Chapter III. Universal Algebra of a Lie Algebra

1. Definition

Let k be a commutative ring and let $\mathfrak{g}$ be a Lie algebra over k.

Definition 1.1. A *universal algebra* of $\mathfrak{g}$ is a map $\varepsilon : \mathfrak{g} \to U\mathfrak{g}$, where $U\mathfrak{g}$ is an associative algebra, with a unit satisfying the following properties:

1). ε is a Lie algebra homomorphism,

$$(\text{i.e., } \varepsilon \text{ is } k\text{-linear and } \varepsilon[x,y] = \varepsilon x \cdot \varepsilon y - \varepsilon y \cdot \varepsilon x).$$

2). If A is any associative algebra with a unit and $\alpha : \mathfrak{g} \to A$ is any Lie algebra homomorphism, there is a unique homomorphism of associative algebras $\varphi : U\mathfrak{g} \to A$ such that the diagram

$$\begin{array}{ccc} \mathfrak{g} & \xrightarrow{\varepsilon} & U\mathfrak{g} \\ {\scriptstyle\alpha}\downarrow & \swarrow{\scriptstyle\varphi} & \\ A & & \end{array}$$

is commutative [i.e., there is an isomorphism

$$\mathrm{Hom}_{\mathrm{Lie}}(\mathfrak{g}, LA) \cong \mathrm{Hom}_{\mathrm{Ass}}(U\mathfrak{g}, A)$$

where LA is the Lie algebra associated to A, cf. Chap. I, example (iii).]

It is trivial that $U\mathfrak{g}$, if it exists, is unique (up to a unique isomorphism). To prove its existence, we use the *tensor algebra* $T\mathfrak{g}$ of $\mathfrak{g}$, i.e., $T\mathfrak{g} = \sum_{n=0}^{\infty} T^n\mathfrak{g}$, where $T^n\mathfrak{g} = \mathfrak{g} \otimes \cdots \otimes \mathfrak{g} = \bigotimes^n \mathfrak{g}$ for $n \geq 0$. For any associative algebra A with a unit, one has: $\mathrm{Hom}_{\mathrm{Mod}}(\mathfrak{g}, A) = \mathrm{Hom}_{\mathrm{Ass}}(T\mathfrak{g}, A)$.

Now let I be the two-sided ideal of $T\mathfrak{g}$ generated by the elements of the form $[x,y] - x \otimes y + y \otimes x, \quad x, y \in \mathfrak{g}$.

Take $U\mathfrak{g} = T\mathfrak{g}/I$, then we have:

Theorem 1.2. *Let $\varepsilon : \mathfrak{g} \to U\mathfrak{g}$ be the composition $\mathfrak{g} \to T^1\mathfrak{g} \to T\mathfrak{g} \to U\mathfrak{g}$. Then the pair $(U\mathfrak{g}, \varepsilon)$ is a universal algebra of $\mathfrak{g}$.*

In fact, let α be a Lie homomorphism of $\mathfrak{g}$ into an associative algebra A. Since α is k-linear, it extends to a unique homomorphism $\psi : T\mathfrak{g} \to A$. It is clear that $\psi(I) = 0$, hence ψ defines $\varphi : U\mathfrak{g} \to A$, and we have checked the universal property of $U\mathfrak{g}$.

Remark. Let E be a $\mathfrak{g}$-module (i.e., a k-module with a bilinear product $\mathfrak{g} \times E \to E$ such that $[x,y]e = x(ye) - y(x \cdot e)$ for $x, y \in \mathfrak{g}$, $e \in E$). The map $\mathfrak{g} \to \mathrm{End}(E,E)$ which defines the module structure of E is a Lie homomorphism. Hence it extends to an algebra homomorphism $U\mathfrak{g} \to \mathrm{End}(E,E)$ and E becomes a *$U\mathfrak{g}$-left-module*. It is easy to check that one obtains in this

way an *isomorphism* of the category of $\mathfrak{g}$-modules onto the category of $U\mathfrak{g}$-*left-modules*.

Exercise (Bergman). Prove that $U\mathfrak{g} = k \iff \mathfrak{g} = 0$. (Hint: use the adjoint representation.)

2. Functorial properties

1). If $\mathfrak{g} = \varinjlim \mathfrak{g}_\alpha$, then $U\mathfrak{g} = \varinjlim U\mathfrak{g}_\alpha$.

2). If $\mathfrak{g} = \mathfrak{g}_1 \times \mathfrak{g}_2$, where $\mathfrak{g}_1$ and $\mathfrak{g}_2$ commute, then $U\mathfrak{g} = U\mathfrak{g}_1 \otimes U\mathfrak{g}_2$.

3). Let k' be an extension of k and let $\mathfrak{g}' = \mathfrak{g} \otimes_k k'$, then $U\mathfrak{g}' = U\mathfrak{g} \otimes_k k'$.

Proof of 2). Consider the homomorphisms $\varepsilon_i : \mathfrak{g}_i \to U\mathfrak{g}_i$, $i = 1, 2$, $f : \mathfrak{g} \to U\mathfrak{g}_1 \otimes U\mathfrak{g}_2$ given by $f(x) = \varepsilon(x_1) \otimes 1 + 1 \otimes \varepsilon(x_2)$ where $x = x_1 + x_2$ with $x_1 \in \mathfrak{g}_1$, $x_2 \in \mathfrak{g}_2$. The map f is a Lie algebra homomorphism since $\mathfrak{g}_1$ commutes with $\mathfrak{g}_2$. Hence f induces an associative algebra homomorphism $\psi : U\mathfrak{g} \to U\mathfrak{g}_1 \otimes U\mathfrak{g}_2$.

On the other hand we have the homomorphisms $\mathfrak{g}_i \to \mathfrak{g} \to U\mathfrak{g}$, $i = 1, 2$, which induce homomorphisms $\varphi_1 : U\mathfrak{g}_i \to U\mathfrak{g}$ and since $\mathfrak{g}_1$ commutes with $\mathfrak{g}_2$ we have that $\varphi_1(x_1)\varphi_2(x_2) = \varphi_2(x_2)\varphi_1(x_1)$ for all $x_1 \in \mathfrak{g}_1$, $x_2 \in \mathfrak{g}_2$.

Finally take $\varphi : U\mathfrak{g}_1 \otimes U\mathfrak{g}_2 \to U\mathfrak{g}$ given by $\varphi(x_1 \otimes x_2) = \varphi_1(x_1)\varphi_2(x_2)$, then we have $\psi \circ \varphi = \mathrm{id}$ and $\varphi \circ \psi = \mathrm{id}$.

The proof of 1) and 3) are similar.

3. Symmetric algebra of a module

Let $\mathfrak{g}$ be a k-module and define $[x, y] = 0$ for all $x, y \in \mathfrak{g}$. In this case, the universal algebra $U\mathfrak{g}$ of $\mathfrak{g}$ is called the *symmetric algebra* of the k-module $\mathfrak{g}$ and it is denoted by $S\mathfrak{g}$.

We can define $S\mathfrak{g}$ as the largest commutative quotient of $T\mathfrak{g}$, i.e., $S\mathfrak{g} = \sum_{n=0}^{\infty} S^n\mathfrak{g}$ where $S^n\mathfrak{g} = (\bigotimes^n \mathfrak{g})/I$ where I is generated by the elements of the form $a - \sigma a$ where σ is a permutation of $[1, n]$, and $a \in \bigotimes^n \mathfrak{g}$.

We will consider the case where $\mathfrak{g}$ is a free k-module with basis $(e_i)_{i \in I}$.

Let $\varepsilon : \mathfrak{g} \to k[(X_i)_{i \in I}]$ be the homomorphism given by $\varepsilon(e_i) = X_i$ where $k[(X_i)_{i \in I}]$ is the polynomial ring in the indeterminates X_i, $i \in I$. Then $(\varepsilon, k[(X_i)_{i \in I}])$ has the universal property of 1.1, i.e., ε is a k-linear map such that $\varepsilon(x)\varepsilon(y) = \varepsilon(y)\varepsilon(x)$ and if $f : \mathfrak{g} \to A$ is a k-linear map with $f(x)f(y) = f(y)f(x)$ for all $x, y \in \mathfrak{g}$ where A is an associative algebra, then there exists an associative algebra homomorphism $f^* : k[(X_i)] \to A$ such that $f^* \circ \varepsilon = f$. In fact if $P(x_i) \in k[(X_i)]$ then $f^*(P) = P(f(e_i))$. This shows that we can identify $S\mathfrak{g}$ with the polynomial algebra $k[(X_i)_{i \in I}]$.

