Lie Groups: Fall, 2024
Lecture V: The Universal Enveloping Algebra, Free Lie Algebras, and the Baker-Campbell-Hausdorff Formula

September 30, 2024

1 Introduction

The Baker-Campbell-Hausdorff Theorem asserts the existence of a universal power series in two variables in a Lie algebra where the terms are sums iterated brackets of the two variables. For any finite dimensional real Lie algebra L, this series converges absolutely in some neighborhood of $(0,0)$ in $L \times L$. This series defines a local Lie group multiplication on some neighborhood of $0 \in L$. If $L = \mathfrak{g}$ for some Lie group G, the exponential mapping is an isomorphism between this local Lie group on a neighborhood of $0 \in \mathfrak{g}$ and a local Lie group determined by a neighborhood of $e \in G$.

In this lecture we prepare the way to prove this theorem by establishing it in a formal sense, working with power series completions. The homework carries out the appropriate estimates to show that for a finite dimensional real Lie algebra L, for $V \subset L$ a sufficiently small neighborhood of 0, the power series converge absolutely.

Let me begin the discussion by pointing out the issue. We write

$$\exp(a+b) = \sum_{k=0}^{\infty} \frac{(a+b)^k}{k!}.$$

If we are in an algebra where a and b do not commute, then $(a+b)^k$ is the sum of all 2^n strings of a's and b's of length n. The quadratic term is $a^2+ab+ba+b^2$ and the cubic term is $a^3+a^2b+aba+ba^2+ab^2+bab+ba^2+b^3$. It is only when a and b commute that we can rearrange these terms of degree

k to produce

$$\sum_{m=0}^{k} \binom{k}{m} a^m b^{k-m}.$$

In this case the factorials cancel nicely to yield $\exp(a) \cdot \exp(b) = \exp(a+b)$.

In particular, in a non-commutative Lie algebra and Lie group we cannot assert that $\exp(X)\exp(Y) = \exp(X+Y)$. Of course, one can keep track of the switches by replacing YX by $XY - [X,Y]$, so it is not unreasonable that the product can be written as a power series using iterated brackets of X, Y. Indeed, one for this. Working by induction, one can write write an explicit formula to any finite order. In fact there is a general formula which we establish in the problems. But there is a more conceptual way to approach the question using free Lie algebras and bi-algebra structure on their universal enveloping algebra. This lecture introduces all this structure.

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

For this section we fix a field K of characteristic 0. By an *associative algebra* we mean an associative K-algebra with unit. Maps of such are K-linear multiplicative maps sending the unit to the unit. A vector space will mean a K-vector space. A Lie algebra will mean a Lie algebra over K

The Poincaré-Birkhoff-Witt Theorem (PBW Theorem) says that every finite dimensional Lie algebra L is a sub Lie algebra of the Lie algebra given by an associative algebra with the bracket $[A,B] = AB - BA$. There is a universal such associative algebra which is called the *Universal Enveloping Algebra* of the Lie algebra.

2.1 The Construction

Definition 2.1. By a *linear representation* of a Lie algebra L on a vector space V we mean a Lie algebra map $\rho\colon L \to \mathrm{End}(V)$; that is to say a linear map that satisfies

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

N.B. If $K = \mathbb{R}$ (or $\mathbb{C}$) and $\rho\colon G \times V \to V$ is a linear representation of a real (or complex) 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. Consider the tensor algebra

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

(tensor product over K) 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 $x, y \in L$. By construction $U(L)$ is an associative algebra and there is a natural map $L \to U(L)$ which 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 $L \to \mathrm{End}(V)$ 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 $\hat{\rho}\colon T(L) \to A$ extending ρ. Since ρ is a map of Lie algebras, $\hat{\rho}$ sends every defining relation for $U(L)$ to zero in A. Hence, $\hat{\rho}$ a map of associative algebras $U(L) \to A$. This is the unique map of associative algebras $U(L) \to A$ extending ρ, showing that $U(L)$ has the universality property stated above.

