1.3 Some commutator calculations

If G is a group and $x,\, y \in G$, we will write x^y for the conjugation $y^{-1}xy$ and (x, y) for the commutator $x^{-1}y^{-1}xy$. For $x, y, z \in G$ we have

(1) $xy = yx^y$.

(2) $x^y = x(x, y)$.

(3) $(x, x) = 1$.

(4) $(y, x) = (x, y)^{-1}$.

(5) $(x, yz) = (x, z)(x, y)^x$.

(6) $(xy, z) = (x, z)^y(y, z)$.

(7) $(x^y, (y, z))(y^z, (z, x))(x^z, (x, y)) = 1$.

(1) to (4) are trivial. For (5), from (2) we have

$$x(x, yz) = x^{yz} = (x^y)^z = [x(x, y)]^z = x^z(x, y)^z = x(x, z)(x, y)^z.$$

Cancelling x on the left gives the desired identity. For (6),

$$xy(xy, z) = (xy)^z = x^z y^z = x(x, z)y(y, z) = xy(x, z)^y(y, z),$$

and then we cancel xy. For (7),

$$\begin{aligned}(x^y, (y, z)) &= (y^{-1}x^{-1}y)(z^{-1}y^{-1}zy)(y^{-1}xy)(y^{-1}z^{-1}yz)\\ &= y^{-1}x^{-1}yz^{-1}y^{-1}zxz^{-1}yz\\ &= (yzy^{-1}xy)^{-1}(zxz^{-1}yz).\end{aligned}$$

Similarly

$$\begin{aligned}(y^z, (z, x)) &= (zxz^{-1}yz)^{-1}(xyx^{-1}zx),\\ (z^x, (x, y)) &= (xyx^{-1}zx)^{-1}(yzy^{-1}xy),\end{aligned}$$

so

$$(x^y, (y, z))(y^z, (z, x))(z^x, (x, y)) = 1.$$

The significance of these identities is that if $A, B \trianglelefteq G$ are normal subgroups, then their commutator

$$(A, B) := \{(a, b) : a \in A, b \in B\}$$

is again normal, and we have

$$(A, (B, C)) \subset (B, (C, A))(C, (A, B)).$$

by (7).

1.4 Filtered groups

A *filtration* on a group G is a map $w\colon G \to \mathbb{R}_+ \cup \{+\infty\}$ such that

(i) $w(1) = +\infty$.

(ii) $w(xy^{-1}) \geq \inf\{w(x), w(y)\}$.

(iii) $w((x,y)) \geq w(x) + w(y)$.

Taking $x = 1$ in (iii) we have $w(y^{-1}) \geq w(y)$, and since y is arbitrary symmetry implies $w(y) = w(y^{-1})$. For $\lambda \in \mathbb{R}_+$ define

$$G_\lambda = \{x \in G : w(x) \geq \lambda\}$$
$$G_\lambda^+ = \{x \in G : w(x) > \lambda\}.$$

By (3) these are subgroups: if $x, y \in G_\lambda$ then

$$w(xy^{-1}) \geq \inf\{w(x), w(y)\} \geq \lambda,$$

and identically for G_λ^+. In fact, if $x \in G_\lambda$ and $y \in G$ then

$$x^y \equiv x \pmod{G_\lambda^+}.$$

Indeed by (2) from the previous section, this identity may be more tractably written as $x^{-1}x^y = (x,y) \in G_\lambda^+$ and this follows from (iii):

$$w((x,y)) \geq w(x) + w(y) \geq \lambda + w(y) > \lambda.$$

In particular, $x^y \equiv \pmod{G}_\lambda$ so we have shown that G_λ is a normal subgroup of G and the same holds for

$$G_\lambda^+ = \bigcup_{\mu > \lambda} G_\mu.$$

We now use this filtration to define a Lie algebra structure on a filtered group. For $\alpha \geq 0$, define

$$\operatorname{gr}_\alpha G := G_\alpha / G_\alpha^+$$

and

$$\operatorname{gr} G := \sum_\alpha \operatorname{gr}_\alpha G.$$

First remark that $\operatorname{gr}_\alpha G$ is abelian. Indeed if $\overline{x}, \overline{y} \in \operatorname{gr}_\alpha G$ have representatives $x, y \in G_\alpha$ then

$$w((x,y)) \geq w(x) + w(y) \geq 2\alpha > \alpha,$$

so $\overline{(x,y)} = 1$ in $\operatorname{gr}_\alpha G$. Thus we will write $\operatorname{gr}_\alpha G$ additively.

