Lie Groups: Fall, 2022
Lecture IIIA: The Universal Enveloping Algebra, Free Lie Algebras, and the Baker-Campbell-Hausdorff Formula

September 27, 2022

1 PBW Theorem concerning the Universal Enveloping Algebra of a Lie Algebra

The Poincaré-Birkhoff-Witt Theorem (PBW Theorem). says that every finite dimensional Lie algebra is a sub Lie algebra of the Lie algebra coming from an associative algebra. There is a universal such associative algebra which is called the *Universal Enveloping Algebra* of the Lie algebra.

1.1 The Construction

Definition 1.1. By a *linear representation* of a Lie algebra L on a vector space V we mean a linear map $\rho\colon L \to \mathrm{End}(V)$ that satisfies

$$\rho([X,Y]) = \rho(X)\rho(Y) - \rho(Y)\rho(X).$$

N.B. If $\rho\colon G \times V \to V$ is a linear representation of a Lie group G, then the differential at the identity $D_e\rho\colon \mathfrak{g} \to End(V)$ is a linear representation of the Lie algebra $\mathfrak{g}$ of G.

Let $(L, [\cdot,\cdot])$ be a Lie algebra over a field. Consider the tensor algebra

$$T(L) = \sum_{n=0}^{\infty} \otimes^n L$$

with the usual (associative) multiplication defined by juxtaposition of tensors. This is the free associative algebra generated by L in the sense that given an associative algebra A and a linear map $\psi\colon L \to A$ there is a unique extension of ψ to a map of associative algebras $T(L) \to A$.

We define the *universal enveloping algebra of* L, denoted $U(L)$ to be the quotient of $T(L)$ by the two-sided ideal generated by $(x \otimes y - y \otimes x - [x, y])$ for all $x, y \in L$. Then $U(L)$ is an associative algebra and the natural map $L \to U(L)$ is a homomorphism of Lie algebras when $U(L)$ is given the $AB - BA$ Lie bracket coming from its associative multiplication.

Clearly, any linear representation of the Lie algebra $L \to \mathrm{End}(V)$ on a real vector space extends to a unique algebra homomorphism $U(L) \to \mathrm{End}(V)$. Indeed, $L \to U(L)$ is the universal solution to the problem of mapping L to the Lie algebra determined by an associative algebra. For, if we have an associative algebra A and a linear map $\rho\colon L \to A$ with $\rho([X, Y]) = \rho(X)\rho(Y) - \rho(Y)\rho(X)$ then there is a unique map $T(L) \to A$ extending ρ. Since ρ is a map of Lie algebras, it sends every defining relation for $U(L)$ to zero in A. Hence, it factors to give a map of associative algebras $U(L) \to A$. This is the unique map of associative algebras $U(L) \to A$ extending ρ. There is a subtlety here. Is $L \to U(L)$ an injective linear map? If it has a kernel, this means that every linear representation of L has a non-trivial kernel. In fact, the PBW Theorem says that this is not the case.

Theorem 1.2. *(PBW) Let L be a finite dimensional Lie algebra. We denote its underlying vector space by V. The natural increasing filtration on $T(V)$ defined by $F_n(T(V)) = \sum_{k=0}^{n} \otimes^k V$ induces an increasing filtration of $U(L)$, also denoted $\{F_n\}$. This is a multiplicative filtration in the sense that the multiplication induces a map $F_n \otimes F_m \mapsto F_{n+m}$. It follows that there is a graded algebra structure on the associated graded*

$$F(L) = \oplus_{n=0}^{\infty} F_n(U(L)/F_{n-1}(U(L)).$$

There is an isomorphism between this graded algebra and the polynomial algebra on L which is the identity on $L = \otimes^1 L$.