Of course, we have not yet ruled out that for some non-zero Lie algebra L, we have $U(L) = 0$. The PBW Theorem does rule this out for it says that the natural map $L \to U(L)$ is injective. By the universal property of $L \to U(L)$, any linear representation of L on a vector space V extends to a map of algebras $U(L) \to \mathrm{End}(V)$. So, if the map $L \to U(L)$ has a non-zero kernel L_0, this would mean that every linear representation of L vanishes on L_0. One special case is that the adjoint representation $\mathrm{ad}_L\colon L \to \mathrm{End}(L)$ extends to an algebra map $U(L) \to \mathrm{End}(L)$.

Theorem 2.2. *(PBW) Let L be a finite dimensional Lie algebra. The tensor algebra on the vector space L has a natural increasing filtration on the tensor algebra $T(L)$ defined by $F_n(T(L)) = \sum_{k=0}^{n} \otimes^k L$. We define an increasing filtration of $U(L)$ by setting $F_n(U(L))$ equal to the image of $F_n(T(L))$ under the natural map. This is a multiplicative filtration in the sense that the multiplication induces a map $F_n(U(L)) \otimes F_m(U(L)) \mapsto F_{n+m}(U(L))$. Let*

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

be the associated graded algebra. The map $L \to F_1(U(L))$ induces with a linear map $\mu_L \colon L \to Gr_1^F(U(L))$. The map μ_L extends to an isomorphism of associative graded algebras from the polynomial algebra, $P(L)$, on L to $Gr_*^F(U(L))$. In particular, the natural map $\mu_L \colon L \to U(L)$ is an injection.

Corollary 2.3. *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.*

Remark 2.4. If L is an abelian Lie algebra (meaning the bracket is identically zero), then $U(L)$ is the quotient of $T(L)$ by the two-sided ideal generated by $xy - yx$. In this case $U(L)$ is naturally isomorphic to the polynomial algebra $P(L)$.

Proof. (of the theorem) We fix a K-basis $\{X_i\}_{i \in I}$ for L and choose a total ordering for the basis, or equivalently a total ordering on the index set I. The images of the X_i in the polynomial algebra $P(L)$ generated by L are a multiplicative basis. To avoid confusion we use the notation z_i for the element in $P(L)$ determined by X_i, so that $P(L)$ is the algebra of all polynomials on the $\{z_i\}_{i \in I}$.

For a finite sequence $J = \{j_1, \ldots, j_k\}$ of elements of I and $i \in I$, the notation $i \preceq J$ 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 denote by z_J the product $z_{j_1} \cdots z_{j_r}$ in $P(L)$. Of course, this element depends only on the set J not the ordering of its elements.

Our goal is to define an action of the Lie algebra L on the vector space $P(L)$.

$$\sigma \colon L \otimes P(L) \to P(L);$$

that is to say a map of Lie algebras from $L \to \mathrm{End}(P(L))$ (where $\mathrm{End}(P(L))$ refers to the endomorphism algebra of the vector space $P(L)$). We do this by induction on the degree p of the polynomial. Let $P^{\leq i}(L)$ denote the subspace of polynomials of degree $\leq i$. The inductive hypothesis for p is that we have a map $\sigma_p \colon L \otimes P^{\leq p}(L) \to P^{\leq (p+1)}(L)$ satisfying the following:

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

$B(p)$: For 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 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, $\sigma_p|_{P^{\leq(p-1)}(L)} = \sigma_{p-1}$.)

We construct the maps σ_p by induction on the $P^{\leq p}(L)$. For $p = 0$, it follows from Condition $A(0)$ that $\sigma_0(X_i)1 = z_i$. Extending by K-linearity defines σ_0 as required since $B(0)$ is equivalent to $A(0)$ and $C(0)$ is vacuous.

Now suppose that for some $p \geq 1$ we have defined $\sigma_{p-1} \colon L \otimes P^{\leq(p-1)}(L) \to P^{\leq p}(L)$ 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 $i \preceq J$, invoking Condition $A(p)$ we define $\sigma_p(X_i)z_J = z_i z_J$. Otherwise, re-odering J we have $J = (k, K)$ with $k \preceq K$ and $k < i$. By Condition $A(p-1)$ we have $z_J = \sigma_{p-1}(X_k)z_K$. We invoke Condition $C(p)$ to equate

