First Isomorphism Theorem" for semigroups:

  Let $\varphi : S_1 \to S_2$ be a homomorphism of semigroups. Define a congruence relation by that $x \sim y$ if and only if $\varphi(x) = \varphi(y)$, then $S_2 \cong S_1/\sim$
- Natural Homomorphism $S/\sim_1 \to S/\sim_2$, $\sim_1 \subset \sim_2$.

Lecture 2

§2 Generators and Relations

Free Semigroups

Given a set X, the *free semigroup generated by* X is

$$X^* \coloneqq \{\text{all words of elements in } X\}.$$

It enjoys the following universal property:

Proposition 2.1. Universal Property of Free Semigroup

Let S be a semigroup and $\varphi\colon X \to S$ be a mapping, then φ uniquely extends to a homomorphism $\varphi\colon X^* \to S$.

It is easy to observe the following: let S be a semigroup and $\mathscr{A} = \{a_i\}_{i\in I}$ be a set of generators of S, then the inclusion $\mathscr{A} \hookrightarrow S$ induces a surjective homomorphism $\varphi\colon \mathscr{A}^* \to S$.

Therefore $S \cong \mathscr{A}^*/\sim$, where $\sim$ is defined by $a \sim b$ if and only if $\varphi(a) = \varphi(b)$.

Generated Congruence

Let $\sim$ be a congruence on S, then it can be seen as a subset of $S \times S$. Let $R \subset S \times S$, we say that R generates $\sim$ if $\sim$ is the smallest congruence that contains R, or equivalently,

$$R = \bigcap \text{all congruences that contains } R.$$

The congruence generated by R always exists, noticing that $S \times S$ is itself a congruence.

Presentation

Let S be a semigroup and is isomorphic to the quotient of a free semigroup X^*, $S \cong X^*/\sim$. Let $R = \{a_k \times b_k\}_k$ be a subset of $X^* \times X^*$ that generates $\sim$, then we say that

$$S = \langle X \mid a_k = b_k \rangle,$$

and S is *presented* by generators X and relations R.

If $|X| < \infty$ and $|R| < \infty$, then we say that S is *finitely presented*. It turns out that the notion of finitely presented is independent of the choice of generators, see proposition 2.3.

Undecidable Word Problem

In general, given a presentation, it is algorithmatically undecidable to tell whether two words are equal under that presentation.

Extension Problem

Let $S = \langle s_i, i \in I \rangle$ and T be two semigroups. Let $\{t_i\}_{i\in I} \subset T$ and consider the map

$$\varphi\colon \{s_i\}_{i\in I} \to \{t_i\}_{i\in I}\colon s_i \mapsto t_i.$$

Question

When is φ extendable to a homomorphism?

The answer is quite simple. Write $S = \langle x_i, i \in I \mid a_k = b_k \rangle$ as is presented by $X = \{x_i\}_{i\in I}$ and $R = \{a_k \times b_k\}_k$, where $x_i \mapsto s_i$ gives a natural map $X^* \to S$, then

Proposition 2.2. Characterization of Extension

φ extends to a homomorphism if and only if it preserves all the defining relations, in other words, $a_k(\boldsymbol{t}) = b_k(\boldsymbol{t})$.

Proof. Just notice that we have naturally $S \cong X^*/\sim$ and $X^*/\sim_1 \hookrightarrow T$, and $\sim \subset \sim_1$ gives a homomorphism $X^*/\sim \to X^*/\sim_1$. □

Finitely Presented

Let $S = \langle s_1, \cdots, s_m \rangle = \langle s'_1, \cdots, s'_k \rangle$ be generated by two different sets of generators.

Proposition 2.3. Finitely Presented is Well-defined

If S is finitely presented in $s_1, \cdots, s_m$, then it is also finitely presented in $s'_1, \cdots, s'_k$.

Proof. Write $S = \langle x_1, \cdots, x_m \mid a_1 = b_1, \cdots, a_n = b_n \rangle$ with $x_i \mapsto s_i$. Since s_i and s'_i generate S, we have

$$s_i = c_i(\boldsymbol{s}'), \quad 1 \le i \le m,$$

$$s'_j = d_j(\boldsymbol{s}), \quad 1 \le j \le k,$$

where $c_i(\boldsymbol{s})$ denotes some algebraic combination of s_j's, and similarly is $d_j(\boldsymbol{s})$ defined. These give rise to two kinds of relations that are satisfied in S:

$$a_l(\boldsymbol{c}(\boldsymbol{s}')) = b_l(\boldsymbol{c}(\boldsymbol{s}')), \quad 1 \le l \le n, \tag{I}$$

$$s'_j = d_j(\boldsymbol{c}(\boldsymbol{s}')), \quad 1 \le j \le k. \tag{II}$$

Let $Y^* = \langle y_i, \cdots, y_k \rangle$ be a free semigroup, the homomorphism $Y^* \to S \colon y_j \mapsto s'_j$ gives $S \cong Y^*/\sim_1$. Under the convention $y_j \rightsquigarrow s'_j$, the two kinds of relations above are all included by $\sim_1$, hence if we let them generate a relation $\sim_2$ on Y^*, then $\sim_2 \subset \sim_1$, giving rise to a homomorphism $Y^*/\sim_2 \to Y^*/\sim_1$.

On the other hand, the first kind of relations, with proposition 2.2, defines a homomorphism

$$S \cong X/\sim \to Y^*/\sim_2 \colon s_i \mapsto x_i \mapsto c_i(\boldsymbol{y}').$$

It remains only to show that these two homomorphisms are the inverse to each other.

Exercise 2.1

Complete the rest of the proof.

It is easy to see that $S \cong X/\sim \to Y^*/\sim_2 \to Y^*/\sim_1 \cong S$ is the identity on S. For the other direction, we have

$$\begin{array}{ccccc} Y^*/\sim_2 & \to & Y^*/\sim_1 \cong S \cong X/\sim & \to & Y^*/\sim_2 \\ y_j & \mapsto & y_j \rightsquigarrow s'_j \rightsquigarrow d_j(\boldsymbol{x}) & \mapsto & d_j(\boldsymbol{c}(\boldsymbol{y})) \overset{\text{by (II)}}{=\!=} y_j \end{array}$$ □

Therefore, as we claimed before, the notion of finitely presented does not depend on the choice of the finite set of generators.

Length-lex (Lexicographical) Order

Consider a free semigroup $X^* = \langle x_i, i \in I \rangle$ whose index I is ordered. We define an order on X^* by the following: for any two elements

$$v = x_{i_1} \cdots x_{i_n}, \quad w = x_{j_1} \cdots x_{j_m},$$

1. if $n > m$ or $m < n$, then $v > w$ or $w < v$ respectively;
2. else $n = m$, then compare i_1 and j_1: if $i_1 > j_1$ then $v > w$; if $i_1 = j_1$, then compare i_2 and j_2 and so on.

This order on X^* is called *the length-lex order*, or *the lexicographical order*.

Definition 2.1. Minimality Condition

An ordered set satisfies *the minimal condition* if there does not exists an infinite descending chain $a_1 > a_2 > \cdots$ in it.

Theorem 2.4. Lexicographical Order Inherits Minimality Condition

If $X = (X, >)$ satisfies the minimal condition, then the length-lex order on X^* also satisfies the minimality condition.

Proof. Suppose that

$$v_1 > v_2 > \cdots$$

is an infinitely descending chain in X^*, then the length of v_i's forms a descending chain

$$\ell(v_1) \geq \ell(v_2) \geq \cdots$$

in $\mathbb{N}$. Thus there must exists n such that $\ell(v_n) = \ell(v_{n+1}) = \cdots$. Then $v_n > v_{n+1} > \cdots$ is an infinitely descending chain in X^* where all elements have the same length. The first letters of this sequences gives $x_{i_1} \geq x_{i_2} \geq \cdots$, which stabilizes since X has the minimality condition. Cut the sequence again and then consider the second letters, and so on. After $\ell(v_n)$ steps, we see that the original chain must stabilize, contradicting the assumption. □

Lecture 3

Free Semigroup Algebras

Let F be a field and S be a semigroup, we can consider the *semigroup algebra on* S,

$$FS \coloneqq \{\alpha_1 s_1 + \cdots \alpha_n s_n\}.$$

Remark 2.1

Unlike building a ring on an abelian group, here the operation of the semigroup induces the multiplication in the algebra, instead of the summation.

Given any set X, we can consider the *free associative F-algebra on the set of free generators X*

$$F\langle X\rangle \coloneqq FX^* \coloneqq \{\alpha_1\omega_1 + \cdots + \alpha_n\omega_n \mid \alpha_i \in F, \omega_i \in X^*\}.$$

Note that $F\langle X\rangle$ is a unitary algebra, since the empty word lies in X^*. It enjoys the following universal property:

Proposition 2.5. Universal Property of Free Associative F-algebra

For any F-algebra A and an arbitrary mapping $\varphi\colon X \to A$, φ uniquely extends to a homomorphism of F-algebras $\varphi\colon F\langle X\rangle \to A$.

Consider the case where A is generated by $\{a_j\}_{j\in J}$, $X \coloneqq \{x_j\}_{j\in J}$ and $\varphi\colon x_j \mapsto a_j$, then the induced homomorphism $\varphi\colon F\langle X\rangle \to A$ is surjective, giving $A \cong F\langle X\rangle/I$ where $I = \ker\varphi$.

Let $R \subset F\langle X\rangle$, we say that R generates I as an ideal if I is the smallest ideal containing R, or equivalently,

$$I = \left\{ \sum_j a_j\iota_j b_j \mid a_j, b_j \in F\langle X\rangle, \iota_j \in R \right\}.$$

Similar as the presentation of semigroups, when $A \cong F\langle X\rangle/I$ and I is generated by R, we may write

$$A = \langle X \mid R = 0\rangle.$$

Also similarly, finitely presented is defined and can be proved to be well-defined.

§3 Reduction to Irreducible Elements

The same undecidable word problem exists for presentations of algebras, but we do have some algorithm under certain conditions:

Keep the notations above. Now suppose X is (totally) ordered and satisfies the minimality condition. For an element $r \in R$, it can be written as

$$r = \alpha_1\omega_1 + \cdots + \alpha_n\omega_n,$$

with $\alpha_i \neq 0$ and $\omega_i \in X^*$ are distinct words. The maximal element among $\omega_1, \cdots, \omega_n$, $\bar{r} \coloneqq \omega_i = \max(\omega_1, \cdots, \omega_n)$, is called *the leading monomial of* r, and its coefficient α_i is called *the leading coefficient of* r. In A, we have $r = 0$, which gives the following relation:

$$\alpha_i\omega_i = -\sum_{j\neq i}\alpha_j\omega_j \quad \Rightarrow \quad \omega_i = -\sum_{j\neq i}\frac{\alpha_j}{\alpha_i}\omega_j.$$

Definition 3.1. Reducible Word

A word $v \in X^*$ is called *reducible* if it contains a leading monomial $\bar{r}$ of some $r \in R$ as a *subword*, i.e. $v = v'\bar{r}v''$ for some $v', v'' \in X^*$.

Therefore if v is reducible, then in A, $v = \sum_k \alpha_k u_k$, $\alpha_k \in F$, $u_k < v$. This means that v can be *reduced* into a sum of "smaller" words.

Definition 3.2. Irreducible Word

A word is *irreducible* if it is not reducible.

Let us denote by $Ir \subset X^*$ the set of all irreducible words.

Proposition 3.1. Irreducible Words Span the Algebra

Ir spans A.

Proof. The result follows immediately by the reduction of reducible words and the minimality condition.□

Gröebuer-Shirshov Bases

We wish to use Ir as a basis of A, which would solve our undecidable word problem in this case completely since we can reduce any element in A to a linear combination of irreducible words within finitely many steps. Hence the following question arises:

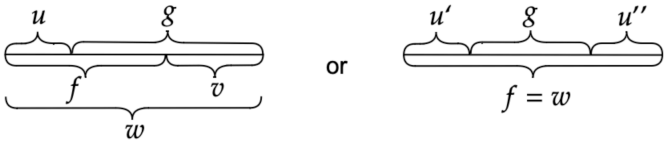
Question

When is Ir linearly independent in A?

Definition 3.3. Composition of Words

We say that two words v, w *admit a composition* if the end of one of these words equals the beginning of the other one, e.g. $v = x_3x_5x_1$ and $w = x_5x_1^2x_4$ admit a composition, or one of these words is a subword of the other one.

Suppose that $f, g \in R$ and the leading coefficients of f and g are 1. Suppose that $\bar{f}, \bar{g}$ admit a composition, i.e. they can be pieced together to get a word w as the following illustrates:



We define the composition of f and g, $(f, g)_w$, as

$$(f, g)_w = fv - ug \text{ or } f - u'gu'',$$

respectively to the two cases in the picture. Note that there may be several different choices of w associated to f and g, for example, $axyxb$ and $axyxyxb$ are both available w's of $f = axyx$ and $g = xyxb$.

Theorem 3.2. Gröebuer-Shirshov Bases

Ir is a basis of A if and only if for any two relations $f, g \in R$ that admit a composition, all these compositions $(f, g)_w$ reduce to 0 (in the free algebra) (instead of a nontrivial linear combination of irreducible elements).

By "reduce" we mean to substitute reducible words by the sum of smaller words given by the relations in R and obtain a new element in $F\langle X\rangle$ which is "closer" to a sum of irreducible words, and so on, until we get a linear combination of irreducible words (or 0).

The basis Ir found in this way is called the *Gröebuer-Shirshov basis*.

Remark 3.1. Reformulation of Reduction

The procedure of reduction can be reformulated as the following: An element $r \in R$ is of the form $r = \bar{r} - r'$. A word $v\bar{r}u$ is equal to $v(r + r')u = vru + vr'u$ in the free algebra, which lives in $vr'u + I(R)$. By passing through such equalities, any element is a linear combination of irreducible words whose leading monomial is smaller than the original one of the element, plus $I(R)$. The "old version" of reduction differs with this reformulation in the sense that it erases elements in $I(R)$. So far we can conclude that a necessary condition for R to be *closed under composition* (i.e. for any $f, g \in R$, $(f, g)_w$ reduces to 0) is that for any $f, g \in R$, we can write in the free algebra $(f, g)_w = \sum_j \alpha_j v_j r_j u_j$ with $\alpha_j \in F$, $v_j, u_j \in X^*$, $r_j \in R$ and $v_j \bar{r}_j u_j < w$ for all j.

Exercise 3.1

Show that the necessary condition above for R to be closed under composition is in fact sufficient.

Remark 3.2

A consequence of this theorem is that it does matter how we reduce words, since different non-trivial linear combination of irreducible words would be distinct.

One direction is easy: $(f, g)_w$ is always 0 in A since $f = g = 0$ in A and the procedure of reduction gives equality in A. Hence if $(f, g)_w$ does not reduce to 0, then we obtain that a nontrivial linear combination of irreducible words equals 0 in A, indicating that Ir is not linearly independent in A.

The other direction is straightforward, but is long and uninteresting. Before the complete proof, let us go through some examples of applications of the theorem:

Applications of Gröebuer-Shirshov Bases

Example 3.1

Consider $\langle x, y \mid yx - xy = 1\rangle$ with order $x < y$. There is nothing to compose, and the irreducible words are words of the form $x^i y^j$, which form a basis by the theorem. This algebra is isomorphic to the *Weyl algebra*, the algebra generated by $y = \frac{d}{dt}$ and $x = t$ seen as linear operators on the space of differentiable functions.

Example 3.2

Consider $\langle x, y, z \mid [x, y] = z, [z, y] = 2y, [z, x] = -2x\rangle$ with order $x < y < z$, where $[a, b] \coloneqq ab - ba$. The relations give the following reductions

$$\begin{cases} yx \to xy - z \\ zy \to yz + 2y \\ zx \to xz - 2x \end{cases}$$

Now that the first two elements admit a composition: $zy - yz - 2y$ and $yx - xy + z$. The composition is

$$(zy - yz - 2y)x - z(yx - xy + z) = -yzx - 2yx + zxy - z^2.$$

The reduction goes:

$$\begin{aligned} &\to -y(xz - 2x) - 2(xy - z) + (xz - 2x)y - z^2 \\ &= -yxz + 2yz - 4xy + 2z + xzy - z^2 \\ &\to -(xy - z)z + 2(xy - z) - 4xy + 2z + x(yz + 2y) - z^2 = 0. \end{aligned}$$

Since these are the only two elements in the relation that admit a composition, we see by theorem 3.2 that the irreducible words $x^i y^j z^k$ form a basis.

Lecture 4

Example 3.3

This example shows that the order on X matters. Consider $\langle x, y \mid y^2x - xyx = 0\rangle$. When $x < y$, the leading word is y^2x and there is no nontrivial composition with itself; when $y < x$, the leading word is xyx, which admits a nontrivial composition with itself, $xyxyx$. The first order, by the theorem, shows that the irreducible words (in that order) form a basis of A. However, for the second order we have nontrivial composition

$$(f, f)_{xyxyx} = (xyx - y^2x)yx - xy(xyx - y^2x) = -y^2xyx + xy^3x.$$

The relation gives reduction

$$xyx \to y^2x,$$

Hence the above composition reduces to

$$-y^4x + xy^3x,$$

which is a nontrivial linear combination of irreducible words, showing that the irreducible words (in this order) are not linearly independent.

The next example is about Lie algebras. Let us briefly recall some definitions.

Definition 3.4. Lie Algebra

A *Lie algebra* L is a vector space with a bilinear operation $[\cdot, \cdot]\colon L \times L \to L$ that satisfies

(1) (Antisymmetry) $[a, b] = -[b, a]$;

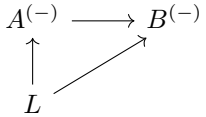
(2) (Jacobi identity) $[[a, b], c] + [b, c], a] + [[c, a], b] = 0$,

for any $a, b, c \in L$.

Clearly, a *homomorphism of Lie algebras* is defined as a linear map that commutes with the brackets.

Definition 3.5. Representation of Lie Algebra

A *representation of a Lie algebra* L is a homomorphism of Lie algebras $\varphi\colon L \to A^{(-)}$, where A is an associative algebra and $A^{(-)}$ is the Lie algebra with the bracket $[a, b] = ab - ba$. A *homomorphism of representations of* L is a homomorphism of Lie algebras $A^{(-)} \to B^{(-)}$ that makes the following triangle commutes.



A representation of L is called *universal* if it is initial in the category of representations of L.

Note that for a Lie algebra L there may not exists an algebra A such that $L = A^{(-)}$.

Lemma 3.3. Image of Lie Algebra Generates Universal Enveloping Algebra

Let $u\colon L \to U^{(-)}$ be a universal representation of L. Then U is generated by $u(L)$ as an associative algebra.

Proof. Let $\langle u(L)\rangle$ be the associative algebra generated by $u(L)$, then $u\colon L \to \langle u(L)\rangle^{(-)}$ is also a representation of L. The universality gives a homomorphism of representations of L, $U^{(-)} \to \langle u(L)\rangle^{(-)}$, which is identical to the identity when restricted on $\langle u(L)\rangle$; in particular it is surjective. Note that the inclusion $\langle u(L)\rangle^{(-)} \hookrightarrow U^{(-)}$ is also a homomorphism of representations of L. The composition $U^{(-)} \to \langle u(L)\rangle^{(-)} \hookrightarrow U^{(-)}$ gives a homomorphism of representations of L, which must be the identity on $U^{(-)}$, implying that $U^{(-)} \to \langle u(L)\rangle^{(-)}$ is injective. Therefore $U = \langle u(L)\rangle$. □

The initial property gives immediately that

Proposition 3.4. Uniqueness of Universal Enveloping Algebra

If a universal representation exists then it is unique up to isomorphism.

Also, in this case, the existence is always true:

Proposition 3.5. Existence of Universal Enveloping Algebra

The universal representation of a Lie algebra L always exists.

Proof. Let $\{e_i\}_{i\in I}$ be a basis of L (for infinite-dimensional L, use the Hamel basis), then we have

$$[e_i, e_j] = \sum_k \gamma_{ij}^k e_k,$$

for some $\gamma_{ij}^k \in F$ for any $i, j \in I$. Write $X = \{x_i\}_{i\in I}$ and consider

$$U \coloneqq \langle X \mid x_i x_j - x_j x_i - \sum_k \gamma_{ij}^k x_k = 0\rangle.$$

The homomorphism $\varphi\colon L \to U^{(-)}$ defined by $e_i \mapsto x_i$ gives the universal representation, as one can verify. □

Note that φ is never surjective, because $\varphi(L) = \operatorname{span}_F\{x_i \mid i \in I\} \subsetneq U$ since $U \ni x_i x_j \notin \operatorname{span}_F\{x_i \mid i \in I\}$. Let us call the unique U in the universal representation of L as *the universal enveloping algebra of the Lie algebra L.*

Example 3.4

Example. Consider the famous Lie algebra $sl_2(F) \coloneqq \{2 \times 2 \text{ matrices with zero trace}\}$. It has basis

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$[e, f] = h, \quad [f, h] = 2f, \quad [e, h] = -2e.$$

Its universal enveloping algebra is exactly $\langle x, y, z \mid [x, y] = z, [y, z] = 2y, [x, z] = -2x\rangle$, which has been discussed in example 3.2.

Poincaré-Birkhoff-Witt Theorem

Let $R = \{x_i x_j - x_j x_i - \sum_k \gamma_{ij}^k x_k \mid i, j \in I\}$ and introduce any order on I with the minimality condition, which always exists provided the Axiom of Choice.

Theorem 3.6. Poincaré-Birkhoff-Witt

R is closed with respect to compositions, i.e., for any $f, g \in R$, $(f, g)_w$ reduces to 0. And the irreducible words $\{x_{i_1} \cdots x_{i_k} \mid i_1 \le \cdots \le i_k\}$ form a basis of $U = \langle X \mid R = 0\rangle$.

Proof. See the next lecture. □

Corollary 3.6.1. Lie Algebra Embeds Into Universal Enveloping Algebra

Let $\varphi\colon L \to U^{(-)}$ be the universal representation of L, then φ is an embedding.

Proof. For any $\sum_i \alpha_i e_i \neq 0$ in L, it is mapped by φ to $\sum_i \alpha_i x_i$ in U, which is a non-trivial linear combination of irreducible words, hence is nonzero by Poincaré-Birkhoff-Witt Theorem. □

Lecture 5

Proof of the Poincaré-Birkhoff-Witt Theorem

We first show that R is closed with respect to compositions. Write $\{x_i, x_j\} \coloneqq \sum_k \gamma_{ij}^k x_k$, then $\mathrm{span}_F\{x_i \mid i \in I\}$ with this bracket is a Lie algebra that is isomorphic to L. In particular, the bracket $\{\cdot,\cdot\}$ satisfies the Jacobi identity. Any two elements in R that admits a nontrivial composition are of the form

$$x_j x_i - x_i x_j - \{x_j, x_i\},$$
$$x_k x_j - x_j x_k - \{x_k, x_j\},$$

where $i < j < k$. Their composition is

$$\begin{aligned}
&(x_k x_j - x_j x_k - \{x_k, x_j\})x_i - x_k(x_j x_i - x_i x_j - \{x_j, x_i\}) \\
=&- x_j x_k x_i - \{x_k, x_j\}x_i + x_k x_i x_j + x_k\{x_j, x_i\} \\
\to&- x_j(x_i x_k + \{x_k, x_i\} - \{x_k, x_j\}x_i + (x_i x_k + \{x_k, x_i\})x_j + x_k\{x_j, x_i\} \\
=&- x_j x_i x_k - x_j\{x_k, x_i\} - \{x_k, x_j\}x_i + x_i x_k x_j + \{x_k, x_i\}x_j + x_k\{x_j, x_i\} \\
\to&- \{x_j, x_i\}x_k - x_j\{x_k, x_i\} + x_i\{x_k, x_j\} + \{x_k, x_i\}x_j + x_k\{x_j, x_i\} \\
=&[\{x_i, x_j\}, x_k] + [\{x_k, x_i\}, x_j] + [\{x_j, x_k\}, x_i] \\
\to&\{\{x_i, x_j\}, x_k\} + \{\{x_k, x_i\}, x_j\} + \{\{x_j, x_k\}, x_i\} = 0,
\end{aligned}$$

where $[\cdot,\cdot]$ denotes the commutator, i.e., $[a, b] = ab - ba$. This ends the first part. Keep in mind that the above procedure, though looks complicated, can be done automatically by a computer with only one single click of button.

The following displays one reason why the universal enveloping is important. Recall that if L is a Lie algebra and V is a vector space, then a homomorphism $L \to \mathrm{Lin}_F(V)^{(-)} \colon a \mapsto T_a$ defines an action of L on V, and we have $T_{[a,b]} = T_a T_b - T_b T_a$. Lifting this homomorphism to $U^{(-)} \to \mathrm{Lin}_F(V)^{(-)}$, we then obtain a homomorphism of associative algebras $U \to \mathrm{Lin}_F(V)$.

Joke

Associative algebras are in general easier to deal with than Lie algebras. However, the above procedure is still a trade off of difficulties: even if L is finitely dimensional, U is infinite dimensional. This displays *the law of conservation of difficulty*.

The rest of the Poincaré-Birkhoff-Witt Theorem follows from the other direction of the theorem 3.2, which we are now going to prove.

Proof of the Groöebuer-Shirshov Bases Theorem

Before the proof, let us briefly recall the statement of the theorem:

Theorem. Groöebuer-Shirshov Bases

Ir is a basis in A if and only if for any two relations $f, g \in R$ that admit a composition, all these compositions $(f, g)_w$ reduce to 0.

Proof. Recall that the necessity has been proved right after theorem 3.2. For the other direction, we show that for any nonzero $f \in I(R)$ the leading monomial $\bar{f}$ is reducible, hence a nontrivial linear combination of irreducible words is never zero modulo $I(R)$, implying the linear independence of irreducible words.

For any $f \in I(R)$, we can write $f = \sum_i \alpha_i u_i r_i v_i$ for finitely many $\alpha_i \in F \setminus \{0\}$, $u_i, v_i \in X^*$ and $r_i \in R$; note that u_i and v_i are words, while r_i's are linear combinations of words.

Note that we have $\overline{u_i r_i v_i} = u_i \bar{r}_i v_i$. Write $w \coloneqq \max_i\{\overline{u_i r_i v_i}\}$ and define for convention

$$S \coloneqq \{i \in I \mid w = \overline{u_i r_i v_i}\}.$$

If $\#S = 1$, then $\bar{f} = w$ is reducible and we are done.

If $\#S > 1$, it may occur that $\sum_{i\in S} \alpha_i w = 0$ so that $\bar{f} \neq w$. To resolve this problem we use induction on $(w, \#S) \in \overline{I(R)} \times \mathbb{N}^*$, where $\overline{I(R)}$ denotes the set of leading monomials of elements in $I(R)$ and $\overline{I(R)} \times \mathbb{N}^*$ is equipped with the lexicographical order that compares w firstly and then $\#S$, which satisfies the minimality condition.

Since for $\#S = 1$ the statement is true no matter what w is, the initial condition is satisfied and we can proceed by induction, supposing that $\#S > 1$ and that the statement is true for all pairs less than $(w, \#S)$.

Now that $\#S > 1$, so there exists $i \neq j$ with $u_i \bar{r}_i v_i = u_j \bar{r}_j v_j = w$. We have

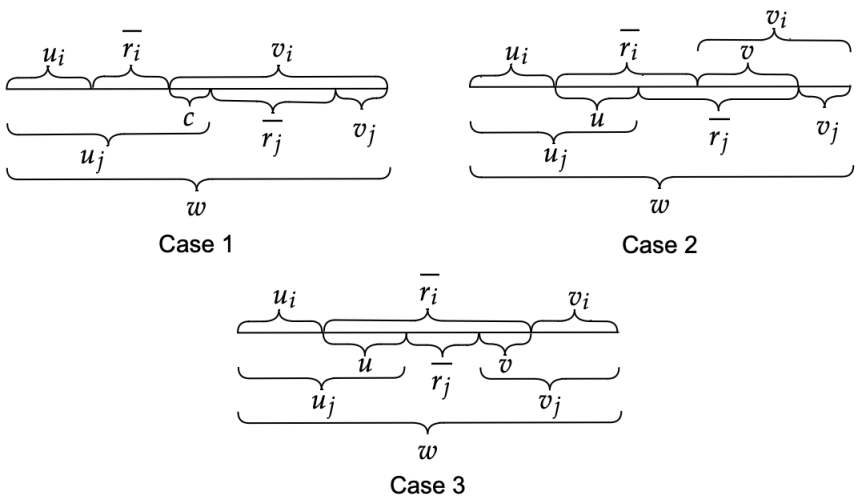
$$\alpha_i u_i r_i v_i + \alpha_j u_j r_j v_j = (\alpha_i + \alpha_j) u_i r_i v_i + \alpha_j (u_j r_j v_j - u_i r_i v_i).$$

The left hand side attributes 2 elements to S. For the right hand side, $(\alpha_i + \alpha_j) u_i r_i v_i$ attributes 1 or 0 denpending on whether $\alpha_i + \alpha_j$ is zero. If we can show that $u_j r_j v_j - u_i r_i v_i$ is of the form $u_j r_j v_j - u_i r_i v_i = \sum_k \beta_k u'_k r'_k v'_k$ with $\max_k\{\overline{u'_k r'_k v'_k}\} < w$, then by replacing the left hand side by the right hand side in the summation $f = \sum_i \alpha_i u_i r_i v_i$, we obtain a new summation that expresses f where either $\#S$ is smaller (while w is fixed) or w is smaller (while whatever $\#S$ becomes), consequently we are done by the induction.