If we assume that I is totally ordered, then $S\mathfrak{g}$ has for basis the set of monomials $e_{i_1} \cdots e_{i_n}$, $i_1 \leq i_2 \leq \cdots \leq i_n$, $n \geq 0$.

4. Filtration of $U\mathfrak{g}$

Let $\mathfrak{g}$ be a Lie algebra over k, and let $U\mathfrak{g}$ be the universal algebra of $\mathfrak{g}$. We define a filtration of $U\mathfrak{g}$ as follows: Let $U_n\mathfrak{g}$ be the submodule of $U\mathfrak{g}$ generated by the products $\varepsilon(x_1)\cdots\varepsilon(x_m)$, $m \leq n$, where $x_i \in \mathfrak{g}$. We have

$$U_0\mathfrak{g} = k$$
$$U_1\mathfrak{g} = k \oplus \varepsilon(\mathfrak{g})$$

and $U_0\mathfrak{g} \subset U_1\mathfrak{g} \subset \cdots \subset U_n\mathfrak{g} \subset U_{n+1}\mathfrak{g} \subset \cdots$.

Now we define $\operatorname{gr} U\mathfrak{g} = \sum_{n=0}^{\infty} \operatorname{gr}_n U\mathfrak{g}$, where $\operatorname{gr}_n U\mathfrak{g} = U_n\mathfrak{g}/U_{n-1}\mathfrak{g}$.

The map $U_p\mathfrak{g} \times U_q\mathfrak{g} \to U_{p+q}\mathfrak{g}$ given by $(a,b) \mapsto ab$ defines, by passage to quotient, a bilinear map

$$\operatorname{gr}_p U\mathfrak{g} \times \operatorname{gr}_q U\mathfrak{g} \to \operatorname{gr}_{p+q} U\mathfrak{g} \ .$$

We then obtain a structure of *graded algebra* on $\operatorname{gr} U\mathfrak{g}$; with this structure $\operatorname{gr} U\mathfrak{g}$ is called the *graded algebra* associated to $U\mathfrak{g}$. It is associative and has a unit.

Proposition 4.1. *The algebra* $\operatorname{gr} U\mathfrak{g}$ *is generated by the image of* $\mathfrak{g}$ *under the map induced by* $\varepsilon : \mathfrak{g} \to U\mathfrak{g}$.

Proof. Let $\alpha \in \operatorname{gr}_n U\mathfrak{g}$ and let $a \in U_n\mathfrak{g}$ be a representative of α, i.e., $\bar{a} = \alpha$. Now, we have $a = \sum_{m_\mu \leq n} \lambda_\mu \varepsilon(x_1^{(\mu)})\cdots\varepsilon(x_{m_\mu}^{(\mu)})$. Thus we have

$$\alpha = \sum_{m_\mu = n} \lambda_\mu \overline{\varepsilon(x_1^{(\mu)})}\cdots\overline{\varepsilon(x_{m_\mu}^{(\mu)})} \qquad \text{q.e.d.}$$

Theorem 4.2. *The algebra* $\operatorname{gr} U\mathfrak{g}$ *is commutative.*

Proof. Using 4.1 it is enough to prove that $\overline{\varepsilon(x)}$, $\overline{\varepsilon(y)}$ commute in $\operatorname{gr}_2 U\mathfrak{g}$ for all $x, y \in \mathfrak{g}$.

Since ε is a Lie algebra homomorphism we have

$$\varepsilon(x)\varepsilon(y) - \varepsilon(y)\varepsilon(x) = \varepsilon([x,y]) \ ,$$

but $\varepsilon([x,y]) \in U_1\mathfrak{g}$ so $\varepsilon(x)\varepsilon(y) \equiv \varepsilon(y)\varepsilon(x) \bmod U_1\mathfrak{g}$. Therefore

$$\overline{\varepsilon(x)}\,\overline{\varepsilon(y)} = \overline{\varepsilon(y)}\,\overline{\varepsilon(x)} \ .$$

It follows from Theorem 4.2 that the canonical map $\mathfrak{g} \to \operatorname{gr} U\mathfrak{g}$ extends to a homomorphism

$$\iota : S\mathfrak{g} \to \operatorname{gr} U\mathfrak{g}$$

where $S\mathfrak{g}$ is the symmetric algebra of $\mathfrak{g}$ (cf. III.3).

Since $\operatorname{gr} U\mathfrak{g}$ is generated by the image of $\mathfrak{g}$, ι *is surjective.*

Theorem 4.3 (Poincaré-Birkhoff-Witt). *If $\mathfrak{g}$ is a k-free module, then ι is an isomorphism.*

In order to prove the theorem we will prove first two lemmas.

Let $(x_i)_{i\in I}$ be a basis of $\mathfrak{g}$ and choose a total order in I.

Lemma 4.4. *The family of monomials $\varepsilon(x_{i_1})\cdots\varepsilon(x_{i_m})$, $i_1 \le \cdots \le i_m$, $m \le n$, generate $U^n\mathfrak{g}$ (as a k-module).*

Proof. We proceed by induction with respect to n.

For $n = 0$ the statement is trivial.

Suppose now $n > 0$ and take $a \in U^n\mathfrak{g}$. Then its image $\bar{a} \in \mathrm{gr}^n U\mathfrak{g}$ is a polynomial of degree n in the $\overline{\varepsilon(x_i)}$, but this implies a is a linear combination of products $\varepsilon(x_{i_1})\cdots\varepsilon(x_{i_n})$, $i_1 \le \cdots \le i_n$ plus an element $a_1 \in U^{n-1}\mathfrak{g}$.

By the hypothesis of induction a_1 is a linear combination of products $\varepsilon(x_{i_1})\cdots\varepsilon(x_{i_m})$, $i_1 \le \cdots \le i_m$, $m < n$. q.e.d.

Lemma 4.5. *The following statement is equivalent to* 4.3:

The family of monomials $\varepsilon(x_{i_1})\cdots\varepsilon(x_{i_n})$, $i_1 \le \cdots \le i_n$, $n \ge 0$ is a basis of $U\mathfrak{g}$.

For $M = (i_1, \ldots, i_m)$ with $i_1 \le i_2 \le \cdots \le i_m$, write

$$x_M = \varepsilon(x_{i_1})\cdots\varepsilon(x_{i_m}) ,$$

and denote the *length* of M by $\ell(M) = m$. For each $n \ge 0$ the elements x_M with $\ell(M) = n$ lie in $U_n\mathfrak{g}$, and their images $\bar{x}_M$ in $\mathrm{gr}_n U\mathfrak{g} = U_n\mathfrak{g}/U_{n-1}\mathfrak{g}$ are the images, under the map $\iota : S^n\mathfrak{g} \to \mathrm{gr}_n U\mathfrak{g}$, of the monomial basis elements of $S^n\mathfrak{g}$. Thus, the injectivity of ι is equivalent to the non-existence of a relation

$$\sum_{\ell(M)=n} c_M x_M \equiv 0 \pmod{U_{n-1}\mathfrak{g}}$$

with some $c_M \neq 0$. By Lemma 4.4 this is the same as the non-existence of a relation

$$\sum_{\ell(M)=n} c_M x_M = \sum_{\ell(M)<n} c_M x_M ,$$

with some c_M on the left not zero. But any non-trivial k-linear dependence relation among the x_M can be put in the latter form. Hence Lemma 4.5 is true, and we can now proceed to prove Theorem 4.3 in the new form.

