PBW THEOREM

JEREMY BOOHER

1. Reminders

Definition 1. The universal enveloping algebra for a Lie group $\mathfrak{g}$ is an algebra $U(\mathfrak{g})$ with map $\iota : \mathfrak{g} \to U(\mathfrak{g})$ such that for any map of Lie algebras $\phi : \mathfrak{g} \to A$ there is a unique map of algebras $\phi' : U(\mathfrak{g}) \to A$ with $\phi = \phi' \circ \iota$.

The representing object is the tensor algebra modulo the ideal generated by $x \otimes y - y \otimes x - [x, y]$ with the obvious map. We will prove

Theorem 2 (Poincaré-Birkhoff-Witt)**.** *For a Lie algebra $\mathfrak{g}$, Sym$(\mathfrak{g}) \simeq \mathrm{gr}(U(\mathfrak{g}))$.*

Note that $\mathfrak{g}$ need not be finite dimensional, and the characteristic of the base field may be nonzero.

2. PBW

Let $\{x_1, \ldots, x_n, \ldots\}$ be an ordered basis for $\mathfrak{g}$. Let y_i be the image of x_i in $U(\mathfrak{g})$ under the canonical map $\iota : \mathfrak{g} \to U(\mathfrak{g})$. For $I = (i_1, \ldots, i_n)$, let y_I denote $y_{i_1} \ldots y_{i_n} \in U(\mathfrak{g})$. Say $I \leq m$ if $i_j \leq m$ for all j. Call I increasing if $i_1 \leq i_2 \leq \ldots \leq y_n$.

Lemma 3. *The set of all y_I with I increasing and $I \leq n$ generates $U_n(\mathfrak{g})$.*

Proof. Let π be a permuation of n elements. I claim that

$$\iota(g_1) \ldots \iota(g_n) - \iota(g_{\pi(1)}) \ldots \iota(g_{\pi(n)}) \in U_{n-1}(\mathfrak{g})$$

which it suffices to check on transpositions flipping i and $i + 1$. Then

$$\iota(g_1) \ldots \iota(g_i)\iota(g_{i+1}) \ldots \iota(g_n) - \iota(g_1) \ldots \iota(g_{i+1})\iota(g_i) \ldots \iota(g_n) = \iota(g_1) \ldots \iota([g_i, g_{i+1}]) \ldots \iota(g_n) \in U_{n-1}$$

Now $U_n(\mathfrak{g})$ is generated by elements of the form $y_J = \iota(x_{j_1}) \ldots \iota(x_{j_n})$ where $J = (j_1, \ldots, j_n)$ is not necessarily increasing. Let π be the permutation with $\pi(j_1) \leq \pi(j_2) \ldots \leq \pi(j_n)$. Then

$$y_J = \iota(x_{j_1}) \ldots \iota(x_{j_n}) = \iota(x_{\pi(j_1)}) \ldots \iota(x_{\pi(j_n)}) + r$$

where the first term is increasing and the second is in $U_{n-1}(\mathfrak{g})$. Then by induction y_J is expressable in terms of y_I with I increasing and $I \leq n$. □

Now let P be the algebra of polynomials in variables $x_1 \ldots x_n \ldots$. To avoid confusion, I'll denote the variables as z_i instead to make clear which algebra the elements lie in. Filter P so P_n is the polynomials of degree at most n. Set $z_I = z_{i_1} \ldots z_{i_n}$ for $I = (i_1 \ldots i_n)$

Lemma 4. *For all n, there exists a unique function $f_n : \mathfrak{g} \otimes P_n \to P$ such that*
(A_n) $f_n(x_i \otimes z_I) = z_i z_I$ for $i \leq I$, $z_I \in P_n$.
(B_n) $f_n(x_i \otimes z_I) = z_i z_I \mod P_q$ for $z_I \in P_q$ and $q \leq n$.
(C_n) $f_n(x_i \otimes f_n(x_j \otimes z_J)) = f_n(x_j \otimes f_n(x_i \otimes z_J)) + f_n([x_i, x_j] \otimes z_J)$ for $z_J \in P_{n-1}$
Furthermore, the restriction of f_n to $\mathfrak{g} \otimes P_{n-1}$ is f_{n-1}.