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

We claim that the right-hand side of this expression is defined by induction and Condition $A(p)$. To see that, notice by Condition $B(p-1)$ we have $\sigma_{p-1}(X_i)z_K = z_i z_K + w$ for some $w \in P^{\leq(p-1)}(L)$ Thus, the term $\sigma_p(X_k)\sigma_{p-1}(X_i)z_K = \sigma_p(X_k)(z_i z_K) + \sigma_{p-1}(X_k)(w)$ is defined by invoking Condition $A(p)$ (for the first summand) and the inductive hypothesis (for the second summad). Of course, $\sigma_{p-1}([X_i, X_k])Z_K$ is defined by induction. By linearity this defines the unique extension of σ_{p-1} to σ_p to linear map $L \otimes P^{\leq p}(L) \to P^{\leq(p+1)}(L)$. We must show this extension satisfies $A(p)$, $B(p)$, and $C(p)$.

Clearly, by construction Conditions $A(p)$ and $B(p)$ hold for σ_p. It remains to show that $C(p)$ holds. Consider a multi-index K with $|K| = p-1$. By construction $C(p)$ holds for $\sigma_p(X_i)\sigma_{p-1}(X_j)z_K$ if $j \preceq K$ and $i \geq j$. By symmetry it holds if $i \preceq K$ and $j \geq i$. Thus, the remaining cases are where $K = (k, M)$ with $k \preceq M$ 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

$$\begin{aligned} X_i X_j z_K = X_i X_j X_k z_M &= X_i X_k X_j z_M + X_i[X_j, X_k]z_M \\ &= (X_k X_i X_j z_M + [X_i, X_k]X_j z_M) + X_i[X_j, X_k]z_M \\ &= (X_k X_j X_i z_M + X_k[X_i, X_j]z_M) + [X_i, X_k]X_j z_M \\ &\quad + X_i[X_j, X_k]z_M \end{aligned}$$

By the symmetric argument we have

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

The first equation minus the second one yields:

$$\begin{aligned} X_iX_jz_K - X_jX_iZ_K =& - X_k[X_i,X_j]z_M + 2X_k[X_i,X_j]z_M \\ &+ [[X_i,X_k],X_j]z_M + [X_i,[X_j,X_k]]z_M \\ =&(X_k[X_i,X_j] + [[X_i,X_k],X_j] + [X_i,[X_j,X_k]])z_M. \end{aligned} \tag{2.1}$$

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 2.1 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 $\sigma\colon U(L) \otimes P(L) \to P(L)$. By definition $\sigma(X_i)z_M = z_iz_M$ modulo $F_{|M|}P(L)$ and hence

$$\sigma(X_{i_1}\cdots X_{i_t})z_M = z_{i_1}\cdots z_{i_t}z_M \quad \text{modulo} \quad F_{|M|+t-1}U(L).$$

We define a vector space map $\varphi\colon U(L) \to P(L)$ be sending $a \in U(L)$ to $\varphi(a) = \sigma(a)\cdot 1$. Then $\varphi\colon U(L) \to P(L)$ is compatible with the gradings by degree. Of course, $P(L)$ is already a graded algebra and hence naturally isomorphic to its associated graded algebra. The associated graded $Gr^F\varphi$ induces a map of graded algebras $Gr^F\varphi\colon Gr^F_*U(L) \to P(L)$ sending the element $X_{i_1}\cdots X_{i_t}$ to the monomial $z_{i_1}\cdots z_{i_t}$. This shows that the map of graded algebras is surjective.

We claim that the map of graded algebras is also injective. Since every element of $F_nU(L)$ is represented by a sum of monomials of degrees $\leq n$ in the X_i. Monomials of degree less than n and any two monomials of degree n 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 given by weakly ordered sequences of the elements X_i. Since the weakly ordered sequences map via $Gr^F(\varphi)$ to a basis for the homogeneous polynomials of degree n, it follows that for each $n \geq 1$ these monomials of degree n are linearly independent and hence are a basis for $F_nU(L)/F_{n-1}U(L)$.

This shows that $Gr^F(\varphi)\colon Gr^F(U(L)) \to P(L)$ is an isomorphism of graded algebras. Hence, $\varphi\colon U(L) \to P(L)$ is a linear isomorphism of the filtered vector spaces that induces an isomorphism of the associated graded algebras.