To do so we can (and will) assume that I is *well-ordered*. Let V be the free k-module with basis $\{z_M\}$ where M runs through the set of all sequences $(i_1, \ldots, i_n)$ with $n \ge 0$ and $i_1 \le i_2 \cdots \le i_n$ as above. If $i \in I$ and $M = (i_1, \ldots, i_n)$, we define $i \le M \iff i \le i_1$, in which case we introduce the notation $iM = (i, i_1, \ldots, i_n)$.

Main lemma. *We can make V into a $\mathfrak{g}$-module in such a way that $x_i Z_M = Z_{iM}$ whenever $i \le M$.*

We shall first define a k-bilinear map $(x,v) \mapsto xv$ of $\mathfrak{g} \times V$ into V, and will then prove that it makes V a $\mathfrak{g}$-module, that is, satisfies

$$xyv - yxv = [x,y]v \ , \qquad \text{for } x,y \in \mathfrak{g}, \text{ and } v \in V. \tag{1}$$

To define xv it suffices to define $x_i Z_M$ for all i and M, and to define $x_i Z_M$ we may assume by induction that $x_j Z_N$ has been defined for all $j \in I$ when $\ell(N) < \ell(M)$ and for $j < i$ when $\ell(N) = \ell(M)$. Moreover we assume that this has been done in such a way that the following holds:

$(*)$ $\quad x_j Z_N$ is a k-linear combination of Z_L's with $\ell(L) \leq \ell(N) + 1$.

We then put

$$x_i Z_M = \begin{cases} Z_{iM} & \text{, if } i \leq M \\ x_j(x_i Z_N) + [x_i, x_j] Z_N & \text{, if } M = jN \text{ with } i > j. \end{cases} \tag{2}$$

This makes sense because, in the second case, $x_i Z_N$ is already defined as a linear combination of Z_L's with $\ell(L) \leq \ell(N)+1 = \ell(M)$, and $[x_i, x_j]$ is a linear combination of x_k. Moreover the condition $(*)$ holds with j and N replaced by i and M.

To check (1) it suffices, by linearity, to show

$$x_i x_j Z_N - x_j x_i Z_N = [x_i, x_j] Z_N \tag{1'}$$

for all i, j and N. Since both sides are skew symmetric and vanish when $i = j$, we may assume $i > j$. If $j \leq N$, then $x_j Z_N = Z_{jN}$ and (1') follows from the second case of our inductive definition (2) above. There remains the case $N = kL$, with $i > j > k$, when (1') becomes

$$x_i x_j x_k Z_L - x_j x_i x_k Z_L = [x_i, x_j] x_k Z_L \ . \tag{ijk}$$

By induction on $\inf(i,j)$, we know this equation does hold if we permute ijk cyclically, that is the equations (jki) and (kij) are correct. On the other hand, by induction on $\ell(N)$ we can assume $xyZ_L = yxZ_L + [x,y]Z_L$ for all $x, y \in \mathfrak{g}$. Thus the right hand side of (ijk) can be rewritten:

$$\begin{aligned} [x_i, x_j] x_k Z_L &= x_k [x_i, x_j] Z_L + [[x_i, x_j], x_k] Z_L \\ &= x_k x_i x_j Z_L - x_k x_j x_i Z_L + [[x_i, x_j], x_k] Z_L \ . \end{aligned}$$

If we substitute this on the right side of (ijk) and then add the three equations $(ijk) + (jki) + (kij)$ we get an equation of the form

$$\sum = \sum + \mathrm{Jac}(x_i, x_j, x_k) Z_L \ .$$

Hence, (ijk) is true, and our main lemma is proved.

Since V is a $\mathfrak{g}$-module, it is also a $U\mathfrak{g}$-left module, cf. Remark at the end of III.1.

In particular we have in V the element $Z_\emptyset$ where $\emptyset$ is the empty set. For all M we have $x_M Z_\emptyset = Z_M$. We will prove this by induction on $\ell(M)$. If

$\ell(M) = 0$ then it is clear because $x_M = 1$. If $\ell(M) > 0$ we write $M = iN$, $i \leq N$. Then $x_M = x_i x_N$ and $x_M Z_\emptyset = x_i x_N Z_\emptyset = x_i Z_N = Z_{iN} = Z_M$.

Finally, suppose we have $\sum c_M x_M = 0$, then

$$0 = \sum c_M x_M Z_\emptyset = \sum c_M Z_M \ ,$$

but this implies $c_M = 0$ for all M. q.e.d.

Corollary 1. *If $\mathfrak{g}$ is a free k-module then $\varepsilon : \mathfrak{g} \to U\mathfrak{g}$ is injective.*

In fact, in this case $\mathfrak{g} \cong \mathrm{gr}_1 U\mathfrak{g}$.

Corollary 2. *Let $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$ where $\mathfrak{g}_1$ and $\mathfrak{g}_2$ are subalgebras of $\mathfrak{g}$ and are free k-modules. Then the map $U\mathfrak{g}_1 \otimes U\mathfrak{g}_2 \to U\mathfrak{g}$ given by $u_1 \otimes u_2 \mapsto u_1 u_2$ is a k-linear isomorphism.*

Proof. Let $(x_i)_{i\in I}$, $(y_j)_{j\in J}$ be a basis of $\mathfrak{g}_1$ and $\mathfrak{g}_2$ respectively, then $\{(x_i),(x_j)\}$ is a basis of $\mathfrak{g}$. Take a total order in $I \cup J$ such that every element of I is less than every element of J. Applying 4.5 we have that the families of monomials $\{\varepsilon(x_{i_1}) \cdots \varepsilon(x_{i_n})\}$, $\{\varepsilon(y_{j_1}) \cdots \varepsilon(y_{j_m})\}$ and $\{\varepsilon(x_{i_1}) \cdots \varepsilon(x_{i_n})\varepsilon(y_{j_1}) \cdots \varepsilon(y_{j_m})\}$ for $i_1 \leq \cdots \leq i_n$ and $j_1 \leq \cdots \leq j_m$ are basis of $U\mathfrak{g}_1$, $U\mathfrak{g}_2$ and $U\mathfrak{g}$ respectively. Thus the map $U\mathfrak{g}_1 \otimes U\mathfrak{g}_2 \to U\mathfrak{g}$ given by $u_1 \otimes u_2 \mapsto u_1 u_2$ is a bijection on the basis of $U\mathfrak{g}_1 \otimes U\mathfrak{g}_2$ and $U\mathfrak{g}$. q.e.d.

Notice that in this case we have also induced an isomorphism

$$\mathrm{gr}\, U\mathfrak{g}_1 \otimes \mathrm{gr}\, U\mathfrak{g}_2 \xrightarrow{\approx} \mathrm{gr}\, U\mathfrak{g}$$

because $\mathrm{gr}\, U\mathfrak{g}_i = S\mathfrak{g}_i$ and $\mathrm{gr}\, U\mathfrak{g} = S\mathfrak{g} \simeq S\mathfrak{g}_1 \otimes S\mathfrak{g}_2$.

5. Diagonal map

Let $\mathfrak{g}$ be a Lie algebra over k and suppose $\mathfrak{g}$ is free as a k-module.

Definition 5.1. The Lie algebra homomorphism $\Delta : \mathfrak{g} \to \mathfrak{g} \times \mathfrak{g}$ given by $x \mapsto (x, x)$ induces a homomorphism of associative algebras

$$\Delta : U\mathfrak{g} \to U\mathfrak{g} \otimes U\mathfrak{g} \ ,$$

which is called the *diagonal map*.

Proposition 5.2. *The diagonal map Δ is characterized by the following two conditions:*

1) *Δ is an algebra homomorphism.*

2) *$\Delta x = x \otimes 1 + 1 \otimes x$ for all $x \in \mathfrak{g}$.*

Notice that we identify $x \in \mathfrak{g}$ with its image in $U\mathfrak{g}$.

Definition 5.3. An element $\alpha \in U\mathfrak{g}$ is called *primitive* if $\Delta\alpha = \alpha \otimes 1 + 1 \otimes \alpha$.

Hence every element $x \in \mathfrak{g}$ is primitive.