Corollary 1.3. *The natural map $L \to U(L)$ is an injection.*

Corollary 1.4. *Every (finite dimensional) Lie algebra is a sub Lie algebra of a Lie algebra given by the $AB - BA$ Lie bracket of an associative algebra.*

Proof. (of the theorem) We fix a basis $\{X_i\}_{i \in I}$ and choose a total ordering for the basis, or equivalently a total ordering on the index set I. For a finite sequence $J = \{j_1, \ldots, j_k\}$ of elements of I and $i \in I$, the notation $i \leq J$ means that $i \leq j_r$ for all $j_r \in J$. We denote the number of elements in the sequence J by $|J|$. We have a corresponding basis for the polynomial algebra on L, denoted $P(L)$. To avoid confusion we use the notation z_i for the variable associated to X_i, so that the algebra is the algebra of polynomials in the $\{z_i\}_{i\in I}$ with a basis being the monomials $z_J = \prod_{j\in J} z_j$ as J ranges over the finite sequences that are weakly ordered: $\{j_1, \ldots, j_t\}$ such that $j_i \leq j_k$ for all $i < k$. Our goal is to define an action $\sigma\colon L \otimes P(L) \to P(L)$. We do this by induction on the degree p of the polynomial. The inductive hypothesis for p is that we have a map $\sigma_p\colon L \otimes P^p(L) \to P(L)^{p+1}$ satisfying the following:

$A(p)$: If $i \leq J$ for $|J| \leq p$, then $\sigma_p(X_i)z_J = z_i z_J$.

$B(p)$: For any J with $|J| = q \leq p$ we have $\sigma_p(X_i)z_J - z_i z_J \in P(L)^q$.

$C(p)$: For any J with $|J| < p$,

$$\sigma_p(X_i)\sigma_p(X_j)z_J - \sigma_p(X_j)\sigma_p(X_i)z_J = \sigma_p([X_i, X_j])z_J.$$

(Also, $\sigma_p|_{P(L)^{p-1}} = \sigma_{p-1}$.)

We construct the maps by induction on the $P(L)^p$. For $p = 0$, it follows from Condition $A(0)$ that $\sigma_0(X_i)1 = z_i$. We take this as our definition of σ_0.

Now suppose that $\sigma_{p-1}\colon L \otimes P^{p-1}(L) \to P(L)^p$ is defined satisfying Conditions $A(p-1)$, $B(p-1)$, and $C(p-1)$. We define σ_p on all monomials z_J with $|J| = p$. If $i \leq J$, the Condition $A(p)$ requires $\sigma_p(X_i)z_J = z_i Z_J$. Otherwise, re-odering J we have $J = (k, K)$ with $k \leq K$ and $k < i$. By the inductive hypothesis $Z_J = \sigma_{p-1}(X_k)Z_K$. We invoke Condition $C(p)$ and define

$$\sigma_p(X_i)z_J = \sigma_p(X_i)\sigma_{p-1}(X_k)Z_K = \sigma_p(X_k)\sigma_{p-1}(X_i)(z_K) + \sigma_{p-1}([X_i, X_k])Z_K.$$

Notice the first term $\sigma_p(X_k)\sigma_{p-1}(X_i)z_K = \sigma_p(X_k)(z_i Z_K) + \sigma_{p-1}(X_k)(w)$ for some $w \in P(L)^{p-1}$. Since $k \leq (i, K)$, $\sigma_p(X_k)(z_i z_K)$ is already defined by $A(p)$. This shows that there is at most one extension of σ_{p-1} to σ_p satisfying $A(p)$, $B(p)$, and $C(p)$.

Clearly, Conditions $A(p)$ and $B(p)$ hold for σ_p. It remains to show that $C(p)$ holds. By construction it holds for $\sigma_p(X_i)\sigma_{p-1}(X_j)z_K$ if $j \leq K$ and $i \geq j$. By symmetry it holds if $i \leq K$ and $j \geq i$. Thus, the remaining cases are where $K = (k, M)$ with $k \leq M$ and $k < i, j$. To simplify the notation we drop σ_p and σ_{p-1} from the notation and simply write the product as a juxtaposition. By induction and the cases where we already know that $C(p)$

holds we have

$$\begin{aligned} X_iX_jz_K &= X_iX_jX_kz_M = X_iX_kX_jz_M + X_i[X_j,X_k]z_M \\ &= X_kX_iX_jz_M + [X_i,X_k]X_jz_M + X_i[X_j,X_k]z_M \\ &= X_kX_jX_iz_M + X_k[X_i,X_j]z_M + [X_i,X_k]X_jz_M + X_i[X_j,X_k]z_M \end{aligned}$$