Therefore, for our purpose, there are now three cases to discuss:

1. $\bar{r}_i$ and $\bar{r}_j$ do not intersect in w, i.e. (up to a permutation) $w = u_i \bar{r}_i c \bar{r}_j v_j$ where $c \in X^*$, hence we have $v_i = c\bar{r}_j v_j$ and $u_j = u_i \bar{r}_i c$.
2. $\bar{r}_i$ and $\bar{r}_j$ intersect, but no one is contained by the other.
3. one of $\bar{r}_i$ and $\bar{r}_j$ is contained by the other.

Below gives the illustration of these three cases along with the notations that we will use later in our proof.



Let us keep the notation that $r_i = \bar{r}_i - r'_i$ (and similarly for r_j) in the following.

For Case 1, we have

$$\begin{aligned} u_j r_j v_j - u_i r_i v_i =& u_i \bar{r}_i c r_j v_j - u_i r_i c \bar{r}_j v_j \\ =& u_i((r_i + r'_i) c r_j - r_i c (r_j + r'_j)) v_j \\ =& u_i (r'_i c r_j - r_i c r'_j) v_j = u_i r'_i c r_j v_j - u_i r_i c r'_j v_j. \end{aligned}$$

Since $\max\{\overline{u_i r_i c r_j v_j}, \overline{u_i r_i c r'_j v_j}\} < \overline{u_i r_i c r_j v_j} = w$, we obtain the result as desired.

The other two cases will be discussed in the next lecture.

Lecture 6

Now for Case 2, keeping the notations in the illustration above, we have

$$\begin{aligned} u_j r_j v_j - u_i r_i v_i =& u_i u r_j v_j - u_i r_i v v_j \\ =& u_i (u r_j - r_i v) v_j \\ =& - u_i (r_i, r_j)_{u\bar{r}_j = \bar{r}_i v} v_j. \end{aligned}$$

Recall that by remark 3.1, we have $(f,g)_w = \sum_k \alpha_k v_k r'_k u_k$ with $v_k \bar{r}_k u_k < w$ for all k for any $f, g \in R$, let $f = r_i$, $g = r_j$ and $w = u\bar{r}_j = \bar{r}_i v$ then we are done.

The argument for Case 3 goes similarly. □

Exercise 3.2

Finish the rest of the proof for Case 3.

Gröebuer-Shirshov Bases For Semigroups

Let us call a finite presentation $\langle X \mid R \rangle$ (either for semigroups or for algebras), where X is equipped with an order with the minimal condition, a *reduction system*.

A reduction system is *confluent*, if for every word in it, the result of reduction of the word is irrelevant to the choice of how the reduction is applied.

With this definition, the theorem which we just proved can be reformulated as that, for algebras, a reduction system is confluent if and only if R is closed under composition. The following lemma by Newman tells essentially the same thing for semigroups:

Lemma 3.7. Newman

For semigroups, a reduction system $S = \langle X \mid u_i = v_i \rangle$ with $u_i > v_i$ is confluent if and only if for any u_i and u_j such that $v' u_j = u_i v''$ for some $v', v'' \in X^*$, $v' v_j$ and $v_i v''$ has the same descendant*, i.e. they reduce to a same word after finitely many steps of reductions.*

Note that in a reduction system of a semigroup, every word can be represented by a irreducible word. The above gives a necessary and sufficient condition for all irreducible words to be different. If this condition is satisfied, then we call the irreducible words as *normal forms*, and we conclude that every word can be reduced to a unique normal form.

§4 Further Applications of Gröebuer-Shirshov Bases

Let us give two general and important examples of algebras where the algorithm of reduction applies: graded algebras and commutative algebras.

Graded Algebras

Let $A = \oplus_{i=1}^{\infty} A_i$ be a graded algebra with $A_0 = F$ and $A = \langle A_1 \rangle$. Furthermore, we assume that $\dim_F A_i < \infty$ for all i, so that A is finitely generated as an algebra (by a basis of A_1).

Note that for homogeneous f and g, their composition (if exists) $(f,g)_w$ is also homogeneous and we always have $\deg(f,g)_w > \max\{\deg f, \deg g\}$, where $\deg 0$ is ∞ by convention.

Let $A = \langle X \mid R = 0 \rangle$ be a finite presentation where every relation in R is homogeneous. Define an order with the minimality condition on X so that we obtain a reduction system. Write $R_0 := R$. Define inductively that

$$R_n \coloneqq R_{n-1} \cup \{(f,g)_w \mid f,g \in R_{n-1}\},$$

i.e., R_n is R_{n-1} union all possible compositions of elements in R_{n-1}. Since R_0 is finite, each R_n is also finite.

Write $R_\infty \coloneqq \bigcup_{i=0}^{\infty} R_i$, then $A = \langle X \mid R = 0\rangle = \langle X \mid R_\infty = 0\rangle$, since if $f,g \in R$, then $(f,g)_w = fv - ug \in I(R)$, where $I(R)$ is the ideal generated by R.

Now that R_∞ is closed under composition, so the Groöebuer-Shirshov bases theorem applies and we see that the set of irreducible words (with respect to R_∞) is a basis of A.

Although R_∞ might contain infinitely many elements, we still have an algorithm for reduction in this case: for any elemen a in A, there must exists $N \in \mathbb{N}$ such that every element in $R_\infty \setminus R_N$ is of degree strictly larger than the degree of any homogeneous component of a, because, by the construction, the minimal degree of elements in $R_\infty \setminus R_n$ strictly increases as n increases. Therefore to reduce a into a linear combination of irreducible words, we need only check the reduction relations in R_N, where there are only finitely many of them.

Commutative Algebras

Let A be a finitely generated commutative algebra, then we can find a surjective homomorphism

$$\varphi\colon F[x_1,\cdots,x_n] \twoheadrightarrow A,$$

so that $A \cong F[x_1,\cdots,x_n]/\ker\varphi$.

Recall the following lemma by Hilbert, which is a standard result in commutative algebra, c.f. Corollary 2.13 in [Kem11]:

Lemma 4.1. Hilbert

Every ideal of a polynomial ring $F[x_1,\cdots,x_n]$ is finitely generated.

Therefore every finitely generated commutative algebra admits a finite presentation, e.g. let $r_1,\cdots,r_m$ be a set of generators of $\ker\varphi$, then $A = \langle x_1,\cdots,x_n \mid r_i = 0\rangle$.

Let us now define the composition of two elements in a polynomial ring $F[x_1,\cdots,x_n]$.

Without loss of generality, let the generators be ordered as $x_1 > \cdots > x_n$. To compare two monomials, we compare firstly their degrees, then the numbers of powers of x_1, and then the numbers of powers of x_2 and so on (i.e. from the largest generator to the smallest generator). Given a polynomial f, we define its *leading monomial* $\bar{f}$ as the largest monomial among all of its monomials.

Now for any two polynomials $f,g \in F[x_1,\cdots,x_n]$, we say that they are composable if their leading monomials have a non-constant common divisor, i.e. an element $d \in F[x_1,\cdots,x_n] \setminus F$ such that $\bar{f} = ad$ and $\bar{g} = bd$ for two elements $a,b \in F[x_1,\cdots,x_n]$. The composition of f and g with respect to this common divisor is thus defined as

$$(f,g)_{abd} \coloneqq bf - ag.$$

Similar to definition 3.1 and definition 3.2, reducible monomials and irreducible monomials are defined, and a similar argument shows that the Groöebuer-Shirshov bases theorem in this case is also true.

Now that given a commutative algebra A along with a finite presentation $A = \langle X \mid R = 0\rangle$, we want to apply the Groöebuer-Shirshov bases theorem to it to obtain a basis of A along with an algorithm of reduction. Like what we did for graded algebras, we would like to consider $R_0 \coloneqq R$ and then add compositions of elements in R_{n-1} to obtain R_n, while seeking for a way to obtain an algorithm. For this pupose, we consider the following proposition:

Proposition 4.2

Among every infinite set of monomials of finite many variables there exists two (distinct) monomials that one divides the other one.

Proof. Let us proceed by induction on the number of variables. The case where there is only one variable is trivial.

Suppose that the statement is true for $n-1$ variables. Let us identify monomials in n variables bijectively to elements in $\mathbb{N}^n$, hence we say that a monomial $(i_1,\cdots,i_n)$ divides another monomial $(j_1,\cdots,j_n)$ if and only if $i_k \le j_k$ for all $k = 1,\cdots,n$. Suppose we have an infinite set of monomials S in

which no monomial divides another one, then among all the corresponding tuples of the monomials we find a tuple whose first index is the smallest, say $(i'_1, \cdots, i'_n)$ with $i'_1 \leq i_1$ for all other tuples $(i_1, \cdots, i_n)$. For any other element $(i_1, \cdots, i_n) \in S$, there must be $i_k < i'_k$ for some $k = 2, \cdots, n$, hence if we define $S_k \coloneqq \{(i_1, \cdots, i_n) \in S \mid i_k < i'_k\}$, then there must be

$$S = \{(i'_1, \cdots, i'_n)\} \cup \left(\bigcup_{k=2}^{n} S_k \right).$$

Therefore one of S_k's must be infinite (since S is). Up to a relabelling let us say that S_2 is infinite. Write $T_l \coloneqq \{(i_1, \cdots, i_n) \in S \mid i_2 = l\}$, then we have

$$S_2 = \bigcup_{0 \leq l \leq i'_2 - 1} T_l.$$

Again, there must exist an $l \in \{0, \cdots, i'_2 - 1\}$ such that T_l is infinite. Now that the second index of elements in T_l is the constant l, hence there exists two n-variable monomials in T_l that one divides the other if and only if there exists two $(n-1)$-variable monomials in $T'_l \coloneqq \{(i_1, i_3, \cdots, i_n) \mid (i_1, i_2, i_3, \cdots, i_n) \in T_l\}$ that one divides the other. By our induction assumption we see that there exists two monomials in $T_l \subset S$ that one divides another, contradicting the definition of S. □

With this proposition, there must exist an $N \in \mathbb{N}$ such that the leading monomial of any element in $R_\infty \setminus R_N$ is divisible by the leading monomial of some element in R_N. Noticing that elements in $R_\infty \setminus R_N$ do not give any new reducible word other than those are given by R_N, we see that R_∞ and R_N define a same set of irreducible words. Therefore every element in A can be reduced to a linear combination of irreducible words using only the relations in R_N. By theorem 3.2, since R_∞ is closed under compositions, the set of irreducible words does give a basis of A. Since we need only the finitely many relations in R_N to operate the reduction, we obtain an algorithm. These fulfill our purpose completely.

Remark 4.1

This result for commutative algebra is called Buchberger's theorem, or Buchberger's algorithm.

Lecture 7

Let us now look at the number of steps that we need to reduce a word in a reduction system of a semigroup.

§5 Dehn function

In a reduction system (whether confluent or not), we say that two words are *equivalent* if they are *congruent*, i.e. they have the same descendent. This means that two congruent words can be transformed to each other by finitely many steps of substituting the relations, say.

$$u = w_1 \sim w_2 \sim \cdots \sim w_r = v.$$

Given two congruent words u and v, we denote by $\|u \times v\|$ the length of the smallest chain of substitution that we need to go through to transform u into v.

The *Dehn function* $D\colon \mathbb{N} \to \mathbb{R}_{>0}$ of a reduction system of a semigroup is now defined by

$$D(n) \coloneqq \max\{\|u \times v\| \mid u, v \text{ are equivalent with lengths no more than } n\}.$$

The maximum always exists since there are only finitely many u and v for a fixed n. Clearly, the Dehn function gives a measurement of the complexity of a system.

Given two functions $f, g\colon \mathbb{N} \to \mathbb{R}_{>0}$, we say that f is *asymptotically less or equal to* g, denoted as $f \preceq g$, if there exists $C \in \mathbb{N}$ such that

$$f(n) < Cg(Cn), \quad \forall n \in \mathbb{N}.$$

If $f \preceq g$ and $f \succeq g$, then we say that f and g are *asymptotically equivalent*, denoted by $f \sim g$.

Theorem 5.1. Dehn Functions Are Asymptotically Equivalent

Given two finite presentations of a semigroup, their corresponding Dehn functions are asymptotically equivalent.

Proof. Let $\langle X \mid R\rangle \cong \langle Y \mid R'\rangle$ be two finite presentations, then any generator $y_i \in Y$ can be written as a word $y_i(x)$ in X. Since there are only finitely many generators in Y, we can find an upper bound C of the lengths $\{\text{length}_X(y) \mid y \in Y\}$. For any two words $u = v$ with lengths less than n in Y, their lengths in X are thus less than Cn. Also, for each relation in R', it may be achieved by finitely many compositions of relations in R; since there are only finitely many relations, we may enlarge our C so that the number of needed compositions for each relation is always less than C. These give us

$$\|u \times v\|_Y \leq CD_X(Cn).$$

Take the maximum of the left hand side and then we obtain $D_Y(n) \leq CD_X(Cn)$. □

Therefore we may think Dehn function as an equivalence class of functions corresponding to a semigroup, regardless of the choice of presentations.

In a confluent reduction system, let us denote by $\gamma_{\min}(v)$ the minimum time of reduction that is needed to reduce v to its normal form and by $\gamma_{\max}(v)$ the maximum time of reduction. Note that by considering the time of reduction, we are requiring that each step gives a smaller word so that we cannot substitute a same relation back and forth, hence the maximum time is well-defined.

By linking two equivalent words with their normal form, which is a same irreducible word in a confluent reduction system, we see that

$$\|u \times v\| \leq \gamma_{\min}(u) + \gamma_{\min}(v).$$

Note that we have by definition that

$$\gamma_{\min}(u) = \|u \times \tilde{u}\|,$$

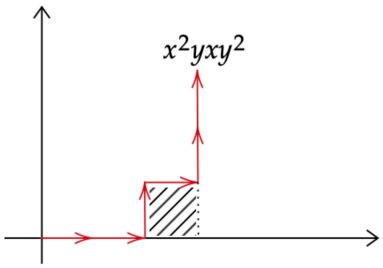
where $\tilde{u}$ denotes the normal form of u. Since the normal form of a word has length no longer than that word, we see that the Dehn function is asymptotically equivalent to the function $\gamma\colon \mathbb{N} \to \mathbb{R}_{>0}$ defined by

$$\gamma(n) \coloneqq \max\{\gamma_{\min}(u) \mid \text{length}(u) \leq n\}.$$

Let us take a look at some examples where we can compute the Dehn function.

Example 5.1

Consider $\langle x, y \mid xy = yx\rangle$ with $x < y$. Clearly the irredcubile words are words of the form $x^i y^j$. We can relate words in this system with 2-dimensional graphics: given any word, we start from the origin, and read each letter in the word from the left to the right. Each time we read x, we go right for a unit length, and each time we read y, we go up for a unit length. For example, the word x^2yxy^2 relates to the following graph:



Below the graph and the horizontal axis there is a region which is shadowed in the illustration. Let us call the area of the shadowed region the *area of the word*, say, the area of x^2yxy^2 is 1. We see that the irreducible words are exactly the words of 0 area, and each time we apply the reduction $yx \to xy$ to a word, its area goes down by 1. Therefore, in this case, $\gamma_{\max}(u) = \gamma_{\min}(u) = \text{Area}(u)$. For a word of length n, its area is at maximum $\left(\frac{n}{2}\right)^2$, therefore we conclude that the Dehn function of this system is asymptotically equivalent to n^2.

Example 5.2

Consider $\langle x_1, \cdots, x_m \mid x_i x_j = x_j x_i, 1 \leq i, j \leq m\rangle$ with $x_1 < \cdots < x_m$. Write a word as $v = x_{i_1} \cdots x_{i_n}$, we can define its area as

$$\text{Area}(v) \coloneqq \#\{(k,l) \in \{1, \cdots, n\}^2 \mid x_{i_k} > x_{i_l},\ k < l\}.$$

For example, $\text{Area}(x_3x_2x_2x_1) = 3 + 1 + 1 = 5$. Again, each reduction reduces the area by 1, hence the Dehn function is asymptotically equivalent to the maximum area. For a word of length n, its area is bounded above by n^2 by definition. Also, since $\text{Area}(x_m^{n/2}x_1^{n/2}) = \left(\frac{n}{2}\right)^2$, the maximum area is bounded below by $\left(\frac{n}{2}\right)^2$. These conclude that the Dehn function of this system is asymptotically equivalent to n^2.

In general, if we can define the area of the words in a system in a way that the irreducible words are exactly those of area 0 and each step of reduction reduces an area that is bounded uniformly, i.e., irrelevant to whatever the word is, then we can estimate the Dehn function of the system. For example, if in a reduction system the reductions always reduce an area more than ε_1 and less than ε_2, i.e. for any word wuw' and relation $u = v$ with $u > v$ in that system, we have

$$\text{Area}(wvw') + \varepsilon_1 < \text{Area}(wuw') < \text{Area}(wvw') + \varepsilon_2,$$

then the number K of required steps to reduce a word w is among $[\text{Area}(w)/\varepsilon_2, \text{Area}(w)/\varepsilon_1]$. Let $A(n) \coloneqq \max_{\text{length}(w) \leq n}\{\text{Area}(w)\}$, then the Dehn function is estimated by

$$D(n) \in [A(n)/\varepsilon_2, A(n)/\varepsilon_1].$$

In particular, the Dehn function is asymptotically equivalent to the maximum-area function A.

Remark 5.1

Although one can define the Dehn function similarly for reduction systems of algebras in a naive way, the situation for algebras is more complicated and has been remained undecided. Say, let us define $D(n)$ as the maximum of the minimum number of steps of reduction for reducing an element in which each monomial has length no more than n. Since there are only finitely many generators, there are only finitely many monomials of length no more than n and finitely many ways to reduce even if we take the factorization into consideration, so the Dehn function is well-defined. However, when it comes to decide a way to reduce, the question becomes complicated: if we reduce the monomials term by term, then we may be too slow (at least m^n) compared to the real Dehn function; if we want to factorize firstly and then reduce, then how should we decide how to factorize? For example, the algebra $\langle x, y \mid x^2 = 0\rangle$. It is very simple: once you see x^2 you kill that term. However, it may still cost about m^n steps to reduce if we look term by term, which means that the Dehn function defined above fails to measure the complexity of the system. Maybe this can be solved when the quantum computer comes out, where we will be able to reduce all the monomials at the same time for one single step.

Now let us end this chapter and move on to talk about the free groups.

§6 Free Groups

[Zelmanov introduced the definition, uniqueness and construction of free group which are not taken down here; for these things, see for example Chapter II, Section 5.1~5.3 of [Alu09].]

In short, let $X = \{x_i\}_{i\in I}$ be a set of generators and $Y = \{y_i\}_{i\in I}$ is another set of generators with the same index set I, then the following semigroup

$$\langle X, Y \mid x_i y_i = 1, y_i x_i = 1,\ i \in I\rangle,$$

is the free group generated by X, and it is a confluent system as one can verify using our previous theory. The normal forms in the above semigroup are called *reduced forms* of the free group generated by X. Explicitly, the reduced forms are

$$x_1^{\varepsilon_1} \cdots x_n^{\varepsilon_n},$$

where $\varepsilon_i = \pm 1$, $x_i \in X$ and we require that no subword is of the form xx^{-1} or $x^{-1}x$. We will keep this notation for reduced forms.

Let us denote the free group generated by X by $F(X)$. If $|X| = m < \infty$, then we may write $F(X) = F(m)$. It is obvious that the free groups are always isomorphic if their sets of generators have a same cardinality.

Question

For different cardinalities of the index, can the associated free groups be isomorphic? For example, is a free group generated by n elements isomorphic to a free group generated by m elements if $n \neq m$?

The answer is yes, for both finite and infinite cases. Here we only talk about the finite cases. We may answer this question firstly for the abelian case:

[Zelmanov introduced the definition, uniqueness and construction of free abelian group which are not taken down here; for these things, see for example Chapter II, Section 5.4 of [Alu09].]

Proposition 6.1

Two finitely generated free abelian groups are isomorphic if and only if they are freely generated by a same number of elements

Proof. Recall that two finitely generated free abelian groups must be of the form $\mathbb{Z}^n$ and $\mathbb{Z}^m$. If they are isomorphic, say $\varphi\colon \mathbb{Z}^n \to \mathbb{Z}^m$ is an isomorphism, then we have

$$\varphi^{-1}(2\mathbb{Z}^m) = 2\varphi^{-1}(\mathbb{Z}^m) = 2\mathbb{Z}^n.$$

Therefore $\mathbb{Z}^m/2\mathbb{Z}^m \cong \mathbb{Z}^n/\varphi^{-1}(2\mathbb{Z}^m) = \mathbb{Z}^n/2\mathbb{Z}^n$ by the first isomorphism theorem, see the diagram below.

$$\begin{array}{ccc} \mathbb{Z}^n & \xrightarrow{\quad\varphi\quad} & \mathbb{Z}^m \\ \downarrow & & \downarrow \\ \mathbb{Z}^n/2\mathbb{Z}^n = \mathbb{Z}^n/\varphi^{-1}(2\mathbb{Z}^m) & \xrightarrow[\cong]{} & \mathbb{Z}^m/2\mathbb{Z}^m \end{array}$$

Since $\mathbb{Z}^n/2\mathbb{Z}^n \cong \bigoplus_{i=1}^n \mathbb{Z}/2\mathbb{Z}$, we thus obtain $2^m = 2^n$, concluding that there must be $m = n$. □

Lecture 8

Let us recall some constructions in group theory.

- Given a group G, recall that the elements of the form $[a,b] = a^{-1}b^{-1}ab$ are called the commutators. Notice that $ab = ba[a,b]$, hence in some way the commutator measures how far the elements a and b are from being commutative. Note also that there is $[a,b]^{-1} = [b,a]$.
- Recall that $[G,G]$ the subgroup generated by all commutators in G is normal.
- Recall that $\mathrm{Aut}(G)$ is a group. For each element $x \in G$, we can define an automorphism $g \mapsto x^{-1}gx$ on G; automorphisms of such form are called *inner-automorphisms*. It is easy to verify that the set of inner-automorphisms forms a normal subgroup of $\mathrm{Aut}(G)$; let us denote it as $\mathrm{InAut}(G)$. Let us call the quotient group $\mathrm{Aut}(G)/\mathrm{InAut}(G)$ the group of outter-automorphisms, which we denote by $\mathrm{OutAut}(G)$.
- For an arbitrary normal group of G, the only thing we can say is that it is invariant under the inner-automorphisms, but not all automorphisms. However, the commutator subgroup $[G,G]$ is invariant under all automorphisms on G (just check the generators).
- Recall that for any normal subgroup H of G such that G/H is abelian, we have $[G,G] \subset H$. Hence we have that

$$[G,G] = \bigcap_{H \lhd G,\ G/H \text{ abelian}} H.$$

We are now ready to present the following lemma:

Lemma 6.2

$F(m)/[F(m),F(m)]$ is the free abelian group of rank m, i.e. the free abelian group generated by m elements.

Proof. Firstly we check that the quotient preserves that different x_i's in X to be distinct, i.e. we need to check that $x_i^{-1}x_j \notin [F(m),F(m)]$ for any $i \neq j$. Indeed, noticing that for any element in $[F(m),F(m)]$, the number of appearances of x_i must be equal to the number of appearances of x_i^{-1}. Hence we are done for this part.

The rest of the proof is a straightforward verification of the fact that $F(m)/[F(m),F(m)]$ satisfies the universal property of the free abelian groups. □

The result that $F(m) \cong F(n)$ if and only if $m = n$ thus follows from the above lemma, the proof of proposition 6.1 and the proposition itself.

Schreier Theorem

We now proceed to another result, that all subgroups of free groups are free.

Recall that for any subgroup H, we have $G = \bigsqcup_i Hg_i$. Two elements $x, y \in G$ lies in a same coset Hg_i if and only if $xy^{-1} \in H$. In each coset Hg_i, let us select one representative with only one restriction that for H we choose the identity 1; for any element $g \in G$, we denote by $\bar{g}$ the representative that we chose in the coset Hg. Let S be the set of all representatives.

Remark 6.1

It is easy to verify that $\overline{\bar{a}\bar{b}} = \overline{ab}$ for any $a, b \in G$.

Let X generate G and consider the elements of the form $sx^{\varepsilon}(\overline{sx^{\varepsilon}})^{-1}$ where $s \in S$ and $x \in X$. Note that $sx^{\varepsilon}(\overline{sx^{\varepsilon}})^{-1} \in H$.

Lemma 6.3

The set $\{sx(\overline{sx})^{-1} \mid s \in S, x \in X\}$ generates H.

Proof. Notice that we have that, since $s = \overline{\overline{sx^{-1}}x}$ (using remark 6.1),

$$\left(sx^{-1}\left(\overline{sx^{-1}}\right)^{-1}\right) \cdot \left(\overline{sx^{-1}}x\left(\overline{\overline{sx^{-1}}x}\right)^{-1}\right) = 1.$$

It follows that $sx^{-1}\left(\overline{sx^{-1}}\right)^{-1} \in \{sx(\overline{sx})^{-1} \mid s \in S, x \in X\}$.

Also, inspired by the above equality, we have the following algorithm: For any element $h \in H$, it has reduced form $h = x_{i_1}^{\varepsilon_1} \cdots x_{i_n}^{\varepsilon_n}$. Since $\overline{h} = 1$, we have

$$\begin{aligned} h =& x_{i_1}^{\varepsilon_1} \cdots x_{i_n}^{\varepsilon_n} \\ =& \left(1 \cdot x_1^{\varepsilon_1}\left(\overline{1 \cdot x_1^{\varepsilon_1}}\right)^{-1}\right)\left(\overline{x_{i_1}^{\varepsilon_1}} \cdot x_{i_2}^{\varepsilon_2}\left(\overline{\overline{x_{i_1}^{\varepsilon_1}} \cdot x_{i_2}^{\varepsilon_2}}\right)^{-1}\right) \cdots \left(\overline{x_{i_1}^{\varepsilon_1} \cdots x_{i_{n-1}}^{\varepsilon_{n-1}}} \cdot x_{i_n}^{\varepsilon_n}\left(\overline{x_{i_1}^{\varepsilon_1} \cdots x_{i_n}^{\varepsilon_n}}\right)^{-1}\right), \end{aligned}$$

as desired. □

Corollary 6.3.1

Let G be finitely generated group. Let $H < G$ be a subgroup of G with $|G : H| < \infty$. Then the group H is finitely generated.

Proof. Since $|G : H| < \infty$, S is finite. Since G is finitely generated, X is finite. Therefore the generating set of H given by lemma 6.3 is also finite. □

More precisely, if $|X| = m$ and $|G : H| = n$, then H is generated by no more than mn elements.

We now move back to discussions about free groups. In the following, let H be a subgroup of $F(m)$.

Definition 6.1. Schreier System

We say that a set of representatives S of cosets of H in $F(m)$ is a Schreier system if for any $s \in S$, the reduced form $s = x_{i_1}^{\varepsilon_1} \cdots x_{i_k}^{\varepsilon_k}$ satisfies that $x_{i_1}^{\varepsilon_1} \cdots x_{i_j}^{\varepsilon_j} \in S$ for each $j = 1, \cdots, k$.

Lemma 6.4. Existence of Schreier System

For any subgroup $H < F(X)$, a Schreier system exists.

Proof. For any $g \in F(X)$ with reduced form $g = x_{i_1}^{\varepsilon_1} \cdots x_{i_k}^{\varepsilon_k}$, we say that k is the length of g. The length of the identity is defined to be 0. Given a coset $C = Hg$, the length of C is defined as $\min\{\text{length}(a) \mid a \in C\}$. We now proceed by induction on lengths.

Note that $\text{length}(H) = 0$, and H is the only coset with length 0. Suppose that for every coset C of length less than n, we have selected a representative $\bar{C}$ such that $\text{length}(\bar{C}) = \text{length}(C)$ and $\bar{C}$ satisfies the Schreier condition, i.e. if $\bar{C} = x_{i_1}^{\varepsilon_1} \cdots x_{i_k}^{\varepsilon_k}$ then $x_{i_1}^{\varepsilon_1} \cdots x_{i_j}^{\varepsilon_j} \in \overline{Hx_{i_1}^{\varepsilon_1} \cdots x_{i_j}^{\varepsilon_j}}$. Note that $\text{length}(Hx_{i_1}^{\varepsilon_1} \cdots x_{i_j}^{\varepsilon_j}) = j$, otherwise $\bar{C}$ can be replaced by a shorter word. For a coset C of length n, there exists some reduce form of length n, $x_{i_1}^{\varepsilon_1} \cdots x_{i_{n-1}}^{\varepsilon_{i_{n-1}}} x_{i_n}^{\varepsilon_n} \in C$, and we define

$$\bar{C} \coloneqq \overline{x_{i_1}^{\varepsilon_1} \cdots x_{i_{n-1}}^{\varepsilon_{i_{n-1}}}} x_{i_n}^{\varepsilon_n},$$

whose reduced form is of length n and satisfies the Schreier condition as desired. □

Theorem 6.5. Shreiez

Let $H < F(X)$ and S be a Schreier system corresponding to H. Then $\{sx(\overline{sx})^{-1} \neq 1 \mid s \in S,\ x \in X\}$ is a set of free generators of H.

