25 §17: The PBW Theorem

The following are lecture notes based on §17 of Humphreys, covering universal enveloping algebras, the PBW theorem, and some consequences. (Speaker: Josh Swanson.)

Comment *Outline:*

1. *Notation, overview*
2. *Universal properties;* $\mathfrak{T}(L), \mathfrak{S}(L), \mathfrak{U}(L)$
3. *Associated graded objects, algebras; Humphreys' PBW theorem*
4. *Usual PBW theorem, more consequences*
5. *PBW proof*

Note: I've taken the opportunity to discuss some categorical notions that aren't strictly necessary (e.g. representable functors, adjoints, associated graded algebras, Serre spectral sequence).

Notation *In this section,* R *is a commutative ring,* $\mathbb{F}$ *is an arbitrary field,* L *is a Lie algebra over* $\mathbb{F}$ *(possibly infinite dimensional),* V *is an* $\mathbb{F}$*-vector space with ordered basis* $\{v_1, v_2, \ldots\}$ *(which need not be countable, despite the notation), and tensor products are always as* $\mathbb{F}$*-modules, so over* $\mathbb{F}$*. Many statements are true in much greater generality.*

Definition *An (associative unital)* R*-algebra* A *is:*

1. *An* R*-module*
2. *with a bilinear, associative product* $A \times A \to A$
3. *with a two-sided identity* 1.

It is graded if $A = \oplus_{i=0}^{\infty} A^i$ *and the product is of the form* $A^i \times A^j \to A^{i+j}$*. Note that* R*-algebra homomorphisms are* $\mathbb{F}$*-linear, multiplicative, and send* 1 *to* 1.

Comment *Overview: a standard source of Lie algebras is by taking commutators in an algebra. All Lie algebras at least inject into such an algebra, the universal enveloping algebra* $\mathfrak{U}(L)$*. Indeed,* $\mathfrak{U}(L)$ *is roughly obtained by taking formal products of elements of* L *modulo* $[x, y] = xy - yx$*. If* L *is abelian, one expects* $\mathfrak{U}(L)$ *to just be a polynomial ring (it is). The PBW theorem says this is true in general at least on the level of* $\mathbb{F}$*-modules, i.e. given an ordered basis* $\{v_1, v_2, \ldots\}$ *for* L*,* $\mathfrak{U}(L)$ *has a basis of "monomials"* $v_{i_1} \cdots v_{i_\ell}$ *where* $i_1 \leq \cdots \leq i_\ell$*. The bracket structure of* L *is hiding in the multiplicative structure of* $\mathfrak{U}(L)$*.*

Why would you expect this? Any "monomial" $v_{i_1} \cdots v_{i_n}$ *can be reordered at the cost of introducing lower degree terms by replacing* $v_i v_j$ *with* $[v_i, v_j] - v_j v_i$*. The monomials thus span, though the hard part is showing they're independent.*

First, a sample of the power of a PBW basis:

Corollary 110 *The map $L \to \mathfrak{U}(L)$ is injective.*

Proof This map sends v_i in L to the "monomial" v_i in $\mathfrak{U}(L)$, and by the PBW theorem these v_i are linearly independent. ∎

Comment *Let A be an $\mathbb{F}$-algebra. We have "natural" (i.e. functorial in A) bijections*

$$\begin{aligned}\mathrm{Hom}_{\mathbb{F}\text{-}linear}(V, A) &\leftrightarrow \mathrm{Hom}_{Sets}(\{v_1, v_2, \ldots\}, A)\\ &\leftrightarrow \mathrm{Hom}_{\mathbb{F}\text{-}algebra}(\mathbb{F}\langle v_1, v_2, \ldots\rangle, A).\end{aligned}$$

