

Math 128 Lecture 3

Beginning of an algebraic proof of the Campbell-Baker-Hausdorff formula: the universal enveloping algebra.

The universal enveloping algebra.

A **universal algebra** of a Lie algebra L is a map $\epsilon : L \rightarrow UL$ where UL is an associative algebra with unit such that

1. ϵ is a Lie algebra homomorphism, i.e. it is linear and

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

2. If A is any associative algebra with unit and $\alpha : L \rightarrow A$ is any Lie algebra homomorphism then there exists a unique homomorphism ϕ of associative algebras such that

$$\alpha = \phi \circ \epsilon.$$

It is clear that if UL exists, it is unique up to a unique isomorphism. So we may then talk of *the* universal algebra of L . We will call it the universal enveloping algebra and sometimes put in parenthesis, i.e. write $U(L)$.

$U(L)$ of the Lie algebra of a group.

In case $L = \mathfrak{g}$ is the Lie algebra of left invariant vector fields on a group G , we may think of L as consisting of left invariant first order homogeneous differential operators on G . Then we may take UL to consist of all left invariant differential operators on G . In this case the construction of UL is intuitive and obvious. The ring of differential operators \mathcal{D} on any manifold is filtered by degree: \mathcal{D}^n consisting of those differential operators with total degree at most n . The quotient, $\mathcal{D}^n/\mathcal{D}^{n-1}$ consists of those homogeneous differential operators of degree n , i.e. homogeneous polynomials in the vector fields with function coefficients. For the case of left invariant differential operators on a group, these vector fields may be taken to be left invariant, and the function coefficients to be constant. In other words, $(UL)^n/(UL)^{n-1}$ consists of all symmetric polynomial expressions, homogeneous of degree n in L . This is the content of the Poincaré-Birkhoff-Witt theorem. In the algebraic case we have to do some work to get all of this. We first must construct $U(L)$.

Universal constructions: tensor product.

Let E_1, \dots, E_m be vector spaces and (f, F) a multilinear map $f : E_1 \times \dots \times E_m \rightarrow F$. Similarly (g, G) . If ℓ is a linear map $\ell : F \rightarrow G$, and $g = \ell \circ f$ then we say that ℓ is a morphism of (f, F) to (g, G) . In this way we make the set of all (f, F) into a *category*. Want a universal object in this category; that is, an object with a unique morphism into every other object. So want a pair (t, \mathcal{T}) where \mathcal{T} is a vector space, $t : E_1 \times \dots \times E_m \rightarrow \mathcal{T}$ is a multilinear map, and for every (f, F) there is a unique linear map $\ell_f : \mathcal{T} \rightarrow F$ with

$$f = \ell_f \circ t$$

Uniqueness. By the universal property $t = \ell'_t \circ t'$, $t' = \ell_t \circ t$ so $t = (\ell'_t \circ \ell_t) \circ t$, but also $t = t \circ \text{id}$. So $\ell'_t \circ \ell_t = \text{id}$. Similarly the other way. Thus (t, \mathcal{T}) , if it exists, is unique up to a unique morphism. This is a standard argument valid in any category proving the uniqueness of “initial elements”.

Existence of tensor products.

Existence. Let M be the free vector space on the symbols x_1, \dots, x_m , $x_i \in E_i$. Let N be the subspace generated by all the $(x_1, \dots, x_i + x'_i, \dots, x_m) - (x_1, \dots, x_i, \dots, x_m) - (x_1, \dots, x'_i, \dots, x_m)$ and all the

$$(x_1, \dots, ax_i, \dots, x_m) - a(x_1, \dots, x_i, \dots, x_m)$$

for all $i = 1, \dots, m, x_i, x'_i \in E_i, a \in k$. Let $\mathcal{T} = M/N$ and

$$t((x_1, \dots, x_m)) = (x_1, \dots, x_m)/N.$$

This is universal by its very construction. QED

We introduce the notation

$$\mathcal{T} = \mathcal{T}(E_1 \times \cdots \times E_m) =: E_1 \otimes \cdots \otimes E_m.$$

The universality implies an isomorphism

$$(E_1 \otimes \cdots \otimes E_m) \otimes (E_{m+1} \otimes \cdots \otimes E_{m+n}) \cong E_1 \otimes \cdots \otimes E_{m+n}.$$

The tensor product of two algebras.