By the symmetric argument we have

$$X_jX_iX_kz_M = X_kX_jX_iz_M + [X_j,X_k]X_iz_M + X_j[X_i,X_k]z_M.$$

The first equation minus the second one yields:

$$\begin{aligned} X_iX_jz_K - X_jX_iz_K =& X_k[X_i,X_j]z_M[+[X_i,X_k],X_j]z_M \\ &+ X_i[X_j,X_k]z_M - [X_j,X_k]X_iz_M - X_j[X_i,X_k]z_M \\ =&([[X_i,X_k],X_j] + [X_i,[X_j,X_k]] + X_k[X_i,X_j])z_M. \end{aligned} \tag{1.1}$$

The Jacobi identity tells us that

$$[[X_i,X_k],X_j]+[[X_i,[X_j,X_k]]] = -[[X_k,[X_i,X_j]] = [X_i,X_j]X_k - X_k[X_i,X_j].$$

Thus, Eequation 1.1 becomes

$$X_iX_jz_K - X_jX_iz_K = [X_i,X_j]X_kz_M = [X_i,X_j]z_K.$$

This completes the proof of property $C(p)$ and hence completes the inductive proof of the existence of the action $\sigma\colon L\otimes P(L)\to P(L)$ with properties $A(p), B(p), C(p)$ for all $p\geq 0$.

By Condition $C(p)$ for all p, the map resulting map $\sigma\colon L\to End(P(L))$ is a map of Lie algebras and hence extends to an action of $U(L)\otimes P(L)\to P(L)$. By definition $\sigma(X_i)z_M = z_iz_M$ modulo $F_{|M|}P(L)$ and hence

$$X_{i_1}\cdots X_{i_t}z_M = z_{i_1}\cdots z_{i_t}z_M \quad \text{modulo} \quad F_{|M|+t}U(L).$$

We define a map $U(L)\to P(L)$ be sending $a\in U(L)$ to $\varphi(a)=a\cdot 1$. Then $\varphi\colon U(L)\to P(L)$ is compatible with the gradings by degree. We define the associated graded $FU(L)$ to the increasing filtration by degree $F_nU(L)$. The associated graded to φ, denoted $F(\varphi)$, induces a map of graded algebras $F(\varphi)\colon FU(L)\to P(L)$ sending the element $X_{i_1}\cdots X_{i_t}$ to the monomial $z_{i_1}\cdots z_{i_t}$. This map of graded algebras is clearly surjective.

We claim that it is also injective. Since every element of $F_nU(L)$ is represented by a sum of monomials of degrees $\leq n$ and any two monomials that involve exactly the same X_i each the same number of times, just in different orders, are equal modulo $F_{n-1}U(L)$ it follows that $F_nU(L)/F_{n-1}U(L)$ is a quotient of the vector space generated by the monomials of length n on weakly ordered sequences. Since these elements map via $F(\varphi)$ to a basis for the homogeneous polynomials of degree n, it follows that the these monomials of degree n are a basis for $F_nU(L)/F_{n-1}U(L)$, and hence that $F(\varphi)$ is an isomorphism of graded algebras. □

1.2 The Co-multiplication of $U(L)$

We give $U(L) \otimes U(L)$ the product associative algebra structure.

We define a map $c\colon L \to U(L) \otimes U(L)$ by $c(x) = x \otimes 1 + 1 \otimes x$..

Proposition 1.5. *c extends uniquely to an algebra map $c\colon U(L) \to U(L) \otimes U(L)$. The map c is a co-associative and co-commutative and has a co-unit*

Proof. We define c on the tensor algebra $T^*(L)$ by multiplicativity, giving an algebra map $T^*(L) \to U(L) \otimes U(L)$. It descends to an algebra map $U(L) \to U(L) \otimes U(L)$ because $c(x)c(y) - c(y)c(x) = c([x, y])$ for all $x, y \in L$.

