POINCARÉ-BIRKHOFF-WITT THEOREM

Steven Sy

February 17, 2006

I. BACKGROUND

Definition: Let $\mathcal{I}(V) = \bigoplus_{i=0}^{\infty} T^i(V)$, where $T^i(V)$ is the k-fold tensor product of V with itself.
product on $\mathcal{I}(V)$: $(v_1 \otimes \cdots \otimes v_k)(w_1 \otimes \cdots \otimes w_n) = v_1 \otimes \cdots \otimes v_k \otimes w_1 \otimes \cdots \otimes w_m \in T^{k+m}(V)$

Definition: Let $I = (x \otimes y - y \otimes x)_{\mathcal{I}(V)}$

Definition: Let $\mathcal{S}(V) = \mathcal{I}(V)/I$. $\mathcal{S}(V)$ is called the **symmetric algebra**

Definition: Let $\sigma : \mathcal{I}(V) \to \mathcal{S}(V)$ be the canonical map.

Comments:

1. $\mathcal{S}(V)$ inherits the grading from $\mathcal{I}(V)$: $\mathcal{S}(V) = \bigoplus_{i=0}^{\infty} S^i(V)$.

2. For $\{x_1, \cdots x_n\}$, a fixed basis of V, $\mathcal{S}(V) \cong \mathbb{F}[x_1, \cdots, x_n]$

Definition: Let $J = (x \otimes y - y \otimes x - [xy])_{\mathcal{I}(V)}$

Definition: Let $\mathcal{U}(L) = \mathcal{I}(L)/J$. $\mathcal{U}(L)$ is the **universal enveloping algebra**.

Definition: Let $\pi : \mathcal{I}(L) \to \mathcal{U}(L)$ be the canonical map.

Definition: Let $T_m = T^0 \oplus \cdots \oplus T^m$, a **filtration** on $\mathcal{I}(L)$

Definition: Let $U_m = \pi(T_m)$ and $U_{-1} = 0$

Definition: Let $G^m = U_m/U_{m-1}$ and $G = \bigoplus_{m=0}^{\infty} G^m$

Definition: Let $\eta_m : U_m \to G^m$ be the canonical map.

Definition: Let $\phi_m = \eta_m \circ \pi|_{T_m} : T^m \to G^m$.

Remark: ϕ_m is surjective, being the composition of two surjective maps.

Definition: Let $\phi : \mathcal{I}(L) \to G$ by $\phi[(a_m)_{m=0}^{\infty}] = (\phi_m(a_m))_{m=0}^{\infty}$. Then ϕ is surjective.

Definition: (Product in G): for $x \in G^m$, $y \in G^n$, define $xy = \phi_{m+n}(\tilde{x} \otimes \tilde{y})$, where $\phi_m(\tilde{x}) = x$ and $\phi_m(\tilde{y}) = y$.

Claim: $\phi : \mathcal{I}(L) \to G$ is an algebra homomorphism

Proof:

Let $\tilde{x} \in T^m$ and $\tilde{y} \in T^n$. Then $\phi(\tilde{x})\phi(\tilde{y}) = \phi_m(\tilde{x})\phi_n(\tilde{y})$.

By definition of multiplication in G, $\phi_m(\tilde{x})\phi_n(\tilde{y}) = \phi_{m+n}(\tilde{x} \otimes \tilde{y}) = \phi(\tilde{x} \otimes \tilde{y})$.

Thus $\phi(\tilde{x})\phi(\tilde{y}) = \phi(\tilde{x} \otimes \tilde{y})$. This completes the proof.

Proposition: $\phi(I) = 0$

Proof:

Since $I = (x \otimes y - y \otimes x)_{\mathcal{I}(V)}$, it suffices to show that $\phi(x \otimes y - y \otimes x) = \bar{0}$.

Now $x \otimes y - y \otimes x \in T_2$, so $\pi(x \otimes y - y \otimes x) \in U_2$.

Thus $\phi_2(x \otimes y - y \otimes x) = \eta_2(\pi(x \otimes y - y \otimes x)) \in G^2 = U_2/U_1$.

Hence $\phi_2(x \otimes y - y \otimes x) = a + U_1$ for some $a \in U_2$.

Thus $\pi(x \otimes y - y \otimes x) + U_1 = a + U_1$, so $\pi(x \otimes y - y \otimes x) - a \in U_1$.

Then $(x \otimes y - y \otimes x + J) - a \in U_1$, so $[xy] + J - a \in U_1$.

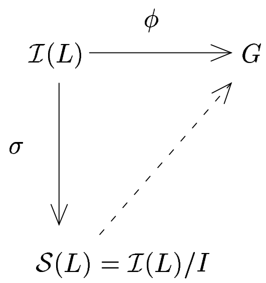
Now $[xy] + J \in U_1$, so $a \in U_1$.