Obviously, the composition of $Gr^F\varphi$ following the natural map $L \to F_1(U(L))$ is the identity from L to polynomials of degree 1 in L. Thus, the inverse of $Gr^F\varphi$ is an isomorphism of graded algebra $P(L) \to Gr^F_*(U(L))$ extending the given map $\mu\colon L \to Gr^F_1(U(L))$. □

Corollary 2.5. *Give $P(L)$ the increasing filtration associated to the grading by degrees. Let $\psi = \varphi^{-1}\colon P(L) \to U(L)$. Then ψ is a filtered, linear isomorphism whose associated graded*

$$Gr^F(\psi)\colon P(L) \to Gr^F(U(L))$$

is an isomorphism of graded algebras.

Remark 2.6. This result should not be surprising: We saw that if L is abelian the $U(L) \cong P(L)$ as graded algebras. Also, the identity $xy - yx = [x, y]$ is a correction of the commutative relation $xy - yx = 0$ by a term of lower order. Thus, it is not too surprising that these two relations induce isomorphic graded algebras. Still it is not completely formal since one makes essential use of the Jacobi identity in the proof.

2.2 The Bi-Algebra Structure on $U(L)$ and Its Primitive Elements

Now we introduce the structure that allows us to tell when an element in $U(L)$ lies in L.

Not only is $U(L)$ an associative algebra, it also has a natural co-multiplication. We define a map $c\colon L \to U(L) \otimes U(L)$ by setting $c(x) = x \otimes 1 + 1 \otimes x$.

Proposition 2.7. *c extends uniquely to an algebra map $c\colon U(L) \to U(L) \otimes U(L)$. The map c is an algebra homomorphism that is co-associative and co-commutative and has a co-unit. Thus, by definition $U(L)$ with this co-multiplication is a co-associative, co-commutative bi-algebra with a co-unit.*

Proof. Direct computation shows that $c(x)c(y) - c(y)c(x) = c([x, y])$ for all $x, y \in L$. Thus, $c\colon L \to U(L) \otimes U(L)$ is a map of Lie algebras from L to the $AB - BA$ Lie algebra structure on $U(L) \otimes U(L)$. Thus, c extends to an algebra map $c\colon U(L) \to U(L) \otimes U(L)$. 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 a *co-commutative* co-multiplication. 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 of L. Since L generates $U(L)$ as an algebra and c is an algebra map, it follows that this equation holds for all $u \in U(L)$, which is the definition of a *co-associative* co-multiplication. Finally, the *co-unit* of c is the map $U(L) \to K$ of unital algebras that sends $x \in L$ to zero for all $x \in L$. □

Definition 2.8. An element $x \in U(L)$ is *primitive* if $c(x) = x \otimes 1 + 1 \otimes x$

Lemma 2.9. *The primitive elements form a real vector subspace of $U(L)$ containing $L \subset U(L)$.*

Proof. Exercise. □

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. It is a homework problem to show the following:

Claim 2.10. *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.*

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

Proposition 2.11. *The primitive elements in $U(L)$ for the co-multiplication are exactly the elements in L.*

Proof. We define an increasing filtration $F_n[U(L) \otimes U(L)] = \sum_{i+j \leq n} F_i(U(L)) \otimes F_j(U(L))$. Then $c\colon U(L) \to U(L) \otimes U(L)$ preserves the filtration and hence induces a co-multiplication $c' = Gr^F(c)$ on $Gr^F_*(U(L))$, which is a homomorphism of algebras with every element in degree 1 being primitive. Thus, under the identification of $Gr^F_* U(L)$ 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 $a \in F_n(U(L))$ and has non-trivial projection to $F_n(U(L))/F_{n-1}(U(L))$. We shall show that $n = 1$. Let $\overline{a} \in F_n(U(L))/F_{n-1}(U(L))$ be the image of a. Since $\overline{a} \in Gr^F_*(U(L))$ is primitive, under the identification of $Gr^F_*(U(L))$ with $P(L)$, the element $\overline{a}$

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 $\lambda \in K$. 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 $a \in L$. □

3 Free (Non-Associative) Algebras and Free Lie Algebras

Definition 3.1. By an *algebra* (over K) we mean a K-vector space V with a multiplication, which is a K-linear map $\mu\colon V \otimes_K V \to V$. A map of algebras is a K-linear map preserving the multiplications.