If A and B are algebras, they are they are vector spaces, so we can form their tensor product as vector spaces. We define a product structure on $A \otimes B$ by defining

$$(a_1 \otimes b_1) \cdot (a_2 \otimes b_2) := a_1 a_2 \otimes b_1 b_2.$$

It is easy to check that this extends to give an algebra structure on $A \otimes B$. In case A and B are associative algebras so is $A \otimes B$, and if in addition both A and B have unit elements, then $1_A \otimes 1_B$ is a unit element for $A \otimes B$. We will frequently drop the subscripts on the unit elements, for it is easy to see from the position relative to the tensor product sign the algebra to which the unit belongs. In other words, we will write the unit for $A \otimes B$ as $1 \otimes 1$. We have an isomorphism of A into $A \otimes B$ given by

$$a \mapsto a \otimes 1$$

when both A and B are associative algebras with units. Similarly for B . Notice that

$$(a \otimes 1) \cdot (1 \otimes b) = a \otimes b = (1 \otimes b) \cdot (a \otimes 1).$$

In particular, an element of the form $a \otimes 1$ commutes with an element of the form $1 \otimes b$.

The tensor algebra of a vector space.

Let V be a vector space. The **tensor algebra** of a vector space is the solution of the universal problem for maps α of V into an associative algebra: it consists of an algebra TV and a map $\iota : V \rightarrow TV$ such that ι is linear, and for any linear map $\alpha : V \rightarrow A$ where A is an associative algebra there exists a unique algebra homomorphism $\psi : TV \rightarrow A$ such that $\alpha = \psi \circ \iota$. We set

$$T^n V := V \otimes \cdots \otimes V \quad n - \text{factors.}$$

We define the multiplication to be the isomorphism

$$T^n V \otimes T^m V \rightarrow T^{n+m} V$$

obtained by “dropping the parentheses,” i.e. the isomorphism given at the end of the last subsection. Then

$$TV := \bigoplus T^n V$$

(with $T^0 V$ the ground field) is a solution to this universal problem, and hence the unique solution.

Construction of the universal enveloping algebra.

If we take $V = L$ to be a Lie algebra, and let I be the two sided ideal in TL generated the elements $[x, y] - x \otimes y + y \otimes x$ then

$$UL := TL/I$$

is a universal algebra for L . Indeed, any homomorphism α of L into an associative algebra A extends to a unique algebra homomorphism $\psi : TL \rightarrow A$ which must vanish on I if it is to be a Lie algebra homomorphism.

Extension of a Lie algebra homomorphism to its universal enveloping algebra.

If $h : L \rightarrow M$ is a Lie algebra homomorphism, then the composition

$$\epsilon_M \circ h : L \rightarrow UM$$

induces a homomorphism

$$UL \rightarrow UM$$

and this assignment sending Lie algebra homomorphisms into associative algebra homomorphisms is functorial.

Universal enveloping algebra of a direct sum.

Suppose that: $L = L_1 \oplus L_2$, with $\epsilon_i : L_i \rightarrow U(L_i)$, and $\epsilon : L \rightarrow U(L)$ the canonical homomorphisms. Define

$$f : L \rightarrow U(L_1) \otimes U(L_2), \quad f(x_1 + x_2) = \epsilon_1(x_1) \otimes 1 + 1 \otimes \epsilon_2(x_2).$$

This is a homomorphism because x_1 and x_2 commute. It thus extends to a homomorphism

$$\psi : U(L) \rightarrow U(L_1) \otimes U(L_2).$$

Also,

$$x_1 \mapsto \epsilon(x_1)$$

is a Lie algebra homomorphism of $L_1 \rightarrow U(L)$ which thus extends to a unique algebra homomorphism

$$\phi_1 : U(L_1) \rightarrow U(L)$$

and similarly $\phi_2 : U(L_2) \rightarrow U(L)$. We have

$$\phi_1(x_1)\phi_2(x_2) = \phi_2(x_2)\phi_1(x_1), \quad x_1 \in L_1, x_2 \in L_2$$

since $[x_1, x_2] = 0$.