If A is commutative, we likewise have

$$\begin{aligned}\mathrm{Hom}_{\mathbb{F}\text{-}linear}(V, A) &\leftrightarrow \mathrm{Hom}_{Sets}(\{v_1, v_2, \ldots\}, A)\\ &\leftrightarrow \mathrm{Hom}_{comm.\ \mathbb{F}\text{-}algebra}(\mathbb{F}[v_1, v_2, \ldots], A).\end{aligned}$$

In both cases, we've identified a functor of the form $\mathrm{Hom}_{\mathcal{C}}(C, -): \mathcal{C} \to Sets$ up to natural isomorphism. This is called a representable functor. One version of Yoneda's lemma gives that any other object D satisfying these properties is itself isomorphic to C. In this sense, we've found universal properties.

(Alternative explanation: these C are initial objects in a certain slice category, so are unique up to suitably unique isomorphism.)

Definition *Coordinate-free definitions for these spaces are common. Define:*

$$\begin{aligned}&T^iV := V^{\otimes i} \qquad (\text{so } T^0V := \mathbb{F})\\ &T^nV \times T^mV \to T^{n+m}V : (v_1 \otimes \cdots \otimes v_n) \cdot (w_1 \otimes \cdots \otimes w_m)\\ &\qquad := v_1 \otimes \cdots \otimes v_n \otimes w_1 \otimes \cdots w_m\\ &\mathfrak{T}(V) := \oplus_{i=0}^{\infty} T^iV\end{aligned}$$

which makes $\mathfrak{T}(V)$ a graded $\mathbb{F}$-algebra, the tensor algebra of V.

Note: any ring morphism $Q \to R$ allows us to turn R-algebras into Q-algebras. Since $\mathbb{Z}$ is the initial object in the category of rings, any R-algebra is a $\mathbb{Z}$-algebra. But a $\mathbb{Z}$-algebra is precisely a ring. So R-algebras are rings.

Easy to see: $\mathfrak{T}(V) \cong \mathbb{F}\langle v_1, v_2, \ldots\rangle$ (non-commutative) directly. Can also amplify up the universal property of tensor products and apply Yoneda's lemma as above. Similarly:

Definition *Let*

$$\begin{aligned}I &:= \text{two-sided ideal in } \mathfrak{T}(V) \text{ generated by } x \otimes y - y \otimes x\\ \mathfrak{S}(V) &:= \text{commutative, graded } \mathbb{F}\text{-algebra } \mathfrak{T}(V)/I\end{aligned}$$

which gives the symmetric algebra $\mathfrak{S}(V)$. (Warning: graded-commutative differs from commutative, graded.)

We have $\mathfrak{S}(V) \cong \mathbb{F}[v_1, v_2, \ldots]$.

Definition *Let*

$$J := \text{two-sided ideal in } \mathfrak{U}(L) \text{ generated by } x \otimes y - y \otimes x - [x, y]$$
$$\mathfrak{U}(L) := \mathbb{F}\text{-algebra } \mathfrak{T}(V)/J$$

Note: J is typically not homogeneous, so $\mathfrak{U}(L)$ is almost never graded. If L is abelian, then $J = I$, so $\mathfrak{U}(L) = \mathfrak{S}(L) \cong \mathbb{F}[v_1, v_2, \ldots]$ as $\mathbb{F}$-algebras. The PBW basis is obvious in this case.

Comment *The natural map $i\colon L \to \mathfrak{U}(L)$ given by $L = \mathfrak{T}(L)_1 \hookrightarrow \mathfrak{T}(L) \twoheadrightarrow \mathfrak{U}(L)$ is a Lie algebra homomorphism:*

$$i([x, y]) = [x, y] = x \otimes y - y \otimes x = xy - yx = i(x)i(y) - i(y)i(x) = [i(x), i(y)].$$

Given any algebra A, there is a bijection (functorial in A)

$$\mathrm{Hom}_{\textit{Lie alg.}}(L, A) \leftrightarrow \mathrm{Hom}_{\mathbb{F}\textit{-algebra}}(\mathfrak{U}(L), A)$$

which is the universal property of $\mathfrak{U}(L)$, again specifying it uniquely. Indeed, this says that $L \mapsto \mathfrak{U}(L)$ is left adjoint to the functor sending an algebra to its Lie algebra under the commutator. (One can prove this bijection from the corresponding universal properties for $\mathfrak{T}(L)$ and quotients.)