Theorem 5.4. *Assume k is torsion free (as a $\mathbf{Z}$-module) and $\mathfrak{g}$ is a free k-module. Then the set of primitive elements of $U\mathfrak{g}$ coincides with $\mathfrak{g}$.*

Case 1. $\mathfrak{g}$ abelian. In this case $U\mathfrak{g}$ can be identified with the ring of polynomials $k[(X_i)]$ in variables X_i corresponding to the basis elements x_i of $\mathfrak{g}$. The diagonal map can be interpreted as a homomorphism $\Delta : k[X_i] \to k[(X_i'), (X_i'')]$ where $X_i' \sim X_i \otimes 1$ and $X_i'' \sim 1 \otimes X_i$, and is then given by $\Delta f(X_i', X_i'') = f(X_i' + X_i'')$, because it sends X_i to $X_i' + X_i''$ for each i. Thus the primitive elements $f(x) \in k[(X_i)]$ are those which satisfy $f(X_i' + X_i'') = f(X_i') + f(X_i'')$. If f is additive in this sense, then so is each homogeneous component f_n. If f is homogeneous of degree n and additive then

$$2^n f(X_i) = f(2X_i) = f(X_i + X_i) = 2f(X_i),$$

so $(2^n - 2)f = 0$. Since k is $\mathbf{Z}$-torsion free, we must have $f = 0$ if $n \neq 1$. Thus the only additive polynomials are the linear homogeneous ones.

Case 2. The general case. The map $\Delta : U\mathfrak{g} \to U\mathfrak{g} \otimes U\mathfrak{g}$ induces a map

$$\operatorname{gr}\Delta : \operatorname{gr} U\mathfrak{g} \to \operatorname{gr}(U\mathfrak{g} \otimes U\mathfrak{g}) \simeq \operatorname{gr} U(\mathfrak{g} \oplus \mathfrak{g}) \simeq \operatorname{gr} U\mathfrak{g} \otimes \operatorname{gr} U\mathfrak{g}$$

(see end of III.4). On the other hand, we have $\operatorname{gr} U\mathfrak{g} \simeq S\mathfrak{g}$, and the corresponding map $S\mathfrak{g} \to S\mathfrak{g} \otimes S\mathfrak{g}$ is the same as the one discussed in the first case, as one sees by looking at its effect on elements of the form $\bar{x} \in \operatorname{gr}_1 U\mathfrak{g}$ coming from elements $x \in \mathfrak{g}$.

Let $x \in U_n\mathfrak{g}$, and let $\bar{x}$ denote its image in $\operatorname{gr}_n U\mathfrak{g}$. If x is primitive, then $\bar{x}$ is primitive for $\operatorname{gr}\Delta$, hence, if $n > 1$, we have $\bar{x} = 0$ by case 1. Iterating this, we conclude $x \in U_1\mathfrak{g}$, that is, $x = \lambda + y$, with $\lambda \in k$, $y \in \mathfrak{g}$. Then

$$\Delta x = \lambda + y \otimes 1 + 1 \otimes y$$
$$x \otimes 1 + 1 \otimes x = \lambda + y \otimes 1 + \lambda + 1 \otimes y\ .$$

Thus, if x is primitive, then $2\lambda = \lambda$, hence $\lambda = 0$, and $x \in \mathfrak{g}$.

Exercises

1. Let $PU\mathfrak{g}$ denote the set of primitive elements of $U\mathfrak{g}$. Show that $PU\mathfrak{g}$ is stable under $[\ ,\ ]$, that is, if x and $y \in PU\mathfrak{g}$, so is $xy - yx$.
2. Suppose $pk = 0$ for some prime number p, and suppose $\mathfrak{g}$ is free, with basis $(x_i)_{i \in I}$. Show
   a) $PU\mathfrak{g}$ is stable under the map $y \mapsto y^p$.
   b) The elements $(x_i^{p^\nu})$, $i \in I$, $\nu \geq 1$, form a k-basis for $PU\mathfrak{g}$.
   c) If x and y are in $\mathfrak{g}$, then $(x + y)^p - x^p - y^p \in \mathfrak{g}$.

Chapter IV. Free Lie Algebras

In this chapter, k denotes a commutative and associative ring, with a un
All modules and algebras are taken over k.

1. Free magmas

Definition 1.1. A set M with a map

$$M \times M \to M$$

denoted by $(x, y) \mapsto xy$ is called a *magma.*

Let X be a set and define inductively a family of sets X_n $(n \geq 1)$ a follows:

1) $X_1 = X$

2) $X_n = \coprod_{p+q=n} X_p \times X_q$ $(n \geq 2)$ (= disjoint union).

Put $M_X = \coprod_{n=1}^{\infty} X_n$ and define $M_X \times M_X \to M_X$ by means of

$$X_p \times X_q \to X_{p+q} \subset M_X \ ,$$

where the arrow is the canonical inclusion resulting from 2).

The magma M_X is called the *free magma on* X. An element w of M_X called a non-associative word on X. Its length, $\ell(w)$, is the unique n such tha $w \in X_n$.

Theorem 1.2. *Let N be any magma, and let $f : X \to N$ be any map. The there exists a unique magma homomorphism $F : M_X \to N$ which extends f.*

Proof. Define F inductively by $F(u, v) = F(u) \cdot F(v)$ if $u, v \in X_p \times X_q$.

Properties of the free magma M_X:

1) M_X is generated by X.

2) $m \in M_X - X \iff m = u.v$, with $u, v \in M$; and u, v are uniquel determined by m.

2. Free algebra on X

Let A_X be the k-algebra of the free magma M_X. An element $\alpha \in A_X$ is a finite sum $\alpha = \sum_{m \in M_X} c_m m$, with $c_m \in k$; the multiplication in A_X extend the multiplication in M_X.

Definition 2.1. The algebra A_X is called the *free algebra* on X.

This definition is justified by the following:

IS
ISL

Theorem 2.2. *Let B be a k-algebra and let $f : X \to B$ be a map. There exists a unique k-algebra homomorphism $F : A_X \to B$ which extends f.*

Proof. By 1.2, we can extend f to a magma homomorphism $f' : M_X \to B$, where B is viewed as a magma under multiplication. This map extends by linearity to a k-linear map $F : A_X \to B$. One checks easily that F is an algebra homomorphism. The uniqueness of F follows from the fact that X generates A_X.

Remark. A_X is a graded algebra, the homogeneous elements of degree n being those which are linear combinations of words $m \in M_X$ of length n.

3. Free Lie algebra on X

Let I be the two-sided ideal of A_X generated by the elements of the form aa, $a \in A_X$ and $J(a,b,c)$, where $a,b,c \in A_X$ ($J(a,b,c) = (ab)c + (bc)a + (ca)b$).

Definition 3.1. The quotient algebra A_X/I is called the *free Lie algebra* on X.

This algebra will be denoted by $L_X(k)$, or simply L_X.

Functorial properties.

1) If $f : X \to X'$ is any map, then there exists a unique map $F : L_X \to L_{X'}$ such that $F|_X = f$.

1′) If $\{X_\alpha, i_\alpha^\beta\}$ is a direct system and $X = \varinjlim X_\alpha$ then

$$\varinjlim L_{X_\alpha} = L_X \ .$$

2) Let k' be an extension of k, then

$$L_X(k') = L_X(k) \otimes_k k' \ .$$

3) I is a graded ideal of A_X, which implies L_X has a natural structure of graded algebra.

Proof. Let $I^\#$ be the set of $a \in A_X$ such that every homogeneous component of a belongs to I. Then $I^\#$ is a two-sided ideal and $I^\# \subset I$.

Now let $x \in A_X$, $x = \sum_{n=1}^\infty x_n$, x_n homogeneous. Then

$$x \cdot x = \sum x_n^2 + \sum_{n<m} (x_n x_m + x_m x_n) \ ,$$

but $x_n^2 \in I$, $x_n x_m + x_m x_n = (x_n + x_m)^2 - x_n^2 - x_m^2 \in I$, so that $x \cdot x \in I^\#$. For three elements, $x = \sum x_n$, $y = \sum y_n$, and $z = \sum z_n$ we have $J(x,y,z) = \sum_{l,m,n} J(x_l, y_m, z_n) \in I^\#$. Thus $I^\# = I$, q.e.d.