Proof. We show firstly that if a reduction happens when we put sx^{ε} and $(\overline{sx^{\varepsilon}})^{-1}$ together for some $s \in S$, $x \in X$ and $\varepsilon = \pm 1$, then there must be $sx^{\varepsilon}(\overline{sx^{\varepsilon}})^{-1} = 1$. Let the reduced form of s be $s = x_{i_1}^{\varepsilon_1} \cdots x_{i_n}^{\varepsilon_n}$ and that of $\overline{sx^{\varepsilon}}$ be $\overline{sx^{\varepsilon}} = x_{j_1}^{\delta_1} \cdots x_{j_m}^{\delta_m}$, then

$$sx^{\varepsilon}(\overline{sx^{\varepsilon}})^{-1} = x_{i_1}^{\varepsilon_1} \cdots x_{i_n}^{\varepsilon_n} x^{\varepsilon} x_{j_m}^{-\delta_m} \cdots x_{j_1}^{-\delta_1},$$

hence a reduction happens if either $x_{i_n}^{\varepsilon_n} x^{\varepsilon} = 1$ or $x^{\varepsilon} x_{j_m}^{-\delta_m} = 1$. If $x_{i_n}^{\varepsilon_n} x^{\varepsilon} = 1$, then $sx^{\varepsilon} = x_{i_1}^{\varepsilon_1} \cdots x_{n-1}^{\varepsilon_{n-1}} \in S$ by the Schreier condition, hence $sx^{\varepsilon} = \overline{sx^{\varepsilon}}$, consequently $sx^{\varepsilon}(\overline{sx^{\varepsilon}})^{-1} = 1$. If $x^{\varepsilon} x_{j_m}^{-\delta_m} = 1$, then $\overline{sx^{\varepsilon}} = x_{j_1}^{\delta_1} \cdots x_{j_{m-1}}^{\delta_{m-1}} x^{\varepsilon}$. Write $s_1 \coloneqq x_{j_1}^{\delta_1} \cdots x_{j_{m-1}}^{\delta_{m-1}}$, then $s_1 \in S$ by the Schreier condition, and we have $sx^{\varepsilon}(\overline{sx^{\varepsilon}})^{-1} = ss_1^{-1}$. Therefore it suffices to show that $ss_1^{-1} \in H$, so that s and s_1 live in a same coset and thus $s = s_1$. Indeed, we have

$$H \ni sx^{\varepsilon}(\overline{sx^{\varepsilon}})^{-1} = sx^{\varepsilon}x^{-\varepsilon}s_1^{-1} = ss_1^{-1},$$

done.

The rest of the proof will be given in the next lecture.

Lecture 9

Continue of the Proof of Theorem 6.5. Consider a nontrivial reduced product in $\{sx(\overline{sx})^{-1} \neq 1 \mid s \in S,\ x \in X\}$,

$$s_1 x_1^{\varepsilon_1} \left(\overline{s_1 x_1^{\varepsilon_1}}\right)^{-1} \cdots s_k x_k^{\varepsilon_k} \left(\overline{s_k x_k^{\varepsilon_k}}\right)^{-1},$$

where by reduced we mean that no two adjacent $s_i x_i^{\varepsilon_i} \left(\overline{s_i x_i^{\varepsilon_i}}\right)^{-1}$ cancel. It suffices to show that cancellations won't touch $x_i^{\varepsilon_i}$'s, so that no nontrivial reduced product is equal to 1, which implies the freeness.

By what we have shown in the beginning of the proof, no cancellation happens in the middle of $s_i x_i$ nor $x_i \left(\overline{s_i x_i^{\varepsilon_i}}\right)^{-1}$, the only chance for anything to get cancelled lies in the middle of $\left(\overline{s_i x_i^{\varepsilon_i}}\right)^{-1} s_{i+1}$. We will check that such cancellation will not kill any of $x_i^{\varepsilon_i}$ and $x_{i+1}^{\varepsilon_{i+1}}$, which finishes the proof. It suffices to do this for $i = 1$ for simplicity of notation.

Let us check $x_1^{\varepsilon_1}$ firstly, assuming that $x_2^{\varepsilon_2}$ is not killed. the only possibility is that s_2 has reduced form $s_2 = \left(\overline{s_1 x_1^{\varepsilon_1}}\right) x_1^{-\varepsilon_1} \cdots$. By the Schreier condition, $\left(\overline{s_1 x_1^{\varepsilon_1}}\right) x_1^{-\varepsilon} \in S$. Since

$$\overline{\left(\overline{s_1 x_1^{\varepsilon_1}}\right) x_1^{-\varepsilon_1}} = \overline{s_1 x_1^{\varepsilon_1} x^{-\varepsilon_1}} = \overline{s_1} = s_1,$$

we obtain that $\left(\overline{s_1 x_1^{\varepsilon_1}}\right) x_1^{-\varepsilon} = s_1$. Therefore $s_1 x_1^{\varepsilon_1} \left(\overline{s_1 x_1^{\varepsilon_1}}\right)^{-1} = 1$, contradiction.

Then let us check x_2^{ε}, assuming that $x_1^{\varepsilon_1}$ is not killed. Again, the only chance is that $\left(\overline{s_1 x_1^{\varepsilon_1}}\right)^{-1} = \cdots x_2^{-\varepsilon_2} s_2^{-1}$, hence $\overline{s_1 x_1^{\varepsilon_1}} = s_2 x_2^{\varepsilon_2} \cdots$. By the Schreier condition, this means that $s_2 x_2^{\varepsilon_2} \in S$, therefore $s_2 x_2^{\varepsilon_2} \left(\overline{s_2 x_2^{\varepsilon_2}}\right)^{-1} = 1$, contradiction.

Finally, let us check that $x_1^{\varepsilon_1}$ and $x_2^{\varepsilon_2}$ cannot be killed simultaneously. If they are killed simultaneously, then $x_1^{\varepsilon_1} \left(\overline{s_1 x_1^{\varepsilon_1}}\right)^{-1} s_2 x_2^{\varepsilon_2} = 1$. Hence $s_2 x_2^{\varepsilon_2} = \overline{s_1 x_1^{\varepsilon_1}} x_1^{-\varepsilon_1}$. Therefore

$$s_1 \left(\overline{s_2 x_2^{\varepsilon_2}}\right)^{-1} = s_1 \left(\overline{\left(\overline{s_1 x_1^{\varepsilon_1}}\right) x_1^{-\varepsilon_1}}\right)^{-1} = s_1 s_1^{-1} = 1,$$

contradicting that the product is reduced. □

We now obtain a set of free generators of H along with an algorithm, but the following question remains:

Question

What is the cardinality of $\{sx(\overline{sx})^{-1} \neq 1 \mid s \in S,\ x \in X\}$? That is to say, how many $sx\left(\overline{sx}\right)^{-1}$ is equal to 1?

Suppose that $|X| = m$. Assuming the Schreier condition, the answer to the latter question is $n - 1$, so that H is freely generated by $mn - n + 1$ elements. (Recall that n is the cardinality of S.)

In fact, if $sx\left(\overline{sx}\right)^{-1} = 1$, then a cancellation must happen in the middle of either sx or $x\left(\overline{sx}\right)^{-1}$. Conversely, we have shown that if a cancellation happens then $sx\left(\overline{sx}\right)^{-1} = 1$. For each nontrivial element s in S, its reduced form is $s = x_1^{\varepsilon_1} \cdots x_k^{\varepsilon_k}$. If $\varepsilon_k = 1$, then $\left(\overline{sx_k^{-1}}\right) x_k s^{-1} = 1$. If $\varepsilon_k = -1$, then $sx_k\left(\overline{sx_k}\right)^{-1} = 1$. By a simple argument using the Schreier condition, one sees that these two cases do not give repeated pairs $sx\left(\overline{sx}\right)^{-1} \leftrightsquigarrow (s, x) \in S \times X$. Therefore among the pairs $(s, x) \in S \times X$, there are exactly $n - 1$ pairs that lead to $sx\left(\overline{sx}\right)^{-1} = 1$.

A consequence of this result is that subgroups of a free group may have larger rank than the original free group. In fact, a subgroup of a finitely generated free group may have rank infinity, see the second set of exercises.

Note that this algorithm for computing generators of subgroups of free groups can be applied to any (finitely generated) group: by sending the free generators to generators of the group, we obtain an epimorphism $F(m) \twoheadrightarrow G$. For any subgroup $H < G$, the preimage of H gives a subgroup of $F(m)$. Apply the algorithm to that subgroup of $F(m)$ to compute its generators, and then bring the generators to G, and we obtain generators of H.

Cayley Graph of Group

Consider a group $G = \langle a_1, \cdots, a_m \rangle = \langle X \rangle$, the *cayley graph* of G with respect to generators a_i's , denoted as $\mathrm{Cay}(G, X)$, is constructed by the following:

> For each element in G we assign a vertice, thus vertices in the graph are identified with elements in G. For each pair of vertices g and $a_i g$, we connect them with an edge.

In $\mathrm{Cay}(G, X)$, every element $g = a_1^{\varepsilon_1} \cdots a_k^{\varepsilon_k}$ is connected to the identity via

$$g - (a_2^{\varepsilon_2} \cdots a_k^{\varepsilon_k}) - \cdots - a_k^{\varepsilon_k} - 1.$$

Therefore the cayley graph is always connected.

By claiming that each edge has length 1 and that the distance between any two vertices is the minimal length that one has to go through the edges, the cayley graph becomes a metric space. The space can be also assigned a norm of length, by $\mathrm{length}(g) = d(g, 1)$.

A *cycle* in a graph is a loop without self-intersection. A graph without any cycle is called a *tree*. It is easy to see the following characterization of freeness:

Proposition 6.6

$Cay(G, X)$ is a tree if and only if G is a free group on free generators X.

Proof. It is easy to observe that a cycle exists if and only if a reduced word is trivial. □

For any element $a \in G$, we can consider its action on the cayley graph M of G, by sending a vertice g to ga, and edges $g - a_i g$ to $ga - a_i ga$ accordingly. Let us denote this action by R_a, then clearly R_a preserves the distance, hence it is an isometry on M. Also, we have $R_a R_b = R_{ba}$. Therefore the map $a \mapsto R_a$ gives an embedding $G^{op} \hookrightarrow \mathrm{Isom}(M)$. Note that every isometry on a graph preserves vertices, because the vertices are exactly the points with integer lengths.

In general, we can define group actions on any metric space M: an action of G on M is a group homomorphism $G \to \mathrm{Isom}(M)$.

With such action, we have another perspective of freeness. Recall that a *fixed-point free action* is an action where the identity is the only element whose action has a fixed point.

Theorem 6.7. Serre

G is free if and only if there is a fixed-point free action of G on a tree.

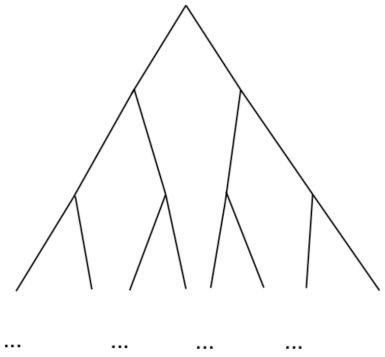
Note that we do not require the tree in the theorem to be a cayley graph of some group. It can be an arbitrary graph that is a tree.

If G acts freely on a tree, then any of its subgroup also acts freely on that tree. Therefore we obtain again that every subgroup of a free group is free, from this perspective.

Below gives an example of a group acting on a tree:

Example 6.1. Rooted Tree

Consider the *rooted tree*, which is illustrated below:



The dots indicate that the tree grows down indefinitely. The set of isometries on this tree is the set of reflections of its branches (let the length of the top point be zero and check the length of each vertex). The group of isometries naturally gives a group that acts on this tree. Every isometry must preserve the top point since it is the only point jointed with only two edges, hence no subgroup of this group of isometries is free.

We will talk about the rooted tree in detail in section 9.

Lecture 10

Let us talk a little bit more about the word problem before proceeding.

Let us consider finitely generated groups. Say $G \cong F(m)/N$, $R \subset N$ generates N so that $G = \langle X \mid R = 1\rangle$. The problem to determine whether two different words in $F(m)$ are equal when brought to G by quotienting by N is the famous *word problem*. The research into the word problem significantly contributed to the development of computer science, in the sense that it made it clear what an algorithm is. It was proved by P. Novikov (1959) that there exists a finitely presented group for which no computer can ever exist that can decide whether an arbitrary word is equal to 1.

Still, we can have some discussion about this. Let us focus on the reduction of elements in N. Note that we have

$$N = \{(\tau_1^{g_1})^{\pm 1} \cdots (\tau_k^{g_k})^{\pm 1} \mid \tau_i \in R, g_i \in F(m)\},$$

where $\tau^g := g^{-1}\tau g$, the conjugation. Note that the conjugation indeed follows the rule of exponentiation: $(\tau^{g_1})^{g_2} = g_2^{-1}g_1^{-1}\tau g_1 g_2 = \tau^{g_1 g_2}$.

Recall the definition of the Dehn function for semigroups in section 5. For groups, we can also define the Dehn function: for any $h \in N$, let $\|h\|$ denote the minimal possible k such that $h = (\tau_1^{g_1})^{\pm 1} \cdots (\tau_k^{g_k})^{\pm 1}$. Then

$$D(n) := \max\{\|h\| \mid h \in N, h \in B(n)\},$$

where $B(n)$ is the ball of radius n with center at 1, i.e. $h \in B(n)$ means that h is a word of length no more than n.

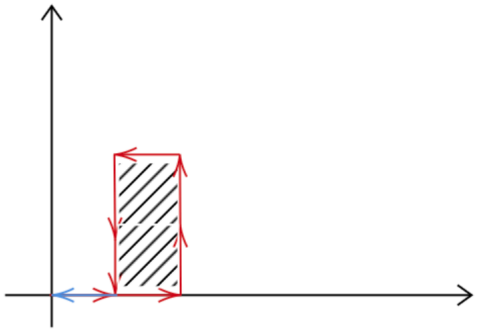
Exercise 6.1

Prove that the Dehn functions for different finite systems of a same object are asymptotically equivalent.

In fact, noticing that a step of reduction $\tau \to 1$ is the same as a step toward writing h as the form $h = (\tau_1^{g_1})^{\pm 1} \cdots (\tau_k^{g_k})^{\pm 1}$, one sees this immediately.

Example 6.2

Let us consider the group $\langle x, y \mid x^{-1}y^{-1}xy = 1\rangle$. Like what we did in example 5.1, for each word, we start from the origin on the two-dimensional plane and look from left to the right. For each x, we go right by a unit length, for x^{-1} we go left, y we go up and for y^{-1} we go down. Thus any $x^{-1}y^{-1}xy$ would give a unit square. For example, the word $x^2y^2x^{-1}y^{-2}x^{-1}$ corresponds to the following graph:



So let us set $\text{Area}(x^2y^2x^{-1}y^{-2}x^{-1}) := 2$. For this particular example, each replacement

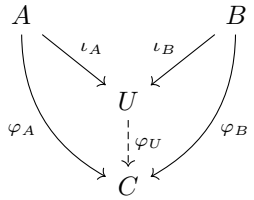
$yx^{-1}y^{-1} \to x^{-1}$ reduces the area by 1, hence it takes two steps to reduce the word $x^2y^2x^{-1}y^{-2}x^{-1}$ to its normal form. It follows that the Dehn function is asymptotically no less than $O(n^2)$, because it will take $n^2/16$ steps (cancellation of inverse elements is not counted) to reduce the square given by $x^{n/4}y^{n/4}x^{-n/4}y^{-n/4}$. Conversely, no element of length n can encircle a larger area than this square. Therefore the Dehn function is asymptotically equivalent to $O(n^2)$.
More precisely, if one wishes to stick to the definition, then the procedure is translated as the following:

$$\begin{aligned} x^2y^2x^{-1}y^{-2}x^{-1} =& x^2y^2(x^{-1}y^{-1}xy)y^{-1}x^{-1}y^{-1}x^{-1} \\ =& x^2y^2(x^{-1}y^{-1}xy)y^{-1}(x^{-1}y^{-1}xy)y^{-1}x^{-1}x^{-1} \\ =& (x^{-1}y^{-1}xy)^{y^{-2}x^{-2}}(x^{-1}y^{-1}xy)^{y^{-1}x^{-2}}. \end{aligned}$$

§7 Free Products

Free Products of Algebras

Let F be a field. Let A and B be two F-algebras (with identity, and not necessarily commutative). The *free product* of A and B is the coproduct of A and B in the category of F-algebras, i.e. it is an F-algebra U along with two homomorphisms $\iota_A\colon A \to U$ and $\iota_B\colon B \to U$ that satisfies the universal property that for any other F-algebra C along with two homomorphisms $\varphi_A\colon A \to C$ and $\varphi_B\colon B \to C$, there exists a unique homomorphism $\varphi_U\colon U \to C$ such that $\varphi_U \circ \iota_A = \varphi_A$ and $\varphi_U \circ \iota_B = \varphi_B$. See the diagram below.



Assume for now that the free product always exists. Similar as we did for the universal enveloping of Lie algebras, let U denote the free product of A and B with $\iota_A\colon A \to U$ and $\iota_B\colon B \to U$, then U is generated by the images of A and B, i.e. $U = \langle \iota_A(A), \iota_B(B) \rangle$.

The uniqueness of free product is, again, permitted by the universal property. We now prove that it always exists:

Proposition 7.1

For any two F-algebras A and B, their free product exists.

Proof. Write $A = \langle X \mid R_A(X) = 0 \rangle$ and $B = \langle Y \mid R_B(Y) = 0 \rangle$. Define

$$U = \langle X \sqcup Y \mid R_A(X) = 0, R_B(Y) = 0 \rangle,$$

then U is the free product of A and B with the obvious ι_A and ι_B, as one can check. □

We can equip U with a more concrete system. Let $\{1, a_i \mid i \in I\}$ be a basis of A and $X = \{x_i\}_{i\in I}$. We have for any $i, j \in I$,

$$a_ia_j = \gamma_{ij}^0 1 + \sum_k \gamma_{ij}^k a_k,$$

for some $\gamma_{ij}^k \in F$. Then the set $R_A(X) = \{x_ix_j - \gamma_{ij}^0 1 - \sum_k \gamma_{ij}^k x_k \mid i, j \in I\}$ is closed under composition: we have

$$\begin{aligned} 0 &= \left(a_ia_j - \gamma_{ij}^0 - \sum_k \gamma_{ij}^k a_k\right) a_l - a_i \left(a_ja_l - \gamma_{jl}^0 - \sum_k \gamma_{jl}^k a_k\right) \\ &= \cdots (\text{progress of reduction}) = (\text{linear combination of } a_k\text{'s }) \end{aligned}$$

The basis condition then forces the coefficients to be all zeros. Replace the a's by x's and we see that $R_A(X)$ is closed under composition. Similarly $R_B(Y)$ is defined. Since no relation from $R_A(X)$ admit any composition with relations in $R_B(Y)$, $R_A(X)$ and $R_B(Y)$, put together, is still closed under composition.

Therefore the map $\iota_A\colon A \to U$, sending nontrivial linear compositions of $\{1, a_i\}$ to nontrivial linear compositions of $\{1, x_i\}$ which are irreducible in $\langle X \sqcup Y \mid R_A(X), R_B(Y)\rangle$, is injective. The same is true for ι_B. We thus obtain the following lemma.

Lemma 7.2

ι_A and ι_B are embeddings.

Notation 7.1

Let $A * B$ denote the free product of A and B.

With the above system, we have shown that the irreducible words of $A * B$ form a basis, which are exactly

$$\{1, c_1 \cdots c_k \mid (c_l \in \{a_i\} \text{ and } c_{l+1} \in \{b_j\}) \text{ or } (c_l \in \{b_j\} \text{ and } c_{l+1} \in \{a_i\}) \text{ for each } l = 1, \cdots, k-1\},$$

where $\{1, a_i\}$ and $\{1, b_j\}$ are the chosen basis of A and B respectively.

The free product for an arbitrary (set-valued) collection of F-algebras is defined similarly; it is just the coproduct in the categorical viewpoint. All constructions and arguments above pass immediately, giving exactly the same results. Also, we have the law of associativity, i.e. $(A * B) * C \cong A * (B * C)$; indeed, this law would follow immediately from a diagram chasing.

Free Products of Groups

Similarly, the free product of groups are defined as the coproduct in the category of groups. All arguments in the previous section pass to groups easily (including the presentation for the free product) except for the concrete system given after proposition 7.1. But we can still find a such system by bringing the question back to algebras:

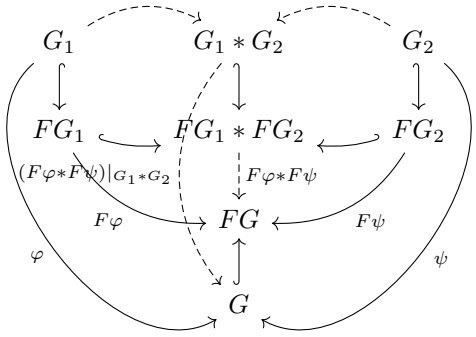
Let F be a field. For any groups G_1, G_2 and G, group homomrphisms $\varphi\colon G_1 \to G$ and $\psi\colon G_2 \to G$, we have homomorphisms of F-algebras

$$F\varphi\colon FG_1 \to FG, \quad F\psi\colon FG_2 \to FG.$$

The free product of F-algebras gives us a unique homomorphism $F\varphi * F\psi\colon FG_1 * FG_2 \to FG$ such that the diagram commutes. Since G_1 and G_2 are basis of FG_1 and FG_2 respectively, our previous result tells that the set

$$\{1, c_1 \cdots c_k \mid (c_l \in G_1 \setminus \{1\},\ c_{l+1} \in G_2 \setminus \{1\}) \text{ or } (c_l \in G_2 \setminus \{1\},\ c_{l+1} \in G_1 \setminus \{1\}),\ l = 1, \cdots, k-1\},$$

forms a basis of $FG_1 * FG_2$. Noticing that the set along with the multiplication is a group, the claim is that this group (with the obvious embeddings of G_1 and G_2) is exactly the free product of G_1 and G_2; let us denote it by $G_1 * G_2$. Indeed, the uniqueness of the morphism that makes the coproduct diagram commutes is permitted by the above description of $G_1 * G_2$. With natural inclusions $G \subset FG$, the existence follows from the fact that $(Ff)|_G = f$ for any group homomorphism f whose domain is G: restrict $F\varphi * F\psi$ to $G_1 * G_2$ and we see by the (apple-looking) commutative diagram below that the image of the restricted map should live in G.



Again, the free product for an arbitrary (set-valued) collection of groups is defined similarly, and all results pass over.

Example 7.1

We have immediately that $F(m) = \langle a_1 \rangle * \cdots * \langle a_m \rangle$, where each $\langle a_i \rangle$ is the infinite cyclic group generated by a_i.

Lecture 11

Ping-Pong Lemma

From what we have got, we can easily see that if $G = \langle G_1, G_2 \rangle$ and every element of G can be written uniquely as an interchanging product of nonidentical elements[1] from G_1 and G_2, then $G \cong G_1 * G_2$ canonically. This observation leads to the following lemma:

Lemma 7.3. Ping-Pong

Consider a group G acting on a set X. Let G_1 and G_2 be two different subgroups with $|G_1| \geq 3$ and $|G_2| \geq 2$. Let X_1 and X_2 be two disjoint subsets of X such that

$$(G_1 \setminus \{1\})X_1 \subset X_2, \quad (G_2 \setminus \{1\})X_2 \subset X_1,$$

then $\langle G_1, G_2 \rangle \cong G_1 * G_2$.

Proof. It suffices to show that any interchanging product of nonidentical elements in G_1 and G_2 is not equal to 1. Let a denote elements in G_1 and b denotes elements in G_2, then the interchanging products can be divided into four cases:

Case 1. $a_1b_1a_2b_2 \cdots a_{n-1}b_{n-1}a_n$. Since $a_1b_1a_2b_2 \cdots a_{n-1}b_{n-1}a_nX_1 \subset X_2$, we are done.

Case 2. $b_1a_1b_2a_2 \cdots b_{n-1}a_{n-1}b_n$. Since $b_1a_1b_2a_2 \cdots b_{n-1}a_{n-1}b_nX_2 \subset X_1$, we are done.

Case 3. $a_1b_1a_2b_2 \cdots a_nb_n$. Since $|G_1| \geq 3$, there exists $a \in G_1$ such that $a \neq 1$ and $a \neq a_1$. If $a_1b_1a_2b_2 \cdots a_nb_n = 1$, then its conjugation by a is also equal to 1, i.e. $a^{-1}a_1b_1a_2b_2 \cdots a_nb_na = 1$. However, this cannot be true because the conjugation is in the form of Case 1.

Case 4. $b_1a_1 \cdots b_na_n$. Conjugate by an element $a \in G_1 \setminus \{1, a_n\}$ and we are back in Case 1. □

As an important application of Ping-Pong Lemma, let us consider $\mathrm{SL}(n, \mathbb{Z}) = \{A \in M_n(\mathbb{Z}) \mid \det(A) = 1\}$. It is in fact a group, because for any invertible matrix $A = (a_{ij})$, we have the formula $A^{-1} = \frac{1}{\det(A)}((-1)^{ij} \det(A_{ij}))^T$, which is in $\mathrm{SL}(n, \mathbb{Z})$ provided that $a_{ij} \in \mathbb{Z}$ and $\det(A) = 1$.

Theorem 7.4

We have an embedding $F(2) \hookrightarrow \mathrm{SL}(2, \mathbb{Z})$.

Such embedding is not canonical, though.

Proof. Consider the action of $SL(2, \mathbb{Z})$ on $\mathbb{C}^2$ by matrix multiplication. The subgroups $G_1 \coloneqq \left\langle \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \right\rangle = \left\{ \begin{pmatrix} 1 & 2n \\ 0 & 1 \end{pmatrix} \mid n \in \mathbb{Z} \right\}$ and $G_2 \coloneqq \left\langle \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \right\rangle = \left\{ \begin{pmatrix} 1 & 0 \\ 2n & 1 \end{pmatrix} \mid n \in \mathbb{Z} \right\}$ are both cyclic. By example 7.1, we have $F(2) = G_1 * G_2$, hence it suffices to show that $G_1 * G_2 \cong \langle G_1, G_2 \rangle \subset SL(n, \mathbb{Z})$, which will be done using the Ping-Pong Lemma.

Indeed, consider the subsets $X_1 = \{(z_1, z_2)^T \in \mathbb{C} \mid |z_2| > |z_1|\}$ and $X_2 = \{(z_1, z_2)^T \in \mathbb{C} \mid |z_1| > |z_2|\}$. Then $\begin{pmatrix} 1 & 2n \\ 0 & 1 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} z_1 + 2nz_2 \\ z_2 \end{pmatrix}$. Since

$$|z_1 + 2nz_2| \geq 2|n||z_2| - |z_1| > |z_2|,$$

provided that $|z_2| > |z_1|$ and $n \neq 0$, we see that $(G_1 \setminus \{1\})X_1 \subset X_2$. Similarly $(G_2 \setminus \{1\})X_2 \subset X_1$ is seen. □

[1] i.e., elements that are not equal to 1.

A related theorem that is more general and much more difficult to prove is posted below without proof:

Theorem 7.5. J.Tits Alternative

Let H be a finitely generated subgroup of $GL(n, F)$. Then either $F(2)$ embeds into H or H contains a normal subgroup with $|H : N| < \infty$ and N is solvable.

Definition 7.1

A group G is *residually finite* if there exists a family of homomorphisms $\varphi_i \colon G \to G_i$ with $|G_i| < \infty$ and $\bigcap_i \ker \varphi_i = (1)$.

It is obvious that every subgroup of a residually finite group is again residually finite.

$F(2)$ is residually finite, as a result that $SL(n, \mathbb{Z})$ is residually finite: consider the homomorphisms $SL(n, \mathbb{Z}) \to SL(n, \mathbb{Z}/m\mathbb{Z})$ and we are done.

Recall that, let p be a prime number, G is a finite p-group if $|G| = p^s$. A group G is *residually-p* if there exists a family of homomorphisms $\varphi_i \colon G \to G_i$ where each G_i is a finite p-group and $\bigcap_i \ker \varphi_i = (1)$.

Again, every subgroup of a residually-p group is residually-p.

$F(2)$ is residually-p for any prime p: we have $F(2) = \left\langle \begin{pmatrix} 1 & p \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ p & 1 \end{pmatrix} \right\rangle \subset \mathrm{SL}(2, \mathbb{Z}, p)$, where $\mathrm{SL}(n, \mathbb{Z}, p)$ is defined by

$$\mathrm{SL}(n, \mathbb{Z}, m) \coloneqq \{A \in \mathrm{SL}(n, \mathbb{Z}) \mid A = I \bmod m\}.$$

The following exercise thus implies what we claimed:

Exercise 7.1

Show that $\mathrm{SL}(2, \mathbb{Z}, p)/\mathrm{SL}(2, \mathbb{Z}, p^s)$ is a finite p-group.