Comment *There is a natural bijection*

$$\{L\text{-modules}\} \leftrightarrow \{\mathfrak{U}(L)\text{-modules}\}.$$

Equivalently, there is a natural bijection

$$\mathrm{Hom}_{\textit{Lie alg.}}(L, \mathfrak{gl}(V)) \leftrightarrow \mathrm{Hom}_{\mathbb{F}\textit{-algebra}}(\mathfrak{U}(L), \mathrm{End}(V)).$$

In this sense representations of L are the same as representations of $\mathfrak{U}(L)$. (It's actually an isomorphism between the categories of L-modules and $\mathfrak{U}(L)$-modules.)

Next, we build up associated graded objects and algebras.

Definition *Let $\mathcal{A}$ be an abelian category with $A \in \mathcal{A}$; think the category of $\mathbb{F}$-algebras. (We essentially just want short exact sequences to make sense.) Given a filtration*

$$0 := F_{-1}A \hookrightarrow F_0A \hookrightarrow F_1A \hookrightarrow \cdots \hookrightarrow A,$$

the associated graded object is the sequence of cokernels $(G^iA := F_iA/F_{i-1}A)_{i=0}^{\infty}$. Assuming countable coproducts exist, these are often assembled into $\oplus_{i=0}^{\infty} G^iA$ and sometimes are given further structure. These quotients are typically much easier to deal with than A alone, yet they frequently contain important information about A.

Example *Let* $F^m\mathfrak{T}(L) := \oplus_{i=0}^m T^i L$, *which gives a filtration of* $\mathfrak{T}(L)$ *(in the category of* $\mathbb{F}$*-modules). The graded pieces are just* $G^m\mathfrak{T}(L) \cong T^m L$. *Assembling these cokernels together yields the* $\mathbb{F}$*-module* $\oplus_{i=0}^\infty T^i L$. *It has a natural multiplication corresponding to* $T^n L \times T^m L \to T^{n+m} L$. *Hence we've made an associated graded algebra from* $\mathfrak{T}(L)$ *and this filtration, which is just isomorphic to* $\mathfrak{T}(L)$.

Example (The Serre spectral sequence) *Start with a Serre fibration* $X \to B$, *with the goal of computing the integral singular cohomology* $H^*(X)$ *of the topological space* X. *One takes* B *to be a CW-complex, which comes with a filtration by its* n*-skeleta. This filtration (eventually) induces a filtration on each* $H^n(X)$.

In fact, one can recover the associated graded objects of these filtrations through an involved process using the Serre spectral sequence. The very rough idea is to create a "book" with pages $1, 2, \ldots$ *and groups at each lattice point* $\mathbb{Z}^2$ *in the plane, together with "diagonals" which form chain complexes; this is a spectral sequence. One "turns the page" by taking homology. In nice cases, the objects at each fixed lattice point will "stabilize" in some finite number of pages, resulting in the "*E^∞*" page. Then the associated graded object of* $H^n(X)$ *is formed by looking at the* n*th antidiagonal on the* E^∞ *page.*

In practice, there are often many zeros in the first few pages and on the E^∞ *page, which allows one to not only recover the associated graded objects, but potentially also identify* $H^n(X)$ *on the nose in terms of the rest of the fibration (e.g. there may be a single non-zero term on each antidiagonal on the* E^∞ *page).*

The associated graded algebra of a filtered algebra:

Definition *A filtered* $\mathbb{F}$*-algebra is an* $\mathbb{F}$*-algebra* A *together with a sequence of vector subspaces*

$$0 =: F_{-1}A \hookrightarrow F_0 A \hookrightarrow F_1 A \hookrightarrow \cdots \hookrightarrow A$$

where $\cup_{i=0}^\infty F_i A = A$ *and* $F_n A \cdot F_m A \subset F_{n+m} A$.