Let S be a set. (We are primarily interested in the case when S has cardinality 2.) By induction on $i \geq 1$ 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 two types of elements: those of the form $(x_1 \cdot (x_2 \cdot x_3))$ and those of the form $((x_1 \cdot x_2) \cdot x_3)$ for $x_1, x_2, x_3 \in S$. S_4 has the following types of elements:

$$\begin{aligned} S_1 \times S_3 &: \quad (x_1, ((x_2, x_3), x_4)), (x_1, (x_2, (x_3, x_4))) \\ S_2 \times S_2 &: \quad ((x_1, x_2), (x_3, x_4)) \\ S_3 \times S_1 &: \quad (((x_1, x_2), x_3), x_4), ((x_1, (x_2, x_3)), x_4) \end{aligned}$$

We set $S_\infty = \coprod_{n=1}^{\infty} S_n$, and we denote by $F(S)$ to be the K-vector space generated by S_∞. The multiplication of $x \in S_i$ and $y \in S_j$ is the element $(x, y) \in S_i \times S_j \subset S_{i+j}$. We extend this multiplication on the basis elements S_∞ by bilinearity to a multiplication on $F(S)$. The freeness of $F(S)$ is captured in the following property.

Lemma 3.2. *Given an algebra A and a set function $\psi\colon S \to A$, There is a unique extension of the function of ψ to a map of algebras $\hat{\psi}\colon F(S) \to A$.*

Proof. Exercise. □

The grading on S_∞ induces a grading on $F(S)$ making $F(S)$ a graded algebra.

Definition 3.3. Let S be a set. The *free Lie algebra generated by* S, denoted $FL(S)$, is the quotient of $F(S)$ by the two-sided ideal generated by $Q(a.a) = a \cdot a$ and $J(a,b,c) = a \cdot (b \cdot c) + c \cdot (a \cdot b) + b \cdot (c \cdot a)$ for $a, b, c \in S$. The Lie bracket on $FL(S)$ is induced from the multiplication on $F(S)$. Since the relations are homogeneous with respect to the grading on $F(S)$, there is an induced grading on $FL(S)$ that makes it a graded Lie algebra.

We have imposed by fiat the skew symmetry and Jacobi identity on the multiplication in $FL(S)$ (and the multilinearity over K comes from the fact that $F(S)$ is an algebra over K). Thus, $FL(S)$ is indeed a Lie algebra over K. The proper way now to write an element of this quotient is to replace $(a \cdot b)$ by the Lie bracket $[a, b]$. For example, $(a \cdot ((b \cdot c) \cdot d))$ is written as $[a, [[b, c], d]]$. Then each each element of S_∞ is a legitimate expression in the Lie algebra generated by $FL(S)$. The fact that it is a free Lie algebra generated by S is the content of the next proposition.

Proposition 3.4. *Given a Lie algebra L and a set function $\varphi\colon S \to L$, there is a unique extension of φ to a homomorphism of Lie algebras $F(\varphi)\colon FL(S) \to L$.*

Proof. First use the universal property of $F(S)$ to define an algebra map $F(\varphi)\colon F(S) \to L$ 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 $F(\varphi)$ factors through the quotient $FL(S)$, and thus defines a Lie algebra homomorphism $L(\varphi)\colon FL(S) \to L$. Uniqueness of the is clear since S generates $F(S)$ and hence $FL(S)$. □

Demote by $T(S)$ the tensor algebra on the K vector space with basis S. As we have seen this is the free associative algebra (with unit) generated by S meaning that if A is any associative algebra with unit and $S \to A$ is a set function, then there is a unique unital algebra map $T(S) \to A$ extending the given map $S \to A$. It has a multiplicative grading by setting the elements of S to be homogeneous of degree 1. This is the usual grading on $T(S)$

Proposition 3.5. *Let S be a set.*

1. *The inclusion $S \to T(S)$ extends uniquely to a map of Lie algebras $\psi_S\colon FL(S) \to T(S)$, where the Lie algebra structure on $T(S)$ is the $XY - YX$ bracket $T(S)$. This map preserves the gradings.*

2. *The map ψ_S extends uniquely to a map of associative algebras*

$$\hat{\psi}_S\colon U(FL(S)) \to T(S).$$

This map also preserves the gradings

3. *The map $\hat{\psi}_S$ is an isomorphism of graded, associative algebras and identifies the universal enveloping algebra of $FL(S)$ with the tensor algebra $T(S)$.*