Solution. By Corollary 5.3 in [Hun80], it suffices to show that every element in $SL(2, \mathbb{Z}, p)/SL(2, \mathbb{Z}, p^s)$ has order a power of p. Indeed, every element in $SL(2, \mathbb{Z}, p)$ is of the form $I + p \cdot A$ for some matrix $A \in SL(n, \mathbb{Z})$, and the binomial theorem implies that

$$(I + p \cdot A)^{p^s} = I + p^s \cdot (\text{something}).$$

Remark 7.1

With a similar argument (along with the Lagrange's theorem), one sees that any finite residually-p group is in fact a p-group.

§8 Wreath Products

Recall that the Cartesian product $\prod_{i \in I} G_i$ of an arbitrary family of groups G_i index by I can be thought as a subset of functions from I to $\bigcup_{i \in I} G_i$.

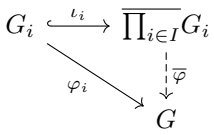
Example 8.1

The universal property of the Cartesian product implies that residually-p groups are precisely groups that are embeddable into a Cartesian product of a family of finite p-groups.

The direct product $\overline{\prod_{i \in I}} G_i$ is the subgroup of $\prod_{i \in I} G_i$ where for each element $(g_i)_{i \in I}$ there are only finitely many components g_i that are not equal to 1. Every G_i embeds naturally into $\overline{\prod_{i \in I}} G_i$, and G_i and G_j commute for any $i \neq j$ seen as subgroups of $\overline{\prod_{i \in I}} G_i$.

Except the universal property as a final object that the direct product inherits as a subobject of Cartesian product, the finiteness makes the direct product a quotient of free product, which means that it enjoys a universal property as an initial object. To be explicit, the direct product is free product quotient the relation such that G_i and G_j are made commutative whenever $i \neq j$, hence the universal property is that, for any family $\{\varphi_i\}_{i \in I}$ of homomorphisms $\varphi_i \colon G_i \to G$ such that $\varphi_i(G_i)$ and $\varphi_j(G_j)$

commute in G whenever $i \neq j$, there exists a unique homomorphism $\overline{\varphi}\colon \overline{\prod_{i\in I} G_i} \to G$ such that $\varphi_i = \overline{\varphi} \circ \iota_i$ for each i where $\iota_i\colon G_i \hookrightarrow \overline{\prod_{i\in I} G_i}$ is the natural embedding. See the diagram below.



Using the presentation of free product, we thus obtain a presentation for the direct product: Suppose that each G_i is presented by $G_i = \langle X_i \mid R_i(X_i) = 1\rangle$, then

$$\overline{\prod_{i\in I}} G_i = \langle \sqcup_{i\in I} X_i \mid R_i(X_i) = 1,\ x_i x_j = x_j x_i,\ i \neq j\rangle.$$

Note that there is no similar thing that holds for the Cartesian product, since the Cartesian product is not generated by the groups G_i's.

Remark 8.1

Similarly one can define the Cartesian product and direct product for algebras. However, note that the direct product of infinitely many algebras does not contain a multiplicative identity.

Lecture 12

Recall that so far we can conclude the normal forms in $\overline{\prod_{i\in I}} G_i$, which is explicitly, for any $x \in \overline{\prod_{i\in I}} G_i$, we have unique expression (up to permutation) that $x = g_{i_1} \cdots g_{i_k}$ for distinct i's with $g_{i_j} \in G_{i_j}$ for each $j = 1, \cdots, k$. Let A and B be two groups and B acts on A via a left action $B \to \mathrm{Aut}(A)$, which induces a right action of the opposite group B^{op} on A. The *semidirect product of B acting on A*, denoted as $A \rtimes B$ or $B \ltimes A$, is defined as the group whose underlying set is the set of pairs

$$A \rtimes B \coloneqq \{b \cdot a \mid b \in B,\ a \in A\} = B \times A,$$

and the multiplication is given by, for any $b_1 a_1$ and $b_2 a_2$, that

$$(b_1 \cdot a_1)(b_2 \cdot a_2) \sim b_1^{\mathrm{op}} b_2^{\mathrm{op}} \left((b_2^{\mathrm{op}})^{-1} a_1 b_2^{\mathrm{op}}\right) a_2 \sim b_1 b_2 \cdot (a_1^{b_2} a_2),$$

where $a_1^{b_2}$ denotes the element in A obtained by the action of b_2 on a_1. The opposite is taken because the conjugation by b_2 in the form $b_2^{-1} a_1 b_2$ is a right action, instead of left. With the opposite taken, the notation is compatible with conjugation in the sense that the following are all equal to $(b_1^{\mathrm{op}} b_2^{\mathrm{op}})^{-1} a b_1^{\mathrm{op}} b_2^{\mathrm{op}}$:

$$\left(a^{b_1}\right)^{b_2} \overset{\text{left action}}{=\!=\!=} a^{b_2 b_1} = a^{(b_2 b_1)^{\mathrm{op}}} = a^{b_1^{\mathrm{op}} b_2^{\mathrm{op}}} \overset{\text{right action}}{=\!=\!=} \left(a^{b_1^{\mathrm{op}}}\right)^{b_2^{\mathrm{op}}}.$$

In other words, we have an isomorphism

$$A \rtimes B \cong \langle A, B^{\mathrm{op}} \mid (b^{\mathrm{op}})^{-1} a b^{\mathrm{op}} = a^b,\ a \in A,\ b \in B \rangle,$$

given by $b \cdot a \mapsto b^{\mathrm{op}} a$. Note that A is normal seen as a subgroup of $A \rtimes B$.

Now let us consider the set of all functions from B to A, denoted by $\mathrm{Fun}(B, A)$. Endowed with the point-wise multiplication, it is a group isomorphic to the Cartesian product $A^B = \prod_B A$. Consider the (left) action of B on $\mathrm{Fun}(B, A)$ defined by, for any element $b \in B$ and $f \in \mathrm{Fun}(B, A)$, the action of b brings f to the function $f^b \colon B \to A$ given by $f^b(b') = f(bb')$ for any $b' \in B$.

With this action, the *wreath product of A by B*, denoted as $A \wr B$, is then defined by

$$A \wr B \coloneqq \mathrm{Fun}(B, A) \rtimes B.$$

Note that there is a natural embedding $A \hookrightarrow \mathrm{Fun}(B, A) \hookrightarrow A \wr B$ by sending elements $a \in A$ to the function $\dot{a} \colon B \to A$ defined by $\dot{a}(1_B) = a$ and $\dot{a}(b) = 1_A$ whenever $b \neq 1_B$. Let us denote the image of this embedding by $\dot{A}$.

For this lecture, we are more concerned about the *restricted wreath product*, which is defined by

$$A \bar{\wr} B \coloneqq \overline{\prod_B} A \rtimes B = \left\{ b \cdot f \mid b \in B,\ f \in \overline{\prod_B} A \subset \mathrm{Fun}(B, A) \right\}.$$

The natural embedding above also restricts to this case.

A basic but important observation is that elements in $\dot{A}^b$ commutes with $\dot{A}^{b'}$ whenever $b \neq b'$.

Remark 8.2

Note that since in $A \wr B$ the conjugation by B is not trivial as long as both A and B are not, the wreath product (either restricted or not) of two nontrivial groups is never abelian.

For the restricted wreath product, we have the following generating theorem:

Theorem 8.1

Suppose that $B = \langle b_1, \cdots, b_m \rangle$ and $A = \langle a_1, \cdots, a_n \rangle$, then $A \bar{\wr} B$ is generated by $b_1, \cdots, b_m, \dot{a}_1, \cdots, \dot{a}_n$.

Proof. Since every element in $A\bar{\wr}B$ is of the form bf, it suffices to show that $b_1, \cdots, b_m, \dot{a}_1, \cdots, \dot{a}_n$ generate $\overline{\prod_B} A$. Indeed, we have

$$f = \prod_{b \in B} \dot{f(b)}^{b^{-1}},$$

where only finitely many $\dot{f(b)}$ is nontrivial. Since $\dot{f(b)} \in \dot{A} = \langle \dot{a}_1, \cdots, \dot{a}_n \rangle$, we are done. □

Therefore, the restricted wreath product of any two finitely generated groups is finitely generated. However, it does not preserve the property of being finitely presented:

Theorem 8.2

The group of restricted wreath product $\mathbb{Z}\bar{\wr}\mathbb{Z} = \langle a \rangle \bar{\wr} \langle b \rangle$ is not finitely presented.

Before the proof of the theorem, let us firstly look at the following lemma:

Lemma 8.3

Let $G = \langle X \mid R = 1 \rangle$ be a presentation of group G where $|X| < \infty$. If G is finitely presented, then there exists a finite subset R_0 of R such that $G = \langle X \mid R_0 = 1 \rangle$.

Proof of Lemma 8.3. By proposition 2.3, we have shown that if G is finitely presented, then there exists a finite set of relations S such that

$$G = \langle X \mid S = 1 \rangle.$$

For each element $s \in S$, since R and S generate a same normal subgroup of the free group $F(X)$, there exists $r_1, \cdots, r_t \in R$ and words $g_1, \cdots, g_t$ such that

$$s = r_1^{g_1} \cdots r_t^{g_t}.$$

Since $|S| < \infty$, collect the $r_1, \cdots, r_t$ for each $s \in S$ and we obtain the desired R_0. □

Proof of Theorem 8.2. By theorem 8.1 we know that $\mathbb{Z}\bar{\wr}\mathbb{Z}$ is generated by b and $\dot{a}$. The fact that $\dot{A}^b$ and $\dot{A}^{b'}$ commutes whenever $b \neq b'$ thus gives an epimorphism

$$\langle x, y \mid [x^{y^i}, x^{y^j}] = 1,\ i, j \in \mathbb{Z} \rangle \to \mathbb{Z}\bar{\wr}\mathbb{Z},\ x \mapsto \dot{a},\ y \mapsto b,$$

where $x^{y^i} = y^{-i} x y^i$ and $[x^{y^i}, x^{y^j}] = (x^{y^j} x^{y^i})^{-1} x^{y^i} x^{y^j}$. This epimorphism is in fact an isomorphism: any word in the presentation can be written uniquely (up to permutation) as

$$y^k (x^{y^{i_1}})^{m_1} \cdots (x^{y^{i_j}})^{m_j},$$

where $i_1, \cdots, i_j$ are distinct, which is sent to $b^k (\dot{a}^{b^{i_1}})^{m_1} \cdots (\dot{a}^{b^{i_j}})^{m_j}$. Note that $b^k (\dot{a}^{b^{i_1}})^{m_1} \cdots (\dot{a}^{b^{i_j}})^{m_j} = 1$ only if $k = 0$, since $\mathbb{Z}\bar{\wr}\mathbb{Z} = \bigsqcup_{k \in \mathbb{Z}} \left(b^k \cdot \overline{\prod_B} \langle \dot{a} \rangle \right)$ and the only coset containing the identity is $\overline{\prod_B} \langle \dot{a} \rangle$. However, for $(\dot{a}^{b^{i_1}})^{m_1} \cdots (\dot{a}^{b^{i_j}})^{m_j} = 1$, the only chance is that $m_1 = \cdots = m_j = 0$, as one can see by evaluating it at elements in $\langle b \rangle$. Therefore we obtain a presentation

$$\mathbb{Z}\bar{\wr}\mathbb{Z} = \langle x, y \mid [x^{y^i}, x^{y^j}] = 1,\ i, j \in \mathbb{Z} \rangle.$$

Note that since we have

$$[x^{y^i}, x^{y^j}] = [x^{y^{i-j}}, x]^{y^j},$$

and

$$[a, b] = 1 \ \Leftrightarrow [b, a] = 1,$$

there is in fact

$$\mathbb{Z}\bar{\wr}\mathbb{Z} = \langle x, y \mid [x^{y^i}, x] = 1,\ i \in \mathbb{N}^* \rangle.$$

By lemma 8.3, it now suffices to show that there does not exist a finite subset S of $\{[x^{y^i}, x] \mid i \in \mathbb{N}^*\}$ such that $\mathbb{Z}\bar{\wr}\mathbb{Z} = \langle x, y \mid S = 1 \rangle$. The existence of such S would give that

$$\mathbb{Z}\bar{\wr}\mathbb{Z} = \langle x, y \mid [x^{y^i}, x] = 1,\ i = 1, \cdots, n \rangle,$$

for some n that is large enough, hence it suffices to show that for any $n \in \mathbb{N}^*$, the relations $[x^y, x], \cdots, [x^{y^{n-1}}, x]$ cannot give $[x^{y^n}, x]$, so that no finite subset of $\{[x^{y^i}, x] \mid i \in \mathbb{N}^*\}$ would generate it. For this purpose we need only construct a specific group where we have $[x^{y^i}, x] = 1$ for $i = 1, \cdots, n-1$, while $[x^{y^n}, x] \neq 1$.

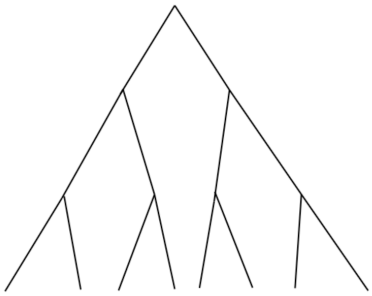
Let G be any nonabelian group, which means that there exists $u, v \in G$ such that $uv \neq vu$. Consider $G \wr \mathbb{Z} = G \wr \langle b \rangle$ and the function $f \colon \langle b \rangle \to G$ given by $f(1) = u$, $f(b^n) = v$, and $f(b^k) = 1_G$ for any $k \neq 0, n$. Now that in $G \wr \mathbb{Z}$, we have $[f^{b^i}, f] = 1$ for $i = 1, \cdots, n-1$ while $[f^{b^n}, f] \neq 1$ because $f^{b^n}(1) = v$. □

Remark 8.3

With a similar argument, one sees that the restricted wreath product of any infinite group B acting on nontrivial A is not finitely presented. In fact, since $|A^B| \geq 2^{\mathbb{N}} = |\mathbb{R}|$, we would have that $A \wr B$ is uncountable.

Exercise 8.1

Show that the group of isometries on the finite rooted tree



is isomorphic to the iterated wreath product $C_2 \wr (C_2 \wr C_2)$, where $C_2 = \mathbb{Z}/2\mathbb{Z}$.

Lecture 13

Let us end this section with the following theorem:

Theorem 8.4. Krasner–Kaloujnine

Let $A \lhd G$ and $B = G/A$. Then there exists a natural embedding $G \hookrightarrow A \wr B$.

Proof. Note that B is a set of cosets. Let us denote the image of an element $g \in G$ in B by $\bar{g}$. Note that adding bar is a homomorphism since it is identical to the quotient map. Let $s\colon B \to G\colon b \mapsto b^s$ be any function of choosing representatives. Since elements in $A \wr B$ are of the form $b \cdot f$, it is natural to expect that the embedding $G \hookrightarrow A \wr B$ is of the form $g \mapsto \bar{g} f_g$ for some $f_g \in \mathrm{Fun}(B, A)$. Assuming this form, then for it to be a homomorphism we would need exactly that $f_{1_G} = 1_{\mathrm{Fun}(B,A)}$ and

$$\overline{g_1 g_2} f_{g_1 g_2} = (\overline{g_1} f_{g_1})\,(\overline{g_2} f_{g_2}) = \overline{g_1 g_2} f_{g_1}^{\overline{g_2}} f_{g_2}.$$

Therefore, it suffices to define f_g in the way that $f_{1_G} = 1_{\mathrm{Fun}(B,A)}$ and $f_{g_1 g_2} = f_{g_1}^{\overline{g_2}} f_{g_2}$. Let us define

$$f_g(b) \coloneqq ((\bar{g}b)^s)^{-1}\, g b^s.$$

It is easy to see that $((\bar{g}b)^s)^{-1}\, g b^s \in A$, since $\bar{g}b = \overline{\bar{g}b^s} = \overline{gb^s}$. Hence $f_g\colon B \to A$ is well-defined. Clearly $f_{1_G} = 1_{\mathrm{Fun}(B,A)}$. For $f_{g_1 g_2} = f_{g_1}^{\overline{g_2}} f_{g_2}$, we have

$$f_{g_1}^{\overline{g_2}}(b) f_{g_2}(b) = f_{g_1}\,(\overline{g_2}b)\, f_{g_2}(b) = ((\overline{g_1 g_2}b)^s)^{-1}\, g_1(\overline{g_2}b)^s\,(g_1(\overline{g_2}b)^s)^{-1}\, g_2 b^s = ((\overline{g_1 g_2}b)^s)^{-1}\, g_1 g_2 b^s = f_{g_1 g_2}(b).$$

Therefore the homomorphism is well-defined.

To show that it is an embedding, suppose that $g \in G$ is mapped to $1_{A \wr B}$. Then $\bar{g} f_g = 1_{A \wr B}$, which forces that $\bar{g} = 1_B$, hence $g \in A$. Also, we have $f_g = 1_{\mathrm{Fun}(B,A)}$, which means that

$$((\bar{g}b)^s)^{-1} g b^s = 1,$$

for any $b \in B$. Since $\bar{g}b = b$, we have that $(b^s)^{-1} g b^s = 1$, therefore $g = 1_G$. □

§9 The Burnside Problem

Consider a finitely generated group G where every element $g \in G$ is of finite order. The condition that every $g \in G$ has a finite order is called that G is *torsion*.

The *General Burnside Problem* asks, if a group G is finitely generated and is torsion, then must G be finite?

The answer is: No. We will construct two counterexamples later, each of them are important by its own right.

The following states a less general version of the problem, which is known as The Burnside Problem:

Question. The Burnside Problem

If a group G is finitely generated and there exists $n \in \mathbb{N}^*$ such that $g^n = 1_G$ for all $g \in G$, then must G be finite?

For $n = 2$, this is trivial: since $g = g^{-1}$ for any $g \in G$, G is abelian. Since every element in G has an order no larger than 2, every element is a word without any repetition of alphabets, hence $|G| \leq 2^m$, where m is the number of generators of G.

However, the problem is highly nontrivial for any $n \geq 3$: Burnside himself proved the statement for the case $n = 3$, Sanov proved it for the case $n = 4$ and M. Hall proved it for the case $n = 6$. The problem

for $n = 5$, however, remains open till now. In 1968, Novikov-Adian proved it for the cases of any odd $n \geq 4381$, which used a simultaneous induction on more than 100 indices and has a length of more than 300 pages; we will not go through that proof.

Noticing that any finite group can be embedded into the general linear group $\mathrm{GL}(n, F)$ for any field F (consider the inclusion $G \hookrightarrow FG$, where FG is the free vector space generated by G over F), Burnside proved the following restricted statement:

Theorem 9.1. Burnside

Every finitely generated torsion subgroup of $\mathrm{GL}(n, \mathbb{C})$ *is finite.*

We will prove this theorem, admitting the following lemma which is again by Burnside. Let $V = \mathbb{C}^n$, we will consider $\mathrm{GL}(V)$ instead of $\mathrm{GL}(n, \mathbb{C})$, so that we do not specify any basis.

Lemma 9.2. Burnside

If a subset S *of* $\mathrm{GL}(V)$ *acts irreducibly on* V*, i.e. there is no non-trivial subspace of* V *that is invariant under all elements in* S*, then* $\mathbb{C}S \coloneqq \mathrm{span}_{\mathbb{C}} S = \mathrm{End}_{\mathbb{F}}(V)$.

The statement is easily verified for $S = \mathrm{GL}(V)$ since $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right) = \frac{1}{2}\left(\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right) + \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)\right)$ and so on.

Observe that, if $W \subset V$ is invariant under all elements in S, then S acts naturally on V/W, and if we choose a basis of W and then extend it to a basis of V, then elements in S would have matrix representations of the form

$$\begin{pmatrix} * & * \\ 0 & * \end{pmatrix}.$$

With this observation, we state and prove the following two lemmas, which will be useful in our proof of Burnside's Theorem:

Lemma 9.3

For any subset $S \subset \mathrm{GL}(V)$*, there exists a finite chain of subspaces*

$$\{0\} = V_0 \subsetneq V_1 \subsetneq V_2 \subsetneq \cdots \subsetneq V_k = V,$$

such that all V_i*'s are* S*-invariant and* S *acts irreducibly on each fraction* V_{i+1}/V_i.

Proof. If S acts irreducibly on V then there is nothing to prove. If not, let V_1 be a nontrivial S-invariant subspace of minimal dimension. If S acts irreducibly on V/V_1, then we are done; if not, let $V_2 \subset V$ be the preimage of a nontrivial S-invariant subspace of minimal dimension in V/V_1, and repeat this procedure. We will be done in at most $\dim V = n$ steps. □

Lemma 9.4

In proving Burnside's theorem, it suffices to consider $G < \mathrm{GL}(V)$ *that acts irreducibly on* V.

Proof. Suppose that the theorem has been proved for any G that acts irreducibly on V. Suppose now that we are given a finitely generated torsion subgroup $G \subset \mathrm{GL}(V)$ whose action on V is not necessarily irreducible. Let $\{0\} = V_0 \subsetneq V_1 \subsetneq V_2 \subsetneq \cdots \subsetneq V_k = V$ be a chain as in lemma 9.3 with $S = G$. Choose a basis of V_1 and extend it to a basis of V_2, and so on until we obtain a basis of $V_k = V$. Then the matrix of an element $g \in G$ with respect to this basis is of the form

$$M(g) = \begin{pmatrix} M(g|_{V_1}) & * & 0 & 0 & 0 \\ 0 & M(g|_{V_2/V_1}) & * & 0 & \\ 0 & 0 & M(g|_{V_3/V_2}) & * & \vdots \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & \cdots & M(g|_{V/V_{k-1}}) \end{pmatrix}.$$

Consider the map $G \to M_n(\mathbb{C})$ given by

$$g \mapsto \begin{pmatrix} M(g|_{V_1}) & 0 & 0 & 0 & 0 \\ 0 & M(g|_{V_2/V_1}) & 0 & 0 & \\ 0 & 0 & M(g|_{V_3/V_2}) & 0 & \vdots \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & \cdots & M(g|_{V/V_{k-1}}) \end{pmatrix},$$

it is clear that it is a group homomorphism. Its image is a finite set, because by our assumption we know that the choice of each $M(g|_{V_{i+1}/V_i})$ is finite. The map is in fact an embedding, because if $g \mapsto 1$, then we have

$$M(g) = \begin{pmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = I + U,$$

where $U \in M_n(\mathbb{C})$ is upper-triangular. Recall that the order of g should be finite, we have

$$I + dU + \binom{d}{2} U^2 \cdots + U^d = (I+U)^d = I,$$

for some positive integer d, hence $dU + \binom{d}{2} U^2 \cdots + U^d = (I+U)^d = 0$. If $U \neq 0$, then consider the upper diagonal in N which is the nearest to the diagonal among those whose entries are not identically zero. However, the particular upper diagonal in U^k for any $k \geq 2$ must be identically zero, giving that the upper diagonal itself must be zero, a contradiction. Therefore we must have $U = 0$, concluding the proof. □

Proof of Theorem 9.1. By lemma 9.4, we may assume that G acts irreducibly. Let us prove the theorem firstly with an additional assumption, and then remove the assumption. The additional assumption is that the orders of elements of G are bounded uniformly, i.e. there exists $d \in \mathbb{N}^*$ such that $g^d = 1$ for all $g \in G$. With this assumption, we then see that the diagonal of the Jordan form of an element $g \in G$ consists of d-th roots of the unity. Therefore, there are only finitely many available choices for the trace $\mathrm{Tr}(g)$ for elements $g \in G$, i.e. the image of the trace $\mathrm{Tr}\colon G \to \mathbb{C}\colon g \mapsto \mathrm{Tr}(g)$ is finite.[2] By lemma 9.2, since G acts irreducibly, G spans $\mathrm{End}_{\mathbb{C}}(V)$, so we can find elements $g_1, \cdots, g_{n^2} \in G$ that form a basis of $\mathrm{End}_{\mathbb{C}}(V)$. Consider the map

$$G \to \mathbb{C}^{n^2} : g \mapsto (\mathrm{Tr}(gg_1), \cdots, \mathrm{Tr}(gg_{n^2})),$$

then its image is again finite, since there are only finitely many choices for each entry. Now it suffices to show that this map is injective. Indeed, since Tr is linear, if $\mathrm{Tr}(gg_i) = \mathrm{Tr}(g'g_i)$ for all $i = 1, \cdots, n^2$, then $\mathrm{Tr}((g-g')g_i) = 0$ for all $i = 1, \cdots, n^2$. Since $g_1, \cdots, g_{n^2}$ spans $\mathrm{End}_{\mathbb{C}}(V)$, we thus see that $\mathrm{Tr}((g-g')M) = 0$ for any $M \in \mathrm{End}_{\mathbb{C}}(V)$. Since $\mathrm{Tr}((a_{ij})E_{kl}) = a_{lk}$, the only chance that this happens is $g - g' = 0$, concluding the injectivity.

The rest of the proof will be given in the next lecture. □

[2] In fact, by the proof of the preceding Lemma, the Jordan form of g must be diagonal; but this cannot tell the finiteness of G, because the Jordan basis for each element may be different. Note though that the trace is invariant under the choice of basis.

Lecture 14

Continue of the Proof of Theorem 9.1. Let us now remove the assumption. We will use some field theory.

Lemma 9.5

For any fixed $n \geq 1$, there exists a sufficiently large N such that any torsion matrix from $\mathrm{GL}(n, \mathbb{Q})$ has order no more than N. It follows that $A^{N!} = 1$ for all torsion matrix $A \in \mathrm{GL}(n, \mathbb{Q})$.

Proof of Lemma 9.5. Let $A \in \mathrm{GL}(n, \mathbb{Q})$ be torsion. Brought to $\mathbb{C}$, our preceding results tell that any torsion matrix A must be diagonalizable over $\mathbb{C}$, and the entries λ_i's in the diagonal would all be roots of the unity. Suppose that λ_i has multiplicative order d, i.e. $\lambda_i^d = 1$, then basic field theory gives that $[\mathbb{Q}(\lambda_i) : \mathbb{Q}] \coloneqq \dim_{\mathbb{Q}} \mathbb{Q}(\lambda_i) = \varphi(d)$, where $\varphi \colon \mathbb{N}^* \to \mathbb{N}^*$ is the Euler function, i.e. $\varphi(d)$ is the number of integers less than d and are coprime to d. On the other hand, the characteristic polynomial of A tells that the λ_i's are roots of a polynomial of degree n, hence $[\mathbb{Q}(\lambda_i) : \mathbb{Q}] \leq n$. Therefore, for any λ_i on the diagonal, its multiplicative order is no more than the integer N such that $\varphi(N + m) > n$ for all $m \in \mathbb{N}^*$. Such N exists, because φ is asymptotically increasing in the sense that $\varphi(n)/n^{1-\delta} \to \infty$ as $n \to \infty$ for any $\delta > 0$. □

Let X be any set, we can consider the field $\mathbb{Q}(X)$ of all rational polynomials with variables in X, i.e. the field generated by elements of the form $\frac{f(x_1, \cdots, x_m)}{g(y_1, \cdots, y_k)}$, with $x_1, \cdots, x_m, y_1, \cdots, y_k \in X$, $f, g \in \mathbb{Q}[X]$ and $g \neq 0$. Note that since $\mathbb{Q}$ is an infinite field, by an induction on the number of variables one sees that for any polynomial $f \in \mathbb{Q}[X]$ that is not identically zero, there exists $\alpha_1, \cdots, \alpha_m \in \mathbb{Q}$ such that $f(\alpha_1, \cdots, \alpha_m) \neq 0$. Moreover, given any finite set of nonzero polynomials $\{f_i\}_{i=1}^s$ over $\mathbb{Q}$, their multiplication $\prod_{i=1}^s f_i$ is nonzero, hence there exists scalar αs in $\mathbb{Q}$ such that $f_i(\alpha) \neq 0$ for all $i = 1, \cdots, s$.

With the same N as above, we have the following lemma

Lemma 9.6

For any fixed $n \geq 1$, any torsion matrix from $\mathrm{GL}(n, \mathbb{Q}(X))$ has order dividing $N!$.

Proof of Lemma 9.6. Suppose $A^{N!} \neq I_n$, then we get a system of inequalities. Substitute scalar α's in $\mathbb{Q}$ into the indeterminants such that all the inequalities along with those that the dominator is nonzero hold and we obtain a contradiction to the preceding lemma. □

Finally, let G be a finitely generated subgroup of $GL(n, \mathbb{C})$, then there exists a finitely generated subfield $L = \mathbb{Q}(\alpha)$, where α is the set of entries of generators of G. By basic field theory, there exists a transcendental field $K = \mathbb{Q}(X)$ for some set X such that there is a chain of inclusions

$$\mathbb{Q} \subset K \subset L,$$

and $[L : K] < \infty$. Write $s \coloneqq [L : K]$, then L naturally embeds into $M_s(K)$ by considering the right multiplication $L \hookrightarrow \mathrm{End}_K(L) \colon a \mapsto R_a$. With this embedding, we see that G can be seen as a subgroup of $\mathrm{GL}(ns, K)$ via the chain $G < \mathrm{GL}(n, L) < \mathrm{GL}(ns, K)$. If G is torsion, then the preceding lemma applies and we obtain the condition that the orders of elements of G are uniformly bounded. □

Remark 9.1

In general, it is true that every finitely generated torsion subgroup of $\mathrm{GL}(n, F)$ for any field F is finite. Our proof for $F = \mathbb{C}$ can be divided into two parts, one is that every subgroup where the orders of elements are uniformly bounded is finite, the other is that every finitely generated torsion subgroup satisfies the condition that the orders of elements are uniformly bounded. Our argument for the former only applies when the field has zero characteristic.