4) The homogeneous component L_X^1 has basis X and the homogeneous component L_X^2 has for basis the family of elements $[x,y]$, $x < y$, $x,y \in X$, where we have chosen a total order on X.

Proof. Clearly X generates L_X and $[X,X]$ generate L_X^2 ($[X,X] = \{\, [x,y],\ x < y,\ x,y \in X \,\}$). Consider the module $E = k^{(X)}$ and the Lie algebra $E \oplus \bigwedge^2 E = \mathfrak{g}$ (example ii′ of Chapter I). The canonical map $X \to \mathfrak{g}$ induces a Lie algebra homomorphism $L_X \to \mathfrak{g}$, and the composition $L_X^1 \oplus L_X^2 \to L_X \to \mathfrak{g}$ is an isomorphism q.e.d.

4. Relation with the free associative algebra on X

Definition 4.1. Let $E = k^{(X)}$ be the free k-module with basis X. Then the free associative algebra on X, denoted by Ass_X, is the tensor algebra TE of E.

(Elements of Ass_X may be called "associative but non-commutative" polynomials in the elements of X.)

Theorem 4.2. *Let $\phi : L_X \to \mathrm{Ass}_X$ and $\Phi : UL_X \to \mathrm{Ass}_X$ be the maps induced by the map $X \to \mathrm{Ass}_X$. Then:*

1) *The map Φ is an isomorphism.*

2) *The map ϕ is an isomorphism of L_X onto the Lie subalgebra of Ass_X generated by X.*

3) *L_X and its homogeneous components L_X^n are free k-modules.*

4) *If X is finite and $\mathrm{Card}\, X = d$ then L_X^n is free of finite rank $\ell_d(n)$ and*

$$(*) \qquad \sum_{m|n} m\ell_d(m) = d^n$$

Remark. The formula $(*)$ determines $\ell_d(n)$ by induction on n. In fact,

$$n\ell_d(n) = d^n - \sum_{\substack{m|n \\ m<n}} m\ell_d(m)\ .$$

(More precisely, let μ be the *Möbius function*, defined by:

$$\sum_{n=1}^{\infty} \mu(n) n^{-s} = 1/\zeta(s) = \prod_p (1 - p^{-s})\ .$$

One has:

$$n\ell_d(n) = \sum_{m|n} \mu(m) d^{n/m}\ .\quad)$$

Proof of Theorem 4.2.

1) is clear: the map $X \to UL_X$ defines a homomorphism Ψ of Ass_X into UL_X, and $\Phi \circ \Psi = 1$, $\Psi \circ \Phi = 1$.

Note also that ϕ maps L_X *onto* the Lie subalgebra of Ass_X generated by X, so that (2) is equivalent to saying that ϕ is *injective*. Note also that (3) $\Rightarrow$ (2); for, if L_X is free over k, the Birkhoff-Witt theorem shows that $L_X \to UL_X$ is injective, and we can identify UL_X with Ass_X.

The rest of the proof is divided into four steps:

First step: Assume k is a field and X is finite.

Choose a homogeneous basis $(\gamma_i)_{i\in I}$ of L_X and a total order of I.

Put $d_i = \deg(\gamma_i)$.

Now the Birkhoff-Witt theorem implies that the family of elements

$$\gamma^e = \gamma_{i_1}^{e_{i_1}} \cdots \gamma_{i_s}^{e_{i_s}} \qquad \text{with } i_1 < \cdots < i_s$$

is a basis of $UL_X = \mathrm{Ass}_X$ and we have $\deg(\gamma^e) = \sum e_{i_j} d_{i_j}$.

Let $a(n)$ be the rank of Ass_X^n, then $a(n)$ is equal to the number of families (e_i) such that $n = \sum e_i d_i$.

This last statement is equivalent to the fact that the formal power series $A(t) = \sum a(n)t^n$ may be expressed in the form

$$A(t) = \prod_{i\in I} \frac{1}{1 - t^{d_i}}$$

because $\prod_{i\in I} \frac{1}{1-t^{d_i}} = \prod_{i\in I}(1 + t^{d_i} + t^{2d_i} + \cdots)$ and the coefficient of t^n in this product is precisely the number of families (e_i) such that $\sum e_i d_i = n$.

Now, for any positive integer m we have that in the product $\prod_{i\in I} \frac{1}{1-t^{d_i}}$ the number of factors such that $d_i = m$ is the rank $\ell_d(m)$ of L_X^m, i.e.,

$$A(t) = \prod_{m=1}^{\infty} \frac{1}{(1 - t^m)^{\ell_d(m)}} .$$

On the other hand, since Ass_X is the free associative algebra on X the family of monomials $x_{i_1} \cdots x_{i_n}$, $x_{i_\nu} \in X$ is a basis of Ass_X^n.

This implies that $a(n) = d^n$ and therefore

$$A(t) = \sum d^n t^n = \frac{1}{1 - dt}$$

i.e.,

$$\prod_{m=1}^{\infty} \frac{1}{(1 - t^m)^{\ell_d(m)}} = \frac{1}{1 - dt} .$$

From the equality $\log \frac{1}{1-t} = \sum_{n=1}^{\infty} \frac{1}{n} t^n$ we conclude that

$$\sum_{m,\nu} \frac{1}{\nu} \ell_d(m) t^{m\nu} = \sum_{n=1}^{\infty} \frac{1}{n} d^n t^n$$

and hence, for each n, we have $\frac{1}{n}d^n = \sum_{m\nu=n} \frac{1}{\nu}\ell_d(m)$, i.e.,

$$d^n = \sum_{m|n} m\ell_d(m)$$

which proves (4) in this case.

Second Step: *Assume $k = \mathbf{Z}$ and X is a finite set.*

We will use the following lemma.

Lemma 4.3. *If E is a finitely generated $\mathbf{Z}$-module and $\dim(E \otimes_{\mathbf{Z}} \mathbf{F}_p)$ over $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$ is independent of p, for all primes p, then E is a $\mathbf{Z}$-free module with rank equal to the dimension of $E \otimes_{\mathbf{Z}} \mathbf{F}_p$ over $\mathbf{F}_p$.*

This lemma is an easy consequence of the structure theorem of abelian groups.

Now, since $L_X^n(\mathbf{Z}) \otimes_{\mathbf{Z}} \mathbf{F}_p = L_X^n(\mathbf{F}_p)$ and $\dim(L_X^n(\mathbf{F}_p)) = \ell_d(n)$ which is independent of p, it follows that L_X^n is $\mathbf{Z}$-free with rank $\ell_d(n)$.

This proves the theorem in this case.

Third Step: *Assume $k = \mathbf{Z}$ and X is an arbitrary set.*

Let $\{Y_\alpha\}$ be the family of finite subsets of X, then $X = \varinjlim_\alpha Y_\alpha$.

We first prove (2).

Using the second case, we have that the map

$$\phi_\alpha : L_{Y_\alpha} \to \mathrm{Ass}_{Y_\alpha}$$

is injective for all α.

Now $\phi = \varinjlim_\alpha \phi_\alpha$ and the inductive limit of a family of injective maps is injective. This proves (2).

In particular (2) implies that L_X and L_X^n are $\mathbf{Z}$-submodules of Ass_X, which is free, so L_X and L_X^n are free for all n.

This proves the theorem in the third case.

Fourth Step: *General case.*

The equality $L_X^n(k) = L_X^n(\mathbf{Z}) \otimes_{\mathbf{Z}} k$ together with the third case imply $L_X^n(k)$ is k-free, i.e., (3) and therefore (2) holds.

On the other hand $\mathrm{rk}\, L_X^n(k) = \mathrm{rk}\, L_X^n(\mathbf{Z})$ thus, if X is finite, (4) holds. q.e.d.

5. P. Hall families

Definition 5.1. Let X be a set. A *P. Hall family* in M_X, the free magma on X, is a totally ordered subset H of M_X such that:

(1) $X \subset H$.

(2) If $u, v \in H$ with $\ell(u) < \ell(v)$ then $u < v$.