Thus $\phi_2(x \otimes y - y \otimes x) = a + U_1 = U_1 = \bar{0}$.

This completes the proof.

Now by the proposition, we have that $I \subseteq \ker \phi$.

Thus the map ϕ factors through σ as in the diagram below:



Thus there exists a unique algebra homomorphism $\omega : \mathcal{S}(L) \to G$ defined by $\omega(a + I) = \phi(a)$.

Claim: ω is surjective

Proof:

> Let $g \in G$. Since ϕ is surjective, there exists $a \in \mathcal{I}(L)$ such that $\phi(a) = g$. Then $a + I \in \mathcal{S}(L)$ and $\omega(a + I) = \phi(a) = g$.
>
> Thus $\omega : \mathcal{S}(L) \to G$ is a surjective algebra homomorphism.

Poincaré-Birkhoff Witt Theorem: $\omega : \mathcal{S}(L) \to G$ is an algebra isomorphism

Thus to prove the theorem, it remains to show that ω is injective. It relies on various lemmas, and some extra development.

II. PRELIMINARY DISCUSSION

Let $(x_\lambda)_{\lambda\in\Omega}$ be an ordered basis of L. Let $I = \{x_\lambda\}_{\lambda\in\Omega}$.

Then there exists a natural isomorphism $\gamma : \mathbb{F}[I] \to \mathcal{S}(L)$.

Let $\gamma(x_\lambda) = z_\lambda \in S^1$.

For each sequence $\Sigma = (\lambda_1, \cdots, \lambda_m)$, define $z_{_\Sigma} = z_{\lambda_1} \cdots z_{\lambda_m} \in S^m$ and $x_{_\Sigma} = x_{\lambda_1} \otimes \cdots \otimes x_{\lambda_m} \in T^m$.

By commutativity of $\mathcal{S}(L)$, $\{z_{_\Sigma}\}_{\Sigma \text{ increasing}}$ is a basis of $\mathcal{S}(L)$.

Define a filtration on $\mathcal{S}(L) = \bigoplus_{i=0}^{\infty} S^i(L)$: Let $S_m = S^0(L) \oplus \cdots \oplus S^m(L)$.

Definition: For $\lambda \in \Omega$ and Σ a sequence, we say that $\lambda \le \Sigma$ if $\lambda \le \mu$ for all $\mu \in \Sigma$.

III. LEMMA A

Lemma A: For each $m \in \mathbb{W}$, there exists a unique linear function $\mathscr{f}_m : L \otimes S_m \to \mathcal{S}(L)$ such that the following conditions hold:

(A_m) $\mathscr{f}_m(x_\lambda \otimes z_{_\Sigma}) = z_\lambda z_{_\Sigma}$ for $\lambda \le \Sigma$

(B_m) For $k \le m$ and $z_{_\Sigma} \in S_k$, $\mathscr{f}_m(x_\lambda \otimes z_{_\Sigma}) - z_\lambda z_{_\Sigma} \in S_k$

(C_m) For all $z_{_T} \in S_{m-1}$, $\mathscr{f}_m(x_\lambda \otimes \mathscr{f}_m(x_\mu \otimes z_{_T})) = \mathscr{f}_m(x_\mu \otimes \mathscr{f}_m(x_\lambda \otimes z_{_T})) + \mathscr{f}_m([x_\lambda x_\mu] \otimes z_{_T})$

Then furthermore $\mathscr{f}_m|_{L\otimes S_{m-1}} = \mathscr{f}_{m-1}$.

Preliminary Comments:

1. The expressions in the equation in (C_m) are well-defined if (B_m) is proven:

   $z_T \in S_{m-1} \subseteq S_m$, so $\mathscr{f}_m(x_\mu \otimes z_{_T}) \in \mathcal{S}(L)$ is well-defined.

   Now $z_T \in S_{m-1}$, so by (B_m), $\mathscr{f}_m(x_\mu \otimes z_{_T}) - z_\mu z_{_T} \in S_{m-1} \subseteq S_m$.

   Since $z_\mu z_T \in S_m$, we have that $\mathscr{f}_m(x_\mu \otimes z_{_T}) \in S^m$.

   Similarly, $\mathscr{f}(x_\lambda \otimes z_{_T}) \in S^m$.

2. If the first part of the Lemma is proven, then both ξ_{m-1} and $\xi_m|_{L\otimes S_{m-1}}$ satisfy the given conditions, so by the uniqueness in the first part of the Lemma,

$$\xi_m|_{L\otimes S_{m-1}} = \xi_{m-1}$$

Proof of Lemma A:

We prove this by induction on m.

Base Case ($m = 0$)

$\Sigma = \emptyset$, so $z_\Sigma = z_\emptyset = 1$ and $S_0 = \mathbb{F}$.