Proof. By the universal property of the free Lie algebra $FL(S)$, the inclusion of $S \to T(S)$ extends uniquely to a Lie algebra homomorphism $\psi_S \colon FL(S) \to T(S)$, when $T(S)$ is equipped with the $XY - YX$ bracket. By the universal property of $U(FL(S))$, this map extends uniquely to an algebra homomorphism $\hat{\psi}_S \colon U(FL(S)) \to T(S)$. In particular, $\hat{\psi}_S$ is the identity on the natural inclusions of S into $U(FL(S))$ and into $T(S)$.

On the other hand, the universal property of $T(S)$ implies that the inclusion $S \to U(FL(S))$ extends to an algebra homomorphism $\rho_S \colon T(S) \to U(FL(S))$. Both $\rho_S \circ \hat{\psi}_S$ and $\hat{\psi}_S \circ \rho_S$ are the identity on S and hence by the uniqueness part of the universal properties of $T(S)$ and $U(FL(S))$ both compositions are the identity. Thus, they are inverse isomorphisms and each preserves the gradings. □

4 Formal completions of $T(S)$ and $FL(S)$

The reason for introducing the completions of the tensor algebra $T(S)$ and the Lie algebra $FL(S)$ is so that our power series will have meaning, without having to worry about convergence issues. In this algebra context no convergence is possible without completing.

We give $U(FL(S)) \otimes U(FL(S))$ the grading defined by

$$(U(FL(S) \otimes U(FL(S)))^n = \oplus_{i+j=n} U^i(FL(S)) \otimes U^j(FL(S))$$

then the co-multiplication preserves the grading. Hence, image of this co-multiplication transported by the isomorphism $\hat{\psi} \colon U(FL(S)) \to T(S)$ defines a co-multiplication that preserves the grading on $T(S)$.

According to Proposition 2.11, the co-multiplication of $U(FL(S)$ has as primitives $FL(S) \subset U(FL(S))$. Taking the image of this co-multiplication in $T(S)$, we have:

Corollary 4.1. *There is a co-associative, co-commutative co-multiplication c' on $T(S)$ making it a bi-algebra with a co-unit. The primitive elements are exactly $FL(S) \subset T(S)$. The co-multiplication is compatible with the grading, so that an element $a_1 + \cdots + a_k \in T(S)$ with $a_i \in T^i(S)$ is contained in $FL(S)$ if and only if each $a_i \in FL(S)$.*

4.1 The Completions $\widehat{T}(S)$ and $\widehat{FL}(S)$

Define a decreasing filtration $\mathcal{F}^n(T(S)) = \oplus^{k\geq n} T^k(S)$. This is a decreasing filtration given by the powers of the ideal $F^1(T(S))$. We form the completion $\widehat{T}(S)$ of $T(S)$ with respect to the powers of this ideal. It is identified with $\widehat{T}(S) = \prod_{n=0}^{\infty} T^n(S)$ with the topology being the product topology of the discrete topologies on each factor. We let $\widehat{FL}(S)$ be the closure of $FL(S) \subset T(S)$ in $\widehat{T}(S)$. Then

$$\widehat{FL}(S) = \left\{ \prod_{n=0}^{\infty} FL^n(S) \right\}.$$

with the product topology.

Corollary 4.2. *Let $B^n(S) = \prod_{i+j=n} T^i(S) \otimes T^j(S)$ and set $\widehat{B}(S) = \prod_{n\geq 0} B^n(S)$ with the product topology. The co-multiplication in Corollary 4.1 induces a continous map $\hat{c}\colon \widehat{T}(S) \to \widehat{B}(S)$ sending $\sum_n a_n$ to $\sum_n c'(a_n)$. Let $\delta'(\sum_n a_n) = \sum_n a_n \otimes 1$ and $\delta''(\sum_n a_n) = \sum_n 1 \otimes a_n$. These are continuous maps of $\widehat{T}(S) \to \widehat{B}(S)$. Then an element $\sum_n a_n$ is primitive for $\hat{c}$, i.e.,*

$$\hat{c}(\sum_n a_n) = \delta'(\sum_n a_n) + \delta''(\sum_n a_n),$$

if and only if $a_n \in FL(S)$ for all $n \geq 1$; i.e., if and only if $\sum_n a_n \in \widehat{FL}(S)$.