Since $c(x)$ is symmetric under interchange of factors for all $x \in L$, the image of c is symmetric under this interchange. This is the definition of *co-commutative*. Similarly, for all $x \in L$ we have

$$(1 \otimes c) \circ c(x) = x \otimes 1 \otimes 1 + 1 \otimes x \otimes 1 + 1 \otimes 1 \otimes x = (c \otimes 1) \circ c(x),$$

from which it follows that $(1 \otimes c) \circ c = (c \otimes 1) \circ c$ on all elements, which is the definition of *co-associative*. Finally, the *co-unit* of c is the map $U(L) \to \mathbb{R}$ of unital algebras that sends $x \in L$ to zero for all $x \in L$. □

Definition 1.6. An element $x \in U(L)$ is *primitive* if $c(x) = x \otimes 1 + 1 \otimes x$

Lemma 1.7. *The primitive elements form a real vector subspace of $U(L)$.*

Proof. Exercise. □

Clearly, every $x \in L \subset U(L)$ is primitive. We define the *standard co-multiplication* c_0 on the polynomial algebra $P(V)$. It is characterized by $c_0(v) = v \otimes 1 + 1 \otimes v$ and c_0 is a homomorphism of associative, commutative algebras.

Claim 1.8. *In the polynomial algebra $P(V)$ (over a field of characteristic zero) the only primitive elements for the standard co-multiplication are the elements on V.*

Proof. In the polynomial algebra every homogeneous polynomial of degree n in $P(L)$ is a linear combination of n^{th} powers of elements in L. In $P(L)$ with its standard co-multiplication c' the general homogeneous element of degree n can be written as

$$c_0(\sum_i \lambda_i a_i^n) = \sum_i \lambda_i \left(\sum_k \binom{n}{k} a_i^k \otimes a_i^{n-k} \right).$$

Thus, the term of bi-degree (1,n-1) in $c_0(\sum_i \lambda_i a_i^n)$ is $n \sum_i \lambda_i a_i \otimes a_i^{n-1}$. Hence, for $n > 1$ if this element is primitive then $n \sum_i \lambda_i a_i \otimes a_i^{n-1} = 0$. But the image of this element under the multiplication map is $n \sum_i \lambda_i a_i^n = 0$, implying that the element is zero. This shows that the only primitive elements in $P(L)$ are of degree 1 and hence are elements of L. □

There is an analogous proposition for $U(L)$.

Proposition 1.9. *The primitive elements in $U(L)$ for the co-multiplication are exactly the elements in L.*

Proof. We define an increasing filtration $F_n[U(L)\otimes U(L)] = \sum_{i+j\leq n} U^i(L)\otimes U^j(L)$. Then $c\colon U(L) \to U(L) \otimes U(L)$ preserves the filtration and hence induces a co-multiplication $c' = F^*(c)$ on $FU(L)$, which is a homomorphism of algebras with every element in degree 1 being primitive. Thus, under the identification of $FU(L)$ with $P(L)$ the co-multiplication c' becomes the standard co-multiplication c_0 on polynomials.

Suppose that $a \in U(L)$ is primitive and non-zero. Since no multiple of the identity is primitive, there is $n \geq 1$ such that $a \in F_nU(L)$ and has non-trivial projection to $F_nU(L)/F_{n-1}U(L)$. We show that $n = 1$. Let $\overline{a} \in U(L)_n/U(L)_{n-1}$ be the image of a. Since $\overline{a} \in FU(L)$ is primitive, under the identification of $FU(L)$ with $P(L)$ $\overline{a}$ is identified with a primitive element for c_0. it follows from the previous claim that $\overline{a} = 0$ unless $n = 1$. But by construction $\overline{a} \neq 0$. This implies that $n = 1$. Thus, a is the sum of an element in L and a multiple of the identity: $a = x + \lambda 1$ where $x \in L$ and λ is in the ground field. But $c(x + \lambda 1) = x \otimes 1 + 1 \otimes x + \lambda 1 \otimes 1$, so that this element is primitive if and only if $\lambda = 0$ and consequently, if and only if $a \in L$. □

1.3 Free Lie Algebras

Let S be a set. (We are primarily interested in the case when S has cardinality 2.) By induction on i we define sets S_i. We begin with $S_1 = S$.