(3) Let $u \in M_X - X$ and let $u = vw$ be the unique decomposition of u where $v, w \in M_X$. Then $u \in H$ if and only if the following two conditions are satisfied:

(a) $v \in H$, $w \in H$ and $v < w$,

(b) either $w \in X$ or $w = w'w''$ with $w' \in H$, $w'' \in H$ and $w' \leq v$.

Lemma 5.2. *There exists a P. Hall family for any set X.*

Proof. We define by induction $H^n = H \cap X_n$. We take $H^1 = X$, and choose a total order on X. Suppose now $H^1, \ldots, H^{n-1}$ have been defined and totally ordered in such a way that (1), (2), (3) hold for elements of length $\leq n-1$. The set H^n is then defined without ambiguity by condition (3); we choose any total order on H^n, and put $u < v$ if $u \in H^i$ $(i \leq n-1)$ and $v \in H^n$. This completes the induction process, and it is clear that $H = \bigcup H^n$ is a P. Hall family.

Example. Let $X = \{x, y\}$, with $x \neq y$. We can take $H^1, \ldots, H^5$ as follows:

$$
\begin{aligned}
H^1 &= \{x, y\}, \qquad x < y \\
H^2 &= \{x \cdot y\} \\
H^3 &= \{x \cdot (x \cdot y), y \cdot (x \cdot y)\}, \qquad x \cdot (x \cdot y) < y \cdot (x \cdot y) \\
H^4 &= \{x(x(xy)), y(x(xy)), y(y(xy))\} \\
H^5 &= \{x(x(x(xy))), y(x(x(xy))), y(y(x(xy))), y(y(y(xy))), \\
&\qquad (xy)(x(xy)), (xy)(y(xy))\}
\end{aligned}
$$

Theorem 5.3. *If H is a P. Hall family in M_X, then the canonical images of the elements $h \in H$ in L_X make up a basis of L_X.*

Let $h \in H$ and denote by $\bar{h}$ its image in L_X. Theorem 5.3 is equivalent to:

(1) The family $\{\bar{h}\}$, $h \in H$, generates L_X.

(2) The elements $\{\bar{h}\}$, $h \in H$, are linearly independent.

We prove here only the (easier) part (1). For a proof of (2), the reader may look in M. Hall, *The Theory of Groups*, p. 170–171, or E. Witt, *Die Unterringe der freien Lieschen Ringe*, Math. Zeit., 1956; M. Hall's proof is based on a counting argument; Witt's proof is better (but longer). (See also Bourbaki, LIE II, §2, n° 11.)

Proof of (1). Let L'_X be the k-module generated by $\bar{h}$; since L'_X contains X, it will be enough to show that L'_X is a *Lie algebra*, i.e., that $h_1, h_2 \in H$ implies that $[\bar{h}_1, \bar{h}_2]$ is in L'_X.

We will carry the proof by a double induction, first on the length of h_1 + length of h_2 (which is the length n of $h_1 h_2$) and finally for a given n, by

decreasing induction on $\mathrm{Inf}(h_1, h_2)$; in order that this induction process work we will assume that X is *finite*; the general case will follow by passing to an inductive limit.

We may suppose $h_1 < h_2$ (otherwise we use the relations $[\bar{h}_1, \bar{h}_2] = -[\bar{h}_2, \bar{h}_1]$ and $[\bar{h}, \bar{h}] = 0$).

First Case. Let $h_2 \in X$, then $h_1 \in X$ since $h_1 < h_2$, so we have $h_1 h_2 \in H$ and therefore $\overline{h_1 h_2} = [\bar{h}_1, \bar{h}_2]$, q.e.d.

Second Case. $h_2 \notin X$. Put $h_2 = h_3 h_4$, $h_3, h_4 \in H$ and $h_3 < h_4$.

We have the following subcases:

a) $h_3 \leq h_1$ and then $h_1(h_3 h_4) \in H$, so

$$[\bar{h}_1, \bar{h}_2] = [\bar{h}_1, [\bar{h}_3, \bar{h}_4]] = \overline{h_1(h_3 h_4)} \, .$$

b) $h_1 < h_3 < h_4$. Using the Jacobi identity we get

$$[\bar{h}_1, [\bar{h}_3, \bar{h}_4]] = [\bar{h}_3, [\bar{h}_1, \bar{h}_4]] - [\bar{h}_4, [\bar{h}_1, \bar{h}_3]] \, .$$

Now length of $h_1 h_4 <$ length of $h_1 h_2$, hence we can apply the induction hypothesis, i.e., $[\bar{h}_1, \bar{h}_4] = \sum c_\alpha \bar{h}_\alpha$ where $h_\alpha \in H$.

From this equality we get $\ell(h_\alpha) = \ell(h_1) + \ell(h_4)$ which implies $\ell(h_\alpha) > \ell(h_1)$, hence $h_\alpha > h_1$. Since we have $h_1 < h_3$, we obtain $\mathrm{Inf}(h_3, h_\alpha) > h_1 = \mathrm{Inf}(h_1, h_2)$.

Applying the induction hypothesis we see that $[\bar{h}_3, \bar{h}_\alpha]$ is a linear combination of $\bar{h}$'s with $h \in H$.

Similarly, replacing h_3 by h_4, we see that $[\bar{h}_4, [\bar{h}_1, \bar{h}_3]]$ is also a linear combination of $\bar{h}$'s with $h \in H$. q.e.d.

6. Free groups

(In this section, we take $k = \mathbf{Z}$.)

Let X be a set and let F_X be the free group on X. Let F_X^n be the descending central series of F_X, defined by $F_X^1 = F_X$ and $F_X^n = (F_X, F_X^{n-1})$, for $n > 1$.

The associated graded group is, as we know, a Lie algebra, given by

$$\mathrm{gr}\, F_X = \sum_{n=1}^{\infty} \mathrm{gr}^n F_X \, , \qquad \mathrm{gr}^n F_X = F_X^n / F_X^{n+1} \, .$$

In particular, $\mathrm{gr}^1 F_X = F_X/(F_X, F_X)$, that is, $\mathrm{gr}^1 F_X$ is the free abelian group on X.

Theorem 6.1. *The canonical map $X \to \mathrm{gr}^1 F_X$ induces an isomorphism of Lie algebras*

$$\phi_1 : L_X \xrightarrow{\sim} \mathrm{gr}\, F_X \, .$$

Corollary 6.2. *The groups F_X^n / F_X^{n+1} are free $\mathbf{Z}$-modules and if X is finite with $\mathrm{Card}\, X = d$, then $\mathrm{rk}(F_X^n / F_X^{n+1}) = \ell_d(n)$.*

Now consider the free associative algebra Ass_X on X; let Ass_X^n the component of degree n of Ass_X. The *completion* $\widehat{\mathrm{Ass}}_X$ of Ass_X is defined as the infinite product $\prod_{n=0}^{\infty} \mathrm{Ass}_X^n$. An element $f \in \widehat{\mathrm{Ass}}_X$ can be represented by a formal series $f = \sum_{n=0}^{\infty} f_n$, with $f_n \in \mathrm{Ass}_X^n$.

Define a homomorphism $\theta : F_X \to \widehat{\mathrm{Ass}}_X^*$ by $\theta(x) = 1 + x$ where $\widehat{\mathrm{Ass}}_X^*$ is the multiplicative group of the invertible elements of $\widehat{\mathrm{Ass}}_X$ (it is clear that $1 + x$ is invertible in $\widehat{\mathrm{Ass}}_X$, so it is in the multiplicative group $\widehat{\mathrm{Ass}}_X^*$).

For any positive integer n, define $\hat{\mathfrak{m}}^n \subset \widehat{\mathrm{Ass}}_X$ as

$$\hat{\mathfrak{m}}^n = \left\{ f \mid f = \sum_{n=0}^{\infty} f_m \text{ and } f_0 = f_1 = \cdots = f_{n-1} = 0 \right\},$$

and put ${}'F_X^n = \theta^{-1}(1 + \hat{\mathfrak{m}}^n)$. Then $g \in F_X$ is in ${}'F_X^n$ if and only if $\theta(g) = 1 + \sum_{m \geq n} \psi_n$.