A Counterexample to General Burnside Problem

Let us move on to construct counterexamples of the General Burnside Problem. From now on we will not assume that every algebra has an identity element.

Let A be an algebra over a field F. An element in A is *nilpotent* if there exists $n \geq 1$ such that $a^n = 0$. The algebra A is *nilpotent* if there exists $N \geq 1$ such that $A^N = (0)$, which means that the product of any N elements in A is zero. A is called a nil algebra if every element in a is nilpotent.

In 1941, A. Kurosh formulated the following question, which turns out to be related to the General Burnside Problem:

Question. Kurosh Problem

Suppose A is finitely generated and nil, must there be that A is nilpotent and A is finite dimensional over F?

H. M. Wedderburn proved the following

Theorem 9.7. H. M. Wedderburn

If $\dim_F A < \infty$ *and* A *is nil, then* A *is nilpotent.*

Conversely, if A is finitely generated and nilpotent with $A^n = (0)$, then A is spanned by the products of its generators of length less than n, hence A is finite dimensional. Therefore, under the setup of Kruosh Problem, finite dimensionality and nilpotency of A are equivalent.

For an algebra without unit, we can consider the direct sum

$$\hat{A} \coloneqq A \oplus F \cdot 1 = \{(a, \alpha \cdot 1) \mid a \in A, \alpha \in F\},$$

with the obvious structure of algebra. The algebra $\hat{A}$ can be seen as A added the unit, and is called the *unital hull* of A.

Let $a \in A$ be a nilpotent, then $1+a \in \hat{A}$ is invertible, with inverse given by the finite sum $(1+a)^{-1} = 1 - a + a^2 - a^3 + a^4 - \cdots$. Suppose that $\operatorname{char} F = p > 0$, then $1 + a$ has a finite multiplicative order: suppose $a^n = 0$, choose k such that $p^k \geq n$ and then

$$(1+a)^{p^k} = 1 + \binom{p^k}{1} a + \cdots + \binom{p^k}{p^k - 1} a^{p^k - 1} + a^{p^k} = 1,$$

since each $\binom{p^k}{i}$ is divisible by p for $1 \leq i \leq p^k - 1$.

With these observations, we are now able to relate the Kurosh Problem with the General Burnside Problem:

Proposition 9.8

If there exists a counterexample to the KuroshProblem with $\operatorname{char} F = p > 0$ *, then there exists a counterexample to the General Burnside Problem.*

Proof. Let A be a counterexample to the Kurosh Problem, say $A = \langle a_1, \cdots a_m \rangle$ is a nil algebra over F with $\operatorname{char} F = p > 0$, and A is not nilpotent. Consider the multiplicative subgroup G of $\hat{A}$ generated by $1 + a_1, \cdots, 1 + a_m$. The observation above tells that G is torsion. If G is not a counterexample to the General Burnside Problem, then we have $|G| = d < \infty$. Then, any product $(1+a_{i_1}) \cdots (1+a_{i_d})$ of length d of generators $1 + a_1, \cdots, 1 + a_m$ would be equal to a shorter product of these generators, because the list in G,

$$1, 1 + a_{i_1}, (1 + a_{i_1})(1 + a_{i_2}), (1 + a_{i_1}) \cdots (1 + a_{i_d}),$$

has $d + 1$ elements, hence there must be

$$(1 + a_{i_1}) \cdots (1 + a_{i_t}) = (1 + a_{i_1}) \cdots (1 + a_{i_t})(1 + a_{i_{t+1}}) \cdots (1 + a_{i_{t+l}}),$$

for some t and l, which forces $(1+a_{i_{t+1}}) \cdots (1+a_{i_{t+l}}) = 1$. Since $(1+a_{i_{t+1}}) \cdots (1+a_{i_{t+l}})$ is a sub-product of $(1 + a_{i_1}) \cdots (1 + a_{i_d})$, replace it by 1 and we obtain a shorter product. Expand the bracket and move the terms, we then obtain

$$a_{i_1} \cdots a_{i_d} = \sum_{k<d} (\text{some coefficient in } \mathbb{Z}/p\mathbb{Z}) a_{j_1} \cdots a_{j_k}.$$

Therefore, since the choice of a_{i_k}'s are arbitrary, every product of generator a_i's of length no less than d can be expressed as a linear combination of products of generators of length less than d, hence A is spanned by products of generators of length less than d, which are finitely many. This concludes that $\dim_F A < \infty$, contradicting to Wedderburn's Theorem. □

Therefore, to construct a counterexample to the General Burnside Problem, it suffices to construct a counterexample to the Kurosh Problem. We will construct it using graded algebras.

Lecture 15

Let $A = \oplus_{i=0}^{\infty} A_i$ be graded with $A_i A_j \subset A_{i+j}$. The Hilbert series $H_A(t)$ of A is then defined as the formal power sum

$$H_A(t) \coloneqq \sum_{i=0}^{\infty} \dim_F(A_i) t^i.$$

Example 9.1

Consider the noncommutative polynomial algebra $F\langle X\rangle \coloneqq F \cdot 1 + (\sum_{i=1}^{m} Fx_i) + \left(\sum_{i,j} F \cdot x_i x_j\right) + \cdots$ with $|X| = m$. The corresponding Hilbert series is

$$H_{F\langle X\rangle}(t) = 1 + mt + m^2 t^2 + \cdots = \sum_{i=0}^{\infty} (mt)^i = \frac{1}{1-mt}.$$

Example 9.2

Consider the polynomial algebra $F[X]$ with $|X| = m$. The corresponding Hilbert series is

$$H_{F[X]}(t) = \sum_{i=0}^{\infty} \binom{m+k-1}{k} t^k.$$

Recall that an ideal I of a graded algebra A is said to be *homogeneous* if for any element in I, all homogeneous components of that element live in I. Equivalently, $I = (I \cap A_0) + (I \cap A_1) + \cdots$. Also, if I is generated by a set of homogeneous elements, then I is homogeneous.

Let I be a homogeneous ideal of graded algebra A, then the quotient algebra A/I is graded with $(A/I)_i = A_i/(I \cap A_i)$ for all $i \in \mathbb{N}$.

For two formal series $\sum_{i=0}^{\infty} a_i t^i$ and $\sum_{i=0}^{\infty} b_i t^i$, we say that $\sum_{i=0}^{\infty} a_i t^i \geq \sum_{i=0}^{\infty} b_i t^i$, if $a_i \geq b_i$ for all $i \in \mathbb{N}$.

Consider an alphabet X with $|X| = m$. Let R be a set of homogeneous relations in $F\langle X\rangle$, then we have a graded algebra $A = \langle X \mid R = 0\rangle = F\langle X\rangle/I(R)$. By replacing elements in X with their linear combinations of other elements we may assume that every element in R has degree greater than 2. By replacing R with its linear span, we may consider each homogeneous part R_i of R as a linear space and write $\dim_F R_i = r_i$. We then define the series $H_R(t)$ to be

$$H_R(t) = r_2 t^2 + r_3 t^3 + \cdots.$$

Under this setup, we have the following theorem:

Theorem 9.9. Golod-Shafarevich Inequality

We have the following inequality of series

$$(1 - mt + H_R(t)) \cdot H_A(t) \geq 1.$$

A consequence of this theorem is that, suppose that we found a number $0 < t_0 < 1$ such that

(1) $H_R(t)$ converges at t_0,

(2) $1 - mt_0 + H_R(t_0) < 0$,

then the algebra A is infinite dimensional. For this consequence, notice that if A is finite dimensional, then $H_A(t)$ is a polynomial, while we have, if $H_A(t_0)$ converges, then

$$(1 - mt_0 + H_R(t_0))H_A(t_0) \geq 1.$$

Since $H_A(t) \geq 1$, this is impossible. Hence $H_A(t_0)$ cannot converge, so that $H_A(t)$ cannot be a polynomial. Therefore A must be infinite dimensional.

This consequence will be of vital importance in our construction of the counterexample to the Kurosh Problem: all that's left is to construct a nil algebra A who has finite codimension in $F\langle X\rangle/I(R)$, along with a t_0 as above.

Proof of Theorem 9.9. Write $\dim_F A_n = a_n$, then we have by our construction of A that $a_0 = 1$ and $a_1 = m$. Also, we have that, though not canonically,

$$F\langle X\rangle_n \cong I_n \oplus F\langle X\rangle_n/I_n = I_n \oplus A_n.$$

Since $\dim_F F\langle X\rangle_n = m^n$, we have $\dim_F I_n = m^n - a_n$. Moreover, let us fix for each $n \in \mathbb{N}$ a subspace B_n in $F\langle X\rangle_n$ such that $F\langle X\rangle_n = I_n \oplus B_n$, then we have $\dim B_n = a_n$.

Note that $I_n \coloneqq I(R)_n$ is spanned by uR_iv, where u, v are words such that

$$\deg(u) + i + \deg(v) = n.$$

Notice that if v is not empty, then $uR_iv \subset I_{n-1}X$. If $v = 1$, then $\deg u = n - i$ and $uR_iv = uR_i \subset B_{n-i}R_i + I_{n-1}X$, where we are using the decomposition that $u \in B_{n-i} \oplus I_{n-i}$. As u, v varies in all words such that $\deg(u) + i + \deg(v) = n$, we obtain for any $n \geq 2$ that

$$I_n \subset I_{n-1}X + \sum_{i=2}^{n} B_{n-i}R_i,$$

since $r_0 = r_1 = 0$. Take the dimension and we obtain the inequality that

$$\begin{aligned} m^n - a_n = \dim_F I_n &\leq \dim_F \left(I_{n-1}X + \sum_{i=2}^{n} B_{n-i}R_i \right) \\ &\leq (m^{n-1} - a_{n-1}) \cdot m + \sum_{i=2}^{n} a_{n-i}r_i. \end{aligned}$$

Therefore

$$a_n - ma_{n-1} + \sum_{i=2}^{n} a_{n-i}r_i \geq 0,$$

for all $n \geq 2$. Since

$$\begin{aligned} (1 - mt + H_R(t)) \cdot H_A(t) &= \left(1 - mt + \sum_{i=2}^{\infty} r_i t^i\right)\left(1 + mt + \sum_{i=2}^{\infty} a_i t^i\right) \\ &= 1 + 0 \cdot t + \sum_{n=2}^{\infty} \left(a_n - ma_{n-1} + \sum_{i=2}^{n} r_i a_{n-i}\right) t^n, \end{aligned}$$

we are done. □

Let us now construct the promised counterexample to the General Burnside Problem. Let $|X| = m \geq 2$, then any $t_0 \in \left(\frac{1}{m}, 1\right)$ satisfies that $1 - mt_0 < 0$. Hence we may find sufficiently large $N \in \mathbb{N}^*$ such that

$$1 - mt_0 + \sum_{k \in \mathbb{N}} t_0^{N+k} < 0.$$

Let F be a countable field of characteristic $p > 0$, then $F\langle X\rangle$ is also countable. Consider the subalgebra (without the identity) generated by all elements of degree no less than 1, $F\langle X\rangle_{\geq 1}$, then we may list its elements as

$$F\langle X\rangle_{\geq 1} = \{f_1, f_2, \cdots\}.$$

Define a sequence of integers $n_1, n_2, \cdots$ as the following: let $n_1 = N$, then all homogeneous components of $f_1^{n_1}$ has degree no less than N. Let n_2 be strictly larger than the maximal degree of all (nonzero)

homogeneous components of $f_1^{n_1}$, and define recursively n_3 and so on. For any $i \geq 2$, all homogeneous components of $f_i^{n_i}$ has degree no less than n_i, which is strictly larger than the degree of any homogeneous component of $f_{i-1}^{n_{i-1}}$.

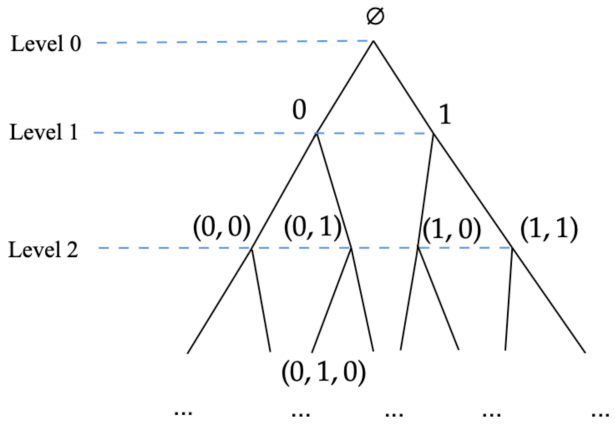
Let R be the set of all linear combinations of all the homogeneous components of $f_i^{n_i}$'s, then $\dim_F R_i \leq 1$ for all $i \in \mathbb{N}$. Consider the algebra $A \coloneqq F\langle X\rangle_{\geq 1}/I(R)$ which is of codimension 1 in $F\langle X\rangle/I(R)$. A is nil, since every element in A is of the form $f_i + I(R)$ for some i, thus is a nilpotent since $f_i^{n_i} \in I(R)$. It remains only to show that $F\langle X\rangle/I(R)$ is infinite dimensional, which follows from that

$$1 - mt_0 + H_R(t_0) \leq 1 - mt_0 + \sum_{k\in\mathbb{N}} t_0^{N+k} < 0,$$

along with the previously mentioned consequence of theorem 9.9.

Another Counterexample to General Burnside Problem

Recall that in example 6.1 we have defined the Grigorchuk group. In order to study it closer, let us introduce some notations. Firstly let us assign each vertex a level, from top to below, starting from zero. Also, we label each vertices in the following way, that the top vertex is labelled with the empty label and whenever we go down along the edge from a labelled vertex we label the target vertex with a label obtained by adding 0 or 1 to the right of the previous vertex's label respectively, according to that the edge is on the left or the right. See the illustration below.



Let T_n denote the subtree consisting of everything of level no more than n. Let L_n be the set of all vertices of level n, hence $|L_n| = 2^n$. Since the automorphisms (isometries) preserve the level, $\mathrm{Aut}(T_n)$ acts naturally on L_n. Such action is effective, hence the behaviour on L_n determines an element in $\mathrm{Aut}(T_n)$: Write $\bar{0} = 1$ and $\bar{1} = 0$, then every element $a \in \mathrm{Aut}(T_n)$ is determined by the formula

$$a(i_1, i_2, \cdots, i_n) = (\hat{i}_1, \hat{i}_2, \cdots, \hat{i}_n),$$

where $(i_1, i_2, \cdots, i_n) \in \{0, 1\}^n$ and the hats indicate that there may or may not be a bar over each i_j. For example, the automorphism that exchanges 0 and 1 while "preserves" everything else is given by

$$a(i_1, i_2, \cdots, i_n) = (\bar{i}_1, i_2, \cdots, i_n).$$

For this moment, let us generalize the wreath product a little bit. Let A and B be two groups and B acts on a set X. The wreath product of A by B acting on X is then defined by

$$A \wr_X B \coloneqq \mathrm{Fun}(X, A) \rtimes B,$$

where the action of B on $\mathrm{Fun}(X, A)$ is given by $f^b(x) \coloneqq f(bx)$ for any $f \in \mathrm{Fun}(X, A)$, $x \in X$ and $b \in B$. The rest should be all similar to the wreath product defined previously, and we have

$$|A \wr_X B| = |B| \cdot |\mathrm{Fun}(X, A)| = |B| \cdot |A|^{|X|}. \tag{9.1}$$

Noticing that we have by restriction of automorphisms a surjective map $\mathrm{Aut}(T_{k+1}) \to \mathrm{Aut}(T_k) \colon g \mapsto \bar{g}$, let us define a map (recall that $\mathrm{Aut}(T_k)$ acts on L_k)

$$\mathrm{Aut}(T_{k+1}) \to (\mathbb{Z}/2\mathbb{Z}) \wr_{L_k} \mathrm{Aut}(T_k) \colon g \mapsto \bar{g} \cdot f_g,$$

where f_g is defined by, for each $(i_1, \cdots, i_k) \in L_k$, $f_g(i_1, \cdots, i_k) = 0 \in \mathbb{Z}/2\mathbb{Z}$ if g does not reflect the two edges below $(i_1, \cdots, i_k)$, and $f_g(i_1, \cdots, i_k) = 1 \in \mathbb{Z}/2\mathbb{Z}$ if g does. Clearly this map is bijective, and the verification that it is a homomorphism is straightforward. Therefore we obtain a recursive formula that determines $\mathrm{Aut}(T_n)$:

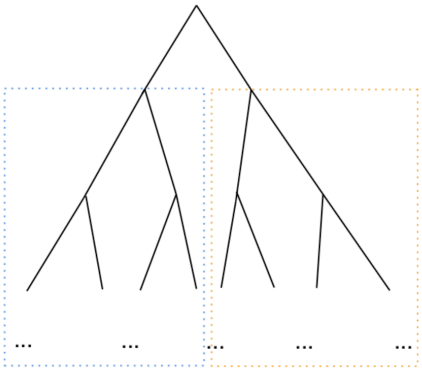
$$\mathrm{Aut}(T_{k+1}) \cong (\mathbb{Z}/2\mathbb{Z}) \wr_{L_k} \mathrm{Aut}(T_k) \colon g \mapsto \bar{g} \cdot f_g.$$

Moreover, with eq. (9.1), we can compute out $|\mathrm{Aut}(T_k)|$ explicitly for each k: we have $|\mathrm{Aut}(T_1)| = 2^1$, hence $|\mathrm{Aut}(T_2)| = 2^1 \cdot 2^2$. Inductively, suppose that $|\mathrm{Aut}(T_{k-1})| = 2^{1+2^2+\cdots+2^{k-2}}$, then

$$|\mathrm{Aut}(T_k)| = 2^{1+2^2+\cdots+2^{k-2}} \cdot 2^{2^{k-1}} = 2^{1+2^2+\cdots+2^{k-1}} = 2^{2^k-1}.$$

Therefore $\mathrm{Aut}(T_k)$ is a 2-group. Noticing that the restriction homomorphism $\mathrm{Aut}(T) \to \mathrm{Aut}(T_k)$ is to forget everything of level greater than k, it follows that $\mathrm{Aut}(T)$ is a residually 2-group, since $\bigcap_{k\in\mathbb{N}} \mathrm{Stab}(k) = 1$, where $\mathrm{Stab}(k)$ denotes the stabilizer subgroup that fixes L_k.

Finally, let us note that we have $\mathrm{Stab}(1) \cong \mathrm{Aut}(T) \times \mathrm{Aut}(T)$ via the map that sends an automorphism $\varphi \in \mathrm{Stab}(1)$ to the pair of restrictions $(\varphi|_{T'}, \varphi|_{T''})$, where T' and T'' are the subtrees rooted at 0 and 1, encircled by the blue and orange squares respectively in the following picture:



Lecture 16

Let us build our counterexample. Recall the isomorphism that $\mathrm{Stab}(1) \cong \mathrm{Aut}(T) \times \mathrm{Aut}(T)$, so that we may write every element $\varphi \in \mathrm{Stab}(1)$ as $\varphi = (\varphi', \varphi'')$. Also, note that we have for any $\varphi, \psi \in \mathrm{Stab}(1)$,

$$\varphi\psi = (\varphi'\psi', \varphi''\psi'').$$

Let $a \in \mathrm{Aut}(T)$ be the automorphism that reflects 0 and 1 while "preserving" everything else, namely we have $a(i_1, \cdots, i_k) = (\bar{i}_1, \cdots i_k)$ for any vertex $(i_1, \cdots i_k)$. Define $b, c, d \in \mathrm{Stab}(1)$ by

$$b = (a, c), \quad c = (a, d), \quad d = (1, b),$$

where the pairs on the right hand sides are understood as an element in $\mathrm{Stab}(1)$ via the isomorphism $\mathrm{Stab}(1) \cong \mathrm{Aut}(T) \times \mathrm{Aut}(T)$. This definition can be understood via an induction on the level of vertices: the action of b, c, d on vertices of level no more than 1 is trivial. Suppose that the actions of b, c, d on vertices of level $k-1$ are defined, then for any vertex $(i_1, \cdots, i_k)$ we have, take b for an example,

$$b(i_1, \cdots, i_k) = \begin{cases} (0, a(i_2, \cdots, i_k)) & i_1 = 0 \\ (1, c(i_2, \cdots, i_k)) & i_1 = 1 \end{cases}$$

Since $(i_2, \cdots, i_k)$ is of level $k-1$, this defines the action of b on vertices of level k; it is easy to verify that actions defined in this way give a well-defined automorphism. Similarly c and d are defined.

Our desired counterexample then, would be $G = \langle a, b, c, d \rangle$, the subgroup of $\mathrm{Aut}(T)$ generated by a, b, c, d. Of course G is finitely generated. What remains to check is that

- G is torsion;
- G is infinite.

Let us look into G step by step, lemma by lemma.

Lemma 9.10

$a^2 = b^2 = c^2 = d^2 = 1$.

Proof. It is obvious that $a^2 = 1$. For b^2, c^2 and d^2, it suffices to show that their actions on all vertices are trivial, and again we do this by an induction on the level of verices. Since $b^2, c^2, d^2 \in \mathrm{Stab}(1)$, clearly they fix all vertices of level no more than 1. Suppose that b^2, c^2, d^2 fix all vertices of level $k-1$, then we have, take b for an example,

$$b^2(i_1, \cdots, i_k) = \begin{cases} (0, a^2(i_2, \cdots, i_k)) & i_1 = 0 \\ (1, c^2(i_2, \cdots, i_k)) & i_1 = 1 \end{cases}$$

Hence the induction assumption tells that b^2 acts trivially on all vertices of level k. Similarly we are done for c^2 and d^2. □

Lemma 9.11

$bc = cd = d$, $bd = db = c$, $dc = cd = b$.

A consequence of this lemma along with the preceding one is that $\langle b, c, d \rangle \cong (\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/2\mathbb{Z})$.

Proof. Again we proceed by induction on level of vertices. Suppose that all three equality holds up to their actions on vertices of level $k-1$, i.e. we have $bc(i_1, \cdots, i_{k-1}) = d(i_1, \cdots, i_{k-1})$ and etc. For any vertice $(i_1, \cdots, i_k)$ of level k, we have, take bc for an example,

$$bc(i_1, \cdots, i_k) = \begin{cases} (0, a^2(i_2, \cdots, i_k)) = (i_1, \cdots, i_k) & i_1 = 0 \\ (1, cd(i_2, \cdots, i_k)) & i_1 = 1 \end{cases}$$

Since $cd(i_2, \cdots, i_k) = b(i_2, \cdots, i_k)$ by our induction assumption, we conclude that $bc(i_1, \cdots, i_k) = d(i_1, \cdots, i_k)$. Similarly the rest equalities follow. Since the equalities are trivial for vertices of level no more than 1, the induction applies and we are done. □

Consider the intersection of Stab(1) with G. We will write $\mathrm{Stab}_G(1) \coloneqq \mathrm{Stab}(1) \cap G$. Note that $a \notin \mathrm{Stab}_G(1)$ and $b, c, d \in \mathrm{Stab}_G(1)$, so that $\mathrm{Stab}_G(1)$ is a proper subgroup of G.

Lemma 9.12

$\mathrm{Stab}_G(1) = \langle b, c, d, aba, aca, ada \rangle$.

Proof. Since $\mathrm{Stab}_G(1)$ is a subgroup of G, every element in it is a product of a, b, c, d. Using the relations given by the preceding two lemmas, it suffices to show that the normal form of any element in $\mathrm{Stab}_G(1)$ can be written as a product of b, c, d, aba, aca, ada.

A product of a, b, c, d lies in $\mathrm{Stab}_G(1)$ if and only if it involves an even number of a, since a interchanges the vertices 0 and 1 while b, c, d all preserve them. By the preceding two lemmas, the normal form of elements in $\mathrm{Stab}_G(1)$ must be obtained by pluging b, c, d into the gaps between the even number of a's, one and only one for each gap, and there may be at most one of b, c, d on each of the most left or right sides. For example *babacada* is a normal form. Since there are an even number of a's, we can always add brackets so that the normal form is a product of b, c, d, aba, aca, ada. Explicitly, we start from left to right and count the a's, labeling each a with the number that we count. For a's labelled by an odd number, we put a left bracket on its left; for those labelled by an eve number, we put a right bracket on its right. For example, we have $b(aba)c(ada)$. □

Observe that we have the following fact, which one could verify via an easy computation of their action on vertices:

Lemma 9.13

$aba = (c, a)$, $aca = (d, a)$, $ada = (b, 1)$. □

With the help of $\mathrm{Stab}_G(1)$, we can then show that G is infinite.

Lemma 9.14

G is an infinite group.

Proof. Let us consider the following map:

$$\begin{array}{ccccccc} \mathrm{Stab}_G(1) & \subset & \mathrm{Stab}(1) & \cong & \mathrm{Aut}(T) \times \mathrm{Aut}(T) & \to & \mathrm{Aut}(T) \\ \varphi & & \mapsto & & (\varphi', \varphi'') & \mapsto & \varphi'' \end{array}$$

Since the projection $\mathrm{Aut}(T) \times \mathrm{Aut}(T) \to \mathrm{Aut}(T)$ is a homomorphism, this map is a homomorphism from $\mathrm{Stab}_G(1)$ to $\mathrm{Aut}(T)$. Let us check its image. The images of the generators of $\mathrm{Stab}_G(1)$ are easily computed using our previous results:

$$\begin{aligned} b &\mapsto c, \quad aba \mapsto a, \\ c &\mapsto d, \quad aca \mapsto a, \\ d &\mapsto b, \quad ada \mapsto 1. \end{aligned}$$

Therefore the map is a homomorphism from $\mathrm{Stab}_G(1)$ onto G. Since $\mathrm{Stab}_G(1)$ is a proper subgroup of G, it is impossible for G to be finite. □

To show that G is torsion, we will use induction on length of words. Recall that in a reduction system a length of an element is the minimal length of words representing that element. Here for an element in G, we consider its length in the reduction system where the generators are a, b, c, d. We have proved that elements of length no more than 1 are torsion.

Noticing that conjugate elements have the same order, we need only consider elements up to conjugation.

Lemma 9.15

For $g \in G$ with $\text{length}(g) = 2$, we have $g^{16} = 1$.

Proof. By lemma 9.10 and lemma 9.11, the elements with length 2 up to conjugation are ab, ac and ad. We have

$$(ad)^2 = (ada)d = (b,1)(1,b) = (b,b) \quad \Rightarrow \quad (ad)^4 = 1,$$

$$(ac)^2 = (aca)c = (d,a)(a,d) = (da,ad) \quad \Rightarrow \quad (ac)^8 = 1,$$

$$(ab)^2 = (aba)b = (c,a)(a,c) = (ca,ac) \quad \Rightarrow \quad (ab)^{16} = 1.$$

These finish the proof. □

According to our previous observation, every element in G can be written as one of the following four reduced forms:

$$\begin{aligned} &ax_1ax_2a\cdots ax_ka && \text{(I)}\\ &x_1ax_2a\cdots ax_ka && \text{(II)}\\ &ax_1ax_2a\cdots ax_k && \text{(III)}\\ &x_1ax_2a\cdots ax_k && \text{(IV)} \end{aligned}$$

where $x_1, \cdots, x_k \in \{b,c,d\}$. For type (II), conjugation by x_1 makes it become type (III). For type (I), conjugation by a makes its length decrease by 2 and become type (IV); for type (IV), conjugation by x_1 either makes it become type (III) (if $x_1 \neq x_k$) or make it become type (I) with length decreased by 2 (if $x_1 = x_k$). Therefore, every element is conjugate to either an element of type (III) or the generators. Note that elements of type (III) has length $2k$.

Lemma 9.16

G is torsion.

Proof. It remains only to show that every element of type (III) $g = ax_1a\cdots ax_k$ is torsion. Let us do induction on $\text{length}(g)$ (instead of k); we have shown that the statement is true for all elements with length no more than 2. For any element g of type (III) with length $2k > 0$, let us make the induction assumption that all elements with length less than $2k$ are torsion, and show that g must also be torsion.

Suppose that k is even, then g lives in $\text{Stab}_G(1)$ since there is an even number of a's. Write $k = 2m$, then we may break g into a product of $(ax_ia)x_{i+1}$'s:

$$g = [(ax_1a)x_2]\cdot[(ax_3a)x_4]\cdots[(ax_{2m-1}a)x_{2m}].$$

Each $(ax_ia)x_{i+1}$ lives in $\text{Stab}_G(1)$ and can be sent to $(\tilde{x}_i\bar{x}_{i+1}, \tilde{y}_i\bar{y}_{i+1}) \in \text{Aut}(T)\times\text{Aut}(T)$ via the map $\text{Stab}_G(1) \subset \text{Stab}(1) \cong \text{Aut}(T)\times\text{Aut}(T)$, and we have $\tilde{x}_i, \bar{x}_{i+1}, \tilde{y}_i, \bar{y}_{i+1} \in \{a,b,c,d\}$ based on lemma 9.13. Noticing that $\text{length}(\tilde{x}_i\bar{x}_{i+1}), \text{length}(\tilde{y}_i\bar{y}_{i+1}) \leq 2$, we see that $g = (u,v) \in \text{Aut}(T)\times\text{Aut}(T)$ with $\text{length}(u), \text{length}(v) \leq 2m = k < 2k$. By our induction assumption, both u and v are torsion, hence g must also be torsion.