Proposition 1.4.1

The map $c_{\alpha,\beta}\colon G_\alpha \times G_\beta \to G_{\alpha+\beta}$ defined by $(x,y) \mapsto (x,y)$ descends to a bilinear map $\overline{c}_{\alpha,\beta}\colon \operatorname{gr}_\alpha G \times \operatorname{gr}_\beta G \to \operatorname{gr}_{\alpha+\beta} G$.

Proof. Let $x, x' \in G_\alpha$, $y, y' \in G_\beta$. To obtain a well-defined map on quotients, we wish to show that for $u \in G_\alpha^+$, $v \in G_\beta^+$, we have

$$(xu, y) \equiv (x, y) \bmod G_{\alpha+\beta}^+$$

and

$$(x, yv) \equiv (x, y) \bmod G_{\alpha+\beta}^+.$$

Indeed we have

$$\overline{(xu, y)} = \overline{(x, y)^u} + \overline{(u, y)}$$

by 1.3(6), and since $w((u, y)) \geq w(u) + w(y) > \alpha + \beta$, $\overline{(u, y)} = 0$ so

$$\overline{(xu, y)} = \overline{(x, y)^u} = \overline{(x, y)}.$$

Similarly by 1.3(5),

$$\overline{(x, yv)} = \overline{(x, v)} + \overline{(x, y)^v} = \overline{(x, y)}.$$

For bilinearity, 1.3(6) again gives

$$\overline{(xx', y)} = \overline{(x, y)^{x'}} + \overline{(x', y)} = \overline{(x, y)} + \overline{(x', y)}$$

and similarly

$$\overline{(x, y'y)} = \overline{(x, y)} + \overline{(x, y')^y} = \overline{(x, y)} + \overline{(x, y')}.$$

□

Proposition 1.4.2

The maps $\bar{c}_{\alpha,\beta}$ can be extended by linearity to $c\colon \operatorname{gr} G \times \operatorname{gr} G \to \operatorname{gr} G$, defining a Lie algebra structure on $\operatorname{gr} G$.

Proof. For $\xi \in \operatorname{gr}_\alpha G$, $\eta \in \operatorname{gr}_\beta G$ we will write $[\xi, \eta]$ for $\bar{c}_{\alpha,\beta}(\xi, \eta)$. To show that $[\omega, \omega] = 0$ for $\omega = \sum_\alpha \omega_\alpha \in \operatorname{gr} G$, it is enough to show $[\omega_\alpha, \omega_\beta] = -[\omega_\beta, \omega_\alpha]$ for all α, β. Let $x_\alpha \in G_\alpha$ be such that $\overline{x}_\alpha = \omega_\alpha$. Then

$$[\omega_\alpha, \omega_\beta] = \overline{(x_\alpha, x_\beta)} = \overline{(x_\beta, x_\alpha)^{-1}} = -[\omega_\beta, \omega_\alpha].$$

It suffices to show the Jacobi identity for $\xi \in \operatorname{gr}_\alpha G$, $\eta \in \operatorname{gr}_\beta G$, $\zeta \in \operatorname{gr}_\gamma G$. This follows from 1.3(7): taking representatives $x \in G_\alpha$, $y \in G_\beta$, $z \in G_\gamma$, we have

$$[\xi, [\eta, \zeta]] + [\eta, [\zeta, \xi]] + [\zeta, [\xi, \eta]] = \overline{(x^y, (y, z))(y^z, (z, x))(z^x, (x, y))} = \overline{1},$$

where we have used that $\overline{x^y} = \overline{x} = \xi$, and so on. □

1.5 Integral filtrations

We say a filtration $w\colon G \to \mathbb{R}_+ \cup \{+\infty\}$ is *integral* if its image lies in $\mathbb{N} \cup \{+\infty\}$.

Proposition 1.5.1

Let G be a group. There is a one-to-one correspondence between integral filtrations of G and decreasing sequence $\{G_n\}_{n\in\mathbb{N}}$ of subgroups such that (i) $G_1 = G$ and (ii) $(G_n, G_m) \subset G_{n+m}$.

Proof. For $n \in \mathbb{N}$, $G_n = \{g \in G : w(g) \geq n\}$ is the desired decreasing sequence.