Date: May 8, 2009.

Proof. First, note that condition C_n is actually well defined because $f_n(x_j \otimes z_j)$ is in P_n by condition B_n.

We will proceed by induction. The base case is when $n = 0$, in which case f_0 must map $x_i \otimes 1$ to z_i to satisfy A_0. Then conditions B_0 and C_0 are vacuously satisfied.

Now suppose we have a unique f_{n-1} satisfying A_{n-1}, B_{n-1} and C_{n-1}. We need to define f_n on elements of the form $x_i \otimes z_J$ where J can be of length n. We may as well assume J is increasing since P is commutative. If $i \leq J$, then $f_n(x_i \otimes z_J) = z_i z_J$ in order to fulfil A_n. Now suppose $J = (j, J')$ and $i > j$. Then

$$\begin{aligned} f_n(x_i \otimes z_j z_{J'}) &= f_n(x_i \otimes f_n(x_j \otimes z_{J'})) \\ &= f_n(x_i \otimes f_{n-1}(x_j \otimes z_{J'})) \\ &= f_n(x_j \otimes f_{n-1}(x_i \otimes z_{J'})) + f_{n-1}([x_i, x_j] \otimes z_{J'}) \end{aligned}$$

using the fact that f_n and f_{n-1} agree where they are both defined and trying to satisfy condition C_p. But now $j < i$ and $j \leq J'$ so by property B_{n-1}

$$f_n(x_j \otimes f_{n-1}(x_i \otimes z_{J'})) = f_n(x_j \otimes (z_i z_{J'} + w))$$

where $w \in P_{n-1}$. By property A_n, this equals $z_j z_i z_{J'} + f_{n-1}(x_j \otimes w)$. Thus we should define $f_n(x_i \otimes z_J) = z_i z_J$ when $i \leq J$, and otherwise

$$f_n(x_i \otimes z_j z_{J'}) = z_i z_J + f_{n-1}(x_j \otimes w) + f_{n-1}([x_i, x_j] \otimes z_{J'}) \tag{1}$$

If this satisfies A_n, B_n, and C_n it will be the unique extension of f_{n-1}, for conditions A_n, B_n, and C_n when restricted to P_{n-1} are conditions A_{n-1}, B_{n-1}, and C_{n-1} which are satisfied by a unique f_{n-1}. Property A_n is obviously satisfied, and so is B_n, since the second and third terms are in P_{n-1} by B_{n-1}. It remains to verify C_n.

Now we need to check $f_n(x_i \otimes f_n(x_j \otimes z_J)) = f_n(x_j \otimes f_n(x_i \otimes z_J)) + f_n([x_i, x_j] \otimes z_J)$ for $z_J \in P_{n-1}$. By the way we constructed f_n, C_n is satisfied if $j < i$ and $j \leq J$ since

$$\begin{aligned} f_n(x_i \otimes f_{n-1}(x_j \otimes z_J)) &= f_n(x_i \otimes z_j z_J) \\ &= z_i z_j z_J + f_{n-1}(x_j \otimes w) + f_{n-1}([x_i, x_j] \otimes z_J) \\ &= f_n(x_j \otimes f_{n-1}(x_i \otimes z_J)) + f_{n-1}([x_i, x_j] \otimes z_J) \end{aligned}$$

Furthermore, if we flip the role of i and j since $[x_i, x_j] = -[x_j, x_i]$ this holds as long as $i \leq J'$ and $i < j$. If $i = j$, there is nothing to prove. Thus the only remaining cases are when neither $i \leq J$ or $j \leq J$: $J = (k, K)$ where $k < i, j$. Then by induction ($z_J \in P_{n-1}$)

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

Now $f_n(x_j \otimes z_K) = z_j z_K + w$ where $w \in P_{n-2}$ by B_{n-1}. Then

$$f_n(x_k \otimes f_n(x_j \otimes z_J)) = f_p(x_i \otimes f_n(x_k \otimes (z_j z_K + w))) + f_n(x_i \otimes f_n([x_j, x_k] \otimes z_K)))$$