Suppose otherwise that k is odd, then g^2 has length $4k$ and is still of type (III). Explicitly, we have

$$g^2 = [(ax_1a)x_2]\cdots[(ax_{k-2}a)x_{k-1}]\cdot[(ax_ka)x_1]\cdot[(ax_2a)x_3]\cdots[(ax_{k-1}a)x_k].$$

Again we have $(ax_ia)x_j = (\tilde{x}_i\bar{x}_j, \tilde{y}_i\bar{y}_j) \in \text{Aut}(T)\times\text{Aut}(T)$. If nothing cancels when we multiply the pairs together and obtain $g^2 = (u,v)$, then we would have $\text{length}(u) = \text{length}(v) = 2k$, for which we cannot apply our induction assumption. Therefore we need to show that there must be something that cancels. Note that for any $i \in \{1,\cdots,k\}$, x_i appears twice in the decomposition, once as x_i and once as ax_ia.

If there exists $i \in \{1,\cdots,k\}$ such that $x_i = d$, then we have $x_i = d = (1,b)$ and $ax_ia = (b,1)$, hence the result follows.

If there does not exist i such that $x_i = d$, then $x_i \in \{b,c\}$ for all $i = 1,\cdots,k$. If there exists $i \in \{1,\cdots,k\}$ such that $x_i = c$, then since $c = (a,d)$ and $aca = (d,a)$, replace g with u and v in the above argument and it follows that they are both torsion, so is g.

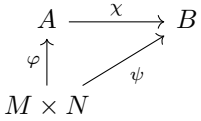
If there does not exists $i \in \{1,\cdots,k\}$ such that $x_i \in \{c,d\}$, then $x_1 = \cdots = x_k = b$, hence $g = (ab)^k$. Recalling that $(ab)^{16} = 1$, we are done. □

Therefore G is a counterexample to the General Burnside Problem.

§10 Tensor Product

Tensor Product for Modules

Let us introduce the tensor product in the most general case. Let R be a ring, M be a right R-module and N be a left R-module. For an abelian group A, a map $\varphi\colon M\times N\to A$ is *balanced* if φ is bilinear over $\mathbb{Z}$ and $\varphi(mr,n)=\varphi(m,rn)$ for any $m\in M$, $n\in N$ and $r\in R$. Consider the category where the objects are pairs $(A,\varphi\colon M\times N\to A)$ where A is an abelian group and $\varphi\colon M\times N\to A$ is a balanced map. A morphism from $(A,\varphi\colon M\times N\to A)$ to $(B,\psi\colon M\times N\to B)$ is a group homomorphism $\chi\colon A\to B$ such that the following diagram commutes:



Let us denote the universal object (recall that by universal we refer to initial) in this category as $(U,u\colon M\times N\to U)$. With the argument we used before for other universal objects, one proves easily that the universal object is unique and that U is generated as an abelian group by $u(M\times N)$. Explicitly, let $X=\{x_{m,n}\mid m\in M,n\in N\}$ be an alphabet indexed by $M\times N$, let $S\subset X$ be defined by $S=\{x_{(m+m),n}-x_{m,n}-x_{m,n},x_{m,n+n}-x_{m,n}-x_{m,n},x_{mr,n}-x_{m,rn}\mid m,m\in M,n,n\in N,r\in R\}$, then

$$U=\mathbb{Z}X/\mathbb{Z}S.$$

Hence the existence of U is also proved.

We define the tensor product of M and N as U, and write $M\otimes_R N\coloneqq U$. Also, elements in U are written in the notation that $m\otimes n\coloneqq u(m,n)$ for $m\in M$ and $n\in N$, hence every element in U is a sum $\sum_i m_i\otimes n_i$. Since u is balanced, we have

$$k(m\otimes n)=ku(m,n)=u(km,n)=(km)\otimes n=u(m,kn)=m\otimes(kn),$$

and

$$(mr)\otimes n=u(mr,n)=u(m,rn)=m\otimes(rn),$$

for any $k\in\mathbb{Z}$, $m\in M$, $n\in N$ and $r\in R$. Note that we have $0=0\otimes 0$ in U, and consequently

$$0\otimes n=(0\cdot 0)\otimes n=0\otimes(0\cdot n)=0\otimes 0=0=m\otimes 0.$$

Example 10.1

Consider the tensor product $(\mathbb{Z}/2\mathbb{Z})\otimes_{\mathbb{Z}}(\mathbb{Z}/3\mathbb{Z})$. For any $m\otimes n\in(\mathbb{Z}/2\mathbb{Z})\otimes_{\mathbb{Z}}(\mathbb{Z}/3\mathbb{Z})$, we have

$$2(m\otimes n)=(2m)\otimes n=0\otimes n=0,$$

and

$$3(m\otimes n)=m\otimes(3n)=m\otimes 0=0.$$

Hence

$$m\otimes n=3(m\otimes n)-2(m\otimes n)=0-0=0.$$

Therefore all elements in $(\mathbb{Z}/2\mathbb{Z})\otimes_{\mathbb{Z}}(\mathbb{Z}/3\mathbb{Z})$ are zero, concluding that $(\mathbb{Z}/2\mathbb{Z})\otimes_{\mathbb{Z}}(\mathbb{Z}/3\mathbb{Z})$ is the trivial group.

Lecture 17

Example 10.2

We have $M \otimes_R R \cong M$ via the map $m \otimes r \mapsto mr$. The surjectivity is obvious. For the injectivity, if $\sum_i m_i \otimes r_i$ is mapped to zero then $\sum_i m_i r_i = 0$. Hence

$$\sum_i m_i \otimes r_i = \sum_i (m_i r_i) \otimes 1 = \left(\sum_i m_i r_i \right) \otimes 1 = 0 \otimes 1 = 0.$$

Similar to this example, we have that if $M = \bigoplus_i M_i$ and $N = \bigoplus_i N_i$, then $M \otimes_R N \cong \bigoplus_{i,j} (M_i \otimes_R N_j)$, where the isomorphism is induced by the balanced map

$$\left(\bigoplus_i M_i \right) \times \left(\bigoplus_j N_j \right) \to \bigoplus_{i,j} (M_i \otimes_R N_j) \colon \left(\sum_i m_i, \sum_j n_j \right) \mapsto \sum_{i,j} m_i \otimes n_j,$$

hence is well-defined. The surjectivity is obvious. The injectivity follows from the direct sum that if $\sum_{i,j} m_i \otimes n_j \in \bigoplus_{i,j} (M_i \otimes_R N_j)$ is zero then $m_i \otimes n_j = 0$ for all pairs (i, j).

Tensor Product for Bimodules

Let R and S be rings. An abelian group M is said to be an (R, S)-*bimodule* if M is a left R-module and a right S-module at the same time, satisfying the associative law that $(rm)s = r(ms)$ for all $r \in R$, $s \in S$ and $m \in M$.

Example 10.3

The additive group of all $m \times n$ matrices over a field F is a $(M_m(F), M_n(F))$-bimodule.

Let R, S and T be rings. Let M be an (R, S)-bimodule and N be an (S, T)-bimodule, then $M \otimes_S N$ is an (R, T)-bimodule, where the scalar multiplication, for example by R, is defined by that

$$r \left(\sum_i m_i \otimes n_i \right) \coloneqq \sum_i (rm_i) \otimes n_i.$$

The well-definedness can be seen from that, the left multiplication by an element $r \in R$ can be seen as a well-defined endomorphism on $M \otimes_S N$ induced by the balanced map

$$M \times N \to M \otimes_S N \colon (m, n) \mapsto (rm, n).$$

For $R = S = F$ (or $S = T = F$) where F is a field, M and N are vector spaces over F and so is $M \otimes_F N$. The previous argument about direct sum that $\left(\bigoplus_i M_i \right) \otimes_R \left(\bigoplus_j N_j \right) \cong \bigoplus_{i,j} (M_i \otimes_R N_j)$ thus tells that if $M = \bigoplus_i Fe_i$ and $N = \bigoplus_j Ff_j$, then

$$M \otimes_F N = \bigoplus_{i,j} (Fe_i \otimes Ff_j).$$

Since $Fe_i \otimes Ff_j = F(e_i \otimes f_j)$, we see that $e_i \otimes f_j$'s form a basis of $M \otimes_F N$. Consequently $\dim_F M \otimes N = (\dim_F M)(\dim_F N)$.

Proposition 10.1

Let F be a field. Let M and N be vector spaces over F. If $u_1, \cdots, u_n \in M$ are linearly independent and $\sum_i u_i \otimes v_i = 0$ for some $v_i \in N$, then $v_1 = \cdots = v_n = 0$.

Note that since for vector spaces we have symmetry $M \otimes_F N = N \otimes_F M$, similar result holds for linearly independent lists in N.

Proof. Extend $u_1, \cdots, u_n$ to a basis of M and let $f_1, \cdots f_m$ be a basis of N. Write for each v_i, $v_i = \sum_j a_{ij} f_j$. We have

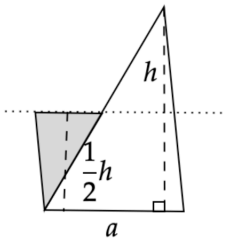
$$0 = \sum_i u_i \otimes v_i = \sum_i u_i \otimes \left(\sum_j a_{ij} f_j \right) = a_{ij} \sum_{i,j} u_i \otimes f_j.$$

Since by the previous argument $u_i \otimes f_j$'s are linearly independent, we have $a_{ij} = 0$ for all i and j, hence the result follow. □

Hilbert's Third Problem

Let us introduce an application of this proposition. For a tetrahedron, we know by calculus that its volume is equal to $\frac{1}{3}hS$, where S is its bottom area and h is its height. Hilbert asked: is there any "Greek" proof to this volume formula? This is known as a variant of Hilbert's Third Problem.

A "Greek" proof refers to a proof of cutting (straightly) an arbitrary tetrahedron (or some copies of the same tetrahedron) into pieces and glue it up to something whose volume can be computed by fundamental methods. For example, in the 2-dimsional case, we cut a triangle along its midline and then glue it up to a parallelogram with half height of the triangle, so that we know the area of the triangle is $\frac{1}{2}ah$. For the 3-dimensional case, we necessarily need to cut and glue a tetrahedron into a parallelepiped.



Hilbert's student, Max Dehn, proved using tensor product, that there does not exist such a "Greek" proof for the volume formula of tetrahedron. He constructed an invariant of polyhedrons under the process of cutting and gluing, which is always zero for parallelepipeds while there exists tetraherons whose corresponding invariant is non-zero. Let us explain.

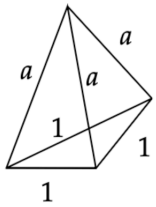
Each polyhedron has finitely many edges $\ell_1, \cdots, \ell_n$. Let $\ell_1, \cdots, \ell_n$ also denote the length of the corresponding edges. Every edge ℓ_i is the intersection of two faces, hence between the faces there lies an angle $\alpha_i \in (0, \pi)$. For each polyhedron with edges $\ell_1, \cdots, \ell_n$ and corresponding angles $\alpha_1, \cdots, \alpha_n$, we assign an element in $\mathbb{R} \otimes_{\mathbb{Q}} \mathbb{R}/\mathbb{Q}\pi$, defined by

$$\sum_{i=1}^{n} \ell_i \otimes (\alpha_i + \mathbb{Q}\pi).$$

This element is called the *Dehn invariant* of that polyhedron. It is invariant under the process of cutting and gluing: every step of cutting creates some pairs of edges of identical length, and for each pair their corresponding angles add up to π, hence the new elements have no contribution to the sum of the invariant. Conversely, every step of gluing eliminates a set of elements that have no contribution to the sum.

It is easy to see that the invariants for parallelepipeds are identically zero, because one can simply cut and glue a polyhedron to a cuboid, where all angles are $\frac{\pi}{2} \in \mathbb{Q}\pi$, giving only zeros in the sum. Now, all what remains is to show that there exists a tetrahedron whose invariant is nonzero.

Let us consider the following family of tetrahedrons, whose bottom is the equilateral triangle whose sides have lengths identically 1, and all other three edges of the polyhedron have identical length a.



Let the angles corresponding to the edges of length a be α and those corresponding to the edges of length 1 be β, then the invariant of such tetrahedron is

$$3\left(a \otimes (\alpha + \mathbb{Q}\pi) + 1 \otimes (\beta + \mathbb{Q}\pi)\right).$$

As 1 is fixed, both α and β are determined by a. Noticing that when a is irrational, a and 1 are linearly independent over $\mathbb{Q}$, hence by proposition 10.1 it suffices to show that there exists irrational a such that either $\alpha + \mathbb{Q}\pi$ or $\beta + \mathbb{Q}\pi$ is nonzero, so that the sum would be nonzero. Let us focus on $\beta + \mathbb{Q}\pi$. We have $\beta + \mathbb{Q}\pi = 0$ if and only if β/π is rational. Let us write $\beta(a)$, indicating that β is a function of a. We notice that $\beta(a)$ is monotonically strictly increasing as a increases, so is $\beta(a)/\pi$. The range a could be is $(\frac{1}{\sqrt{3}}, \infty)$ and the range for β is $(0, \frac{\pi}{2})$, so we conclude that the function $\beta/\pi \colon (\frac{1}{\sqrt{3}}, \infty) \to (0, \frac{1}{2})$ is monotonically strictly increasing, hence is injective. Since there are uncountably many irrational numbers in $(\frac{1}{\sqrt{3}}, \infty)$, the image of β/π on all those irrational numbers cannot only contain rational numbers which are only countable many, concluding the proof.

Tensor Product for Algebras

Let A and B be two F-algebras. Recall that an F-algebra A is a vector space over F equipped with a F-bilinear operator of multiplication $A \times A \to A$. Consider the tensor product $A \otimes_F B$, then it is a vector space over F. It can be made into an F-algebra by defining the multiplication by

$$\left(\sum_i a_i \otimes b_i\right)\left(\sum_j \tilde{a}_j \otimes \tilde{b}_j\right) = \sum_{i,j} a_i \tilde{a}_j \otimes b_i \tilde{b}_j.$$

To see that such multiplication is well-defined, let us show that the left and right multiplications by any element is well-defined. Indeed, the left multiplication by a simple element $a \otimes b$ is the well-defined F-linear map induced by the balanced map

$$A \times B \to A \otimes_F B \colon (\tilde{a}, \tilde{b}) \mapsto a\tilde{a} \otimes b\tilde{b},$$

and similarly is the right multiplication. For left multiplication by non-simple elements, we need to show that if $\sum_i a_i \otimes b_i = 0$ then $(\sum_i a_i \otimes b_i)\, x = 0$ for all $x \in A \otimes_F B$. It suffices to show this for simple elements x, which then follows from the F-linearity of the just-defined right multiplication by simple elements. Similarly right multiplication by non-simple elements is well-defined, so we are done.

Suppose that A and B both have multiplicative identities, say $1_A \in A$ and $1_B \in B$, then obviously $1_A \otimes 1_B$ is the multiplicative identity in $A \otimes_F B$.

Example 10.4

Let A be an F-algebra, then $A \otimes_F M_n(F) \cong M_n(A)$.

If A and B both have multiplicative identities, then there exists natural embeddings

$$A \hookrightarrow A \otimes_F B \colon a \mapsto a \otimes 1_B,$$

and

$$B \hookrightarrow A \otimes_F B \colon b \mapsto 1_A \otimes b.$$

Note that with these embeddings, A and B can be seen as subalgebras of $A \otimes_F B$ and they commute.

Suppose that C is an F-algebra satisfying that

- A and B are subalgebras of C;
- A and B commute in C;

- C is generated by A and B.

Then the balanced map

$$A \times B \to C \colon (a, b) \mapsto ab,$$

induces an F-linear map from $A \otimes_F B$ onto C. Moreover, it is easy to see that this map is a homomorphism of F-algebras. Therefore, $A \otimes_F B$ is the largest F-algebra satisfying the above three conditions.

As a final remark, the tensor product for bimodules and algebras both satisfy the associative law, for example $(A \otimes_F B) \otimes_F C \cong A \otimes_F (B \otimes_F C)$.

Lecture 18

Centroids and Central Simple Algebras

Let A be a ring that is not necessarily unitary or associative. For an element $a \in A$, we denote by R_a the map of right multiplication by a from A to A and by L_a the left. Consider the *multiplication algebra*, $M(A)$, defined by

$$M(A) = \langle R_a, L_a \mid a \in A \rangle,$$

i.e. $M(A)$ is the algebra generated by $\{R_a, L_a \mid a \in A\}$. Note that $M(A)$ is always associative, despite of whether A is associative or not.

If A is associative, then we have $R_a R_b = R_{ba}$, $L_a L_b = L_{ab}$ and $L_a R_b = R_b L_a$ for any $a, b \in A$. Let $R_A \coloneqq \{R_a \mid a \in A\}$ and $L_A \coloneqq \{L_a \mid a \in A\}$ be the rings of right multiplications and left multiplications, then the maps $A^{\mathrm{op}} \to R_A \colon a \mapsto R_a$ and $A \to L_A \colon a \mapsto L_a$ are surjective homomorphisms. Also, we have

$$M(A) = \left\{ \sum_i L_{a_i} R_{b_i} \mid a_i, b_i \in A \right\}.$$

If in addition that A is unitary, then the two maps are also injective, hence $R_A \cong A^{\mathrm{op}}$ and $L_A \cong A$; the unitary condition is required, otherwise a ring with zero multiplication would give a counterexample. Therefore, if A is a unitary associative F-algebra, we thus obtain a surjective homomorphism

$$A \otimes_F A^{\mathrm{op}} \to M(A) \colon \sum a_i \otimes b_i^{\mathrm{op}} \mapsto \sum L_{a_i} R_{b_i}.$$

Recall that a ring A is said to be *prime* if $IJ \neq (0)$ for any two nonzero ideals I and J of A. A ring A is said to be *simple* if there is no nontrivial ideals in A and $A^2 \neq (0)$. Hence a simple ring is automatically prime.

Example 10.5

It is easy to see that for any field F, the ring of n-by-n matrices over F, $M_n(F)$, is simple. Moreover, if a ring A is simple, then $M_n(A)$ is also simple.

Let A be an associative algebra (not necessarily unital or commutative) over a commutative ring R, then the centralizer of $M(A)$ in $\mathrm{End}_R(A)$, called the centroid $\mathrm{Cent}(A)$ of A, is defined by

$$\mathrm{Cent}(A) \coloneqq \{\varphi \in \mathrm{End}_R(A) \mid \varphi R_a = R_a \varphi,\ \varphi L_a = L_a \varphi,\ \forall a \in A\}.$$

Hence for any $\varphi \in \mathrm{Cent}(A)$, we have

$$\varphi(ab) = \varphi(a)b = a\varphi(b),$$

for all $a, b \in A$.

Note that if A is unitary, then $\mathrm{Cent}(A)$ can be identified with the center of A, $Z(A)$, via the map

$$\mathrm{Cent}(A) \to Z(A) \colon \varphi \mapsto \varphi(1).$$

Lemma 10.2

(1) If $A = A^2$, then $\mathrm{Cent}(A)$ is a commutative algebra (over R);

(2) If A is prime, then $\mathrm{Cent}(A)$ is a commutative domain (i.e. integral domain);

(3) If A is simple, then $\mathrm{Cent}(A)$ is a field.

Proof. (1) For any two elements $\varphi, \psi \in \mathrm{Cent}(A)$, we have

$$\varphi\psi(ab) = \varphi(a\psi(b)) = \varphi(a)\psi(b) = \psi(\varphi(a)b) = \psi\varphi(ab),$$

for any $a, b \in A$. Since $A^2 = A$, we thus have $\varphi\psi = \psi\varphi$.

(2) Suppose that $\varphi, \psi \in \mathrm{Cent}(A)$ satisfies that $\varphi\psi = 0$, which gives that

$$0 = \varphi\psi(ab) = \varphi(a)\psi(b),$$

for any $a, b \in A$. Hence $\varphi(A)\psi(A) = (0)$. Noticing that $\varphi(A)$ is an ideal of A since $\varphi(a)b = \varphi(ab) \in \varphi(A)$ for any $a, b \in A$ and so is $\psi(A)$, the primeness of A thus forces that either $\varphi(A) = 0$ or $\psi(A) = 0$, implying that either $\varphi = 0$ or $\psi = 0$.

To show the commutativity, for any two $\varphi, \psi \in \mathrm{Cent}(A)$, we consider $w \coloneqq \varphi\psi - \psi\varphi \in \mathrm{Cent}(A)$. The identity in the proof of (1) gives that $w(A^2) = 0$. Since

$$w(a) \cdot b = w(ab) = 0,$$

for any $a, b \in A$, we see that $w(A) \cdot A = 0$, hence by the primness of A we obtain $w(A) = 0$, concluding that $\varphi\psi - \psi\varphi = w = 0$.

(3) By (2), we see that $\mathrm{Cent}(A)$ is commutative. For any nonzero element $\varphi \in \mathrm{Cent}(A)$, $\varphi(A)$ is a nonzero ideal of simple A, hence $\varphi(A) = A$, thus φ is surjective. Noticing that the kernel of φ is also an ideal of A, we thus have $\ker\varphi = (0)$, hence φ is injective. Therefore there exists a multiplicative inverse φ^{-1} in $\mathrm{End}_R(A)$. To show that $\varphi^{-1} \in \mathrm{Cent}(A)$, since for any generator P of $M(A)$ we have $\varphi P = P\varphi$, there is

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

concluding the proof. □

Remark 10.1

With the same definition, one can also define the centroid for nonassociative algebras, and the proof above passes over since it does not use the associativity. Note that the observation below, which uses the commutativeness of $\mathrm{Cent}(A)$, relies on the associativity.

As a consequence of (3), if a simple F-algebra satisfies that $F = \mathrm{Cent}(A)$, then we say that A is *central simple*. Observe that a simple R-algebra A is also a $\mathrm{Cent}(A)$-module and

$$\mathrm{Cent}(A) \subset \mathrm{End}_{\mathrm{Cent}(A)}(A),$$

because of the commutativeness of $\mathrm{Cent}(A)$, by replacing R with $\mathrm{Cent}(A)$ we can make any simple algebra central simple.

From now on let us consider only associative algebras, since the definition of tensor products of nonassociative algebras is problematic.

Theorem 10.3

Let F be a field. Let A and B be simple F-algebras. Suppose that A is central simple and B is unitary, then $A \otimes_F B$ is simple.
If in addition that B is also central simple, then $A \otimes_F B$ is also central simple.

Before proving the theorem above, let us state the following theorem by Wedderburn and Artin, which reduces the problem of classifying finite-dimensional central simple algebras to the problem of classifying finite-dimensional division algebras. In history, it motivated the research of *division algebras*, i.e. unital algebras (not necessarily commutative or associative) where every nonzero element is invertible, and the later development of quantum mechanics.

Theorem 10.4. Wedderburn-Artin

(1) Any finitely dimensional simple associative unital algebra over a field is isomorphic to $M_n(D)$ for some positive integer n, where D is an associative division algebra over the same field;

(2) Suppose that A is a finitely dimensional associative unital algebra over a field satisfying that, for any ideal $I \lhd A$, the condition that $I^2 = (0)$ is equivalent to that $I = (0)$. Then

$$A \cong M_{n_1}(D_1) \oplus \cdots \oplus M_{n_k}(D_k),$$

for some positive integers $n_1, \cdots, n_k$, where $D_1, \cdots, D_k$ are all associative division algebras over the same field.

We will not prove part (2) of the theorem, whilst part (1) follows from the following two results:

Lemma 10.5. Schur

Let R be an F-algebra. Let V be a irreducible left R-module, i.e. a left R-module that contains no proper submodule. Then the centralizer of R in $\mathrm{End}_F(V)$, defined by $\Delta \coloneqq \{\varphi \in \mathrm{End}_F(V) \mid \varphi(av) = a\varphi(v),\ \forall a \in R,\ \forall v \in V\}$, is an associative division algebra.

Note that the V in the statement of Schur's lemma can also be viewed as a left Δ-module. Since the only difference between associative division algebras and fields is the commutativeness, a lot of results in Linear Algebra carries over to modules over an associative division algebra, including the definition of linear independence, basis, matrix of linear transformations. We may as well call a module over an associative division algebra a vector space.

With the same notation, we have

Theorem 10.6. Jacobson Density

If $v_1, \cdots, v_n \in V$ are linearly independent over Δ, then for any arbitrary elements $w_1, \cdots, w_n \in V$, there exists an element $a \in R$ such that $av_i = w_i$ for all $i = 1, \cdots, n$.

In particular, if $\dim_\Delta V < \infty$, then the map $R \to \mathrm{End}_\Delta(V)$ that sends an element in R to the left multiplication by that element is surjective. Since $\mathrm{End}_\Delta(V) \cong M_{\dim V}(\Delta)$ and $R \to \mathrm{End}_\Delta(V)$ is injective if $V = R$ and R is unital, this proves part (1) of Wedderburn-Artin theorem.

We would not bother proving these two results since it would take more than one hour.

Proof of Theorem 10.3. Now that A is central simple. Noticing that A is a module over $M(A)$ and a submodule of A over $M(A)$ is exactly an ideal of A, A is an irreducible module over $M(A)$. Substituting $R = M(A)$ and $V = A$ in the statement of Schur's lemma, we see that $\Delta = F$.

Let I be a nonzero ideal of $A \otimes_F B$, there there exists a nonzero element $\sum_{i=1}^n a_i \otimes b_i \in I$. By reducing a_i's and b_i's to a linearly independent list, we may assume that $a_1, \cdots, a_n$ and $b_1, \cdots, b_n$ are both linearly independent. In particular, none of them is zero. By Jacobson Density Theorem, for any $a \in A$ there exists an element $P = \sum L_{x_j} R_{y_j} \in M(A)$ such that $P(a_1) = a$ and $P(a_i) = 0$ for all $i = 2, \cdots, n$. Consider the element $\widetilde{P} \coloneqq \sum L_{x_j \otimes 1} R_{y_j \otimes 1} \in M(A \otimes_F B) \tilde{P} \coloneqq \sum L_{x_j \otimes 1} R_{y_j \otimes 1} \in M(A \otimes_F B)$, then $\widetilde{P}(\sum_{i=1}^n a_i \otimes b_i) \in I$ since I is an ideal and

$$\widetilde{P}(\sum_{i=1}^n a_i \otimes b_i) = P(a_1) \otimes b_1 + \cdots + P(a_n) \otimes b_n = a \otimes b_1,$$

concluding that $A \otimes b_1 \subset I$. Consider the ideal $J = \{b \in B \mid A \otimes b \in I\}$ of B, since $0 \neq b_1 \in J$ and B is simple, there must be $J = B$. Therefore $A \otimes_F B \subset I$, finishing the proof of the first half of the theorem.

The proof of the second half of the theorem will be given in the next lecture.

Lecture 19

Let us continue the proof. Suppose that both A and B are central simple, since $1_A \otimes 1_B \in A \otimes_F B$, we have $\mathrm{Cent}(A \otimes_F B) = Z(A \otimes_F B)$. Thus we need to show that $Z(A \otimes_F B) = F \cdot 1_A \otimes 1_B$. Since $F = Z(A) \otimes_F Z(B) \subset Z(A \otimes_F B)$, we need only prove the inclusion in the other direction.

Choose a nonzero element $z = \sum_{i=1}^{n} a_i \otimes b_i \in Z(A \otimes_F B)$, again we may assume that $a_1, \cdots, a_n$ are linearly independent. For any element $b \in B$, we have that

$$0 = [z, 1_A \otimes b] = \sum_{i=1}^{n} a_i \otimes [b_i, b],$$

where the bracket denotes the commutator. By proposition 10.1, we have $[b_i, b] = 0$ for all $i = 1, \cdots, n$, hence $b_1, \cdots, b_n \in Z(B) = F \cdot 1_B$ since b is arbitrary. Therefore we may rewrite z as

$$z = a \otimes 1_B.$$

Again, for any $a' \in A$, we have

$$0 = [a \otimes 1_B, a' \otimes 1_B] = [a, a'] \otimes 1_B,$$

hence $a \in Z(A) = F \cdot 1_A$. Therefore $z = \alpha \cdot 1_A \otimes 1_B$ for some $\alpha \in F$, concluding the proof. □

Brauer Group

From now on let us assume that all algebras are associative and unitary.

Let F be a field and A be a finite dimensional central simple algebra over F, then by part (1) of the Wedderburn-Artin theorem, there exists a positive integer n and a division algebra D over F such that

$$A \cong M_n(D).$$

With this identification, the center of A is exactly

$$Z(A) = \{z \cdot I_n \mid z \in Z(D)\}.$$

Since $Z(A) = F \cdot 1_A$, it follows that $Z(D) = F \cdot 1_A$, hence D is also central simple (note that a division algebra is automatically simple).