Notice that ${}'F_X^1 = F_X$ and ${}'F_X^n \subset {}'F_X^{n-1}$.

Theorem 6.3. ${}'F_X^n = F_X^n$.

We now prove Theorems 6.1 and 6.3.

a) It is clear that $\phi_1 : L_X \to \mathrm{gr}\, F_X$ is *surjective.*

b) $({}'F_X^n)$ *is a filtration of* F_X. In fact, we only have to check

$$({}'F_X^m, {}'F_X^p) \subset {}'F_X^{m+p}.$$

To prove this, take $g \in {}'F_X^m$, $h \in {}'F_X^p$ with $\theta(g) = 1 + G$, $G \in \hat{\mathfrak{m}}^m$, $\theta(h) = 1 + H$, $H \in \hat{\mathfrak{m}}^p$.

We have $gh = hg(g, h)$ and

$$\theta(gh) = 1 + G + H + GH$$
$$\theta(hg) = 1 + G + H + HG.$$

Since θ is a homomorphism we get $\theta(gh) = \theta(hg)\theta((g, h))$, i.e.,

$$(*) \qquad \theta((g, h)) = 1 + (GH - HG) - \text{higher terms.}$$

Therefore $(g, h) \in {}'F_X^{m+p}$.

There is a natural map $\eta : {}'\mathrm{gr}\, F_X \to \mathrm{Ass}_X$ defined as follows: let $\xi \in {}'\mathrm{gr}^n F_X$, let $g \in {}'F_X^n$ be a representative of ξ, and let

$$\theta(g) = 1 + G_n + G_{n+1} + \cdots, \qquad \text{with } G_p \in \mathrm{Ass}_X^p.$$

We define $\eta(\xi)$ by:

$$\eta(\xi) = G_n.$$

It is easy to see that this definition does not depend on the choice of the representative g. Formula $(*)$ shows that $\eta : {}'\mathrm{gr}\, F_X \to \mathrm{Ass}_X$ is a *Lie algebra homomorphism.*

Since $'F_X^n$ is a filtration we know that $F_X^n \subset {}'F_X^n$, which induces a homomorphism $\psi : \operatorname{gr} F_X \to {}'\operatorname{gr} F_X$.

Now let us look at the composition

$$L_X \xrightarrow{\phi_1} \operatorname{gr} F_X \xrightarrow{\psi} {}'\operatorname{gr} F_X \xrightarrow{\eta} \operatorname{Ass}_X$$

where ϕ_1 is surjective and η is injective.

This composition is obviously the map $\phi : L_X \to \operatorname{Ass}_X$ given in the Theorem 4.2 and we know it is injective.

Hence ϕ_1 is injective and therefore is an isomorphism; which proves Theorem 6.1.

This implies now that ψ is injective. Let us prove, by induction, that $F_X^n = {}'F_X^n$.

If $n = 1$ then $F_X^1 = {}'F_X$ by definition.

Now suppose $n > 1$, then we have

$$F_X^n \subset {}'F_X^n \subset F_X^{n-1} = {}'F_X^{n-1}$$

and the injection $\operatorname{gr}^{n-1} F_X \to {}'\operatorname{gr}^{n-1} F_X$ is the canonical map

$$F_X^{n-1}/F_X^n \to F_X^{n-1}/{}'F_X^n ,$$

which implies $F_X^n = {}'F_X^n$. q.e.d.

7. The Campbell-Hausdorff formula

In IV.7 and IV.8, the ground ring k is supposed to be a $\mathbf{Q}$-algebra (for instance, a field of characteristic zero).

Theorem 7.1. *Let X be a set; then the free Lie algebra L_X on X coincides with the set of primitive elements of* Ass_X *(i.e., $L_X = \{ w \in \operatorname{Ass}_X \mid \Delta w = w \otimes 1 + 1 \otimes w \}$, where $\Delta : \operatorname{Ass}_X \to \operatorname{Ass}_X \otimes \operatorname{Ass}_X$ is the diagonal map).*

This follows from a theorem proved in Chapter III, since Ass_X may be identified with UL_X.

Define now, as in IV.6, the completion $\widehat{\operatorname{Ass}}_X$ of Ass_X and the completion $\hat{L}_X$ of L_X by:

$$\widehat{\operatorname{Ass}}_X = \prod_{n=0}^{\infty} \operatorname{Ass}_X^n , \qquad \hat{L}_X = \prod_{n=0}^{\infty} L_X^n .$$

Define similarly the *completed tensor product* $\widehat{\operatorname{Ass}}_X \hat{\otimes} \widehat{\operatorname{Ass}}_X$ by:

$$\widehat{\operatorname{Ass}} \hat{\otimes} \widehat{\operatorname{Ass}}_X = \prod_{p,q} \operatorname{Ass}_X^p \otimes \operatorname{Ass}_X^q .$$

The diagonal map Δ extends to a map $\Delta : \widehat{\operatorname{Ass}}_X \to \widehat{\operatorname{Ass}}_X \hat{\otimes} \widehat{\operatorname{Ass}}_X$ and it is clear that Theorem 7.1 remains valid when Ass_X and $\operatorname{Ass}_X \otimes \operatorname{Ass}_X$ are replaced by their completions.

Theorem 7.2. *Let* $\hat{\mathfrak{m}} \subset \widehat{\mathrm{Ass}}_X$ *be the ideal generated by* X. *Define maps*

$$\exp : \hat{\mathfrak{m}} \to 1 + \hat{\mathfrak{m}} \qquad \textit{and} \qquad \log : 1 + \hat{\mathfrak{m}} \to \hat{\mathfrak{m}}$$

by the usual formulae:

$$\exp(x) = \sum x^n/n! \,, \qquad \log(1+x) = \sum_{n=1}^{\infty} (-1)^{n+1} x^n/n \,.$$

Then $\exp \circ \log = \mathrm{id}$ *and* $\log \circ \exp = \mathrm{id}$.

Proof. Let us prove, for instance, that $\exp(\log(1+y)) = 1+y$ if $y \in \hat{\mathfrak{m}}$. If T is an indeterminate, the formula $\exp(\log(1+T)) = 1+T$ is known to be true in the power series ring $\mathbf{Q}[[T]]$. But, since y belongs to $\hat{\mathfrak{m}}$, there is a well-defined and continuous homomorphism $f : \mathbf{Q}[[T]] \to \widehat{\mathrm{Ass}}_X$ which transforms T into y. Applying f to the equality $\exp(\log(1+T)) = 1+T$, we get $\exp(\log(1+y)) = 1+y$, q.e.d.

Corollary 7.3. *The map* exp *defines a bijection of the set of* $\alpha \in \hat{\mathfrak{m}}$ *with* $\Delta\alpha = \alpha \otimes 1 + 1 \otimes \alpha$ *onto the set of* $\beta \in 1 + \hat{\mathfrak{m}}$ *with* $\Delta\beta = \beta \otimes \beta$.

Proof. Let $\alpha \in \hat{\mathfrak{m}}$ and $\beta = e^{\alpha}$. Since Δ commutes with the exponential map and $\alpha \otimes 1$ commutes with $1 \otimes \alpha$, we obtain

$$\begin{aligned} \Delta\beta = \Delta e^{\alpha} = e^{\Delta\alpha} = e^{\alpha\otimes 1 + 1\otimes\alpha} &= e^{\alpha\otimes 1} e^{1\otimes\alpha} = (\beta \otimes 1)(1 \otimes \beta) \\ &= \beta \otimes \beta \,. \end{aligned}$$

Theorem 7.4 (Campbell-Hausdorff). *Let* $X = \{x, y\}$, $x \neq y$, *then* $e^x e^y = e^z$ *with* $z \in \hat{L}_X$.

Proof. Since $e^x, e^y \in 1 + \hat{\mathfrak{m}}$ we have $e^x e^y \in 1 + \hat{\mathfrak{m}}$ and since the exponential map is a bijection there is one and only one $z \in \hat{\mathfrak{m}}$ such that $e^z = e^x e^y$.