Conversely, given a decreasing sequence, define an integral filtration by $w(x) = \sup_{x \in G_n}\{n\}$. Since every subgroup contains 1, we have $w(1) = +\infty$. Moreover $w(x) = w(x^{-1})$.

Suppose $w(x) = n$, $w(y) = m$, so that $x \in G_n$, $y \in G_m$. Without loss of generality $n \leq m$, so $G_m \subset G_n$. Then $xy^{-1} \in G_n$, so

$$w(xy^{-1}) \geq n = \inf\{w(x), w(y)\}.$$

This formally doesn't make sense if $m = +\infty$, but the argument is the same.

Finally, $(x, y) \in (G_n, G_m) \subset G_{n+m}$ means $w((x, y)) \geq w(x) + w(y)$. □

The following example will ground us back in group theory.

Example 1.5.2. Let $G_1 := G$ and $G_{n+1} := (G, G_n)$. Then $\{G_n\}$ is a decreasing sequence of subgroups satisfying (i) and (ii), called the *descending central series.* To see (ii), the base case is by definition and by induction

$$\begin{aligned}(G_n, G_m) &= ((G, G_{n-1}), G_m)\\ &\subset (G, (G_{n-1}, G_m))(G_{n-1}, (G, G_m))\\ &\subset (G, G_{n+m-1})(G_{n-1}, G_{m+1})\\ &\subset G_{n+m} \cdot G_{n+m}\\ &= G_{n+m}.\end{aligned}$$

The descending central series is in some sense initial among decreasing sequences satisfying (i) and (ii). More precisely, if H_n is such a sequence then $H_n \supset G_n$ for any n. Indeed, the base case is again by definition and we inductively have

$$H_{n+1} \supset (H_1, H_n) \supset (G, G_n) = G_{n+1}.$$

1.6 Filtrations in GL_n

Let k be a field with an ultrametric absolute value; that is, a function $|\cdot| \colon k \to \mathbb{R}_{\geq 0}$ such that

(i) $|x| = 0$ if and only if $x = 0$.

(ii) $|xy| = |x||y|$.

(iii) $|x + y| \leq \max\{|x|, |y|\}$.

For example if $v \colon k \to \mathbb{R} \cup \{+\infty\}$ is a valuation; that is

(i) $v(x) = +\infty$ if and only if $x = 0$.

(ii) $v(xy) = v(x) + v(y)$.

(iii) $v(x + y) \geq \min(v(x), v(y))$.

then for $a > 1$, the function $|x| = a^{-v(x)}$ is an ultrametric absolute value. Here we use the convention $a^{-\infty} = 0$. Let A_v be the valuation ring of k with respect to v, let $\mathfrak{m}_v$ be its maximal ideal, and $k(v) = A_v/\mathfrak{m}_v$ its residue field. Let $n \in \mathbb{N}$. Let

$$G := \{g = (g_{ij}) \in \mathrm{GL}_n(A_v) : g_{ij} \equiv \delta_{ij} \bmod \mathfrak{m}_v\}.$$

Equivalently, $g = 1 + x$ where $x \in M_{n\times n}(\mathfrak{m}_v)$, or

$$G = \ker\{\mathrm{GL}_n(A_v) \to \mathrm{GL}_n(k(v))\},$$

which exhibits G as a group. Note also that a valuation v on k yields a map $v\colon M_n(k) \to \mathbb{R}$ by $v(x_{ij}) = \inf\{v(x_{ij})\}$. This gives a map $w\colon G \to \mathbb{R}_+ \cup \{+\infty\}$ by $w(1+x) = v(x)$.

Proposition 1.6.1

w is a filtration on G.

Proof. Clearly $w(1) = v(0) = +\infty$. Recall that (ii) in the definition of a filtration is equivalent to G_λ being a subgroup of G. If $\mathfrak{a}_\lambda = \{x \in k : v(x) \geq \lambda\}$ then G_λ is the kernel of the canonical homomorphism

$$\mathrm{GL}_n(A_v) \to \mathrm{GL}_n(A_v/\mathfrak{a}_\lambda),$$

hence G_λ is a subgroup. Similarly (iii) is equivalent to $(G_\lambda, G_\mu) \subset G_{\lambda+\mu}$. Consider $g = 1 + x \in G_\lambda$, $h = 1 + y \in G_\mu$. Then

$$\begin{aligned} hg &= 1 + x + y + yx \\ gh &= 1 + x + y + xy, \end{aligned}$$

where xy, $yx \in M_n(\mathfrak{a}_{\lambda+\mu})$. Thus $hg = gh \bmod \mathrm{GL}_n(A_v/\mathfrak{a}_{\lambda+\mu})$, so we have (iii). $\square$