Definition *The associated graded algebra* $\operatorname{gr} A$ *of a filtered algebra* A *is*

- $\operatorname{gr} A := \oplus_{i=0}^\infty G_i A = \oplus_{i=0}^\infty F_i A / F_{i-1} A$ *as an* $\mathbb{F}$*-module,*
- *with product given by* $F_n A/F_{n-1}A \times F_m A/F_{m-1}A \to F_{n+m}A/F_{n+m-1}A$ *induced by* $F_n A \times F_m A \to F_{n+m} A$, *which is well-defined.*

For instance, if a graded algebra is filtered by sums of its graded pieces, $F_n A := \oplus_{i=0}^n A_i$, then $\operatorname{gr} A$ is naturally isomorphic to A as an algebra. In this sense, $\operatorname{gr}\mathfrak{T}(L) \cong \mathfrak{T}(L)$ (as above) and $\operatorname{gr}\mathfrak{S}(L) \cong \mathfrak{S}(L)$.

Proposition 111 $\operatorname{gr} A \cong A$ *as* $\mathbb{F}$*-vector spaces, though unnaturally in general.*

PROOF Successively pick complements so that $F_iA = F_{i-1}A \oplus H_i$ (internal direct sum). Then $F_iA = H_0 \oplus \cdots \oplus H_i$ and $F_iA/F_{i-1}A \cong H_i$, so

$$\begin{aligned}\oplus_{i=0}^{\infty} F_iA/F_{i-1}A &= \oplus_{i=0}^{\infty} H_i = \cup_{n=0}^{\infty} \oplus_{i=0}^{n} H_i \\ &= \cup_{n=0}^{\infty} F_i = A.\end{aligned}$$

Comment *If A is a filtered algebra and $\phi\colon A \twoheadrightarrow B$ is a surjective algebra morphism, then $F^iB := \phi(F^iA)$ makes B a filtered algebra. In particular, if $I \subset A$ is an ideal, then A/I is filtered by $F^i(A/I) := (F^iA + I)/I$. (Easy exercise.)*

We'll drop L from the notation now, so for instance $\mathfrak{U}$ means $\mathfrak{U}(L)$.

Definition *Let*

$$\begin{aligned}F_n\mathfrak{T} &:= \oplus_{i=0}^{n}\mathfrak{T}_i = \oplus_{i=0}^{n} L^{\otimes i} \\ F_n\mathfrak{S} &:= F_n(\mathfrak{T}/I) = \oplus_{i=0}^{n}\mathfrak{S}_i \\ F_n\mathfrak{U} &:= F_n(\mathfrak{T}/J)\end{aligned}$$

which is the filtration on $\mathfrak{T}$ induced by its grading together with its induced filtrations on the quotients $\mathfrak{S}$ and $\mathfrak{U}$.

Intuitively, the nth term of each filtration arises by looking as the "polynomials of degree at most n", possibly modulo some relations. For instance, an element is in $F^n\mathfrak{U}$ if and only if it can be lifted to an at most degree n polynomial in $\mathfrak{T}$ modulo J. (Minor note: Humphreys writes $\mathfrak{G}(L)$ for $\operatorname{gr}\mathfrak{U}(L)$ under this filtration, though we'll stick to $\operatorname{gr}\mathfrak{U}$.)

