Enveloping Algebras and The Poincaré-Birkhoff-Witt Theorem

Rahbar Virk
Department of Mathematics
University of Wisconsin
Madison, WI 53706, USA
virk@math.wisc.edu

A Lie algebra $\mathfrak{g}$ is not an algebra in the sense that the Lie bracket is not associative. We would like to find an associative algebra $U(\mathfrak{g})$ such that the modules for $\mathfrak{g}$ are the same as those for $U(\mathfrak{g})$.

Let $\mathfrak{g}$ be a Lie algebra and let T be the tensor algebra of the vector space $\mathfrak{g}$. Recall that

$$T = T^0 \oplus T^1 \oplus \cdots \oplus T^n \oplus \cdots,$$

where $T^n = \mathfrak{g} \otimes \mathfrak{g} \otimes \cdots \mathfrak{g}$ (n times); in particular $T^0 = k.1$ and $T^1 = \mathfrak{g}$; the product in T is simply tensor multiplication. Let J be the two sided ideal of T generated by the tensors

$$x \otimes y - [x, y],$$

where $x, y \in \mathfrak{g}$. The associative algebra T/J is called the *enveloping algebra* (or sometimes the *universal enveloping algebra*) of $\mathfrak{g}$ and is denoted by $U(\mathfrak{g})$. The composite mapping σ of the canonical mappings $\mathfrak{g} \to T \to U(\mathfrak{g})$ is termed the canonical mapping of $\mathfrak{g}$ into $U(\mathfrak{g})$; observe that for all $x, y \in \mathfrak{g}$ we have that

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

We denote the canonical image in $U(\mathfrak{g})$ of $T^0 + T^1 + \cdots + T^q$ by $U_q(\mathfrak{g})$.

Lemma 1.1. *Let σ be the canonical mapping of $\mathfrak{g}$ into $U(\mathfrak{g})$, let A be an algebra with unity, and let τ be a linear mapping of $\mathfrak{g}$ into A such that*

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

for all $x, y \in \mathfrak{g}$. There exists one and only one homomorphism τ' of $U(\mathfrak{g})$ into A such that $\tau'(1) = 1$ and $\tau' \circ \sigma = \tau$.

Proof. Note that $U(\mathfrak{g})$ is generated by 1 and $\sigma(\mathfrak{g})$ so τ' must be unique. Now let φ be the unique homomorphism of T into A that extends τ and such that $\varphi(1) = 1$. For $x, y \in \mathfrak{g}$, we have

$$\varphi(x \otimes y - y \otimes x - [x, y]) = \tau(x)\tau(y) - \tau(y)\tau(x) - \tau([x, y]) = 0,$$

hence $\varphi(J) = 0$ and thus φ gives us a homomorphism τ' of $U(\mathfrak{g})$ into A such that $\tau'(1) = 1$ and $\tau' \circ \sigma = \tau$. □

Lemma 1.2. *Let $a_1, \ldots, a_p \in \mathfrak{g}$, σ the canonical mapping of $\mathfrak{g}$ into $U(\mathfrak{g})$, and π be a permutation of $\{1, \ldots, p\}$. Then*

$$\sigma(a_1) \cdots \sigma(a_p) - \sigma(a_{\pi(1)}) \cdots \sigma(a_{\pi(p)}) \in U_{p-1}(\mathfrak{g}).$$

Proof. It suffices to prove the statement when π is the transposition of j and $j+1$. But observe that

$$\sigma(a_j)\sigma(a_{j+1}) - \sigma(a_{j+1})\sigma(a_j) = \sigma([a_j, a_{j+1}]).$$

The lemma now follows. □

We now assume $\mathfrak{g}$ to a be finite dimensional Lie algebra and fix a basis $(x_1, \ldots, x_n)$ for $\mathfrak{g}$. We denote the canonical image of x_i in $U(\mathfrak{g})$ by y_i. For every finite sequence $I = (i_1, \ldots, i_p)$ of integers between 1 and n, we set $y_I = y_{i_1} y_{i_2} \ldots y_{i_p}$.