Since $i > k$ and $k \leq j, K$ and $w \in P_{n-2}$, C_n holds for the first term. C_n holds for the second term by induction. Thus this expands as

$$\begin{aligned} f_n(x_i \otimes f_n(x_k \otimes f_n(x_j \otimes z_K))) + f_n(x_i \otimes f_n([x_j, x_k] \otimes z_K))) = f_n(x_k \otimes f_n(x_i \otimes f_n(x_j \otimes z_K)))+ \\ + f_n([x_i, x_k] \otimes f_n(x_j \otimes z_K)) + f_n([x_j, x_k] \otimes f_n(x_i, z_K)) + f_n([x_i, [x_j, x_k]] \otimes z_K) \end{aligned}$$

A similar statement holds if interchange the role of i and j. Then

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

by the Jacobi identity. Thus C_p holds in general, completing the proof. □

Theorem 5. *The y_I for I increasing form a basis for $U(\mathfrak{g})$ as a vector space.*

Proof. Combining the maps for all n, we see there is a bilinear mapping $f : \mathfrak{g} \times P \to P$ such that $f(x_i, z_I) = z_i z_I$ for $i \le I$ and

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

This is a representation ρ of $\mathfrak{g}$ on P with the property that $\rho(x_i)z_I = z_i z_I$. By the universal property of $U(\mathfrak{g})$, there is a map $\psi : U(\mathfrak{g}) \to \mathrm{End}(P)$ with $\psi(y_i)z_I = z_i z_I$ for $i \le I$. Then by induction if $I = (i_1, \ldots, i_n)$ is increasing we see

$$\psi(y_{i_1} \ldots y_{i_n}) \cdot 1 = z_{i_1} \ldots z_{i_n}$$

But the polynomials on the right hand side are linearly independent, so the y_I with I increasing are linearly independent as well. We already showed they generate $U(\mathfrak{g})$.

□

This then implies all the forms of the PBW theorem.

Corollary 6. *The canonical mapping of $\mathfrak{g}$ to $U(\mathfrak{g})$ is injective.*

Using the construction of the universal enveloping algebra as a quotient of the tensor algebra, there is a natural filtration on $U(\mathfrak{g})$ where $U_n(\mathfrak{g})$ is generated by products of the form $x_1 \otimes \ldots \otimes x_p$ where $x_i \in \mathfrak{g}$ and $p \le n$. Remember that $\mathrm{gr}(U(\mathfrak{g})) = \bigoplus_{n=0}^{\infty} G^n$ where $G^n = U_n(\mathfrak{g})/U_{n-1}(\mathfrak{g})$ and $G^0 = k$. Note that $G^1 \simeq \mathfrak{g}$. Multiplication in $U(\mathfrak{g})$ makes this into a commutative ring by the first lemma.

Corollary 7. *Sym*$(\mathfrak{g}) \simeq \mathrm{gr}(U(\mathfrak{g}))$

Proof. Since $\mathrm{gr}(U(\mathfrak{g}))$ is commutative, by the universal property of the symmetric algebra the map $\mathfrak{g} \to \mathrm{gr}(U(\mathfrak{g}))$ extends to a map $\mathrm{Sym}(\mathfrak{g}) \to \mathrm{gr}(U(\mathfrak{g}))$. We know that expressions of the form $x_1^{v_1} \ldots x_n^{v_n} \ldots$ with $\sum v_i \le n$ form a basis for $U_n(\mathfrak{g})$. The ones with sum exactly n form a basis for G^n. Thus elements of this form give a basis for $\mathrm{gr}(U(\mathfrak{g}))$, and the map $\mathrm{Sym}(\mathfrak{g}) \to \mathrm{gr}(U(\mathfrak{g}))$ sends the standard basis for $\mathrm{Sym}(\mathfrak{g})$ to this. Thus the map is an isomorphism. □