Since these filtrations on $\mathfrak{T}$ and $\mathfrak{S}$ are the natural ones induced by their gradings, $\operatorname{gr}\mathfrak{T} \cong \mathfrak{T}$ and $\operatorname{gr}\mathfrak{S} \cong \mathfrak{S}$ as above. However, the structure of $\operatorname{gr}\mathfrak{U}$ is much less clear. One thing is obvious, though: $\operatorname{gr}\mathfrak{U}$ is commutative. To see this, it suffices to check the algebraic generating set $v_1, v_2, \ldots$, which lives in $G_1\mathfrak{U}$. But note that in $G_2\mathfrak{U}$, $v_1v_2 - v_2v_1 = [v_1v_2] \in F_1\mathfrak{U} = 0 \in G_1\mathfrak{U}$.

We have a natural map $\mathfrak{T}_m \hookrightarrow F_m\mathfrak{T} \twoheadrightarrow F_m\mathfrak{U} \twoheadrightarrow G_m\mathfrak{U}$ which assembles to give a natural map $\mathfrak{T} \to \operatorname{gr}\mathfrak{U}$.

Lemma 112 *The natural map $\mathfrak{T} \to \operatorname{gr}\mathfrak{U}$ is a surjective algebra homomorphism with kernel containing I. Hence we have a natural surjection of algebras, $\mathfrak{S} \twoheadrightarrow \operatorname{gr}\mathfrak{U}$.*

PROOF That the map is a surjective algebra homomorphism is clear from how it was defined. It must annihilate I since $xy = yx$ in $\operatorname{gr}\mathfrak{U}$, so $x \otimes y - y \otimes x$ must be in the kernel. ■

Theorem 113 (PBW, Humphreys' version) *The natural map $\mathfrak{S} \to \operatorname{gr}\mathfrak{U}$ is an isomorphism of algebras. More symmetrically, $\operatorname{gr}\mathfrak{S} \cong \operatorname{gr}\mathfrak{U}$, or $\operatorname{gr}\mathfrak{T}/I \cong \operatorname{gr}\mathfrak{T}/J$. (Note that I is obtained from J by taking the "leading terms" of the generators for J; this point of view may be generalized.)*

We'll prove this at the end. First, some consequences. Let $\pi\colon \mathfrak{T} \to \mathfrak{U}$ be the natural projection.

Corollary 114 $\mathfrak{U} \cong \mathbb{F}[v_1, v_2, \ldots]$ *as $\mathbb{F}$-vector spaces.*

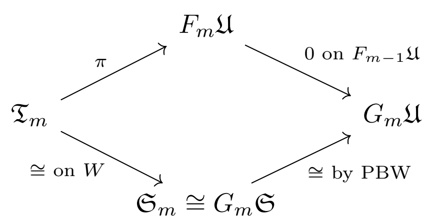
Proof As vector spaces,

$$\mathfrak{U} \cong \operatorname{gr}\mathfrak{U} \cong \mathfrak{S} \cong \mathbb{F}[v_1, v_2, \ldots].$$

The first isomorphism is unnatural, sadly, so the result is as well. The next result is just a more explicit version of this one. ∎

Corollary 115 *Suppose W is a subspace of $\mathfrak{T}_m$ which is sent isomorphically onto $\mathfrak{S}_m$ under the quotient map $\mathfrak{T} \twoheadrightarrow \mathfrak{S}$. Then $F_m\mathfrak{U} = F_{m-1}\mathfrak{U} \oplus \pi(W)$ (internal direct sum) and π is injective on W. (Note: Humphreys seems to forget to note that π is injective on W, though it's implicit in his proofs.)*

Proof Consider the diagram



This is commutative. Following W through the bottom half of the diagram, it is mapped isomorphically onto $G_m\mathfrak{U}$. But then $\pi(W)$ is mapped isomorphically onto $F_m\mathfrak{U}/F_{m-1}\mathfrak{U}$ under the quotient map $F_m\mathfrak{U} \to F_m\mathfrak{U}/F_{m-1}\mathfrak{U}$, so $U_m = U_{m-1} \oplus \pi(W)$. Further, π is injective on W since the bottom composite is an isomorphism on W. ∎