Let D_1 and D_2 be two finitely dimensional central division algebras over F, then by the just proved theorem, $D_1 \otimes_F D_2$ is central simple. Since $\dim(D_1 \otimes_F D_2) = (\dim D_1)(\dim D_2) < \infty$, there exists a positive integer m along with a finitely dimensional central division algebra D_3 such that

$$D_1 \otimes_F D_2 \cong M_m(D_3).$$

Therefore, we may define a binary operator on finitely dimensional central division algebras that send the pair (D_1, D_2) to $D_1 \cdot D_2 \coloneqq D_3$. The following exercise guarantees that the operator is well-defined:

Exercise 10.1

Suppose that $M_n(D) \cong M_{n'}(D')$ as F-algebras, where $n, n' \in \mathbb{N}_{\geq 1}$ and D and D' are both finitely dimensional division algebras over F, then there must be $n = n'$ and $D = D'$.

Proof. If D and D' are both commutative, then we have $D \cdot I_n = Z(M_n(D)) \cong Z(M_{n'}(D')) = D' \cdot I_{n'}$, hence $D = D'$ and consequently $n = n'$. For the general case, since

$$M_n(D/[D, D]) = M_n(D)/[M_n(D), M_n(D)] \cong M_{n'}(D')/[M_{n'}(D'), M_{n'}(D')] = M_{n'}(D'/[D', D']),$$

and $D/[D, D]$ and $D'/[D', D']$ are both commutative, we see that $n = n'$.

Now that $M_n(D) \cong M_n(D')$. It is not hard to see that the following four maps are all isomorphisms of algebras:

$$D^{\mathrm{op}} \to \mathrm{End}_{D-\mathrm{Mod}}(D)\colon d \mapsto (-)\cdot d$$
$$D^{\mathrm{op}} \to \mathrm{End}_{M_n(D)-\mathrm{Mod}}(D^n)\colon d \mapsto (-)\cdot d$$
$$M_n(D)^{\mathrm{op}} \to M_n(D^{\mathrm{op}})\colon A \mapsto A^T$$
$$M_n(D^{\mathrm{op}}) \to \mathrm{End}_{D-\mathrm{Mod}}(D^n)\colon A \mapsto (-)\cdot A$$

Since D^n is the unique simple $M_n(D)$-module up to isomorphism (for uniqueness, see for example Corollary 4.6 in Chapter XVII of [Lan02]), from $M_n(D) \cong M_n(D')$ we see that $D^n \cong (D')^n$ as $M_n(D)$-modules, we thus obtain

$$D^{\mathrm{op}} \cong \mathrm{End}_{M_n(D)-\mathrm{Mod}}(D^n) \cong \mathrm{End}_{M_n(D')-\mathrm{Mod}}((D')^n) \cong (D')^{\mathrm{op}},$$

concluding the proof. □

In other words, $D_1 \cdot D_2$ is the unique division algebra satisfying that

$$D_1 \otimes_F D_2 \cong M_m(D_1 \cdot D_2),$$

for some positive integer m.

Note that $M_n(D) = M_n(F) \otimes_F D$. Since the tensor is associative and commutative, the result of the exercise implies that the operator is also associative and commutative, as we have

$$\begin{aligned}(D_1 \otimes D_2) \otimes D_3 &\cong M_m(D_1 \cdot D_2) \otimes D_3 \\ &= M_m(F) \otimes (D_1 \cdot D_2) \otimes D_3 \\ &\cong M_m(F) \otimes M_n((D_1 \cdot D_2) \cdot D_3) \\ &= M_{mn}(F) \otimes ((D_1 \cdot D_2) \cdot D_3) = M_{mn}((D_1 \cdot D_2) \cdot D_3),\end{aligned}$$

similarly

$$D_1 \otimes (D_2 \otimes D_3) \cong M_{m'n'}(D_1 \cdot (D_2 \cdot D_3)),$$

and

$$M_n(D_3) \cong D_1 \otimes D_2 \cong D_2 \otimes D_1 \cong M_{n'}(D').$$

So far we have seen that the set of all finitely dimensional central division algebras over F along with the operator form a commutative semigroup. In fact, we have the following

Proposition 10.7

The set of all finitely dimensional central division algebras over F along with the operator defined above is an abelian group.

Proof. Since $D \otimes_F F = D$, we have $D \cdot F = D$, i.e. F is the multiplicative identity. For any finitely dimensional central division algebra D, we show that $D \cdot D^{\mathrm{op}} = F$. Since D is central simple, so is D^{op}, hence $D \otimes_F D^{\mathrm{op}}$ is also simple. Recall that we have a surjective homomorphism $D \otimes_F D^{\mathrm{op}} \cong L_D \otimes R_D \to M(D)$. Since $D \otimes_F D^{\mathrm{op}}$ is simple, the surjective homomorphism must also be injective, hence we obtain $D \otimes_F D^{\mathrm{op}} \cong M(D)$. Since D is an irreducible module over $M(D)$ and $\mathrm{Cent}(D) = F$ since D is central, the Jacobson Density Theorem tells that $M(D) \cong \mathrm{End}_F(D) \cong M_k(F)$ where $k = \dim_F D$. Therefore $D \otimes_F D^{\mathrm{op}} \cong M_k(F)$, concluding that $D \cdot D^{\mathrm{op}} = F$. □

The group defined in this way is called *the Brauer group over F*, denoted as $\mathrm{Br}(F)$.

Exercise 10.2

Show that the only finitely dimensional associative division algebra over an algebraically closed field F is F itself. Therefore the Brauer group over an algebraically closed field is trivial.

Proof. Let D be a finitely dimensional associative division algebra over an algebraically closed field F. For any element $x \in D$, consider $F[x] \subset D$. Since x is algebraic over F and F is algebraically closed, there must be $F[x] = F$, hence $x \in F$, concluding the proof. □

Let us end this section with an example displaying that the condition that A is central is necessary in the first half of theorem 10.3. Before the example, let us recall some Galois theory. Let K be a Galois extension over F, hence $\dim_F K < \infty$ and the Galois group $G \coloneqq \mathrm{Aut}_F(K)$, the group of field automorphisms of K that fix F, satisfies that $|G| = \dim_F K =: n$. Let us write $G = \{\sigma_1, \cdots, \sigma_n\}$.

Theorem 10.8

With the notations above, we have $K \otimes_F K \cong \bigoplus_{i=1}^n K$ as vector spaces over F.

Proof. Consider the map $K \times K \to K : (k_1, k_2) \mapsto k_1^{\sigma_i} k_2$, where $k_1^{\sigma_i} \coloneqq \sigma_i(k_1)$. The map is bilinear over F, hence is balanced. Therefore we can define a map $K \otimes_F K \to \bigoplus_{i=1}^n K$ by sending $k_1 \otimes k_2$ to $k_1^{\sigma_1} k_2 \oplus \cdots \oplus k_1^{\sigma_n} k_2$. Since $n^2 = \dim_F K \otimes_F K = \dim_F \bigoplus_{i=1}^n K$, it suffices to show that the map is injective. Let $a_1, \cdots, a_n$ be a basis of K over F and suppose that $\sum_{i=1}^n a_i \otimes b_i$ is in the kernel of the map, which means exactly that

$$\sum_{i=1}^n a_i^{\sigma_j} b_i = 0,$$

for all $j = 1, \cdots, n$. Recall that by proposition 10.1, $\sum_{i=1}^n a_i \otimes b_i = 0$ if and only if $b_1 = \cdots = b_n = 0$, the condition that the map is injective is thus equivalent to that the following n-by-n matrix is nonsingular:

$$\begin{pmatrix} a_1^{\sigma_1} & \cdots & a_n^{\sigma_1} \\ \vdots & \ddots & \vdots \\ a_1^{\sigma_n} & \cdots & a_n^{\sigma_n} \end{pmatrix}.$$

Suppose not, then the rows are linearly dependent, hence there exists $k_1, \cdots, k_n \in K$ which are not all zeros such that

$$\sum_{i=1}^n k_i a_j^{\sigma_i} = 0,$$

for all $j = 1, \cdots, n$. Since $a_1, \cdots, a_n$ is a basis of K over F and σ_i's are F-linear, this implies that

$$\sum_{i=1}^n k_i \sigma_i = 0.$$

However, by Artin's theorem (see for example Theorem 4.1 in Chapter VI of [Lan02]), elements in G are linearly independent over K, a contradiction. □

Example 10.6

Let K be an extension of a field F with $\mathrm{char}\, F = p > 0$ such that there exists an element $k \in K \setminus F$ satisfying $k^{p^s} \in F$ for some positive integer s. Consider the element $k \otimes 1 - 1 \otimes k$ in $K \otimes_F K$, it is nonzero since 1 and k are linearly independent over F, and it is a nilpotent as a consequence of the binomial formula. Therefore K is associative, simple and unital, but $K \otimes_F K$ contains a nonzero nilpotent element $k \otimes 1 - 1 \otimes k$, hence is not simple since the radical ideal of (0) is proper.

§11 Rings of Fractions and Ore Condition

Let us now introduce the construction of fraction rings, which is the noncommutative analogy of localization.

Let R be a ring. An element $r \in R$ is *regular* if it is not a zero divisor, i.e. if $a \in R$ satisfies either $ar = 0$ or $ra = 0$, then $a = 0$. It is easy to see that the set of regular elements is closed under multiplication. Let S be a multiplicative subsemigroup of R consisting of some regular elements of R. An extension ring $\widetilde{R} \supset R$ is a *left ring of fractions of R relative to S* if

(1) all elements in S are invertible in $\widetilde{R}$;

(2) for any element $x \in \widetilde{R}$, there exists $s \in S$ and $r \in R$ such that $x = s^{-1} r$.

A *right ring of fractions of R relative to S* is defined similarly, with the second condition modified as $x = rs^{-1}$.

If an extension is both a left ring of fractions and a right ring of fractions, then it must satisfy that for any $r \in R$ and $s \in S$, there exists $r_1 \in R$ and $s_1 \in S$ such that

$$rs^{-1} = s_1^{-1} r_1.$$

Multiplication on the left by s_1 and on the right by s tells that a necessary condition for a ring R to admit an extension that is a left and right ring of fractions relative to S is that for any $r \in R$ and $s \in S$, there exists $r_1 \in R$ and $s_1 \in S$ such that

$$s_1 r = r_1 s.$$

This condition is called the *Ore condition*. Not all rings (along with the subsemigroup) satisfy the Ore condition:

Example 11.1

Consider an alphabet $X = \{x, y\}$ of two points and let R be the free semigroup algebra $R = F\langle X\rangle$. Let S be the free semigroup $S = X^*$ (in fact, the choice of S does not matter in this case). For $s = x \in S$ and $r = y \in R$, the product $s_1 y$ ends with y for any $s_1 \in S$ while $r_1 x$ ends with x for any $r_1 \in R$, hence there cannot exist $s_1 \in S$ and $r_1 \in R$ such that $s_1 y = r_1 x$.

The Ore condition is in fact also sufficient.

Theorem 11.1

A left and right ring of fractions of R relative to S exists if and only if the pair (R, S) satisfies the Ore condition.

Proof. It remains only to show that the Ore condition is sufficient. Suppose that (R, S) satisfies the Ore condition, we consider the ring

$$\widetilde{R} \coloneqq \langle R, S^{\mathrm{op}} \mid s s^{\mathrm{op}} = 1_R \rangle,$$

which obviously satisfy the first condition. A word in $\tilde{R}$ is an interchanging product of elements in R and S^{op}, e.g., $r_1 s_1^{\mathrm{op}} \cdots s_{n-1}^{\mathrm{op}} r_n$, since both R and S^{op} are closed under multiplication. The Ore condition then tells that for any rs^{op}, there exists $r' \in R$ and $(s')^{\mathrm{op}} \in S^{\mathrm{op}}$ such that

$$rs^{\mathrm{op}} = (s')^{\mathrm{op}} r'.$$

For each word in $\widetilde{R}$, apply this for finite many times and we see that it can be written in both of the forms $s^{\mathrm{op}} r$ and rs^{op}, concluding the proof. □

Lecture 20

Let us give another construction of rings of fractions.

For any subset $M \subset R$, let us consider its *right eliminator*, defined by

$$r(M) \coloneqq \{a \in R \mid Ma = (0)\}.$$

Clearly $r(M)$ is a right ideal of R. Similarly one defines *the left eliminator of* M, which is a left ideal of R.

Let L be a left ideal of R (for which we will write $L \lhd_l R$) and $a \in R$ be an arbitrary element. Consider the following subset

$$La^{-1} \coloneqq \{x \in R \mid xa \in L\},$$

which is again a left ideal of R. Note that we do not require $a \in R$ to be invertible; if a is invertible, then the above notation coincides with the usual one. For instance, we have

$$L0^{-1} = R.$$

Similarly one may define $a^{-1}\varrho$ for a right ideal $\varrho \lhd_r R$, which again is a right ideal.

Definition 11.1

A left ideal $L \lhd_l R$ is called *dense* if for any element $a \in R$, the right eliminator of La^{-1} is zero, i.e., $r(La^{-1}) = (0)$.

We will denote a dense left ideal by $L \lhd_l^{\text{dense}} R$. Note that the zero ideal can never be dense in a nonzero ring, since $r(0) = R$.

The condition that a left ideal L is dense in R is equivalent to the following, which is easier to work with:

> For any element $a \in R$ and $0 \neq b \in R$, there exists an element $x \in R$ such that $xa \in L$ (so that $x \in La^{-1}$) and $xb \neq 0$ (so that $b \notin r(La^{-1})$).

Example 11.2

If a ring R contains a regular element u, then R itself is a dense left ideal of R, since $ub \neq 0$ for any $0 \neq b \in R$.

Lemma 11.2

Suppose that L_1 and L_2 are two dense left ideals of R, then $L_1 \cap L_2$ is also a dense left ideal of R.

Proof. For any element $a \in R$ and $0 \neq b \in R$, since L_1 is dense, there exists an element $x \in R$ such that

$$xa \in L_1, \quad xb \neq 0.$$

Now that $xa \in R$ and $0 \neq xb \in R$, hence by the denseness of L_2, there exists an element $y \in R$ such that

$$y(xa) \in L_2, \quad y(xb) \neq 0.$$

Since L_1 is a left ideal, we also have that $y(xa) \in L_1$. Therefore

$$(yx)a \in L_1 \cap L_2, \quad (yx)b \neq 0,$$

concluding the proof. □

Given a ring R, consider the set $\overline{R} \coloneqq \{f \in \text{L-Mod}_R(L, R) \mid L \lhd_l^{\text{dense}} R\}$, where $\text{L-Mod}_R(L, R)$ denotes the set of all left R-module homomorphisms.

With the above lemma, we can define an equivalence relation on $\overline{R}$: for any two elements $f_1, f_2 \in \overline{R}$, $f_1 \sim f_2$ if they coincide on some dense left ideal, i.e. there exists a dense left ideal $L_3 \subset L_1 \cap L_2$ such that $f_1|_{L_3} = f_2|_{L_3}$, where L_1 and L_2 are the domains of f_1 and f_2 respectively.

Remark 11.1

This construction is very similar to the rational maps in algebraic geometry. However, in this case the domains of maps are not open sets of a topology: although dense left ideals are closed under finite intersections, the union of ideals may not be ideals. Still, one may replace the union by the ideal generated by the union, and obtain something close to a topology.

With this equivalence relation $\sim$, let us consider the set $\widetilde{R} \coloneqq \overline{R}/\sim$. We will define operators on it so that it becomes a ring.

The addition is defined point-wisely, i.e. suppose $f_1, f_2 \in \overline{R}$ with domains L_1 and L_2 respectively, then $f_1 + f_2$ is defined as $f_1 + f_2 \colon L_1 \cap L_2 \to R \colon a \mapsto f_1(a) + f_2(a)$. One can see that it is invariant under the equivalence by replacing beforehead the domains with a sufficiently smaller one, hence it gives a well-defined addition on $\widetilde{R}$.

The multiplication is defined as the composition of maps (restricted appropriately), for which we need the following lemma.

Lemma 11.3

If $L \lhd_l^{dense} R$ and $g \colon L \to R$ is a left R-module homomorphism, then $g^{-1}(L) \lhd_l^{dense} R$ for any $L_1 \lhd_l^{dense} R$.

Proof. For any $a \in R$ and $0 \neq b \in R$, since L is dense, there exists $x \in R$ such that

$$xa \in L, \quad xb \neq 0.$$

Now for $g(xa) \in R$ and $xb \neq 0$, since L_1 is dense, there exists $y \in R$ such that

$$yg(xa) \in L_1, \quad y(xb) \neq 0.$$

Since g is a left R-module homomorphism, we have $g(yxa) = yg(xa)$. Therefore

$$(yx)a \in g^{-1}(L_1), \quad (yx)b \neq 0,$$

concluding the proof. □

With this lemma, for any $f, g \in \overline{R}$, let L be the domain of f, then the product $f \cdot g$ are defined by the composition

$$f \cdot g \coloneqq f \circ g|_{g^{-1}(L)} \in \overline{R}.$$

Again, replacing the domains by a sufficiently smaller one and one sees that this multiplication is invariant under the equivalence, inducing a well-defined multiplication on $\widetilde{R}$.

The following theorem shows that R can be canonically embedded into $\widetilde{R}^{\text{op}}$ such that $\widetilde{R}^{\text{op}}$ contains a subring that is a ring of fractions of R. Note that $\widetilde{R}^{\text{op}}$ itself may not be the ring of fractions of R.

Theorem 11.4

Suppose that a pair (R, S) (with $S \neq \varnothing$) satisfies the Ore condition. Then the map $R \to \widetilde{R}$ that maps an element $a \in R$ to the right multiplication by a gives an embedding of R into $\widetilde{R}^{op}$. Moreover, via this embedding, the elements of S are invertible in $\widetilde{R}^{op}$.

Proof. Since $S \neq \varnothing$, R has regular elements, hence $R \lhd_l^{\text{dense}} R$, so the map from R to R of right multiplication by an element in R does lie in $\widetilde{R}$. Clearly the assignment $R \to \widetilde{R}^{\text{op}}$ described in the statement is a homomorphism of rings. To show that it is injective, let $a, b \in R$ be two elements such that $lb = la$ for all elements $l \in L \lhd_l^{\text{dense}} R$, we need to show that $a = b$. Let $u \in R$ be a regular element. If $b - a \neq 0$, then $u(b - a) \neq 0$. Since L is dense, $u(b - a) \notin r(Lu^{-1})$. Therefore there exists $x \in Lu^{-1}$ such that $xu(b - a) \neq 0$. Since $xu \in L$, this gives a contradiction to the assumption that $lb = la$ for all elements $l \in L$. Therefore the injectivity is proved.

Given an element $s \in S$, we are going to construct an element in $\widetilde{R}^{\mathrm{op}}$ that is the inverse of the right multiplication by s. Consider the set Rs, it is a left R-module and is dense: for any $a \in R$ and $0 \neq b \in R$, there exists by the Ore condition two elements $a_1 \in R$ and $s_1 \in S$ such that

$$s_1 a = a_1 s \in Rs,$$

and since s_1 is regular, we have $s_1 b \neq 0$. Futhermore, the regularity of s implies that every element in Rs can be written as xs uniquely, hence we can define a map

$$s^{-1} \colon Rs \to R \colon xs \mapsto x,$$

where s^{-1} is just a notation. Clearly s^{-1} is a left R-module homomorphism. Since $Rs \lhd_l^{\mathrm{dense}} R$, $s^{-1} \in \widetilde{R}$ and it is obvious the inverse of right multiplication by s. □

Remark 11.2

By reversing the left and right, one may construct a $\widetilde{R}$ similarly where the elements are equivalence classes of right R-module homomorphisms from dense right ideals to R. The dual argument applies and one obtains an embedding $R \hookrightarrow \widetilde{R}$.

Recall that a ring R is called *left Noetherian* if it satisfies maximal chain condition on left ideals, i.e. for any strictly ascending chain of left ideals of R,

$$L_1 \subsetneq L_2 \subsetneq \cdots,$$

there exists $N \in \mathbb{N}^*$ such that $L_N = L_{N+k}$ for any $k \in \mathbb{N}$.

It is not hard to see that the left Noetherian condition is equivalent to that every left ideal in R is finitely generated as left ideals: supposing the Noetherian condition, for any ideal L we construct a chain of left ideals

$$(a_1) \subsetneq (a_1, a_2) \subsetneq \cdots,$$

where $a_1 \in L$ is chosen arbitrarily, and a_{n+1} is defined inductively by a choice of element $a_{n+1} \in L \setminus (a_1, \cdots, a_n)$ as long as $L \setminus (a_1, \cdots, a_n) \neq \varnothing$. The Noetherian condition thus tells that there must be $N \in \mathbb{N}^*$ such that $L = (a_1, \cdots, a_N)$. Conversely, if every left ideal in R is finitely generated, then for any chain

$$L_1 \subsetneq L_2 \subsetneq \cdots,$$

the union $\bigcup_{i=1}^{\infty} L_i$ is also a left ideal of R, which means that it must be finitely generated, say $\bigcup_{i=1}^{\infty} L_i = (a_1, \cdots, a_m)$. Hence there must exists $n_1, \cdots, n_m$ such that $a_j \in L_{n_j}$ for $j = 1, \cdots, m$. Therefore $\bigcup_{i=1}^{\infty} L_i = \bigcup_{j=1}^{m} L_{n_j}$. Take $N \coloneqq \max_{j=1,\cdots,m}\{n_j\}$ and we see that $L_N = L_{N+k}$ for any $k \in \mathbb{N}$.

Recall that a ring is a *domain* if every nonzero element in it is regular. The following theorem gives us a family of domains that satisfies the Ore condition.

Theorem 11.5

A left Noetherian domain R satisfies the Ore condition relative to $S = R \setminus \{0\}$.

Proof. By rephrasing the Ore condition, it suffices to show that for any two elements $a, b \in R \setminus \{0\}$, we have $Ra \cap Rb \neq (0)$. Consider the ideals defined by

$$L_n \coloneqq Rb + Rba + Rba^2 + \cdots + Rba^n,$$

then we have a chain of left ideals

$$L_0 \subset L_1 \subset L_2 \subset \cdots.$$

Hence the Noetherian condition tells that there exists $n \in \mathbb{N}^*$ such that $L_{n-1} = L_n$. Since $b^2 a^n \in L_n = L_{n-1}$, there exists $x_0, x_1, \cdots, x_{n-1} \in R$ such that

$$b^2 a^n = x_0 b + x_1 ba + \cdots + x_{n-1} ba^{n-1}.$$

Let k be the minimal nonnegative integer with $x_k \neq 0$, then

$$b^2 a^n = x_k ba^k + x_{k+1} ba^{k+1} + \cdots + x_{n-1} ba^{n-1}.$$

Since R is a domain, we then have

$$b^2a^{n-k} = x_kb + x_{k+1}ba + \cdots + x_{n-1}ba^{n-k-1},$$

and $x_kb \neq 0$. Therefore

$$0 \neq x_kb = b^2a^{n-k} - (x_{k+1}ba + \cdots + x_{n-1}ba^{n-k-1}) \in Ra \cap Rb,$$

as desired. □

We will show that for a finitely dimensional Lie algebra L, its universal enveloping algebra $U(L)$ is a Noetherian domain. The domain condition is obvious by previous study. For the Noetherian condition, we need the tool of filtrations.

§12 Filtrations and Deformations

Let A be a unitary F-algebra. A *filtration* of A is a family of subspaces $\{V_i\}_{i=0}^{\infty}$ over F in A such that $\bigcup_{i=0}^{\infty} V_i = A$,

$$F \cdot 1 = V_0 \subset V_1 \subset V_2 \subset \cdots,$$

and $V_iV_j \subset V_{i+j}$.

Define the associated graded algebra, $\mathrm{gr}(A)$, by

$$\mathrm{gr}(A) \coloneqq \bigoplus_{k\in\mathbb{N}} (V_k/V_{k-1}),$$

where $V_{-1} \coloneqq (0)$ by convention. The multiplication on $\mathrm{gr}(A)$ is defined by

$$(a_i + V_{i-1})(b_j + V_{j-1}) = a_ib_j + V_{i+j-1} \in V_{i+j}/V_{i+j-1},$$

for any two elements $a_i + V_{i-1} \in V_i/V_{i-1}$ and $b_j + V_{j-1} \in V_j/V_{j-1}$. With this multiplication, we see that $\mathrm{gr}(A)$ is graded by $\mathrm{gr}(A)_i \coloneqq A_i \coloneqq V_i/V_{i-1}$, since $A_iA_j \subset A_{i+j}$.

We say that A is a *deformation* of $\mathrm{gr}(A)$. The deformation is said, in the sense that if we roughly bring each $V_i \setminus V_{i-1} \subset A$ to $V_i/V_{i-1} \subset \mathrm{gr}(A)$ in the obvious way (which does not necessarily give a homomorphism from A to $\mathrm{gr}(A)$), then for an element $a \in V_i \setminus V_{i-1}$, a difference of a nonzero element in V_{i-1} matters in A, but does not matter in $\mathrm{gr}(A)$. Note however that if A is itself graded as $A = \bigoplus_{i\in\mathbb{N}} A_i$ and $V_n \coloneqq \bigoplus_{i=1}^{n} A_i$, then $A = \mathrm{gr}(A)$ via the identification that $V_n/V_{n-1} = A_n$, which is different from the rough mapping.

Proposition 12.1

Let L be a Lie algebra, then $U(L)$ is a deformation of the polynomial algebra whose number of variables (which may also be infinite) is equal to the dimension of L.

Proof. Let $\{e_i\}_{i\in I}$ be a basis of L, then by theorem 3.6, the set $\{e_{i_1}^{k_1} \cdots e_{i_n}^{k_n} \mid k_j \geq 0,\ i_1 < \cdots < i_n\}$ is a basis of $U(L)$. Define

$$V_m \coloneqq \mathrm{span}(e_{i_1}^{k_1} \cdots e_{i_n}^{k_n} \mid\ i_1 < \cdots < i_n,\ k_1 + \cdots + k_n \leq m),$$

for each $m \in \mathbb{N}$ and the associated $\mathrm{gr}(U(L))$, then we have

$$(e_i + V_0)(e_j + V_0) = (e_j + V_0)(e_i + V_0),$$

for any basis elements e_i and e_j, since $e_ie_j - e_je_i = [e_i, e_j] \in L = V_1$. For any $k_1 + \cdots + k_n \leq m$, since

$$e_{i_1}^{k_1} \cdots e_{i_n}^{k_n} + V_{m-1} = (e_{i_1} + V_0)^{k_1} \cdots (e_{i_n} + V_0)^{k_n},$$

we see that $\mathrm{gr}(U(L))$ is generated by $e_i + V_0$'s. The monomials $e_{i_1}^{k_1} \cdots e_{i_n}^{k_n} + V_{m-1}$ are linearly independent as a consequence of theorem 3.6 (that $e_{i_1}^{k_1} \cdots e_{i_n}^{k_n}$'s form a basis of $U(L)$), therefore we conclude that $\mathrm{gr}(U(L))$ is isomorphic to the polynomial algebra generated by $\{e_i + V_0\}_{i\in I}$. □

Lecture 21

Let A be a unitary F-algebra along with a filtration $V_0 \subset V_1 \subset \cdots$. We now prove the following theorem:

Theorem 12.2

(a) If $\mathrm{gr}(A)$ *is a domain, then* A *is a domain.*

(b) If $\mathrm{gr}(A)$ *is left Noetherian, then* A *is left Noetherian.*

Proof. (a) If a and b are two nonzero elements in A, then there exists i and j such that

$$a \in V_i \setminus V_{i-1}, \quad b \in V_j \setminus V_{j-1}.$$

Hence $a + V_{i-1}$ and $b + V_{j-1}$ are both nonzero in $\mathrm{gr}(A)$. Since $\mathrm{gr}(A)$ is a domain, we have

$$ab + V_{i+j-1} = (a + V_{i-1})(b + V_{j-1}) \neq 0,$$

concluding that $ab \notin V_{i+j-1}$, in particular it is nonzero.

(2) Let L be a left ideal of A, consider the chain

$$V_0 \cap L \subset V_1 \cap L \subset V_2 \cap L \subset \cdots.$$

Write $\Gamma_n \coloneqq V_n \cap L/(V_{n-1} \cap L)$, then $\Gamma \coloneqq \bigoplus_{i=1}^{\infty} \Gamma_n$ is a homogeneous left ideal in $\mathrm{gr}(A)$. Since $\mathrm{gr}(A)$ is Noetherian, Γ is finitely generated, say by $\bar{a}_1, \cdots, \bar{a}_s$. By replacing the generators with their homogeneous components, we may assume that $\bar{a}_i$'s are all homogeneous elements with $\bar{a}_i \in \Gamma_{n_i}$; let $a_1, \cdots a_s \in A$ be representatives of $\bar{a}_1, \cdots, \bar{a}_s$. For any element $a \in L$, we have $a \in V_n \setminus V_{n-1}$ for some n, hence $a + V_{n-1} \cap L \in \Gamma_n \subset \Gamma$. Thus there exists $\bar{x}_1, \cdots, \bar{x}_s \in \mathrm{gr}(A)$ such that

$$a + V_{n-1} \cap L = \bar{x}_1 \bar{a}_1 + \cdots + \bar{x}_s \bar{a}_s.$$

Let $x_1, \cdots, x_s \in A$ be representatives of $\bar{x}_i$'s, then

$$a - x_1 a_1 - \cdots - x_s a_s \in V_{n-1} \cap L.$$

Repeat this procedure with a replaced by $a - x_1 a_1 - \cdots - x_s a_s$ and so on after n steps, we see that L is finitely generated, hence is Noetherian. □

Therefore, since $F[X]$ is Noetherian when $|X| < \infty$ by Hilbert's basis theorem (see for example corollary 2.13 in [Kem11]), we conclude that $U(L)$ is Noetherian when L is finitely dimensional.