1.7 The universal enveloping algebra

Let $\mathfrak{g}$ be a Lie algebra over a commutative ring k. Recall that any associative algebra admits a Lie structure by

$$[x, y] = xy - yx.$$

Definition 1.7.1

A *universal enveloping algebra* of $\mathfrak{g}$ is a Lie algebra homomorphism $\epsilon\colon \mathfrak{g} \to U\mathfrak{g}$ where $U\mathfrak{g}$ is an associative algebra with unit, satisfying the following universal property: if A is any associative algebra with unit and $\alpha\colon \mathfrak{g} \to A$ is any Lie algebra homomorphism, then there exists a unique associative algebra homomorphism $\phi\colon U\mathfrak{g} \to A$ such that the following diagram commutes:

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\alpha"] \arrow[d, "\epsilon"'] & A \\ U\mathfrak{g}. \arrow[ur, dashed, "\exists!"'] & \end{tikzcd}$$

As always, $U\mathfrak{g}$ is unique up to unique isomorphism. To prove it exists, consider the tensor algebra

$$T\mathfrak{g} = \sum_{n=0}^{\infty} \mathfrak{g}^{\otimes n}$$

of $\mathfrak{g}$. This satisfies the universal property

$$\mathrm{Hom}_{k\text{-}\mathsf{Mod}}(\mathfrak{g}, A) \cong \mathrm{Hom}_{\mathsf{Alg}_k}(T\mathfrak{g}, A).$$

To upgrade this to the Lie structures, we quotient $T\mathfrak{g}$ be the ideal I generated by

$$[x,y] - x \otimes y + y \otimes x \quad \text{for} \quad x, y \in \mathfrak{g}.$$

Theorem 1.7.2

Let $\epsilon\colon \mathfrak{g} \to U\mathfrak{g}$ be the composition $\mathfrak{g} \to T^1\mathfrak{g} \to T\mathfrak{g} \to T\mathfrak{g}/I$. Then $(T\mathfrak{g}/I, \epsilon)$ is a universal enveloping algebra of $\mathfrak{g}$.

Proof. Let $\alpha\colon \mathfrak{g} \to A$ be a Lie algebra homomorphism. Since it is k-linear, it is a unique homomorphism $\psi\colon T\mathfrak{g} \to A$. Clearly $\psi(I) = 0$ as

$$\psi([x,y] - x \otimes y + y \otimes x) = [\psi(x), \psi(y)] - [\psi(x), \psi(y)] = 0,$$

so it descends to a Lie algebra homomorphism $U\mathfrak{g} \to A$. □

Let E be a k-module with a bilinear map $\mathfrak{g} \times E \to E$ such that $[x,y]e = x(ye) - y(xe)$. We call E a $\mathfrak{g}$-module. Then the natural map $\mathfrak{g} \to \mathrm{End}(E,E)$ is a Lie homomorphism. By the universal property of $U\mathfrak{g}$, it induces an algebra homomorphism $U\mathfrak{g} \to \mathrm{End}(E,E)$, making E a left $U\mathfrak{g}$-module. This association is an equivalence between the category of $\mathfrak{g}$-modules and the category of left $U\mathfrak{g}$-modules.

The next result summarizes functoriality of $U\mathfrak{g}$.

Proposition 1.7.3 (1) If $\mathfrak{g} = \varinjlim \mathfrak{g}_i$ then $U\mathfrak{g} = \varinjlim U\mathfrak{g}_i$.

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

(3) If $k \subset k'$ and $\mathfrak{g}' = \mathfrak{g} \otimes_k k'$, then $U\mathfrak{g}' = U\mathfrak{g} \otimes_k k'$.