Define $\xi_0 : L \otimes S_0 \to \mathcal{S}(L)$ by $\xi_0(x_\lambda \otimes 1) = z_\lambda$, and extend to $L \otimes S_0$ by linearity.

Since $\Sigma = \emptyset$, we have that for all $\lambda \in \Omega$, $\lambda \leq \Sigma$ and $z_\Sigma = z_\emptyset = 1 \in S_0$. Then $\xi_0(x_\lambda \otimes z_\Sigma) = \xi_0(x_\lambda \otimes 1) = z_\lambda = z_\lambda \cdot 1 = z_\lambda z_\Sigma$, so $(A0)$ is satisfied.

Then for $z_\Sigma \in S_0$, i.e. $1 \in S_0$, $\quad \xi_0(x_\lambda \otimes z_\Sigma) - z_\lambda z_\Sigma = \xi_0(x_\lambda \otimes 1) - z_\lambda = z_\lambda - z_\lambda = 0 \in S_0$, so $(B0)$ is satisfied.

$(C0)$ is satisfied vacuously.

If $\tilde{\xi}_0 : L \otimes S_0 \to \mathcal{S}(L)$ is another linear map satisfying $(A0)$, $(B0)$, and $(C0)$, then since for all $\lambda \in \Omega$, $\lambda \leq \Sigma$, we have by $(A0)$, $\tilde{\xi}_0(x_\lambda \otimes z_\Sigma) = z_\lambda z_\Sigma$. Thus $\tilde{\xi}_0(x_\lambda \otimes 1) = z_\lambda = \xi_0(x_\lambda \otimes 1)$, so $\tilde{\xi}_0 = \xi_0$. Thus ξ_0 is unique.

Hence Lemma A is true for $m = 0$.

Induction Step

Assume that the Lemma is true for $m - 1 \in \mathbb{W}$. Thus there exists a unique linear map $\xi_{m-1} : L \otimes S_{m-1} \to \mathcal{S}(L)$ satisfying (A_{m-1}), (B_{m-1}), and (C_{m-1}). We will show that the Lemma is true for m.We first show uniqueness, then existence:

Uniqueness

Assume first that there exists a linear map $\xi_m : L \otimes S_m \to \mathcal{S}(L)$ satisfying (A_m), (B_m), and (C_m).

Since $\{z_\Sigma\}_{\Sigma \text{ increasing and length } m}$ is a basis of S_m, it will suffice to show that ξ_m is uniquely defined on $x_\lambda \otimes z_\Sigma$ for all $\lambda \in \Omega$.

Let $\lambda \in \Omega$, and let Σ be an increasing sequence of length m.

Case I: $\lambda \leq \Sigma$

By (A_m), $\xi_m(x_\lambda \otimes z_\Sigma) = z_\lambda z_\Sigma$. This uniquely defines ξ_m.

Case II: $\lambda \not\leq \Sigma$

Since Σ is increasing, $\Sigma = (\mu, T)$, where $\lambda \not\leq \mu$, but $\mu \leq T$.

Then, by definition, $z_\Sigma = z_\mu z_T$.

Now by (A_{m-1}) (induction hypothesis), we have that
$z_\Sigma = z_\mu z_T = \xi_{m-1}(x_\mu \otimes z_T)$.

Since $z_T \in S_{m-1}$ and $\xi_m|_{L\otimes S_{m-1}} = \xi_{m-1}$, we have that

$$\xi_m(x_\mu \otimes z_T) = z_\Sigma. \qquad (1)$$

By (B_{m-1}) (induction hypothesis), we have that
$\xi_{m-1}(x_\lambda \otimes z_T) - z_\lambda z_T \in S_{m-1}$.

Again, since we have that $\xi_m|_{L\otimes S_{m-1}} = \xi_{m-1}$,
$\xi_m(x_\lambda \otimes z_T) - z_\lambda z_T \in S_{m-1}$.

Let $y = \xi_m(x_\lambda \otimes z_T) - z_\lambda z_T$.

Then

$$\begin{aligned}
\xi_m(x_\mu \otimes y) &= \xi_m\left(x_\mu \otimes (\xi_m(x_\lambda \otimes z_T) - z_\lambda z_T))\right) \\
&= \xi_m\left(x_\mu \otimes \xi_m(x_\lambda \otimes z_T)\right) - \xi_m(x_\mu \otimes z_\lambda z_T) \\
&= \xi_m\left(x_\mu \otimes \xi_m(x_\lambda \otimes z_T)\right) - z_\mu(z_\lambda z_T),
\end{aligned}$$

where the last line is true because $\mu \leq (\mu, T) \leq (\lambda, T)$ [Case I].

Since $y \in S_{m-1}$ and $\xi_m|_{L\otimes S_{m-1}} = \xi_{m-1}$
$\xi_{m-1}(x_\mu \otimes y) = \xi_m\left(x_\mu \otimes \xi_m(x_\lambda \otimes z_T)\right) - z_\mu(z_\lambda z_T)$.