Universal enveloping algebra of a direct sum, 2 .

$$\phi_1(x_1)\phi_2(x_2) = \phi_2(x_2)\phi_1(x_1), \quad x_1 \in L_1, x_2 \in L_2$$

since $[x_1, x_2] = 0$. As the $\epsilon_i(x_i)$ generate $U(L_i)$, the above equation holds with x_i replaced by arbitrary elements $u_i \in U(L_i)$, $i = 1, 2$. So we have a homomorphism

$$\phi : U(L_1) \otimes U(L_2) \rightarrow U(L), \quad \phi(u_1 \otimes u_2) := \phi_1(u_1)\phi_2(u_2).$$

We have

$$\phi \circ \psi(x_1 + x_2) = \phi(x_1 \otimes 1) + \phi(1 \otimes x_2) = x_1 + x_2$$

so $\phi \circ \psi = \text{id}$, on L and hence on $U(L)$ and

$$\psi \circ \phi(x_1 \otimes 1 + 1 \otimes x_2) = x_1 \otimes 1 + 1 \otimes x_2$$

so $\psi \circ \phi = \text{id}$ on $L_1 \otimes 1 + 1 \otimes L_2$ and hence on $U(L_1) \otimes U(L_2)$. Thus

$$U(L_1 \oplus L_2) \cong U(L_1) \otimes U(L_2).$$

Bialgebra structure.

Consider the map $L \rightarrow U(L) \otimes U(L)$:

$$x \mapsto x \otimes 1 + 1 \otimes x.$$

Then

$$\begin{aligned} (x \otimes 1 + 1 \otimes x)(y \otimes 1 + 1 \otimes y) = \\ xy \otimes 1 + x \otimes y + y \otimes x + 1 \otimes xy, \end{aligned}$$

and multiplying in the reverse order and subtracting gives

$$[x \otimes 1 + 1 \otimes x, y \otimes 1 + 1 \otimes y] = [x, y] \otimes 1 + 1 \otimes [x, y].$$

Thus the map $x \mapsto x \otimes 1 + 1 \otimes x$ determines an algebra homomorphism

$$\Delta : U(L) \rightarrow U(L) \otimes U(L).$$

Define

$$\varepsilon : U(L) \rightarrow k, \quad \varepsilon(1) = 1, \quad \varepsilon(x) = 0, \quad x \in L$$

and extend as an algebra homomorphism. Then

$$(\varepsilon \otimes \text{id})(x \otimes 1 + 1 \otimes x) = 1 \otimes x, \quad x \in L.$$

We identify $k \otimes L$ with L and so can write the above equation as

$$(\varepsilon \otimes \text{id})(x \otimes 1 + 1 \otimes x) = x, \quad x \in L.$$

The algebra homomorphism

$$(\varepsilon \otimes \text{id}) \circ \Delta : U(L) \rightarrow U(L)$$

is the identity (on 1 and on) L and hence is the identity. Similarly

$$(\text{id} \otimes \varepsilon) \circ \Delta = \text{id}.$$

A vector space C with a map $\Delta : C \rightarrow C \otimes C$, (called a **comultiplication**) and a map $\varepsilon : C \rightarrow k$ (called a **co-unit**) satisfying

$$(\varepsilon \otimes \text{id}) \circ \Delta = \text{id}$$

and

$$(\text{id} \otimes \varepsilon) \circ \Delta = \text{id}$$

is called a **co-algebra**. If C is an algebra and both Δ and ε are algebra homomorphisms, we say that C is a **bi-algebra** (sometimes shortened to “bigebra”). So we have proved that $(U(L), \Delta, \varepsilon)$ is a bialgebra.

A vector space C with a map $\Delta : C \rightarrow C \otimes C$, (called a **comultiplication**) and a map $\varepsilon : D \rightarrow k$ (called a **co-unit**) satisfying

$$(\varepsilon \otimes \text{id}) \circ \Delta = \text{id}$$

and

$$(\text{id} \otimes \varepsilon) \circ \Delta = \text{id}$$

is called a **co-algebra**. If C is an algebra and both Δ and ε are algebra homomorphisms, we say that C is a **bi-algebra**(sometimes shortened to “bigebra”). So we have proved that $(U(L), \Delta, \varepsilon)$ is a bialgebra.

Also

$$[(\Delta \otimes \text{id}) \circ \Delta](x) = x \otimes 1 \otimes 1 + 1 \otimes x \otimes 1 + 1 \otimes 1 \otimes x = [(\text{id} \otimes \Delta) \circ \Delta](x)$$

for $x \in L$ and hence for all elements of $U(L)$. Hence the comultiplication is coassociative. (It is also co-commutative.)

The Poincaré-Birkhoff-Witt Theorem.