Proof. (1) We have

$$\mathrm{Hom}_{\mathsf{Alg}}(U(\varinjlim \mathfrak{g}_i), A) \cong \mathrm{Hom}_{\mathsf{Lie}}(\varinjlim \mathfrak{g}_i, A) \cong \varinjlim \mathrm{Hom}_{\mathsf{Lie}}(\mathfrak{g}_i, A) \cong \varinjlim \mathrm{Hom}(U\mathfrak{g}_i, A).$$

(2) Consider $\epsilon_i\colon \mathfrak{g}_i \to U\mathfrak{g}_i$ and

$$\begin{aligned} f\colon \mathfrak{g}_1 \times \mathfrak{g}_2 &\longrightarrow U\mathfrak{g}_1 \otimes U\mathfrak{g}_2 \\ (x_1, x_2) &\longmapsto \epsilon_1(x_1) \otimes 1 + 1 \otimes \epsilon_2(x_2). \end{aligned}$$

This is a Lie homomorphism since $\mathfrak{g}_1$ and $\mathfrak{g}_2$ commute, so it induces an algebra homomorphism $\psi\colon U\mathfrak{g} \to U\mathfrak{g}_1 \otimes U\mathfrak{g}_2$.

In the other direction we have $\mathfrak{g}_i \to \mathfrak{g} \to U\mathfrak{g}$, which induce $\phi_i\colon U\mathfrak{g}_i \to U\mathfrak{g}$, and

$$\phi_1(x_1)\phi_2(x_2) \ = \phi_2(x_2)\phi_1(x_1).$$

Thus

$$\begin{aligned} \phi\colon U\mathfrak{g}_1 \otimes U\mathfrak{g}_2 &\longrightarrow U\mathfrak{g} \\ \phi(x_1 \otimes x_2) &\longmapsto \phi_1(x_1)\phi_2(x_2) \end{aligned}$$

is the inverse of ψ.

(3) We have
$$T(g \otimes_k k') \cong T\mathfrak{g} \otimes_k k',$$
and if $U\mathfrak{g} = T\mathfrak{g}/I$ then $U\mathfrak{g}' = T(g \otimes_k k')/I'$ where $I' = I \otimes_k k'$,
$$U(\mathfrak{g} \otimes_k k') \cong U\mathfrak{g} \otimes_k k'.$$

□

1.7.1 The symmetric algebra

We may view any k-module $\mathfrak{g}$ as an abelian Lie algebra; that is $[x, y] = 0$. In this case $U\mathfrak{g}$ is called the symmetric algebra of $\mathfrak{g}$, denoted by $S\mathfrak{g}$. Concretely, it is the quotient of $T\mathfrak{g}$ by the ideal generated by the elements
$$x_1 \otimes \cdots \otimes x_n - x_{\sigma(1)} \otimes \cdots \otimes x_{\sigma(n)},$$
for $\sigma \in S_n$.

A case of special interest is the free k-module with basis $(e_i)_{i\in I}$. Let $\epsilon\colon \mathfrak{g} \to k[X_i]$ be the homomorphism $e_i \mapsto X_i$. Then $(\epsilon, k[X_i])$ has the universal property of a universal enveloping algebra, in the sense that it is k-linear, $\epsilon(x)\epsilon(y) = \epsilon(y)\epsilon(x)$, and for any k-linear $f\colon \mathfrak{g} \to A$ with $f(x)f(y) = f(y)f(x)$ there exists an algebra homomorphism $f^*\colon k[X_i] \to A$ such that $f^* \circ \epsilon = f$. Explicitly, if $P(x_i) \in k[X_i]$ then $f^*(P) = P(f(e_i))$. In this way we identify $S\mathfrak{g}$ with a polynomial algebra $k[X_i]$. If I is totally ordered, then the monomials
$$e_{i_1} \cdots e_{i_n}$$
for $i_1 \leq i_2 \leq \cdots \leq i_n$, $n \geq 0$ form a basis for $S\mathfrak{g}$.

1.7.2 Filtration of $U\mathfrak{g}$

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 $\epsilon(x_1)\cdots\epsilon(x_m)$, for $m \leq n$ and $x_i \in \mathfrak{g}$. For example $U_0\mathfrak{g} = k$, $U_1\mathfrak{g} = k \oplus \epsilon\mathfrak{g}$, and so on. As usual, define $\operatorname{gr}_n U\mathfrak{g} = U_n\mathfrak{g}/U_{n-1}\mathfrak{g}$ and
$$\operatorname{gr} U\mathfrak{g} = \sum_{n=0}^{\infty} \operatorname{gr}_n U\mathfrak{g}.$$
The map
$$\begin{aligned} U_p\mathfrak{g} \times U_q\mathfrak{g} &\longrightarrow U_{p+q}\mathfrak{g} \\ (a, b) &\longmapsto ab \end{aligned}$$
descends to a bilinear map
$$\operatorname{gr}_p U\mathfrak{g} \times \operatorname{gr}_q U\mathfrak{g} \longrightarrow \operatorname{gr}_{p+q} U\mathfrak{g}.$$
This associates to $U\mathfrak{g}$ an associative unital graded algebra $\operatorname{gr} U\mathfrak{g}$.