Thus $\xi_m\left(x_\mu \otimes \xi_m(x_\lambda \otimes z_T)\right) = z_\lambda z_\mu z_T + \xi_{m-1}(x_\mu \otimes y)$, where we switched the order of z_λ and z_μ since $\mathcal{S}(L)$ is a symmetric algebra.

Hence $$\ell_m\left(x_\mu \otimes \ell_m(x_\lambda \otimes z_T)\right) = z_\lambda z_\Sigma + \ell_{m-1}(x_\mu \otimes y). \tag{2}$$

Since we assumed that ℓ_m satisfies (C_m), we have that
$\ell_m\left(x_\lambda \otimes \ell_m(x_\mu \otimes z_T)\right) = \ell_m\left(x_\mu \otimes \ell_m(x_\lambda \otimes z_T)\right) + \ell_m([x_\lambda x_\mu] \otimes z_T)$

Substituting in (1), we get
$\ell_m(x_\lambda \otimes z_\Sigma) = \ell_m\left(x_\mu \otimes \ell_m(x_\lambda \otimes z_T)\right) + \ell_m([x_\lambda x_\mu] \otimes z_T)$.

Substituting in (2), we get
$\ell_m(x_\lambda \otimes z_\Sigma) = z_\lambda z_\Sigma + \ell_{m-1}(x_\mu \otimes y) + \ell_m([x_\lambda x_\mu] \otimes z_T)$.

Since $z_T \in S_{m-1}$ and $\ell_m|_{L \otimes S_{m-1}} = \ell_{m-1}$, we have that
$\ell_m(x_\lambda \otimes z_\Sigma) = z_\lambda z_\Sigma + \ell_{m-1}(x_\mu \otimes y) + \ell_{m-1}([x_\lambda x_\mu] \otimes z_T)$.

This uniquely defines ℓ_m inductively in terms of ℓ_{m-1}.

Hence, $(*)$: $$\begin{cases} \lambda \leq \Sigma, & \ell_m(x_\lambda \otimes z_\Sigma) = z_\lambda z_\Sigma \\ \lambda \not\leq \Sigma, & \ell_m(x_\lambda \otimes z_\Sigma) = z_\lambda z_\Sigma + \ell_{m-1}(x_\mu \otimes y) + \ell_{m-1}([x_\lambda x_\mu] \otimes z_T) \end{cases}$$

Thus, in either case, ℓ_m is uniquely defined on $x_\lambda \otimes z_\Sigma$, by $(*)$.

Existence

We've shown that if ℓ_m exists, then it must be defined as in $(*)$. We now define ℓ_m as in $(*)$, and show that (A_m), (B_m), and (C_m) hold:

Thus define $\ell_m : L \otimes S_m \to \mathcal{S}(L)$ by

$(**)$: $$\begin{cases} \ell_m(x_\lambda \otimes z_\Sigma) = z_\lambda z_\Sigma; \quad \lambda \leq \Sigma \\ \ell_m(x_\lambda \otimes z_\Sigma) = z_\lambda z_\Sigma + \ell_{m-1}(x_\mu \otimes y) + \ell_{m-1}([x_\lambda x_\mu] \otimes z_T), \\ \qquad \text{where } \Sigma = (\mu, T) \text{ and } y = \ell_{m-1}(x_\lambda \otimes z_T) - z_\lambda z_T; \quad \lambda \not\leq \Sigma \end{cases}$$

and extend by linearity.

Now (A_m) holds, by definition of ℓ_m.

By $(**)$, $$\ell_m(x_\lambda \otimes z_\Sigma) - z_\lambda z_\Sigma = \begin{cases} 0; & \lambda \leq \Sigma \\ \ell_{m-1}(x_\mu \otimes y) + \ell_{m-1}([x_\lambda x_\mu] \otimes z_T); & \lambda \not\leq \Sigma \end{cases}$$

We'll now show that (B_m) holds:

Let $k < m$ and $z_{\Sigma} \in S_k$, then $\mathscr{l}_m(x_\lambda \otimes z_\Sigma) - z_\lambda z_\Sigma \in S_k$ by (B_{m-1}) (induction hypothesis).

If $k = m, \quad z_\Sigma \in S_k = S_m$, and $\lambda \leq \Sigma$, then $\mathscr{l}_m(x_\lambda \otimes z_\Sigma) - z_\lambda z_\Sigma = 0 \in S_m$.

Thus it remains to show that if $k = m$, $z_\Sigma \in S_k = S_m$, and $\lambda \not\leq \Sigma$, then $\mathscr{l}_m(x_\lambda \otimes z_\Sigma) - z_\lambda z_\Sigma = \mathscr{l}_{m-1}(x_\mu \otimes y) + \mathscr{l}_{m-1}([x_\lambda x_\mu] \otimes z_T) \in S_m$.