Given S_i for $i < n$. we define $S_n = \coprod_{i+j=n; i,j \geq 1} S_i \times S_j$. We can view S_n as all expressions that are a composition of ordered binary products of pairs of elements. Fox example, $S_2 = (x \cdot y)$ for $x, y \in S_1$. S_3 has elements of the form $x_1 \cdot (x_2 \cdot x_3)$ and $((x_1 \cdot x_2) \cdot x_3)$ for $x_1, x_2, x_3 \in S$. S_4 has elements of the form $((x_1 \cdot x_2) \cdot (x_3 \cdot x_4))$ as well as elements such as $(x_1 \cdot ((x_2 \cdot x_3) \cdot x_4))$, and many others. We set $S_\infty = \coprod S_n$. We define the free non-associative algebra generated by S, denoted $F(S)$ to be the $\mathbb{R}$-vector space spanned by S_∞. The multiplication of $x \in S_i$ and $y \in S_j$ is the element $(x, y) \in S_i \times S_j \subset S_{i+j}$. Given this multiplication on the basis elements we extend by linearity to a multiplication on $F(S)$. The freeness of $F(S)$ is captured in the following property. Given any not necessarily associative algebra A and a set function $S \to A$, it extends uniquely to a map of algebras $F(S) \to A$.

[From the perspective of operads, consider the operad whose operations of order n are the set of rooted trivalent tress whose leaves are numbered $1, \ldots, n$. The composition law at position i of a tree T with r leaves and a tree T' with s leaves is obtained by attaching the root of T' to the i^{th} leaf of T and then renumbering the leaves, starting with the first $i-1$ of T in order, then the s of T' in order and finally the last $r - i - 1$ of T in order. This produces a new rooted trivalent tree with leaves numbers $1, \ldots, s + t - 1$. A magma on a set S is the same thing as the set of all rooted trees (with numbered leaves) together with a function of the leaves of the tree to the generating set of the magma. The product operation for magmum is given by the operad composition. This operad is equivalent to the associahedron: all ways of legitimately associating a product of n elements with a given order in a non-associative, non-commutative algebra.]

We define the free Lie algebra generated by S, denoted $L(S)$ to be the quotient of $F(S)$ by the two-sided ideal generated by $Q(a.a) = a \cdot a$ and $J(a, b, c) = a \cdot (b \cdot c) + c \cdot (a \cdot b) + b \cdot (c \cdot a)$. The proper way now to write an element of this quotient is to replace the parentheses by brackets. So that each element of S_∞ is a legitimate expression in the Lie algebra generated by $L(S)$ and we have imposed by fiat the skew symmetry and Jacobi identity (the multi-linearity over $\mathbb{R}$ comes from the fact that we have an algebra over $\mathbb{R}$). Given any function of S to a Lie algebra L, the function extends uniquely to a homomorphism of Lie algebras $L(S) \to L$. To see this, first use the universal property of $F(S)$ to define an algebra map $\psi \colon F(S) \to L$ extending $S \to L$ and sending the product in $F(S)$ to the bracket in L. Then notice that the generators of the two-sided ideal $Q(a, a)$ and $J(a, b, c)$ map to zero in L since L is a Lie algebra. That implies that the two-sided ideal generated by these elements maps to zero in L and hence $F(S) \to L$ factors through the quotient $L(S)$, and thus defines an algebra homomorphism $L(S) \to L$.

Let $A(S)$ be the free associative algebra generated by S. It has an $\mathbb{R}$-basis consisting of all monomials in S, and multiplication of monomials is juxtaposition: $m_1 \otimes m_2 \mapsto m_1 m_2$. This algebra has the property that if A is any associative algebra and $S \to A$ is a set function, then there is a unique algebra map $A(S) \to A$ extending the given map $S \to A$.