Corollary 116 $i\colon L \to \mathfrak{U}(L)$ *is injective.*

Proof Let $W = \mathfrak{T}_1 = L$, and note that $W = L = \mathfrak{S}_1$ yields the necessary isomorphism. Hence $\pi|_L = i$ is injective on $W = L$. ∎

Corollary 117 (PBW Theorem, standard) *Let $(x_1, x_2, \ldots)$ be an ordered basis of L. Then the elements $x_{i_1} \cdots x_{i_m} = \pi(x_{i_1} \otimes \cdots \otimes x_{i_m})$ for $m \in \mathbb{Z}_{\geq 1}$ with $i_1 \leq \cdots \leq i_m$, together with 1, form a basis for $\mathfrak{U}(L)$, called a PBW basis.*

Proof Let W be the subspace of $\mathfrak{T}_m$ spanned by all such $x_{i_1} \otimes \cdots \otimes x_{i_m}$. Clearly W is mapped isomorphically onto $\mathfrak{S}_m$, so $F_m\mathfrak{U} = F_{m-1}\mathfrak{U} \oplus \pi(W)$ and π is injective on W. The result follows inductively. ∎

Corollary 118 *Let H be a subalgebra of L, and extend an ordered basis $(h_1, h_2, \ldots)$ of H to an ordered basis $(h_1, \ldots, x_1, \ldots)$ of L. Then the homomorphism $\mathfrak{U}(H) \to \mathfrak{U}(L)$ induced by the injection $H \to L \to \mathfrak{U}(L)$ is itself injective, and $\mathfrak{U}(L)$ is a free $\mathfrak{U}(H)$-module with free basis consisting of all $x_{i_1} \cdots x_{i_m}$ with $i_1 \leq \cdots \leq i_m$, along with 1.*

Proof The induced map $\mathfrak{U}(H) \to \mathfrak{U}(L)$ sends $h_{i_1} \cdots h_{i_m}$ to itself, which sends a basis to distinct basis elements. The $\mathfrak{U}(H)$-module structure just multiplies on the left by elements of H, which shows that $\mathfrak{U}(L)$ is a free $\mathfrak{U}(H)$-module with the indicated basis. ∎

We finally turn to proving the PBW theorem.

Definition *Let $(x_\lambda : \lambda \in \Omega)$ be an ordered basis of L, so $\mathfrak{S}(L) \cong \mathbb{F}[z_\lambda : \lambda \in \Omega]$. For each sequence of indexes $\Sigma = (\lambda_1, \ldots, \lambda_m)$, define*

$$\begin{aligned} z_\Sigma &:= z_{\lambda_1} \cdots z_{\lambda_m} \in \mathfrak{S}_m \\ x_\Sigma &:= x_{\lambda_1} \otimes \cdots \otimes x_{\lambda_m} \in \mathfrak{T}_m \\ \lambda \leq \Sigma &\Leftrightarrow \lambda \leq \mu \text{ for all } \mu \in \Sigma \end{aligned}$$

Call Σ increasing if $\lambda_1 \leq \cdots \leq \lambda_m$. Say $\Sigma = \varnothing$ is increasing and $z_\varnothing := 1 =: x_\varnothing$.

Proposition 119 *There exists a unique L-module structure on $\mathfrak{S}(L)$ such that*

$$\begin{aligned} &x_\lambda \cdot z_\Sigma = z_\lambda z_\Sigma && \text{for } \lambda \leq \Sigma \\ &x_\lambda \cdot z_\Sigma \equiv z_\lambda z_\Sigma \pmod{F_m \mathfrak{S}} && \text{if } \Sigma \text{ has length } m. \end{aligned}$$

Proof (It's likely best to skip the proof and instead use the result as indicated afterward, to motivate it.)