It will suffice to show that each term is in S_m.

Since $\{z_\Gamma\}_{\Gamma \text{ increasing and length } m-1}$ is a basis of S_{m-1}, $\quad y = \sum_{j \in J} c_{\Gamma_j} z_{\Gamma_j}$ for some finite set J.

Then

$$\begin{aligned} \mathscr{l}_{m-1}(x_\mu \otimes y) - z_\mu y &= \mathscr{l}_{m-1}\left(x_\mu \otimes \sum_{j \in J} c_{\Gamma_j} z_{\Gamma_j}\right) - z_\mu \left(\sum_{j \in J} c_{\Gamma_j} z_{\Gamma_j}\right) \\ &= \sum_{j \in J} c_{\Gamma_j} \left(\mathscr{l}_{m-1}(x_\mu \otimes z_{\Gamma_j}) - z_\mu z_{\Gamma_j}\right) \end{aligned}$$

Now for each $j \in J$, $\quad \mathscr{l}_{m-1}(x_\mu \otimes z_{\Gamma_j}) - z_\mu z_{\Gamma_j} \in S_{m-1}$, by (B_{m-1}), so $\mathscr{l}_{m-1}(x_\mu \otimes y) - z_\mu y \in S_{m-1} \subseteq S_m$.

Since $z_\mu y \in S_m$, we have that $\mathscr{l}_{m-1}(x_\mu \otimes y) \in S_m$.

Thus the first term in S_m.

Since $[x_\lambda x_\mu] \in L$, and $\{x_\lambda\}_{\lambda \in \Omega}$ is a basis of L, $[x_\lambda x_\mu]$ is some finite linear combination of elements from $\{x_\lambda\}_{\lambda \in \Omega}$. Then by linearity, and by similar reasoning to the above, we have that the second term is in S_m.

Thus (B_m) holds.

We have to show that (C_m) holds for any $z_R \in S_{m-1}$:

If $\mu < \lambda$, and $\mu \leq R$, then by the uniqueness construction, (C_m) holds.

Suppose $\lambda < \mu$ and $\lambda \leq R$. Then "(C_m) with λ and μ roles reversed" holds: $\mathscr{l}_m(x_\mu \otimes \mathscr{l}_m(x_\lambda \otimes z_R)) = \mathscr{l}_m(x_\lambda \otimes \mathscr{l}_m(x_\mu \otimes z_R)) + \mathscr{l}_m([x_\mu x_\lambda] \otimes z_R)$

Then $\mathscr{l}_m(x_\lambda \otimes \mathscr{l}_m(x_\mu \otimes z_R)) = \mathscr{l}_m(x_\mu \otimes \mathscr{l}_m(x_\lambda \otimes z_R)) - \mathscr{l}_m([x_\mu x_\lambda] \otimes z_R)$.

Since $[x_\mu x_\lambda] = -[x_\lambda x_\mu]$ and $\mathscr{l}_m$ is linear, we have (C_m).

Now suppose that $\lambda = \mu$:

Since $\xi_m(x_\lambda \otimes \xi_m(x_\lambda \otimes z_R)) = \xi_m(x_\lambda \otimes \xi_m(x_\lambda \otimes z_R)) + \xi_m([x_\lambda x_\lambda] \otimes z_R)$, (C_m) holds.

Thus (C_m) holds for any λ, μ provided $\lambda \leq R$ or $\mu \leq R$.

Now consider the case when neither $\lambda \leq R$ nor $\mu \leq R$:

Then $R = (\nu, \Psi)$, where $\nu \leq \Psi$, $\nu < \lambda$, and $\nu < \mu$.

Abbreviation: For $x \in L$ and $z \in S_m$, let $\xi_m(x \otimes z) = xz$.

Since $z_\Psi \in S_{m-2}$, by (C_{m-1}), we have

$$x_\mu x_\nu z_\Psi = x_\nu x_\mu z_\Psi + [x_\mu x_\nu] z_\Psi \quad \text{(in abbreviated notation)} \tag{3}$$

Now $z_\Psi \in S_{m-2}$, so by (B_{m-1}), $x_\mu z_\Psi - z_\mu z_\Psi \in S_{m-2}$

Thus let $w = x_\mu z_\Psi - z_\mu z_\Psi \in S_{m-2}$.