We have the relation

$$\begin{aligned} \Delta(e^z) = \Delta(e^x e^y) &= \Delta(e^x)\Delta(e^y) \\ &= (e^x \otimes e^x)(e^y \otimes e^y) \\ &= e^z \otimes e^z \,. \end{aligned}$$

Applying 7.3 we find that z is a primitive element and by $\widehat{7.1}$ $z \in \hat{L}_X$. q.e.d.

Now, let X be an arbitrary set and let $z(x,y)$ denote the element of $\hat{L}_{\{x,y\}} \subset \hat{L}_X$ such that $e^x e^y = e^{z(x,y)}$ for all $x, y \in X$.

We have $z(x,y) = \sum_{n=1}^{\infty} z_n(x,y)$ where $z_n(x,y) \in L_X^n$.

Explicitly, the values of the first three homogeneous components of $z(x,y)$ are

$$z_1(x,y) = x + y$$
$$z_2(x,y) = \tfrac{1}{2}[x,y]$$
$$z_3(x,y) = \tfrac{1}{12}[x,[x,y]] + \tfrac{1}{12}[y,[y,x]]$$

and it is clear that $z(x,0) = x$, $z(0,y) = y$, and $z(z(w,x),y) = z(w,z(x,y))$.

8. Explicit formula

Define a map $\Phi : \mathfrak{m} \to L_X$ ($\mathfrak{m} \subset \mathrm{Ass}_X$) as follows:

$$\Phi(x_1 \cdots x_n) = [x_1,[x_2,\ldots,[x_{n-1},x_n]\cdots] = \mathrm{ad}(x_1)\cdots\mathrm{ad}(x_{n-1})(x_n)$$

where $x_i \in X$.

Now define $\phi : \mathfrak{m} \to L_X$ by $\phi(x_1 \cdots x_n) = \frac{1}{n}\Phi(x_1 \cdots x_n)$.

Theorem 8.1. *The map ϕ is a retraction of $\mathfrak{m}$ onto L_X, i.e., $\phi|_{L_X} = \mathrm{id}_{L_X}$.*

Proof. We have to prove that $\Phi(u) = nu$ if $u \in L_X^n$.

Let $\theta : \mathrm{Ass}_X \to \mathrm{End}(L_X)$ be the algebra homomorphism which extends the Lie algebra homomorphism $\mathrm{ad} : L_X \to \mathrm{End}(L_X)$.

Lemma 8.2. *The relation $\Phi(uv) = \theta(u)\Phi(v)$ holds for $u \in \mathrm{Ass}_X$ and $v \in \mathfrak{m}$.*

Proof of Lemma. Since Φ and θ are linear it is enough to consider the case $u = x_1 \cdots x_n$, $x_i \in X$ and we proceed by induction on n.

If $n = 1$ then it is trivial.

Now suppose $n > 1$, then

$$\begin{aligned}\Phi(x_1 \cdots x_n v) &= \theta(x_1)\Phi(x_2 \cdots x_n v) = \theta(x_1)\theta(x_2 \cdots x_n)\Phi(v)\\ &= \theta(x_1 \cdots x_n)\Phi(v)\ .\end{aligned}$$

This concludes the proof of the lemma.

We now prove that $\Phi(u) = nu$ for $u \in L_X^n$ by induction on n.

If $n = 1$ the property is obvious.

Suppose $n > 1$, then $u = \sum[v_i, w_i]$ and it is enough to prove this when $u = [v,w]$ with $v \in L_X^p$, $w \in L_X^q$, $p + q = n$, $p,q > 0$.

Using the fact that $\theta(v) = \mathrm{ad}\, v$ and $\theta(w) = \mathrm{ad}\, w$ we get

$$\begin{aligned}\Phi([v,w]) &= \Phi(vw - wv) = \theta(v)\Phi(w) - \theta(w)\Phi(v)\\ &= q\theta(v)w - p\theta(w)v\\ &= q[v,w] - p[w,v]\\ &= (q+p)[v,w] = nu\end{aligned}$$ q.e.d.

Finally, we are prepared to give the explicit formula for $z(x,y) = \log(e^x e^y)$ for $x,y \in X$.

As before let us write $z = \sum_{n=1}^{\infty} z_n$ with $z_n \in L_X^n$.

Since $e^x e^y = \left(\sum_{p=0}^{\infty} \frac{x^p}{p!}\right)\left(\sum_{q=0}^{\infty} \frac{y^q}{q!}\right) = 1 + \sum_{p+q\geq 1} \frac{x^p y^q}{p!q!}$ we have

$$z = \log(e^x e^y) = \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} \left(\sum_{p+q\geq 1} \frac{x^p y^q}{p!q!}\right)^m ,$$

so we obtain

$$z = \sum_{p_i+q_i\geq 1} \frac{(-1)^{m+1}}{m} \frac{x^{p_1} y^{q_1} x^{p_2} y^{q_2} \cdots x^{p_m} y^{q_m}}{p_1!q_1! \cdots p_m!q_m!} .$$

Applying the homomorphism Φ to the monomials which appear in this sum we get

$$\Phi(x^{p_1} y^{q_1} \cdots x^{p_m} y^{q_m}) = \mathrm{ad}(x)^{p_1} \mathrm{ad}(y)^{q_1} \cdots \mathrm{ad}(x)^{p_m} \mathrm{ad}(y)^{q_m - 1}(y)$$

if $q_m \geq 1$, and:

$$\Phi(x^{p_1} y^{q_1} \cdots x^{p_m}) = \mathrm{ad}(x)^{p_1} \mathrm{ad}(y)^{q_1} \cdots \mathrm{ad}(x)^{p_m - 1}(x) , \qquad \text{if } q_m = 0.$$

Notice that this is zero if $q_m \geq 2$, or if $q_m = 0$, $p_m \geq 2$. Hence, the only possible non-zero terms are those where $q_m = 1$, or $p_m = 1$, $q_m = 0$.

Hence, using the identity $z_n = \phi(z_n)$, we obtain the *explicit Campbell-Hausdorff formula* (in Dynkin's form):

$$z_n = \frac{1}{n} \sum_{p+q=n} (z'_{p,q} + z''_{p,q}) ,$$

where

$$z'_{p,q} = \sum_{\substack{p_1+\cdots+p_m=p \\ q_1+\cdots+q_{m-1}=q-1 \\ p_i+q_i\geq 1 \\ p_m\geq 1}} \frac{(-1)^{m+1}}{m} \frac{\mathrm{ad}(x)^{p_1} \mathrm{ad}(y)^{q_1} \cdots \mathrm{ad}(x)^{p_m}(y)}{p_1!q_1! \cdots p_m!}$$

and

$$z''_{p,q} = \sum_{\substack{p_1+\cdots+p_{m-1}=p-1 \\ q_1+\cdots+q_{m-1}=q \\ p_i+q_i\geq 1}} \frac{(-1)^{m+1}}{m} \frac{\mathrm{ad}(x)^{p_1} \mathrm{ad}(y)^{q_1} \cdots \mathrm{ad}(y)^{q_{m-1}}(x)}{p_1!q_1! \cdots q_{m-1}!} .$$

Exercises

1. Let X be a finite set, with $\mathrm{Card}(X) = d$. Show that the number of elements of M_X of length n is equal to:

$$2^{n-1} d^n \frac{1 \cdot 3 \cdot 5 \cdots (2n-3)}{n!}$$

2. Show that $L_X^n = [X, L_X^{n-1}]$ for $n \geq 2$.

3. Show that the center of L_X is 0 if $\mathrm{Card}(X) \neq 1$, and that the center of $L_X / \sum_{n>p} L_X^n$ is equal to L_X^p.
4. Let X be a denumerable set with $\mathrm{Card}(X) \geq 2$, and let $\underline{H}$ the set of all Hall families in M_X. Show that $\mathrm{Card}(\underline{H}) = \mathrm{Card}(\mathbf{R})$.
5. Show that the homomorphism $\theta : F_X \to \widehat{\mathrm{Ass}}_X^*$ defined in IV.6 is injective.