Proof. All of this is immediate from the fact that the only primitive elements in $T^n(S)$ are the elements of $FL(S) \cap T^n(S)$. □

The filtration $\mathcal{F}$ on $T(S)$ induces a filtration on $\widehat{T}(S)$, which we also denote by $\mathcal{F}$. The subspace $\mathcal{F}^1(\widehat{T}(S))$ is the maximal ideal of $\widehat{T}(S)$. For $x \in \mathcal{F}^1(\widehat{T}(S))$, set $e(x) = \sum_{n\geq 1} x^n/n!$ and $\ell(x) = \sum_{k\geq 1} (-1)^k x^k/k$. These power series are well defined on the maximal ideal $\mathcal{F}^1(\widehat{T}(S))$, and take values in $\widehat{T}(S)$. The reason is that for $x \in \mathcal{F}^1(\widehat{T}(S))$, for each $n \geq 1$, all but finitely many of the terms in the series for e or ℓ vanish modulo $\mathcal{F}^n(\widehat{T}(S))$. Thus, the infinite sum represents a well-defined element of the inverse limit $\widehat{T}(S)$.

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 $\ell(e(t)) = t$. For each n, this leads to finite number algebraic equations for the coefficients terms of degree n of the composition. These manipulations are valid for composing the power series for e and ℓ in either order applied to $x \in \mathcal{F}^1(\widehat{T}(S)$. The

reason is that, since all homogeneous terms of $\ell((x)$ and $e(x)$ are rational coefficients times a power of x, they commute with each other and with x. Hence, the same manipulations can be carried out for $e(x)$ and $\ell(x)$. Thus, for any $x \in \mathcal{F}^1(\widehat{T}(S))$ we have $\ell(e(x)) = x$ and and $e(\ell(x)) = x$. Hence, for any $x \in \mathcal{F}^1(\widehat{T}(S))$ we have $\log(\exp(x)) = x$ and $\exp(\log(1+x)) = 1+x$.

4.2 Case $S = \{X, Y\}$ and the exponential and logarithm series

Now we specialize to the case when $S = \{X, Y\}$. Consider

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

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

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

Now consider

$$\log(\exp(X)\exp(Y)) = \sum_{m \geq 1} \frac{(-1)^{m-1}}{m} \left(\sum_{r,s \geq 0} \frac{X^r Y^s}{r!s!} \right)^m.$$

Working modulo $\mathcal{F}^n(\widehat{T}(S))$ all but finitely many of the terms vanish and thus there is no issue about convergence of the rearrangement of the coefficients modulo $\mathcal{F}^n(\widehat{T}(S))$ for each n. Applying the discussion above in this context we have

$$\exp(\log(\exp(X)\exp(Y))) = \exp(X)\exp(Y).$$

Clearly

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

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

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

Lastly, we claim that letting $S = \{X, Y, Z\}$

$$\exp(X)(\exp(Y)\exp(Z)) = (\exp(X)\exp(Y))\exp(Z)$$

in $\widehat{T}(S)$. The terms from the left-hand side are of the form $\frac{(X_1^n Y^{n_2})(Z^{n_3})}{n_1! n_2! n_3!}$, whereas the terms from the right-hand side are $\frac{X_1^n (Y^{n_2} Z^{n_3})}{n_1! n_2! n_3!}$. Since $\widehat{T}(S)$ is associative, these terms are equal.

4.3 The Hausdorff Series

Theorem 4.3. *(Hausdorff Series) The series* $H(X,Y) = \log(\exp(X)\exp(Y))$ *in* $\widehat{T}(S)$ *actually lies in* $\widehat{FL}(S)$.

Proof. We have the image $\hat{c}\colon \widehat{T}(S) \to \prod_n B_n$ where $B_n = \sum_{i+j=n} T^i(S) \otimes T^j(S)$ given by $\hat{c}(x) = \sum_n c(x_n)$ where $c(x_n) \in B_n$. Because the only primitive elements in $T(S)$ for c are elements on $FL(S)$, it follows that $\hat{c}(x) = x \otimes 1 + 1 \otimes x$ if and only if $x \in \widehat{FL}(S)$, or equivalently $x_n \in FL_n(S)$ for all $n \geq 1$. To prove the theorem we must show that $\hat{c}(H(X,Y)) = H(X,Y) \otimes 1 + 1 \otimes H(X,Y)$.