Proposition 1.7.4

Let algebra $\operatorname{gr} U\mathfrak{g}$ is generated by the image of $\mathfrak{g}$ under the universal map $\epsilon\colon \mathfrak{g} \to U\mathfrak{g}$.

Proof. Let $\alpha \in \operatorname{gr}_n U\mathfrak{g}$ and let $a \in U_n\mathfrak{g}$ be a representative. Then

$$a = \sum_{m_\mu \leq n} \lambda_\mu \epsilon(x_1^{(\mu)}) \cdots \epsilon(x_{m_\mu}^{\mu}).$$

Thus

$$\alpha = \sum_{m_\mu = n} \lambda_\mu \overline{\epsilon(x_1^{(\mu)})} \cdots \overline{\epsilon(x_{m_\mu}^{\mu})}.$$

□

Proposition 1.7.5

$\operatorname{gr} U\mathfrak{g}$ is a commutative algebra.

Proof. In light of the previous proposition it suffices to show that $\overline{\epsilon(x)}$ and $\overline{\epsilon(y)}$ commute in $\operatorname{gr}_2 U\mathfrak{g}$ for all $x, y \in \mathfrak{g}$. Since ϵ is a Lie homomorphism,

$$\epsilon(x)\epsilon(y) - \epsilon(y)\epsilon(x) = \epsilon([x, y]) \in U_1\mathfrak{g}$$

so $\epsilon(x)\epsilon(y) \equiv \epsilon(y)\epsilon(x) \bmod U_1\mathfrak{g}$. □

By the universal property of the symmetric algebra $S\mathfrak{g}$, the canonical map $\mathfrak{g} \to \operatorname{gr} U\mathfrak{g}$ extends to a homomorphism

$$\iota \colon S\mathfrak{g} \longrightarrow \operatorname{gr} U\mathfrak{g}.$$

By Proposition 1.7.4, ι is surjective. The question of when it is injective brings us our first big theorem.

Theorem 1.7.6 (Poincaré–Birkhoff–Witt)

Let $\mathfrak{g}$ be a free k-module. Then $\iota \colon S\mathfrak{g} \to \operatorname{gr} U\mathfrak{g}$ is an isomorphism.

To prove the Poincaré–Birkhoff–Witt theorem, we will need:

Lemma 1.7.7

?? Let $\mathfrak{g}$ be a free k-module with basis $(x_i)_{i \in I}$, where I is totally ordered. The monomials

$$\epsilon(x_{i_1}) \cdots \epsilon(x_{i_m}) \quad \text{for} \quad i_1 \leq \cdots \leq i_m, m \leq n$$

generate $U^n\mathfrak{g}$ as a k-module.

Proof. By induction on n. The base case $n = 0$ is vacuous. For $n > 0$, let $a \in U^n\mathfrak{g}$. Then $\overline{a} \in \operatorname{gr}^n U\mathfrak{g}$ is a degree n polynomial in the $\overline{\epsilon(x_i)}$, hence a is a linear combination of the $\epsilon(x_{i_1}) \cdots \epsilon(x_{i_n})$ with an element $a_1 \in U^{n-1}\mathfrak{g}$. By the induction hypothesis a_1 is a linear combination of the monomials $\epsilon(x_{i_1}) \cdots \epsilon(x_{i_m})$ for $i_1 \leq \cdots \leq i_m$, $m < n$. □

Lemma 1.7.8

ι is an isomorphism if and only if $U\mathfrak{g}$ has basis

$$\epsilon(x_{i_1})\cdots\epsilon(x_{i_n}) \quad \text{for} \quad i_1 \le \cdots \le i_n, n \ge 0.$$

Proof. Let $M = (i_1, \ldots, i_m)$ an increasing sequence, write $x_M = \epsilon(x_{i_1})\cdots\epsilon(x_{i_m})$, and denote its length by $\ell(M) = m$. For $n \ge 0$, the x_M with $\ell(M) = n$ lie in $U_n\mathfrak{g}$, and their images $\overline{x}_M \in \mathrm{gr}_n U\mathfrak{g}$ are the images of the monomial basis elements of $S^n\mathfrak{g}$ under ι. So injectivity of ι is equivalent to the linear independence of the x_M for $\ell(M) = n$, modulo $U_{n-1}\mathfrak{g}$. That is, there exist no c_M not all 0 such that