Lemma 1.3. *The y_I, for all increasing sequences I of length $\leq p$, generate the vector space $U_p(\mathfrak{g})$.*

Proof. Clearly the vector space $U_p(\mathfrak{g})$ is generated by y_I, for *all* sequences I of length $\leq p$. But now the required statement follows by applying the previous lemma. □

Let P be the algebra $k[z_1, \ldots, z_n]$ of polynomials in n indeterminates $z_1, \ldots, z_n$. For every $i \in \mathbb{Z}_{\geq 0}$, let P_i be the set of elements of P of degree $\leq i$. If $I = (i_1, \ldots, i_p)$ is a sequence of integers between 1 and n, we set $z_I = z_{i_1} z_{i_2} \cdots z_{i_p}$.

Lemma 1.4. *For every integer $p \geq 0$, there exists a unique linear mapping f_p of the vector space $\mathfrak{g} \otimes P_p$ into P which satisfies the following conditions:*

(A_p) *$f_p(x_i \otimes z_I) = z_i z_I$ for $i \leq I$, $z_I \in P_p$;*

(B_p) *$f_p(x_i \otimes z_I) - z_i z_I \in P_q$ for $z_I \in P_q$, $q \leq p$;*

(C_p) *$f_p(x_i \otimes f_p(x_j \otimes z_J)) = f_p(x_j \otimes f_p(x_i \otimes z_J)) + f_p([x_i, x_j] \otimes z_J)$ for $z_J \in P_{p-1}$. [The terms in (C_p) are meaningful by virtue of (B_p)].*

Moreover, the restriction of f_p to $\mathfrak{g} \otimes P_{p-1}$ is f_{p-1}.

Proof. For $p = 0$ the mapping given by $f_0(x_i \otimes 1) = z_i \otimes 1$, satisfies the conditions $(A_0), (B_0)$ and (C_0), furthemore it is clear that the condition (A_0) forces this to be our linear mapping. Proceeding by induction, assume the existence and uniqueness of f_{p-1}. If f_p exists then f_p restricted to $\mathfrak{g} \otimes P_{p-1}$ satisfies $(A_{p-1}), (B_{p-1}), (C_{p-1})$ and is hence equal to f_{p-1}. Thus, to prove our claim it suffices to show that f_{p-1} has a unique linear extension f_p to $\mathfrak{g} \otimes P_p$ which satisfies $(A_p), (B_p), (C_p)$. Note that P_p is generated by z_I for an increasing sequence I of p elements. Thus, we must define $f_p(x_i \otimes z_I)$ for such a sequence I. If $i \leq I$, then the choice is dictated by (A_p). Otherwise we can write I as (j, J) where $j < i$ and $j \leq J$. Then we must have that

$$\begin{aligned} f_p(x_i \otimes z_I) &= f_p(x_i \otimes f_{p-1}(z_j \otimes z_J)) && \text{from } (A_{p-1}) \\ &= f_p(x_j \otimes f_{p-1}(x_i \otimes z_J)) + f_{p-1}([x_i, x_j] \otimes z_J) && \text{from } (C_p). \end{aligned}$$

Now $f_{p-1}(x_i \otimes z_J) = z_i z_J + w$, with $w \in P_{p-1}$ (from (B_{p-1})). Hence

$$\begin{aligned} f_p(x_j \otimes f_{p-1}(x_i \otimes z_J)) &= z_j z_i z_J + f_{p-1}(x_j \otimes w) \qquad \text{from } (A_p) \\ &= z_j z_I + f_{p-1}(x_j \otimes w). \end{aligned}$$