Suppose that V is a vector space made into a Lie algebra by declaring that all brackets are zero. Then the ideal I in TV defining $U(V)$ is generated by $x \otimes y - y \otimes x$, and the quotient TV/I is just the symmetric algebra, SV . So the universal enveloping algebra of the trivial Lie algebra is the symmetric algebra.

For any Lie algebra L define $U_n L$ to be the subspace of UL generated by products of at most n elements of L , i.e. by all products

$$\epsilon(x_1) \cdots \epsilon(x_m), \quad m \leq n.$$

For example,,

$$U_0 L = k, \text{ the ground field}$$

and

$$U_1 L = k \oplus \epsilon(L).$$

We have

$$U_0 L \subset U_1 L \subset \cdots \subset U_n L \subset U_{n+1} L \subset \cdots$$

and

$$U_m L \cdot U_n L \subset U_{m+n} L.$$

The associated graded algebra.

We define

$$\mathrm{gr}_n UL := U_n L / U_{n-1} L$$

and

$$\mathrm{gr} UL := \bigoplus \mathrm{gr}_n UL$$

with the multiplication

$$\mathrm{gr}_m UL \times \mathrm{gr}_n UL \rightarrow \mathrm{gr}_{m+n} UL$$

induced by the multiplication on UL .

If $a \in U_n L$ we let $\bar{a} \in \mathrm{gr}_n UL$ denote its image by the projection $U_n L \rightarrow U_n L / U_{n-1} L = \mathrm{gr}_n UL$. We may write a as a sum of products of at most n elements of L :

$$a = \sum_{m_\mu \leq n} c_\mu \epsilon(x_{\mu,1}) \cdots \epsilon(x_{\mu,m_\mu}).$$

Then \bar{a} can be written as the corresponding homogeneous sum

$$\bar{a} = \sum_{m_\mu = n} c_\mu \overline{\epsilon(x_{\mu,1})} \cdots \overline{\epsilon(x_{\mu,m_\mu})}.$$

The associated graded algebra is commutative.

$$\bar{a} = \sum_{m_\mu = n} c_\mu \overline{\epsilon(x_{\mu,1})} \cdots \overline{\epsilon(x_{\mu,m_\mu})}.$$

In other words, as an algebra, $\text{gr } UL$ is generated by the elements $\overline{\epsilon(x)}$, $x \in L$. But all such elements commute. Indeed, for $x, y \in L$,

$$\epsilon(x)\epsilon(y) - \epsilon(y)\epsilon(x) = \epsilon([x, y]).$$

by the defining property of the universal enveloping algebra. The right hand side of this equation belongs to $U_1 L$. Hence

$$\overline{\epsilon(x)\epsilon(y)} - \overline{\epsilon(y)\epsilon(x)} = 0$$

in $\text{gr}_2 UL$. This proves that $\text{gr } UL$ is *commutative*. Hence, by the universal property of the symmetric algebra, there exists a unique algebra homomorphism

$$\mathbf{w} : SL \rightarrow UL$$

extending the linear map

$$L \rightarrow \text{gr } UL, \quad x \mapsto \overline{\epsilon(x)}.$$

Statement of the PBW theorem.

By the universal

property of the symmetric algebra, there exists a unique algebra homomorphism

$$\mathbf{w} : SL \rightarrow UL$$

extending the linear map

$$L \rightarrow \text{gr } UL, \quad x \mapsto \overline{\epsilon(x)}.$$

Since the $\overline{\epsilon(x)}$ generate $\text{gr } UL$ as an algebra, we know that this map is surjective. The **Poincaré-Birkhoff-Witt theorem** asserts that

$$\mathbf{w} : SL \rightarrow \text{gr } UL \text{ is an isomorphism.} \tag{19}$$

Equivalent statement of the PBW theorem.