We define the action by inducting on the length of Σ; in the base case, $x_\varnothing = 1$ acts as the identity. So, suppose $L \times F_m\mathfrak{S} \to \mathfrak{S}$ has been defined with the two properties above, and further suppose it's an L-module action on $L \times F_m\mathfrak{S}$. Explicitly, suppose

(A_m) $x_\lambda \cdot z_\Sigma = z_\lambda z_\Sigma$ for $\lambda \leq \Sigma, z_\Sigma \in F_m\mathfrak{S}$;

(B_m) $x_\lambda \cdot z_\Sigma - z_\lambda z_\Sigma \in F_m\mathfrak{S}$, $z_\Sigma \in F_m\mathfrak{S}$; and

(C_m) $x_\lambda \cdot (x_\mu \cdot z_T) - x_\mu \cdot (x_\lambda \cdot z_T) = [x_\lambda, x_\mu] \cdot z_T$ for all $z_T \in F_{m-1}\mathfrak{S}$.

Note that (C_m) requires (B_m) to make sense. To extend this to $L \times F_{m+1}\mathfrak{S} \to \mathfrak{S}$, we first impose (A_{m+1}) with no trouble. To define $x_\lambda \cdot z_\Sigma$ when $z_\Sigma \in F_{m+1}\mathfrak{S}$ and $\lambda \not\leq \Sigma$, we turn to C_{m+1}. In this case, take $\Sigma = (\mu, T)$, where we must have $\mu < \lambda$, and $T \in F_m\mathfrak{S}$. From A_m we have $x_\mu \cdot z_T = z_\Sigma$, and using B_m we can write

$$x_\mu \cdot (x_\lambda \cdot z_T) = x_\mu \cdot (z_\lambda z_T - y) = z_\mu z_\lambda z_T - x_\mu \cdot y$$

for some $y \in F_m\mathfrak{S}$, where A_{m+1} was applied in the second equality. Now (C_{m+1}) becomes

$$x_\lambda \cdot z_\Sigma = z_\lambda z_\Sigma - x_\mu \cdot y - [x_\lambda, x_\mu] \cdot z_T,$$

and each term on the right-hand side has already been defined. Hence we are both forced and allowed to define $L \times F_{m+1}\mathfrak{S} \to \mathfrak{S}$ on the remaining part of $L \times F_{m+1}\mathfrak{S}$ by this equality.

Written this way, it is evident that (B_{m+1}) is satisfied Moreover, we've declared (C_{m+1}) to hold when $\mu < \lambda$, $\mu \leq T$. Interchanging μ and λ in (C_{m+1}) negates both sides of the equation, so (C_{m+1}) also holds when $\lambda < \mu, \lambda \leq T$. Further, (C_{m+1}) is trivial when $\lambda = \mu$. Hence it suffices to show (C_{m+1}) holds when neither $\lambda \leq T$ nor $\mu \leq T$.

For that, write $T = (\nu, U)$, where $\nu \leq U, \lambda, \mu$. Our goal is to apply the Jacobi identity, which requires producing iterated commutators. We claim

$$\begin{aligned} x_\lambda \cdot x_\mu \cdot z_T = x_\nu \cdot x_\lambda \cdot x_\mu \cdot z_U &+ [x_\lambda, x_\nu] \cdot x_\mu \cdot z_U \\ &+ [x_\mu, x_\nu] \cdot x_\lambda \cdot z_U + [x_\lambda, [x_\mu, x_\nu]] \cdot z_U. \end{aligned} \tag{*}$$