Then

$$x_\mu z_\Psi = w + z_\mu z_\Psi. \tag{4}$$

Since $w \in S_{m-2}$, (C_{m-1}) implies that

$$x_\lambda x_\nu w = x_\nu x_\lambda w + [x_\lambda x_\nu] w \tag{5}$$

Now $\nu \leq \Psi$ and $\nu < \mu$, so by the uniqueness construction again, (C_m) holds for $z_\mu z_\Psi$:

$$x_\lambda x_\nu (z_\mu z_\Psi) = x_\nu x_\lambda (z_\mu z_\Psi) + [x_\lambda x_\nu](z_\mu z_\Psi) \tag{6}$$

Adding (5) and (6) yields

$$x_\lambda x_\nu (x_\mu z_\Psi) = x_\nu x_\lambda (x_\mu z_\Psi) + [x_\lambda x_\nu](x_\mu z_\Psi) \tag{7}$$

Now

$$x_\lambda (x_\mu z_R) = x_\lambda x_\mu (z_\nu z_\Psi) = x_\lambda x_\mu (x_\nu z_\Psi), \text{ by } (A_m) \tag{8}$$

Substituting (3) into (8), we get

$$x_\lambda(x_\mu z_R) = x_\lambda x_\nu x_\mu z_\Psi + x_\lambda[x_\mu x_\nu]z_\Psi \tag{9}$$

Substituting (7) into (9), we get

$$x_\lambda(x_\mu z_R) = x_\nu x_\lambda x_\mu z_\Psi + [x_\lambda x_\nu](x_\mu z_\Psi) + x_\lambda[x_\mu x_\nu]z_\Psi \tag{10}$$

Claim: $x_\lambda[x_\mu x_\nu]z_\Psi = [x_\mu x_\nu](x_\lambda z_\Psi) + [x_\lambda[x_\mu x_\nu]]z_\Psi$ (11)

Proof:

$[x_\mu x_\nu] \in L$, so $[x_\mu x_\nu] = \sum_{k\in K} c_k x_k$, where K is a finite subset of Ω.

Then

$$\begin{aligned}
x_\lambda[x_\mu x_\nu]z_\Psi &= x_\lambda\left(\sum_{k\in K} c_k x_k\right) z_\Psi \\
&= \sum_{k\in K} c_k x_\lambda x_k z_\Psi \\
&= \sum_{k\in K} c_k\left(x_k x_\lambda z_\Psi + [x_\lambda x_k]z_\Psi\right) \qquad \text{by } (C_{m-1}) \\
&= \left(\sum_{k\in K} c_k x_k\right) x_\lambda z_\Psi + [x_\lambda \sum_{k\in K} c_k x_k]z_\Psi \\
&= [x_\mu x_\nu](x_\lambda z_\Psi) + [x_\lambda[x_\mu x_\nu]]z_\Psi
\end{aligned}$$

This proves (11).

Substituting (11) into (10), we get

$$x_\lambda(x_\mu z_R) = x_\nu x_\lambda x_\mu z_\Psi + [x_\lambda x_\nu](x_\mu z_\Psi) + [x_\mu x_\nu](x_\lambda z_\Psi) + [x_\lambda[x_\mu x_\nu]]z_\Psi \tag{12}$$

Since λ and μ were interchangeable in the above argument, we also get the identity:

$$x_\mu(x_\lambda z_R) = x_\nu x_\mu x_\lambda z_\Psi + [x_\mu x_\nu](x_\lambda z_\Psi) + [x_\lambda x_\nu](x_\mu z_\Psi) + [x_\mu[x_\lambda x_\nu]]z_\Psi \tag{13}$$

Subtracting (13) from (12), we get

$$x_\lambda(x_\mu z_R) - x_\mu(x_\lambda z_R) = x_\nu x_\lambda x_\mu z_\Psi - x_\nu x_\mu x_\lambda z_\Psi + [x_\lambda[x_\mu x_\nu]]z_\Psi - [x_\mu[x_\lambda x_\nu]]z_\Psi$$
$$= x_\nu\,(x_\lambda x_\mu z_\Psi - x_\mu x_\lambda z_\Psi) + [x_\lambda[x_\mu x_\nu]]z_\Psi + [x_\mu[x_\nu x_\lambda]]z_\Psi \quad (14)$$

By (3), using indices λ and μ, we have $x_\lambda x_\mu z_\Psi = x_\mu x_\lambda z_\Psi + [x_\lambda x_\mu]z_\Psi$.

Thus

$$x_\lambda x_\mu z_\Psi - x_\mu x_\lambda z_\Psi = [x_\lambda x_\mu]z_\Psi \quad (15)$$

Substituting (15) into (14), we get

$$x_\lambda(x_\mu z_R) - x_\mu(x_\lambda z_R) = x_\nu[x_\lambda x_\mu]z_\Psi + [x_\lambda[x_\mu x_\nu]]z_\Psi + [x_\mu[x_\nu x_\lambda]]z_\Psi \quad (16)$$

By (11), with relabeled indices, we have

$$x_\nu[x_\lambda x_\mu]z_\Psi = [x_\lambda x_\mu](x_\nu z_\Psi) + [x_\nu[x_\lambda x_\mu]]z_\Psi \quad (17)$$

Substituting (17) into (16), we get

$$x_\lambda(x_\mu z_R) - x_\mu(x_\lambda z_R) = [x_\lambda x_\mu](x_\nu z_\Psi) + ([x_\nu[x_\lambda x_\mu]] + [x_\lambda[x_\mu x_\nu]] + [x_\mu[x_\nu x_\lambda]])\,z_\Psi.$$

Thus, by the Jacobi identity, we get

$$x_\lambda(x_\mu z_R) - x_\mu(x_\lambda z_R) = [x_\lambda x_\mu](x_\nu z_\Psi)$$

Hence, by (A_m) again, we have $x_\lambda(x_\mu z_R) - x_\mu(x_\lambda z_R) = [x_\lambda x_\mu](z_\nu z_\Psi)$.