Claim 1.10. *1. $L(S) = \oplus_{n\geq 1} L(S)_n$ is a graded Lie algebra and $U(L(S))$ inherits a grading from that of $L(S)$. The algebra $A(S)$ is graded by the degree of the monomial.*

2. The inclusion $S \to A(S)$ extends uniquely to a linear map $L(S) \to A(S)$ sending the Lie bracket to the $AB - BA$ bracket in $A(S)$. By the universal property of $U(L(S))$, this map induces an algebra homomorphism $\varphi\colon U(L(S)) \to A(S)$. This map is an isomorphism of algebras.

3. Furthermore, $\varphi\colon U(L(S)) \to A(S)$ is an isomorphism of graded algebras.

Proof. The grading on $F(S)$ is given by the grading on S_∞. The two-sided ideal I whose quotient is $L(S)$ is generated by homogeneous elements so that $L(S)$ inherits a grading. The tensor algebra $T(L(S)))$ inherits a grading from the grading on $L(S)$ and with this grading the defining relations $x \otimes y - y \otimes x = [x, y]$ are homogeneous. Hence, the quotient of the two-sided ideal that these relations generate is a homogeneous ideal and therefore $U(L(S))$ is a graded algebra. The degree of a product in $U(L(S))$ of homogeneous elements in $L(S)$ is the sum of the degrees of these elements. Likewise the algebra $A(S)$ is graded by the degree of monomials in the same way.

By the universal property of $L(S)$ the inclusion of $S \to A(S)$ induces a Lie algebra homomorphism $L(S) \to A(S)$, which in turn by the universal property of $U(L(S))$ induces an algebra homomorphism $\varphi\colon U(L(S)) \to A(S)$. On the other hand, the universal property of $A(S)$ implies that the inclusion $S \to U(L(S))$ extends to an algebra homomorphism $\psi\colon A(S) \to U(L(S))$. Both $\varphi\circ\psi$ and $\psi\circ\varphi$ are the identity on S and hence by the uniqueness part of the universal properties of $A(S)$ and $U(L(S))$ both compositions are the identity. Thus, they are inverse isomorphisms.

Since the generators S have grading one, the map $L(S) \to A(S)$ is a graded map. Then the extension to $U(L(S))$ also preserve the gradings. Thus, φ and ψ are inverse isomorphisms of graded algebras. $\square$

Corollary 1.11. *There is a co-multiplication $c\colon A(S) \to A(S)$ that is a homomorphism of graded algebras and whose space of primitive elements is $(S) \subset A(S)$.*

$A(S)$ is a graded algebra with $A^0(S)$ equal $\mathbb{R}$ with $1 \in \mathbb{R}$ the unit or the algebra. Define $F^n(A(S)) = \oplus_{k\geq n} A_k(S)$. This is a decreasing filtration and $\cap_{n\geq 0} F^n(A(S)) = 0$. Thus, we can form the completion $\hat{A}(S)$ of $A(S)$ with respect to the $F^n(A(S))$. In this case, $\hat{A}(S)$ is simply formal sums of elements in $A(S)$, namely

$$\hat{A}(S) = \left\{ \sum_{n=0}^{\infty} a_n \,|\, \forall n,\ a_n \in A(S)_n \right\}.$$

We let $\widehat{Lie}(S)$ be the closure of $L(S) \subset A(S)$ in $\hat{A}(S)$. Then

$$\widehat{L}(S) = \left\{ \sum_{n=0}^{\infty} x_n \,|\, \forall n,\ x_n \in L(S)_n \right\}.$$

Corollary 1.12. *Let $B_n = \sum_{i+j=n} A(S)_i \otimes A(S)_j$ and set $\hat{B}(S) = \prod_{n\geq 0} B_n$. The co-multiplication in Corollary 1.11 induces a map $\hat{A}(S) \to \hat{B}(S)$ sending $\sum_n a_n$ to $\sum_n c(a_n)$. This makes sense because $c(a_n) \in B_n$. Let $\delta'(\sum_n a_n) = \sum_n a_n \otimes 1$ and $\delta''(\sum_n a_n) = \sum_n 1 \otimes a_n$ be maps of $\hat{A}(S) \to \hat{B}(S)$. Then*

$$c(\sum_n a_n) = \delta'(\sum_n a_n) + \delta''(\sum_n a_n)$$

if and only if $a_n \in L(S)$ for all $n \geq 0$; i.e., if and only if $\sum_n a_n \in \widehat{L}(S)$.