We compute (where terms which will be manipulated to yield the next step are put in parens)

$$\begin{aligned} x_\lambda \cdot x_\mu \cdot (z_T) &= x_\lambda \cdot (x_\mu \cdot x_\nu) \cdot z_U \\ &= (x_\lambda \cdot x_\nu) \cdot x_\mu \cdot z_U + (x_\lambda \cdot [x_\mu, x_\nu]) \cdot z_U \\ &= x_\nu \cdot x_\lambda \cdot x_\mu \cdot z_U + [x_\lambda, x_\nu] \cdot x_\mu \cdot z_U \\ &\quad + [x_\mu, x_\nu] \cdot x_\lambda \cdot z_U + [x_\lambda, [x_\mu, x_\nu]] \cdot z_U. \end{aligned}$$

where we used (A_m) on z_T, (C_m) to commute $x_\mu \cdot x_\nu$, then (C_{m+1}) to commute $x_\lambda \cdot x_\nu$ and (C_m) to commute $x_\lambda \cdot [x_\mu, x_\nu]$. In showing that (C_{m+1}) applies to commute $x_\lambda \cdot x_\nu$ in $(x_\lambda \cdot x_\nu) \cdot x_\mu \cdot z_U$, we must break $x_\mu \cdot z_U$ into $z_\mu z_U + w$ for some $w \in F_{m-1}\mathfrak{S}$ using (B_{m-1}). We then must apply (C_m) to $x_\lambda \cdot x_\nu \cdot w$ and (C_{m+1}) to $x_\lambda \cdot x_\nu \cdot z_\mu z_U$, which is valid since $\nu \leq (\mu, U)$; the result is as stated.

Finally, note that we may interchange λ and μ in (*) and subtract the two resulting equations. The middle two terms cancel, leaving

$$\begin{aligned} x_\lambda \cdot x_\mu \cdot z_T - x_\mu \cdot x_\lambda \cdot z_T &= x_\nu \cdot (x_\lambda \cdot x_\mu) \cdot z_U - x_\nu \cdot (x_\mu \cdot x_\lambda) \cdot z_U \\ &\quad + [x_\lambda, [x_\mu, x_\nu]] \cdot z_U - [x_\mu, [x_\lambda, x_\nu]] \cdot z_U \\ &= (x_\nu \cdot [x_\lambda, x_\mu]) \cdot z_U - [x_\nu, [x_\lambda, x_\mu]] \cdot z_U \\ &= [x_\lambda, x_\mu] \cdot x_\nu \cdot z_U \\ &= [x_\lambda, x_\mu] \cdot z_T, \end{aligned}$$

where we rewrote the remaining terms, applied (C_m) to collapse two terms into $[x_\lambda, x_\mu]$ along with the Jacobi identity to collapse the other two terms, applied (C_m), and finally applied (A_m). ∎

Lemma 120 *The mth homogeneous component of any element of $F_m\mathfrak{T} \cap J$ lies in I.*

Proof Suppose $t \in F_m\mathfrak{T} \cap J$ has mth homogeneous component t_m. Write $t_m = \sum c_\Sigma x_\Sigma$ where Σ ranges over length m tuples and the c_Σ are scalars. By the Proposition, $\mathfrak{S}(L)$ is an L-module, hence a $\mathfrak{U}(L)$-module, so also a $\mathfrak{T}(L)$-module where J acts by 0. Hence $t \cdot 1 = 0 \in \mathfrak{S}(L)$. On the other hand, $t \cdot 1$ is a polynomial whose term of highest degree is $\sum c_\Sigma z_\Sigma$ by property (b). Hence $\sum c_\Sigma z_\Sigma = 0$, forcing $\sum c_\Sigma x_\Sigma \in I$, so $t_m \in I$. ∎

Proof (of Humphreys' PBW theorem) Let $t \in \mathfrak{T}_m$. We must show that $\pi(t) \in F_{m-1}\mathfrak{U}$ implies $t \in I$. Now π maps $F_{m-1}\mathfrak{T}$ onto $F_{m-1}\mathfrak{U}$, so there is some $t' \in F_{m-1}\mathfrak{T}$ for which $\pi(t) = \pi(t')$, so $t - t' \in J$. Now apply the preceding lemma to $t - t' \in F_m\mathfrak{T} \cap J$, which says that the homogeneous component of degree m, namely t itself, is in I. ∎