Thus, $x_\lambda(x_\mu z_R) - x_\mu(x_\lambda z_R) = [x_\lambda x_\mu]z_R$.

Rewriting this, we have

$$\xi_m(x_\lambda \otimes \xi_m(x_\mu \otimes z_R)) = \xi_m(x_\mu \otimes \xi_m(x_\lambda \otimes z_R)) + \xi_m([x_\lambda x_\mu] \otimes z_R).$$

Thus (C_m) holds.

Thus ξ_m exists via the definition given.

This completes the induction step.

Thus Lemma A is proven.

IV. LEMMA B

Lemma B: There exists a representation $\rho : L \to \mathfrak{gl}(\mathcal{S}(L))$ such that the following two conditions hold:

a. $\rho(x_\lambda)z_\Sigma = z_\lambda z_\Sigma$ for $\lambda \leq \Sigma$

b. $\rho(x_\lambda)z_\Sigma \equiv z_\lambda z_\Sigma \pmod{S_m}$, if Σ has length m

Proof:

For each $m \in \mathbb{W}$, let $\mathfrak{f}_m : L \otimes S_m \to \mathcal{S}(L)$ be the unique linear map from Lemma A.

For $x_\lambda \in L$, define $\rho_{x_\lambda} : \mathcal{S}(L) \to \mathcal{S}(L)$ by $a \mapsto \mathfrak{f}_m(x_\lambda \otimes a)$ for $a \in S_{m-1}$, and extend by linearity.

By linearity, $\rho_{[x_\lambda x_\mu]}(a) = \mathfrak{f}_m([x_\lambda x_\mu] \otimes a)$.

Then

$$\rho_{x_\lambda}\rho_{x_\mu}(a) = \rho_{x_\lambda}(\mathfrak{f}_m(x_\mu \otimes a)) = \mathfrak{f}_m(x_\lambda \otimes \mathfrak{f}_m(x_\mu \otimes a)). \qquad (1)$$

By (C_m), we have that

$$\mathfrak{f}_m(x_\lambda \otimes \mathfrak{f}_m(x_\mu \otimes a)) = \mathfrak{f}_m(x_\mu \otimes \mathfrak{f}_m(x_\lambda \otimes a)) + \mathfrak{f}_m([x_\lambda x_\mu] \otimes a). \qquad (2)$$

Substituting (2) into (1), we get

$$\rho_{x_\lambda}\rho_{x_\mu}(a) = \mathfrak{f}_m(x_\mu \otimes \mathfrak{f}_m(x_\lambda \otimes a)) + \mathfrak{f}_m([x_\lambda x_\mu] \otimes a)$$

Thus, $\rho_{x_\lambda}\rho_{x_\mu}(a) = \rho_{x_\mu}\rho_{x_\lambda}(a) + \rho_{[x_\lambda x_\mu]}(a)$.

Thus $\rho_{[x_\lambda x_\mu]} = \rho_{x_\lambda}\rho_{x_\mu} - \rho_{x_\mu}\rho_{x_\lambda}$.

Define $\rho : L \to \mathfrak{gl}(\mathcal{S}(L))$ by $x_\lambda \mapsto \rho_{x_\lambda}$, and extend by linearity.

Then $\rho([x_\lambda x_\mu]) = \rho_{[x_\lambda x_\mu]} = \rho_{x_\lambda}\rho_{x_\mu} - \rho_{x_\mu}\rho_{x_\lambda} = \rho(x_\lambda)\rho(x_\mu) - \rho(x_\mu)\rho(x_\lambda) = [\rho(x_\lambda), \rho(x_\mu)]$.

Thus ρ is a Lie algebra homomorphism, so ρ is a representation.

Since $\rho(x_\lambda)z_\Sigma = \rho_{x_\lambda}(z_\Sigma) = \mathfrak{f}_m(x_\lambda \otimes z_\Sigma)$, (a) and (b) are satisfied by (A_m) and (B_m).

This proves Lemma B.