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

By **??**, this is equivalent t

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

with some nonzero c_M with $\ell(M) = n$. But any nontrivial linear combination takes this form, so the lemma is proven. $\square$

Having reduced to the statement of Lemma 1.7.8, we can now prove the Poincaré–Birkhoff–Witt theorem:

Proof of Theorem 1.7.6. We can and will assume I is well-ordered. Let V be the free module over k generated by $\{z_M\}$ for $M = (i_1, \ldots, i_n)$ increasing with $n \ge 0$. If $i \in I$ and $M = (i_1, \ldots, i_n)$, we say $i \le M$ if $i \le i_1$, in which case we define $iM = (i, i_1, \ldots, i_n)$.

We will define a $\mathfrak{g}$-module on V so that $x_i Z_M = Z_{iM}$ for $i \le M$. Firstly, define a k-bilinear map

$$\mathfrak{g} \times V \longrightarrow V$$

by defining $x_i Z_M$ inductively: we may assume $x_j Z_N$ is defined when $\ell(N) < \ell(M)$ and when $j < i$ and $\ell(N) = \ell(M)$. We may furthermore assume that their definition satisfies the following property:

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

Under these assumption, let

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

It remains to show that this makes V a $\mathfrak{g}$-module, that is

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

It suffices, by linearity, to show that

$$x_i x_j Z_N - x_j x_i Z_N = [x_i, x_j] Z_N.$$

Both sides are skew-symmetric and vanish for $i = j$, so without loss of generality $i > j$. If $j \le N$, then $x_j Z_N = Z_{jN}$ and by the second case of the definition of $x_i Z_M$, we have the desired result. Otherwise if $N = kL$ with $i > j > k$, then we must show that

$$x_i x_j x_k Z_L - x_j x_i x_k Z_L = [x_i, x_j] x_k Z_L.$$

By induction on $\inf(i,j)$ the equation holds under cyclic permutations of ijk. Also, by induction on $\ell(N)$ we have $xyZ_L = yxZ_L + [x,y]Z_L$ for $x, y \in \mathfrak{g}$. Thus

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

By adding together the cyclically-permuted equations, we get an equation of the form

$$\sum = \sum + J(x_i,x_j,x_k)Z_L,$$

where J denote the Jacobi identity, which vanishes, making the equation true.

Consider $Z_\varnothing \in V$; for all M we have $x_M Z_\varnothing = Z_M$. Indeed, we prove this by induction. If $\ell(M) = 0$ then $x_M = 1$ so the result is clear. If $\ell(M) > 0$, then $M = iN$ for some $i \leq N$, and $x_M = x_i x_N$ so that

$$x_M Z_\varnothing = x_i Z_N = Z_{iN} = Z_M.$$

Finally if $\sum c_M x_M = 0$, then

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

which implies $c_M = 0$, as desired. □

Corollary 1.7.9

Let $\mathfrak{g}$ be a free k-module. Then $\epsilon\colon \mathfrak{g} \to U\mathfrak{g}$ is injective.

Proof. $\mathfrak{g} \to S\mathfrak{g}$ is injective. □

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

Corollary 1.7.10

Let $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2$ where $\mathfrak{g}_i$ are subalgebras and free as k-modules. Then the map

$$\begin{aligned} U\mathfrak{g}_1 \otimes U\mathfrak{g}_2 &\longrightarrow U\mathfrak{g} \\ u_1 \otimes u_2 &\longmapsto u_1u_2 \end{aligned}$$

is a k-linear isomorphism.

Proof. Let $(x_i)_{i\in I}$ and $(y_j)_{j\in J}$ be bases for $\mathfrak{g}_1$ and $\mathfrak{g}_2$, respectively. Then $\{(x_i),(y_j)\}$ is a basis for $\mathfrak{g}$. Let $I \cup J$ be totally ordered so that every $x_i < y_j$. By Lemma 1.7.8, the collections of monomials

$$\begin{aligned} &\{\epsilon(x_{i_1})\cdots\epsilon(x_{i_n})\} \\ &\{\epsilon(y_{j_1})\cdots\epsilon(y_{j_m})\} \\ &\{\epsilon(x_{i_1})\cdots\epsilon(x_{i_n})\epsilon(y_{j_1})\cdots\epsilon(y_{j_m})\} \end{aligned}$$

for $i_1 \leq \cdots \leq i_n < j_1 \leq \cdots \leq j_m$, form bases for $U\mathfrak{g}_1$, $U\mathfrak{g}_2$, and $U\mathfrak{g}$, respectively. Thus $U\mathfrak{g}_1 \otimes U\mathfrak{g}_2 \to U\mathfrak{g}$ given by $u_1 \otimes u_2 \mapsto u_1u_2$ is a bijection on the bases. □