The **Poincaré-Birkhoff-Witt theorem** asserts that

$$\mathfrak{w} : SL \rightarrow \text{gr } UL \text{ is an isomorphism.} \quad (19)$$

Suppose that we choose a basis x_i , $i \in I$ of L where I is a totally ordered set. Since

$$\overline{\epsilon(x_i)\epsilon(x_j)} = \overline{\epsilon(x_j)\epsilon(x_i)}$$

we can rearrange any product of $\overline{\epsilon(x_i)}$ so as to be in increasing order. This shows that the elements

$$x_M := \epsilon(x_{i_1}) \cdots \epsilon(x_{i_m}), \quad M := (i_1, \dots, i_m) \quad i_1 \leq \cdots \leq i_m$$

span UL as a vector space. We claim that (19) is equivalent to

Theorem 1 Poincaré-Birkhoff-Witt. *The elements x_M form a basis of UL .*

The **Poincaré-Birkhoff-Witt theorem** asserts that

$$\mathbf{w} : SL \rightarrow \text{gr } UL \text{ is an isomorphism.} \quad (19)$$

Theorem 1 Poincaré-Birkhoff-Witt. *The elements x_M form a basis of UL .*

Proof that (19) is equivalent to the statement of the theorem. For any expression x_M as above, we denote its length by $\ell(M) = m$. The elements $\overline{x_M}$ are the images under \mathbf{w} of the monomial basis in $S_m(L)$. As we know that \mathbf{w} is surjective, equation (19) is equivalent to the assertion that \mathbf{w} is injective. This amounts to the non-existence of a relation of the form

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

with some non-zero coefficients on the left hand side. But any non-trivial relation between the x_M can be rewritten in the above form by moving the terms of highest length to one side. QED

We now turn to the proof of the theorem:

Let V be the vector space with basis z_M where M runs over all ordered sequences $i_1 \leq i_2 \leq \cdots \leq i_n$. (Recall that we have chosen a well ordering on I and that the x_i $i \in I$ form a basis of L .)

Furthermore, the empty sequence, z_\emptyset is allowed, and we will identify the symbol z_\emptyset with the number $1 \in k$. If $i \in I$ and $M = (i_1, \dots, i_n)$ we write $i \leq M$ if $i \leq i_1$ and then let (i, M) denote the ordered sequence (i, i_1, \dots, i_n) . In particular, we adopt the convention that if $M = \emptyset$ is the empty sequence then $i \leq M$ for all i in which case $(i, M) = (i)$. Recall that if $M = (i_1, \dots, i_n)$ we set $\ell(M) = n$ and call it the length of M . So, for example, $\ell(i, M) = \ell(M) + 1$ if $i \leq M$.

Lemma 1 *We can make V into an L module in such a way that*

$$x_i z_M = z_{iM} \quad \text{whenever } i \leq M. \quad (20)$$

Lemma 1 *We can make V into an L module in such a way that*

$$x_i z_M = z_{iM} \quad \text{whenever } i \leq M. \quad (20)$$

Proof of lemma. We will inductively define a map

$$L \times V \rightarrow V, \quad (x, v) \mapsto xv$$

and then show that it satisfies the equation

$$xyv - yxv = [x, y]v, \quad x, y \in g, \quad v \in V, \quad (21)$$

which is the condition that makes V into an L module. Our definition will be such that (20) holds. In fact, we will define $x_i z_M$ inductively on $\ell(M)$ and on i . So we start by defining

$$x_i z_\emptyset = z_{(i)}$$

which is in accordance with (20). This defines $x_i z_M$ for $\ell(M) = 0$.

We want

$$x_i z_M = z_{iM} \quad \text{whenever } i \leq M. \quad (20)$$

and

$$xyv - yxv = [x, y]v, \quad x, y \in g, \quad v \in V, \quad (21)$$

For

$\ell(M) = 1$ we define

$$x_i z_{(j)} = z_{(i,j)} \quad \text{if } i \leq j$$

while if $i > j$ we set

$$x_i z_{(j)} = x_j z_{(i)} + [x_i, x_j] z_\emptyset = z_{(j,i)} + \sum c_{ij}^k z_{(k)}$$

where

$$[x_i, x_j] = \sum c_{ij}^k x_k$$

is the expression for the Lie bracket of x_i with x_j in terms of our basis. These c_{ij}^k are known as the **structure constants** of the Lie algebra, L in terms of the given basis. Notice that the first of these two cases is consistent with (and forced on us) by (20) while the second is forced on us by (21). We now have defined $x_i z_M$ for all i and all M with $\ell(M) \leq 1$. and we have done so in such a way that (20) holds, and (21) holds where it makes sense (i.e. for $\ell(M) = 0$).