Proof. All of this is immediate from the fact that the only primitive elements in $A(S)$ are the elements or $L(S)$. □

We call elements satisfying the equation in the corollary *primitive*.

1.4 Formal Power Series in $\hat{A}(S)$

The reason for introducing the completions to series is so that our power series will have meaning, without having to worry about convergence issues.

Now we consider case when $S = \{A, B\}$. Consider

$$\exp(A) = \sum_{n\geq 0} \frac{A^n}{n!}; \quad \exp(B) = \sum_{n\geq 0} \frac{B^n}{n!}.$$

These formal power series are elements in $\hat{A}(S)$ and as is their product

$$\sum_{n\geq 0}\left(\sum_{i+j=n}\frac{A^iB^j}{i!j!}\right).$$

Now consider

$$\log(\exp(A)\exp(B))=\sum_{n\geq 1}\frac{(-1)^{n-1}}{n}\left(\sum_{r,s\geq 0}\frac{A^rB^s}{r!s!}\right)^n.$$

Let $e(x)=\sum_{n\geq 1}x^n/n!$ and $\ell(x)=\sum_{k\geq 1}(-1)^k x^k/k$. These power series are well defined on $F_1(\hat{A}(S))$, and take values in $\hat{A}(S)$. The reason is that for $x\in F_1(\hat{A}(S))$, all but finitely many of the terms in the series for e or ℓ vanish modulo $F_n(\hat{A}(S))$. Thus, the infinite sum represents a well-defined element of the inverse limit of $\hat{A}(S)/F_n(\hat{A}(S))$, which is $\hat{A}(S)$. Now for $t\in(-1,1)$ the power series for $e(t)$ and $\ell(t)$ are convergent and converge to $\exp(t)-1$ and $\log(1+t)$. Thus, for t sufficiently close to 0 these are inverse functions: we have $e(\ell(t))=t$ and $\ell(e(t))=t$.

This leads to finite algebraic equations for the coefficients of the composed power series. Namely, composing the power series for e and ℓ in either order applied to $x\in F_1(\hat{A}(S)$ and then rearranging the terms, all the coefficients of x^N vanish for $N>1$ and the coefficient of x is 1. Working modulo $F_n(\hat{A}(S))$ all but finitely many of the terms vanish and thus there is no issue about convergence of the rearrangement of the coefficients. Hence, for $x\in F_1(\hat{A}(S))$ we have $e(\ell(x))\equiv x\equiv \ell(e(x))$ modulo $F_n(\hat{A}(S))$ for all n. This means $e(\ell(x))=\ell(e(x))=x$ for all $x\in F_1(\hat{A}(S))$. In particular, in $\hat{A}(S)$ we have

$$\exp(\log(\exp(A)\exp(B)))=\exp(A)\exp(B).$$

Clearly

$$\exp(\log(\exp(0)\exp(A)))=\exp(\log(\exp(A)\exp(0)))=A,$$

and $\exp(A)\exp(-A)=1$ so that

$$\exp(\log(\exp(A)\exp(-A)))=\exp(\log(1))=1.$$

Lastly, we claim that letting $S=\{A,B,C\}$

$$\exp(A)(\exp(B)\exp(C))=(\exp(A)\exp(B))\exp(C)$$

in $\hat{A}(S)$. The terms from the left-hand side are of the form $\frac{(A_1^n B^{n_2})(C^{n_3})}{n_1!n_2!n_3!}$, whereas the terms from the right-hand side are $\frac{A_1^n(B^{n_2}C^{n_3})}{n_1!n_2!n_3!}$. Since $\hat{A}(S)$ is associative, these terms are equal.