The above defines a unique linear extension f_p of f_{p-1} to $\mathfrak{g} \otimes P_p$, and by construction this extension satisfies (A_p) and (B_p). All that remains to show is that f_p when defined this way satisfies (C_p).
Observe that if $j \leq i$ and $j \leq J$ then (C_p) is satisfied by construction. Since $[x_j, x_i] = -[x_i, x_j]$, it is also satisfied if $i < j$ and $i \leq J$. Since (C_p) is trivially satisfied if $i = j$, we see that (C_p) is satisfied if $i \leq J$ or $j \leq J$. Otherwise, $J = (k, K)$, where $k \leq K$, $k < i$ and $k < j$. For the sake of brevity we will write $f_p(x \otimes z) = xz$ for $x \in \mathfrak{g}$ and $z \in P_p$. Then from the induction hypothesis we have that

$$x_i z_J = x_j(x_k z_K) = x_k(x_j z_K) + [x_j, x_k] z_K \qquad (*)$$

Now $x_j z_K$ is of the form $z_j z_K + w$, where $w \in P_{p-2}$. As $k \leq K$ and $k < j$ we can apply (C_p) to $x_i(x_k(z_j z_K))$, and to $x_i(x_k w)$ from the induction hypothesis and hence also to $x_i(x_k(x_j z_K))$; using $(*)$ this then gives us that

$$x_i(x_j z_J) = x_k(x_i(x_j z_K)) + [x_i, x_k](x_j z_K) + [x_j, x_k](x_i z_K) + [x_i, [x_j, x_k]] z_k.$$

Interchanging i and j, this gives us that

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

as required. □

Lemma 1.5. *The y_I, for every increasing sequence I, form a basis for the vector space $U(\mathfrak{g})$.*

Proof. By the previous lemma (whose notation we will also use), there exists a bilinear mapping f of $\mathfrak{g} \times P$ into P such that $f(x_i, z_I) = z_i z_I$ for $i \leq I$ and

$$f(x_i, f(x_i, z_J)) = f(x_j, f(x_i, z_J)) + f([x_i, x_j], z_J),$$

for all i, j, J. Thus, we have a representation ϱ of $\mathfrak{g}$ in P such that $\varrho(x_i) z_I = z_i z_I$ for $i \leq I$. From lemma 1.1 there exists a homomorphism φ of $U(\mathfrak{g}$ into $End((P)$ such that $\varphi(y_i) z_I = z_i z_I$ for $i \leq I$. We thus obtain that, if $i_1 \leq i_2 \leq \cdots \leq i_p$, then

$$\varphi(y_{i_1} y_{i_2} \cdots y_{i_p}) \cdot 1 = z_{i_1} z_{i_2} \cdots z_{i_p}.$$

Thus, the y_I are linearly independent, for I increasing. From lemma 1.3 they generate $U(\mathfrak{g})$ and hence form a basis. □

Proposition 1.6. *The canonical mapping of $\mathfrak{g}$ into $U(\mathfrak{g})$ is injective.*

Proof. This is immediate from the previous lemma. □

We thus have that $\mathfrak{g}$ is embedded in $U(\mathfrak{g})$, so we will identify every element of $\mathfrak{g}$ with its canonical image in $U(\mathfrak{g})$

Theorem 1.7 (Poincareé-Birkhoff-Witt)**.** *Let $(x_1, \ldots, x_n)$ be a basis for the vector space $\mathfrak{g}$. Then the $x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_n^{\lambda_n}$, where $\lambda_1, \ldots, \lambda_n \in \mathbb{N}$, form a basis for $U(\mathfrak{g})$.*

Proof. This is again immediate from the previous lemma. □

Taking the Poincaré-Birkhoff-Witt Theorem into account Lemma 1.1 can be restated as

Lemma 1.8. *Let A be an algebra with unity, τ a linear mapping of $\mathfrak{g}$ into A such that $\tau(x)\tau(y) - \tau(y)\tau(x) = \tau([x,y])$ for all $x, y \in \mathfrak{g}$. Then τ can be uniquely extended to a homomorphism of $U(\mathfrak{g})$ into A which transforms 1 into 1.*

Corollary 1.9. *Let V be a vector space, and $\mathcal{R}$ and $\mathcal{R}'$ the sets of representations of $\mathfrak{g}$ and $U(\mathfrak{g})$ in V respectively. For all $\varrho \in \mathcal{R}$, there exists one and only one $\varrho' \in \mathcal{R}'$ which extends ϱ, and the mapping $\varrho \to \varrho'$ is a bijection of $\mathcal{R}$ onto $\mathcal{R}'$.*