Notice that we also have an isomorphism

$$\operatorname{gr} U\mathfrak{g}_1 \otimes \operatorname{gr} U\mathfrak{g}_2 \xrightarrow{\sim} \operatorname{gr} U\mathfrak{g},$$

as $\operatorname{gr} U\mathfrak{g}_i = S\mathfrak{g}_i$ and $\operatorname{gr} U_\mathfrak{g} = S\mathfrak{g} \cong S\mathfrak{g}_1 \otimes S\mathfrak{g}_2$.

1.7.3 The diagonal map

Let $\mathfrak{g}$ be a free k-module. The diagonal map $\Delta\colon \mathfrak{g} \to \mathfrak{g} \times \mathfrak{g}$ induces a homomorphism of associative algebras

$$\Delta\colon U\mathfrak{g} \longrightarrow U\mathfrak{g} \otimes U\mathfrak{g}.$$

This is uniquely characterized by the following proeprties:

(i) Δ is an algebra homomorphism.

(ii) $\Delta x = x \otimes 1 + 1 \otimes x$ for $x \in \mathfrak{g}$.

We say $\alpha \in U\mathfrak{g}$ is *primitive* if $\Delta\alpha = \alpha \otimes 1 + 1 \otimes \alpha$. In other words, $x \in \mathfrak{g}$ is primitive.

Theorem 1.7.11

Let k be a torsion free $\mathbb{Z}$-module and $\mathfrak{g}$ a free k-module. Then $\mathfrak{g}$ is the set of primitive elements of $U\mathfrak{g}$.

Proof. First suppose $\mathfrak{g}$ is abelian. Then $U\mathfrak{g}$ is the polynomial algebra $k[X_i]$ where the indeterminates X_i correspond to the basis elements x_i of $\mathfrak{g}$. The diagonal is a homomorphism $k[X_i] \to k[X_i', X_i'']$ where X_i' is $X_i \otimes 1$ and X_i'' is $1 \otimes X_i$. Concretely,

$$\Delta f(X_i', X_i'') = f(X_i' + X_i'')$$

because $X_i \mapsto X_i' + X_i''$ for each i. Thus the primitive elements $f(x) \in k[X_i]$ are precisely the elements such that $f(X_i' + X_i'') = f(X_i') + f(X_i'')$. If f satisfies this property, then so does 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 $(2^n - 2)f = 0$. But k is $\mathbb{Z}$-torsion free, so $f = 0$ if $n \neq 1$. So the only additive polynomials are the linear homogeneous polynomials.

In general, $\Delta\colon U\mathfrak{g} \to U\mathfrak{g} \otimes U\mathfrak{g}$ induces

$$\operatorname{gr}\Delta\colon\ \operatorname{gr} U\mathfrak{g} \to \mathfrak{g}(U\mathfrak{g} \otimes U\mathfrak{g}) \cong \mathfrak{g}U(\mathfrak{g} \oplus \mathfrak{g}) \cong \operatorname{gr} U\mathfrak{g} \otimes \operatorname{gr} U\mathfrak{g}.$$

But $\operatorname{gr} U\mathfrak{g} \cong S\mathfrak{g}$, and $\operatorname{gr}\Delta$ agrees with the previous $S\mathfrak{g} \to S\mathfrak{g} \otimes S\mathfrak{g}$.

Let $x \in U_n\mathfrak{g}$ and let $\overline{x} \in \operatorname{gr}_n U\mathfrak{g}$ be its image. If x is primitive, then $\overline{x}$ is primitive for $\operatorname{gr}\Delta$, hence for $n > 1$, we have $\overline{x} = 0$ by the previous case. Iteratively, we conclude that $x \in U_1\mathfrak{g}$, so $x = \lambda + y$ for $\lambda \in k$, $y \in \mathfrak{g}$. 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 $2\lambda = \lambda$, so $\lambda = 0$, showing that $x \in fg$. $\square$