V. LEMMA C

Lemma C: Let $t \in T_m \cap J$. Then the homogeneous component t_m of t of degree m lies in I

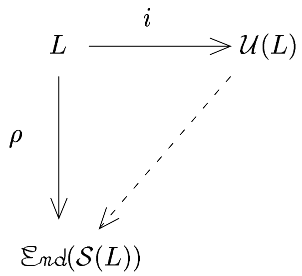
Proof:

Let $t \in T_m \cap J$.

Then $t_m = \sum_i c_{im} x_{\Sigma_m(i)}$, where each $\Sigma_m(i)$ is of length m.

Let $\rho : L \to \mathfrak{gl}(\mathcal{S}(L)) = \mathcal{E}nd(\mathcal{S}(L))$ be as in Lemma B.

Since $(\mathcal{U}(L), i)$ is a universal enveloping algebra, we have the following diagram:



By the Universal Property, there exists a unique homomorphism $\tilde{\rho} : \mathcal{U}(L) \to \mathcal{E}nd(\mathcal{S}(L))$

Define $\tilde{\tilde{\rho}} : \mathcal{I}(L) \to \mathcal{E}nd(\mathcal{S}(L))$ by $\tilde{\tilde{\rho}}(a) = \tilde{\rho}(a + J)$.

Then $\tilde{\tilde{\rho}}$ is an algebra homomorphism and $J \subseteq \ker \tilde{\tilde{\rho}}$.

Since $t \in J$, $\quad \tilde{\tilde{\rho}}(t) = 0. \qquad (1)$

Now $t \in T_m$, so $t = \sum_{i,j} c_{ij} x_{\Sigma_j(i)}$, where Σ_j is a sequence of length j.

$$\text{Then } \tilde{\tilde{\rho}}(t) = \tilde{\tilde{\rho}}\left(\sum_{i,j} c_{ij} x_{\Sigma_j(i)}\right) = \sum_{i,j} c_{ij} \tilde{\tilde{\rho}}(x_{\Sigma_j(i)}) = \sum_{i,j} c_{ij} (\tilde{\tilde{\rho}}(x))_{\Sigma_j(i)} = \sum_{i,j} c_{ij} (\rho(x))_{\Sigma_j(i)}.$$

Thus $\tilde{\tilde{\rho}}(t) \cdot 1 = \sum_{i,j} c_{ij} \left[(\rho(x))_{\Sigma_j(i)} i \cdot 1\right]$.

By Lemma B, part (a), we have $\tilde{\tilde{\rho}}(t) \cdot 1 = \sum_{i,j} c_{ij} z_{\Sigma_j(i)}$.

By (1), we have $\sum_{i,j} c_{ij} z_{\Sigma_j(i)} = 0$.

By linear independence, $\sum_i c_{im} z_{\Sigma_m(i)} = 0$, for Σ of length m. (2)

Now for $\sigma : \mathcal{I}(V) \to \mathcal{I}(V)/I = \mathcal{S}(V)$, we have

$$\sigma(t_m) = \sigma\left(\sum_i c_{im} x_{\Sigma_m(i)}\right) = \sum c_{im} \sigma(x_{\Sigma_m(i)}) = \sum c_{im} z_{\Sigma_m(i)}.$$

By (2), we have that $\sigma(t_m) = 0$. Thus $t_m \in \mathcal{K}er\, \sigma = I$.

This proves Lemma C.

VI. POINCARÉ-BIRKHOFF-WITT THEOREM

Poincaré-Birkhoff Witt Theorem: $\omega : \mathcal{S}(L) \to G$ is an algebra isomorphism

Proof:

By Section I, $\omega : \mathcal{S}(L) \to G$ is a surjective algebra homomorphism.

It remains to show that ω is injective.

Let $\omega(\bar{t}) = 0$. Specifically, let $\bar{t} = t + I \in S^m$ such that $\omega(t + I) = 0$.

Then for $\phi : \mathcal{I}(L) \to G$, $\phi(t) = 0$ for $t \in T^m$ (definition of ω)

Hence $\phi_m(t) = 0$ for $t \in T^m$.

Thus $(\eta_m \circ \pi|_{T_m})(t) = 0$ for $t \in T^m$ (definition of ϕ_m), where
$\pi : \mathcal{I}(L) \to \mathcal{I}(L)/J = \mathcal{U}(L)$ and $\eta_m : U_m \to U_m/U_{m-1} = G^m$.

Hence $\pi(t) \in \mathcal{K}er\, \eta_m = U_{m-1} = \pi(T_{m-1})$.

Thus there exists $t' \in T_{m-1}$ such that $\pi(t) = \pi(t')$.

Then $\pi(t - t') = 0$, so that $t - t' \in \mathcal{ker}\,\pi = J$.

Since $t' \in T_m$ and $t \in T_m$, $t - t' \in T_m$. Thus $t - t' \in T_m \cap J$.

Since t is the homogeneous component of $t - t'$ of degree m, by Lemma C, we have that $t \in I$. Thus $\bar{t} = t + I = 0$.

This completes the proof.