We want $x_i z_M = z_{iM}$ whenever $i \leq M$. (20)

and $xyv - yxv = [x, y]v$, $x, y \in g$, $v \in V$, (21)

So suppose that we have defined $x_j z_N$ for all j if $\ell(N) < \ell(M)$ and for all $j < i$ if $\ell(N) = \ell(M)$ in such a way that

$$x_j z_N \text{ is a linear combination of } z_L \text{'s with } \ell(L) \leq \ell(N) + 1 \quad (*).$$

We then define

$$x_i z_M = \begin{cases} z_{iM} & \text{if } i \leq M \\ x_j(x_i z_N) + [x_i, x_j]z_N & \text{if } M = (jN) \text{ with } i > j. \end{cases} \quad (22)$$

This makes sense since $x_i z_N$ is already defined as a linear combination of z_L 's with $\ell(L) \leq \ell(N) + 1 = \ell(M)$ and because $[x_i, x_j]$ can be written as a linear combination of the x_k as above. Furthermore (*) holds with j and N replaced by M . Furthermore, (20) holds by construction. We must check (21). By linearity, this means that we must show that

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

We defined
$$x_i z_M = \begin{cases} z_{iM} & \text{if } i \leq M \\ x_j(x_i z_N) + [x_i, x_j]z_N & \text{if } M = (jN) \text{ with } i > j. \end{cases} \quad (22)$$

we must show that

$$x_i x_j z_N - x_j x_i z_N = [x_i, x_j]z_N. \quad (\text{a version of 21})$$

If $i = j$ both sides are zero. Also, since both sides are anti-symmetric in i and j , we may assume that $i > j$. If $j \leq N$ and $i > j$ then this equation holds by definition. So we need only deal with the case where $j \not\leq N$ which means that $N = (kP)$ with $k \leq P$ and $i > j > k$. So we have, by definition,

$$\begin{aligned} x_j z_N &= x_j z_{(kP)} \\ &= x_j x_k z_P \\ &= x_k x_j z_P + [x_j, x_k]z_P. \end{aligned}$$

Now if $j \leq P$ then $x_j z_P = z_{(jP)}$ and $k < (jP)$. If $j \not\leq P$ then $x_j z_P = z_Q + w$ where still $k \leq Q$ and w is a linear combination of elements of length $< \ell(N)$.

So we know that (21) holds for $x = x_i, y = x_k$ and $v = z_{(jP)}$ (if $j \leq P$) or $v = z_Q$ (otherwise). Also, by induction, we may assume that we have verified (21) for all N' of length $< \ell(N)$. So we may apply (21) to $x = x_i, y = x_k$ and $v = x_j z_P$ and also to $x = x_i, y = [x_j, x_k], v = z_P$. So

$$x_i x_j z_N = x_k x_i x_j z_P + [x_i, x_k]x_j z_P + [x_j, x_k]x_i z_P + [x_i, [x_j, x_k]]z_P.$$

Proof of Lemma, completed.

We know that

$$x_i x_j z_N = x_k x_i x_j z_P + [x_i, x_k] x_j z_P + [x_j, x_k] x_i z_P + [x_i, [x_j, x_k]] z_P.$$

Similarly, the same result holds with i and j interchanged. Subtracting this interchanged version from the preceding equation the two middle terms from each equation cancel and we get

$$\begin{aligned}(x_i x_j - x_j x_i) z_N &= x_k (x_i x_j - x_j x_i) z_P + ([x_i, [x_j, x_k]] - [x_j, [x_i, x_k]]) z_P \\ &= x_k [x_i, x_j] z_P + ([x_i, [x_j, x_k]] - [x_j, [x_i, x_k]]) z_P \\ &= [x_i, x_j] x_k z_P + ([x_k, [x_i, x_j]] + [x_i, [x_j, x_k]] - [x_j, [x_i, x_k]]) z_P \\ &= [x_i, x_j] z_N.\end{aligned}$$

(In passing from the second line to the third we used (21) applied to z_P (by induction) and from the third to the last we used the antisymmetry of the bracket and Jacobi's equation.) QED

Whew!

Proof of the PBW theorem. We have made V into an L and hence into a $U(L)$ module. By construction, we have, inductively,

$$x_M z_\emptyset = z_M.$$

But if

$$\sum c_M x_M = 0$$

then

$$0 = \sum c_M z_M = \left(\sum c_M x_M \right) z_\emptyset$$

contradicting the fact the the z_M are independent. QED

In particular, the map $\epsilon : L \rightarrow U(L)$ is an injection, and so we may identify L as a subspace of $U(L)$.