Theorem 1.13. *(Hausdorff Series) The series $H(A,B) = \log(\exp(A)\exp(B))$ in $\hat{A}(S)$ lies in $\widehat{L(S)}$.*

Proof. We have the image $\hat{c}\colon \hat{A}(S) \to \prod_n B_n$ where $B_n = \sum_{i+j=n} A^i(S) \otimes A^j(S)$ given by $\hat{c}(x) = \sum_n c(x_n)$ where $c(x_n) \in B_n$. Because the only primitive elements in $A(S)$ for c are elements on $L(S)$, it follows that if $\hat{c}(x) = x \otimes 1 + 1 \otimes x$, then $x \in \widehat{Lie}(S)$, or equivalently $x_n \in L(S)$ for all $n \geq 1$. It remans to show that $\hat{c}(H(A,B)) = H(A,B) \otimes 1 + 1 \otimes H(A,B)$.

Since $A, B \in L(S)$, we have $c(A) = A \otimes 1 + 1 \otimes A$ and $c(B) = B \otimes 1 + 1 \otimes B$. If x is primitive and in the maximal ideal, then $\exp(x)$ is defined, is congruent to 1 modulo the maximal ideal and the multiplicative property of the power series $\exp(x)$ on commuting elements shows that if x is primitive and contained in the maximal ideal then

$$\begin{aligned}\hat{c}(\exp(x)) &= \exp(c(x)) = \exp(x \otimes 1 + 1 \otimes x) \\ &= (\exp(x \otimes 1))(\exp(1 \otimes x)) = \exp(x) \otimes \exp(x).\end{aligned}$$

Such elements are called *group-like*. It is also an easy exercise to show that if x, y in $\hat{A}(S)$ are group-like and congruent to 1 modulo the maximal ideal, then the same is true from xy, and if u is group-like and congruent to 1 modulo the maximal ideal, then $\log(u)$ is primitive and contained in the maximal ideal. Applying these elementary facts tells us that $H(A,B) = \log(\exp(A)\exp(B))$ is primitive in $\hat{A}(S)$ and thus $\log(\exp(A)\exp(B)) \in \widehat{L}(S)$. $\square$

This proves that, formally at least, the series $H(A,B) = \log(\exp(A)\exp(B))$ defines a group structure on $\hat{A}(S)$ with 0 as identity and -1 as inverse. One should think of this as a formal Lie group determined by the Lie algebra L.

1.5 Finite Dimensional Real Lie Algebras and Groups

To show that H leads to a local Lie group structure, we need to see that for any real Lie group $\mathfrak{g}$ the power series $H(A,B)$ has a positive radius of convergence and defines a real analytic function

$$U \times U \to \mathfrak{g}$$

defined in some neighborhood of the identity. If it does, then the formal computations of the local group structure are valid in the neighborhood of convergence meaning that H determines the multiplication of a local Lie group. Here is the statement of the convergence result.

Theorem 1.14. *Let L be a Lie algebra. Then there is a neighborhood $U \subset L$ of the identity invariant under $X \mapsto -X$ such that for all $A, B \in U$ the power series $H(A, B) = \log(\exp(A)\exp(B))$ converges absolutely to an element of L.*

Corollary 1.15. *Set $\Omega \subset U \times U$ equal to $H^{-1}(U)$, define $\theta(X) = -X$ and $m(A, B) = H(A, B) \in U$ for all $(A, B) \in \Omega$. Then $(U, 0, -1, \Omega, m)$ is a local Lie group,* the local Lie group of L with underlying open set U.

Proof. (of the corollary assuming the theorem) We have established the formal properties of the power series showing that it is associative, has 0 as a unit and $-X$ as the inverse for X. On any open set in $\mathfrak{g} \times \mathfrak{g}$ on which the power series converges, these results hold for the analytic function defined by the power series. The corollary now follows immediately. $\square$

The homework consists of a sequence of problems establishing this convergence result