Let us give a few more examples of division rings before the end of this section.

Example 12.1

Similarly to $U(L)$, one sees that $\langle x, y \mid yx - xy = 1 \rangle$ is a deformation of $F[x, y]$ and it is also Noetherian, hence it satisfies the Ore condition.

Example 12.2

For any ring R, the quaternion ring over R is $R \cdot 1 \oplus R \cdot i \oplus R \cdot j \oplus R \cdot k$, where the multiplication is defined by $(a \cdot i)(b \cdot j) = ab \cdot ij$ and similarly for k, modulo the relations $ij + ji = 0$, $ik + ki = 0$, $jk + kj = 0$, $i^2 = j^2 = k^2 = -1$ and $ij = k$, $jk = i$, $ki = j$. Note that for any $\alpha = \alpha_0 \cdot 1 + \alpha_1 i + \alpha_2 j + \alpha_3 k$, define $\bar{\alpha} \coloneqq \alpha_0 \cdot 1 - \alpha_1 i - \alpha_2 j - \alpha_3 k$, then $N(\alpha) \coloneqq \alpha\bar{\alpha} = \alpha_0^2 + \alpha_1^2 + \alpha_2^2 + \alpha_3^2 \in R$. Therefore, if R is a division ring, then the quaternion ring over R is also a division ring: for any

element α, we have $\alpha^{-1} = \bar{\alpha}/N(\alpha)$.

Example 12.3

Let F be a field and $\varphi \in \mathrm{Aut}(F)$ be an automorphism of fields with order n, i.e. $\varphi^n = 1$. Consider the subfield fixed by φ, $F^{\langle\varphi\rangle} \coloneqq \{\alpha \in F \mid \varphi(\alpha) = \alpha\}$, then Artin's theorem tells that $[F\colon F^{\langle\varphi\rangle}] = n$. Consider the set of Laurent series

$$\widetilde{F} \coloneqq \left\{ \sum_{k=k_0}^{\infty} \alpha_k t^k \mid k_0 \in \mathbb{Z} \right\},$$

and define the multiplication noncommutatively by

$$t^{-1}\alpha t = \varphi(\alpha),$$

hence

$$t^p \alpha t^q = \varphi^{-p}(\alpha) t^{p+q}.$$

Since $\varphi^n = 1$, we see that t^n commutes in the multiplication. The center of $\widetilde{F}$ is thus all Laurent series generated by t^n over $F^{\langle\varphi\rangle}$. The fact that $\widetilde{F}$ is a division algebra follows from a similar argument of the proof that elements with nonzero constant terms in the ring of power series are invertible. Finally, note that the dimension of $\widetilde{F}$ over its center is n^2.

Remark 12.1

Note that not every automorphism of fields has finite order. For example, consider $\mathbb{C}(t)$. The automorphism on $\mathbb{C}(t)$ induced by $t \mapsto t\xi$ for some $\xi \in \mathbb{C}$ has finite order if and only if ξ is a root of unity.

§13 Ultraproduct

Let X be an infinite set and $\mathcal{P}(X)$ be the set of all subsets of X. A nonempty system[3] of subsets of X, $\varnothing \neq \mathcal{F} \subset \mathcal{P}(X)$, is called a *filter* if

(1) for any $A \in \mathcal{F}$, any subset B of X containing A is also in $\mathcal{F}$.

(2) if $A, B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$.

(3) $\varnothing \notin \mathcal{F}$.

By (1), it is immediate that $X \in \mathcal{F}$.

Example 13.1

For any nonempty subset $A_0 \subset X$, the filter of all subsets containing A_0, $\mathcal{F} \coloneqq \{A \mid A_0 \subset A \subset X\}$, is called *the principal filter* over A_0.

Example 13.2

The set of cofinite subsets, $\mathcal{F} = \{A \subset X \mid X \setminus A \text{ is finite}\}$, is a filter.

Example 13.3

$\mathcal{F} = \{A \subset X \mid \mathrm{Card}(X \setminus A) < \mathrm{Card}(X)\}$, where $\mathrm{Card}(X)$ denotes the cardinality of X.

Lemma 13.1

A subset $S \subset \mathcal{P}(X)$ is extendible to a filter, i.e. there exists $\mathcal{F} \subset \mathcal{P}(X)$ such that $S \subset \mathcal{F}$, if and only if any finite intersection of elements in S is nonempty.

[3] The word "system" here is just a terminology for "set".

Proof. The only if part is obvious. Now suppose that S satisfies the condition that any finite intersection of elements in S is nonempty, then the system

$$\mathcal{F} \coloneqq \{A \subset X \mid \exists A_1, \cdots, A_n \in S \text{ s.t. } A_1 \cap \cdots \cap A_n \subset A\},$$

is a filter containing S. □

Lecture 22

Observe that if we have an ascending chain of filters $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots$, then the union $\bigcup_{i=1}^{\infty} \mathcal{F}_i$ is also a filter. Hence by Zorn's lemma (see for example Theorem 30 in [SV02]), on any set X, every filter is embeddable in a maximal filter. Note that maximal filter may not be unique, and a filter may be embeddable into several different maximal filters.

The following lemma characterize maximal filters explicitly:

Lemma 13.2

A filter $\mathcal{F}$ is maximal if and only if for any $A \in \mathcal{P}(X)$, either $A \in \mathcal{F}$ or $X \setminus A \in \mathcal{F}$.

Proof. For the if part, if there exists a filter $\mathcal{F}'$ such that $\mathcal{F} \subsetneq \mathcal{F}'$, then there exists $A \subset X$ such that $A \in \mathcal{F}'$ but $A \notin \mathcal{F}$. However, by the assumption we then obtain $X \setminus A \in \mathcal{F} \subset \mathcal{F}'$, hence $\varnothing = A \cap (X \setminus A) \in \mathcal{F}'$, a contradiction. Therefore $\mathcal{F}$ must be maximal.

For the only if part, let $\mathcal{F}$ be a maximal filter and $A \subset X$. Let us consider the following two statements

(1) there exists $B \in \mathcal{F}$ such that $A \cap B = \varnothing$;

(2) there exists $B \in \mathcal{F}$ such that $(X \setminus A) \cap B = \varnothing$.

If at least one of these statements is wrong, without loss of generality say (1) is wrong, then the collection $\mathcal{F} \cup \{A\}$ satisfies the assumption in lemma 13.1, hence it extends to a filter containing $\mathcal{F}$. However, since $\mathcal{F}$ is already maximal, the extended filter is still $\mathcal{F}$, hence we obtain $A \in \mathcal{F}$. Similarly if (2) is wrong, then we obtain $X \setminus A \in \mathcal{F}$.

If the above is not the case, i.e. both statements are correct, noticing that the correctness of statement (1) implies that $X \setminus A \in \mathcal{F}$ and (2) implies that $A \in \mathcal{F}$, we are done. □

Definition 13.1

A maximal filter is called an *ultrafilter*.

Fixing an ultrafilter $\mathcal{F}$, we call sets in $\mathcal{F}$ as *large sets*, and sets not in $\mathcal{F}$ as *small sets*. By the above lemma, a set is either large or small, and the complements of large sets are small sets and vice versa.

Consider a family of groups $\{G_x\}_{x \in X}$ and a filter $\mathcal{F}$ on X. On the Cartesian product $\prod_{x \in X} G_x$ we define a equivalence relation by

$$(g_x)_{x \in X} \sim (g'_x)_{x \in X} \quad \Leftrightarrow \quad \{x \in X \mid g_x = g'_x\} \in \mathcal{F}.$$

Lemma 13.3

The equivalence relation $\sim$ is a congruence, i.e. if $(a_x)_{x \in X} \sim (a'_x)_{x \in X}$ and $(b_x)_{x \in X} \sim (b'_x)_{x \in X}$, then $(a_x b_x)_{x \in X} \sim (a'_x b'_x)_{x \in X}$.

Proof. It follows immediately from the following relation:

$$\{x \in X \mid a_x b_x = a'_x b'_x\} \supset \{x \in X \mid a_x = a'_x\} \cap \{x \in X \mid b_x = b'_x\} \in \mathcal{F}.$$ □

We call the quotient of $\prod_{x \in X} G_x$ modulo the relation $\sim$ *the filter product of $\{G_x\}_{x \in X}$ with respect to $\mathcal{F}$*, denoted as $\prod_{x \in X} G_x / \mathcal{F}$. If $\mathcal{F}$ is an ultrafilter, then we call the filter product as an *ultraproduct*.

By replacing groups with rings, algebras or fields, one obtains the definitions of ultraproducts of these objects similarly.

Philosophically, the fact that $\sim$ is a congruence in the above lemma has nothing to do with the operator, but only with the fact that $\mathcal{F}$ is closed under finite intersection. With this observation, the following theorem is natural:

Theorem 13.4. Wos

A formula holds on an ultraproduct $\prod_{x\in X} G_x/\mathcal{F}$ if and only if it holds on a large set of groups (or rings, algebras, fields, etc) G_x.

The word "formula" needs a definition, but we would not bother to do it here. Basically it refers to algebraic formulas which involves only finitely many variables. The following are several examples of such formulas:

- For any x, y, z, we have $(xy)z = x(yz)$.
- For any x, there exists y such that $xy = 1 = yx$.
- (In rings) For any $x \neq 0$, there exists y such that $xy = 1$.

Take the last formula for example. Consider an ultraproduct of fields $\prod_{x\in X} K_x/\mathcal{F}$, clearly the formula holds for all K_x, which is a large set of K_x's. For any element $0 \neq (k_x)_{x\in X} \in \prod_{x\in X} K_x/\mathcal{F}$, the set $\{x \in X \mid k_x = 0\}$ is small, hence its complement $A \coloneqq \{x \in X \mid k_x \neq 0\}$ is large. Consider the element $(k_x^{-1})_{x\in X}$ where k_x^{-1} is exactly the multiplicative inverse of k_x if $x \in A$, and $k_x^{-1} \coloneqq 0$ if $x \in X \setminus A$, then $(k_x)_{x\in X}(k_x^{-1})_{x\in X} = 1$, since the left hand side equals to the multiplicative identity on the large set A. Therefore, an ultraproduct of fields is again a field.

Noticing that the filter product with respect to the trivial filter $\mathcal{F} = \{X\}$ is the Cartesian product, and the Cartesian product of more than two fields is not a field since for example $(1, 0) \in K^2$ is not invertible, one sees that the condition that $\prod_{x\in X} G_x/\mathcal{F}$ is an ultraproduct is necessary in Wos theorem.

For an ultraproduct of fields $K \coloneqq \prod_{x\in X} K_x/\mathcal{F}$, if there exists a prime number p such that $\{x \in X \mid pK_x = (0)\}$ is a large set, then we know that $p = \operatorname{char} K$. If such prime number p does not exist, then we know that $\operatorname{char} K = 0$. Therefore, let P be the set of all prime numbers and $\mathcal{F}$ be an ultrafilter on it, then the ultraproduct

$$\left(\prod_{p\in P} \mathbb{Z}/p\mathbb{Z}\right)/\mathcal{F},$$

has characteristic zero. Hence we see that an ultraproduct of fields with nonzero characteristics may give a field with characteristic zero.

Theorem 13.5. Malcev

Every group is embeddable into an ultraproduct of its finitely generated subgroups.

Proof. Let G be a group and $S_0(G)$ be the set of all nonempty finite subsets of G. For any $x \in S_0(G)$, we consider the filter $S_x = \{y \in S_0(G) \mid x \subset y\}$. For any $x_1, \cdots, x_n \in S_0(G)$, we have

$$S_{x_1} \cap \cdots \cap S_{x_n} \supset S_{x_1\cup\cdots\cup x_n} \neq \varnothing,$$

hence $\{S_x\}_{x\in S_0(G)}$, which is a system of subsets $S_x \subset S_0(G)$, satisfies the condition of lemma 13.1, thus it extends to a filter $\mathcal{F}$ on $S_0(G)$, which may be taken to be maximal by Zorn's lemma. Consider the ultraproduct $\prod_{x\in S_0(G)} \langle x\rangle/\mathcal{F}$ where $\langle x\rangle$ denotes the subgroup of G generated by x. For any element $g \in G$, we have a large set $S_{\{g\}}$ where for any $x \in S_{\{g\}}$ we have $g \in \langle x\rangle$. Hence we can define a map

$$G \to \prod_{x\in S_0(G)} \langle x\rangle/\mathcal{F} \colon g \mapsto (g \in \langle x\rangle)_{x\in S_{\{g\}}},$$

where the element $(g)_{x\in S_{\{g\}}}$ is defined, since an element in an ultraproduct is determined by its value on a large set of components. It is easy to verify that this map is a group homomorphism and is injective since if $g \neq h$, then the image of g and h disagrees on the large set $S_{\{g,h\}}$. □

Similarly, one sees that the above theorem also holds for rings, algebras and fields.

A group G is called *n-linear* if G is embeddable into $\mathrm{GL}(n, K)$ for some field K.

Corollary 13.5.1. Wos

If every finitely generated subgroup of G is n-linear for some fixed integer n, then G is also n-linear.

Proof. Using the embedding in the previous theorem that $G \hookrightarrow \prod_{x\in X} G_x/\mathcal{F}$ where G_x's are finitely generated subgroups of G, since each G_x is embeddable into $\mathrm{GL}(n, K_x)$ for some field K_x, we then obtain an embedding

$$G \hookrightarrow \prod_{x\in X} \mathrm{GL}(n, K_x)/\mathcal{F}.$$

Noticing that $\prod_{x\in X} \mathrm{GL}(n, K_x)/\mathcal{F} \cong \mathrm{GL}(n, \prod_{x\in X} K_x/\mathcal{F})$ via the natural map, we are done. □

Remark 13.1

Let X be a set of axioms, we say that X is *consistent* if it has a model, i.e. there exists a set with some structures that satisfies all axioms in X. If one defines similarly the ultraproduct for models, then, philosophically, the above Malcev's theorem necessarily implies that, if every finite subsystem of X is consistent, then the whole X is consistent.

Limit

Consider a sequence of real numbers $\{a_n\}_{n\in\mathbb{N}}$. Let $\mathcal{F}$ be a filter on $\mathbb{N}$, then a number a is a called the limit of $\{a_n\}_{n\in\mathbb{N}}$ relative to $\mathcal{F}$ if for any $\varepsilon > 0$ there exists $A \in \mathcal{F}$ such that

$$|a - a_n| < \varepsilon, \quad \forall n \in A.$$

Proposition 13.6

For any filter $\mathcal{F}$, the limit of a sequence with respect to $\mathcal{F}$, if exists, must be unique.

Proof. Suppose that a and b are both limits of $\{a_n\}_{n\in\mathbb{N}}$, then for any $\varepsilon > 0$ there exists $A \in \mathcal{F}$ such that $|a - a_n| < \varepsilon/2$ for all $n \in A$ and $B \in \mathcal{F}$ such that $|b - a_n| < \varepsilon/2$ for all $n \in B$. Since $A \cap B$ is nonempty, it follows that

$$|a - b| < \varepsilon,$$

hence we are done. □

Theorem 13.7

Let $\mathcal{F}$ be an ultrafilter, then every bounded sequence of real numbers has a limit with respect to $\mathcal{F}$.

The proof of the theorem involves the following lemma:

Lemma 13.8

In an ultrafilter, if the union of two sets $A \cup B$ is large, then at least one of A and B must be large.

Proof. If both A and B are not large, then $X \setminus A$ and $X \setminus B$ are large. Hence $(X \setminus A) \cap (X \setminus B) = X \setminus (A \cup B)$ is large, contradicting that $A \cup B$ is large. □

Consequently if any finite union of sets is large, then at least one of the sets must be large. Moreover, if the union is disjoint, then exactly one of the sets is large.

Proof of Theorem 13.7. For any bounded sequence $\{a_n\}_{n\in\mathbb{N}}$, there exists $b \in \mathbb{R}$ such that $\{a_n\}_{n\in\mathbb{N}} \subset [-b, b]$. Fix any positive integer m, since

$$\bigcup_{k=-m}^{m-1} \left\{ n \in \mathbb{N} \mid a_n \in \left[\frac{k}{m} b, \frac{k+1}{m} b \right] \right\} = \mathbb{N} \in \mathcal{F},$$

we see that at least one of the sets $\left\{ n \in \mathbb{N} \mid a_n \in \left[\frac{k}{m} b, \frac{k+1}{m} b \right] \right\}$ is large. By replacing $\{a_n\}_{n\in\mathbb{N}}$ with the subsequence $\{a_n\}_{n \in \left\{ n\in\mathbb{N} \mid a_n \in \left[\frac{k}{m} b, \frac{k+1}{m} b \right] \right\}}$ whose index is a large set and $[-b, b]$ with $\left[\frac{k}{m} b, \frac{k+1}{m} b \right]$ and repeat

the procedure above, one shrinks the interval by $\frac{1}{m}$ each time and sees the existence of the limit from the completeness of $\mathbb{R}$. $\square$

Consider a family of metric spaces $\{(M_n, d_n)\}_{n\in\mathbb{N}}$ where the diameters of metric spaces are bounded uniformly and an ultrafilter $\mathcal{F}$ on $\mathbb{N}$. The previous result allows us to define a premetric on the set $\prod_{n\in\mathbb{N}} M_n/\mathcal{F}$ by setting that

$$d\left((a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}}\right) \coloneqq \lim_{\mathcal{F}} d_n(a_n, b_n).$$

Note that d is only a premetric instead of a metric, since two distinct elements may have distance 0 with respect to d. For instance, consider the ultraproduct $[0,1]^{\mathbb{N}}/\mathcal{F}$, where $\mathcal{F}$ is an ultrafilter containing all cofinite subsets of $\mathbb{N}$. The distance between the element $\varepsilon \coloneqq (1, \frac{1}{2}, \frac{1}{3}, \cdots)$ and the origin is zero, but clearly $\varepsilon \neq 0$. Furthermore, one may define a natural order on $[0,1]^{\mathbb{N}}/\mathcal{F}$, since for any two elements $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ we have

$$\{n \in \mathbb{N} \mid a_n > b_n\} \sqcup \{n \in \mathbb{N} \mid a_n = b_n\} \sqcup \{n \in \mathbb{N} \mid a_n < b_n\} = \mathbb{N}.$$

With respect to this order, we have

$$0 < \varepsilon < \frac{1}{n},$$

for any $n \in \mathbb{N}$.

For any finitely generated group G, choose a finite set of generators and consider its Cayley graph $\mathrm{Cay}(G)$, which is equipped with the natural metric d of graphs. Consider the balls B_n of radius n centered at the identity $1_G \in G$, then the metric spaces $(B_n, d/n)$ have diameters uniformly bounded by 1. Therefore it makes sense to consider an ultraproduct $\prod_{n\in\mathbb{N}} B_n/\mathcal{F}$ equipped with the premetric. Intuitively, as one goes through the index n increasingly, the components of elements in $\prod_{n\in\mathbb{N}} B_n/\mathcal{F}$ become denser and denser, and the overall ultraproduct space is so dense that it is manifold-ish.

Lecture 23

An Elegant Proof of Ax-Grothendieck Theorem

Suppose that $\prod_{x\in X} G_x/\mathcal{F} =: G$ is an ultraproduct (of groups, rings, fields, etc) where for each $x \in X$ there exists an infinite set Y_x and an ultrafilter $\mathcal{F}_x$ on Y_x such that $G_x \cong \prod_{y\in Y_x} G_{xy}/\mathcal{F}_x$ for some G_{xy}'s. Take the disjoint union $Y \coloneqq \bigsqcup_{x\in X} Y_x$ and define an ultrafilter $\widetilde{\mathcal{F}}$ on Y by extending the set $\{\bigcup_{x\in X} U_x \subset Y \mid (U_x)_{x\in X} \in \prod_{x\in X} \mathcal{F}_x\}$, which is extendible since its element-wise intersection with any Y_x is a filter.

Exercise 13.1

Show that $G \cong \prod_{y\in Y} G_{xy}/\widetilde{\mathcal{F}}$.

This tells that the ultraproduct is composable, i.e.

$$\prod_{x\in X}\left(\prod_{y\in Y_x} G_{xy}/\mathcal{F}_x\right)/\mathcal{F} \cong \prod_{y\in\bigsqcup_{x\in X} Y_x} G_{xy}/\widetilde{\mathcal{F}}.$$

With the result of this exercise, we prove the following proposition:

Proposition 13.9

Every commutative domain embeds into an ultraproduct of finite fields.

Firstly, let us briefly recall Hilbert's Nullstellensatz.

Theorem 13.10. Nullstellensatz

(1) *Let F be an algebraically closed field. Let $S = \{f_1(x_1,\cdots,x_n),\cdots,f_k(x_1,\cdots,x_n)\}$ be a finite set of polynomials over F. A point $\alpha = (\alpha_1,\cdots,\alpha_n) \in F^n$ is said to be a zero of S if $f_i(\alpha_1,\cdots,\alpha_n) = 0$ for all $i = 1,\cdots,k$. For a polynomial $g \in F[x_1,\cdots,x_n]$, if every zero of S is also a zero of g, then $g^m = \sum_{i=1}^k f_i h_i$ for some positive integer m and some polynomials $h_i \in F[x_1,\cdots,x_n]$.*

(2) *Every finitely generated commutative domain is a subdirect product of fields. Equivalently, this means that if A is a finitely generated commutative domain, then there exists fields K_i indexed by $i \in I$ and ring homomorphisms $\varphi_i\colon A \to K_i$ for each $i \in I$ such that $\bigcap_{i\in I} \ker\varphi_i = (0)$.*

(3) *A field that is finitely generated as a ring is a finite field.*

Note that in statement (2), by restricting the codomains of φ_i to the subring generated by $\varphi_i(A)$, we may assume that each K_i is finitely generated, hence is finite by statement (3).

With the help of Nullstellensatz, we are able to prove the following lemma.

Lemma 13.11

Let A be a finitely generated commutative domain, then A embeds into an ultraproduct of finite fields.

Proof. Let $\varphi_i\colon A \to K_i$ be the homomorphisms in statement (2) of Nullstellensatz. By adding extra repeated φ_i's, we may assume that I is an infinite set. For any $0 \neq a \in A$, consider the subset of I,

$$I_a \coloneqq \{i \in I \mid \varphi_i(a) \neq 0\},$$

then it is never empty since $\bigcap_{i\in I} \ker \varphi_i = (0)$, and for any finitely many nonzero elements $a_1, \cdots, a_n \in A$,

$$I_{a_1} \cap \cdots \cap I_{a_n} \supset I_{a_1 \cdots a_n} \neq \varnothing,$$

since $\varphi_i(a_1 \cdots a_n) = \varphi_i(a_1) \cdots \varphi_i(a_n)$. Therefore $\{I_a\}_{a\in A}$ extends to an ultrafilter $\mathcal{F}$ on I. Now consider the natural homomorphism $A \to \prod_{i\in I} K_i/\mathcal{F}\colon a \mapsto (\varphi_i(a))_{i\in I}/\mathcal{F}$, it is injective because if $(\varphi_i(a))_{i\in I}/\mathcal{F} = 0$, then $\varphi_i(a) = 0$ for a large set I'. If $a \neq 0$, then I_a is also a large set, so is $I_a \cap I'$, but $\varphi_i(a) \neq 0$ for $i \in I_a \cap I' \subset I_a$ by definition, a contradiction. □

By theorem 13.5 for rings, since every commutative domain A is embeddable into an ultraproduct $\prod_{j\in J} A_j/\mathcal{F}$ of finitely generated commutative domains A_j's, and each A_j is embeddable into an ultraproduct $\prod_{i\in I_j} K_{ji}/\mathcal{F}_j$ of finite fields K_{ji}'s, we obtain by exercise 13.1 that,

$$A \hookrightarrow \prod_{j\in J} A_j/\mathcal{F} \hookrightarrow \prod_{j\in J} \left(\prod_{i\in I_j} K_{ji}/\mathcal{F}_j \right) /\mathcal{F} \cong \prod_{i\in \bigsqcup_{j\in J} I_j} K_{ji}/\widetilde{\mathcal{F}},$$

concluding the proof of proposition 13.9.

With proposition 13.9, we can now give an elegant proof of Ax-Grothendieck theorem, as a piece of art, as the end of this course.

Let F be a field. Let n be an arbitrary positive integer and $P\colon F^n \to F^n$ be a polynomial map, i.e. if we write $P(x_1, \cdots, x_n) = (P_1(x_1, \cdots, x_n), \cdots, P_n(x_1, \cdots, x_n))$, then each P_i is a polynomial over F.

Theorem 13.12. Ax-Grothendieck

If F is either finite or algebraically closed, then the condition that P is injective implies that P is surjective.

Proof. It is obvious for the case where F is finite. Suppose that F is algebraically closed. For P being injective, it means that every zero point $(x, y) = (x_1, \cdots, x_n, y_1, \cdots, y_n) \in F^{2n}$ of the set $\{P_i(x) - P_i(y) \mid i = 1, \cdots, n\}$ satisfies that $x_i = y_i$ for all $i = 1, \cdots, n$. In other words, every zero point of $\{P_i(x) - P_i(y) \mid i = 1, \cdots, n\}$ is a zero point of $x_i - y_i$ for an arbitrary i. Hence by statement (1) of Nullstellensatz, there exists positive integer k_i and polynomial h_{ij}'s such that

$$(x_i - y_i)^{k_i} = \sum_{j=1}^{n} (P_j(x) - P_j(y)) h_{ij}(x, y),$$

for each $i = 1, \cdots, n$, as polynomials in $F[x_1, \cdots, x_n, y_1, \cdots, y_n]$. Conversely, since the equality forces $x = y$ whenever $P(x) = P(y)$, the equation above is equivalent to the injectivity of P. Note that the converse does not rely on the algebraically closedness of F.

Suppose that P is not surjective, which means exactly that there exists $\alpha = (\alpha_1, \cdots, \alpha_n) \in F^n$ such that the system $\{P_i(x) - \alpha_i \mid i = 1, \cdots, n\}$ does not have a zero point. By statement (1) of Nullstellensatz, there exists polynomials $g_i \in F[x_1, \cdots, x_n]$ such that

$$1 = \sum_{i=1}^{n} (P_i(x) - \alpha_i) g_i(x).$$

Note also that this equality is equivalent to the non-surjectivity of P, where the converse does not rely on the algebraically closedness of F.

Therefore, if P is injective and not surjective, then we obtain two systems of equalities of polynomials over F that are equivalent to the injectivity and non-surjectivity respectively. By proposition 13.9, F is embeddable into an ultraproduct of finite fields, say

$$F \hookrightarrow \prod_{i\in I} K_i/\mathcal{F}.$$

Bring via this embedding the polynomials over F to polynomials over $\prod_{i\in I} K_i/\mathcal{F}$, we thus obtain two systems of equalities of polynomials over $\prod_{i\in I} K_i/\mathcal{F}$. By theorem 13.4 (or one can do it by hand), there exists a large set of fields K_i's on which (the projection of) the two systems of equalities both hold, in particular the two systems both hold for at least one certain K_i. Therefore, if we project the coefficients of P_i's to the K_i summand, then we obtain a polynomial map $K_i^n \to K_i^n$ that is injective and not surjective, guaranteed by these two systems of equalities. However, K_i is a finite field, while a contradiction. □

Bibliography

[Hun80] Thomas W. Hungerford. *Algebra*. Vol. 73. Graduate Texts in Mathematics. Reprint of the 1974 original. Springer-Verlag, New York-Berlin, 1980, pp. xxiii+502. ISBN: 0-387-90518-9.

[Lan02] Serge Lang. *Algebra*. third. Vol. 211. Graduate Texts in Mathematics. Springer-Verlag, New York, 2002, pp. xvi+914. ISBN: 0-387-95385-X. DOI: 10.1007/978-1-4613-0041-0. URL: https://doi.org/10.1007/978-1-4613-0041-0.

[SV02] A. Shen and N. K. Vereshchagin. *Basic set theory*. Vol. 17. Student Mathematical Library. Translated from the 1999 Russian edition by Shen. American Mathematical Society, Providence, RI, 2002, pp. viii+116. ISBN: 0-8218-2731-6. DOI: 10.1090/stml/017. URL: https://doi.org/10.1090/stml/017.

[Alu09] Paolo Aluffi. *Algebra: chapter 0*. Vol. 104. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2009, pp. xx+713. ISBN: 978-0-8218-4781-7. DOI: 10.1090/gsm/104. URL: https://doi.org/10.1090/gsm/104.

[Kem11] Gregor Kemper. *A Course in Commutative Algebra*. Graduate Texts in Mathematics. Springer, 2011. ISBN: 9783642035449.