Notice that $\delta'(x)$ and $\delta''(x)$ commute and an element x is primitive if and only if $\hat{c}(x) = \delta'(x) + \delta''(x)$.

Since $\hat{c}$ is an algebra map, for any x in the maximal ideal, we have

$$\hat{c}(\exp(x)) = \exp(\hat{c}(x))$$

and for any y congruent to 1 modulo the maximal ideal we have

$$\hat{c}(\log(y)) = \log(\hat{c}(y)).$$

Since $X, Y \in FL(S)$, they are each primitive elements of $\hat{T}(S)$. Thus,

$$\hat{c}(\exp(X)) = \exp(X \otimes 1 + 1 \otimes X).$$

Since $\exp(X \otimes 1)$ and $\exp(1 \otimes X)$ commute we have

$$\begin{aligned} \hat{c}(\exp(X)) &= \exp(X \otimes 1)\exp(1 \otimes X) \\ &= (\exp(X) \otimes 1)(1 \otimes \exp(X)) \\ &= \delta'(\exp(X))\delta''(\exp(X)). \end{aligned}$$

Elements $x \in \widehat{T}(S)$ congruent to 1 modulo the maximal ideal and satisfying $\hat{c}(x) = \delta'(x) \otimes \delta''(x)$ are called *group-like*. It is an easy exercise to show that if x, y in $\widehat{T}(S)$ are group-like, then so is xy. We have just seen that the exponential map sends primitive elements in $\mathcal{F}^1(\widehat{T}(S))$ to group-like elements.congruent to 1 modulo $\mathcal{F}^1(\widehat{T}(S))$.

Claim 4.4. log *sends group-like elements congruent to* 1 *modulo* $\mathcal{F}^1(\widehat{T}(S))$ *to primitive elements in* $\mathcal{F}^1(\widehat{T}(S))$ *and* exp *and* log *are inverse isomorphisms between these primitive elements and group-like elements.*

Proof. Suppose that x is a group-like element congruent to 1 modulo $\mathcal{F}^1(\widehat{T}(S))$. Then

$$\hat{c}(\log(x)) = \log(\hat{c}(x)) = \log(\delta'(x)(\delta''(x)).$$

Since $\delta'(x)$ and $\delta''(x)$ commute we have

$$\begin{aligned}\log(\delta'(x)(\delta''(x)) &= \log((\delta'(x)) + \log(\delta''(x)) \\ &= \log((x \otimes 1) + \log(1 \otimes x)) \\ &= \log(x) \otimes 1 + 1 \otimes \log(x),\end{aligned}$$

so that $\log(\delta'(x)(\delta''(x))$ is a primitive element of $\mathcal{F}^1(\widehat{T}(S))$.

We already know that log and exp are inverses of each other. □

We now see that X, Y are primitive and hence $\exp(X)$ and $\exp(Y)$ are group-like. This means that $\exp(X) \cdot \exp(Y)$ is group-like. Consequently, $H(X,Y) = \log(\exp(X) \cdot \exp(Y))$ is primitive. By Corollary 4.2 it follows that $H(X,Y) \in \widehat{FL}(S) \subset \hat{T}(S)$. □

The direct computations above translate to $H(0,X) = H(X,0) = X$, proving that 0 is the identity for the multiplication defined by H. Also, $H(X,-X) = 0$, which means that $-X$ is the inverse of X.

Claim 4.5. *$S = \{X, Y, Z\}$ we have $H(X, H(Y,Z)) = H((X,Y),Z)$.*

Proof. We have the following expressions:

$$H(X, H(Y,Z)) = \sum_{r,s,t} \frac{X^r}{r!}\Big(\frac{Y^s Z^t}{s!t!}\Big)$$

$$H(H(X,Y),Z) = \sum_{r,s,t}\Big(\frac{X^r Y^s}{r!s!}\Big)\frac{Z^t}{t!}.$$

Since $\widehat{T(S)}$ is an associative algebra, these expressions are equal. □

This proves the formal version of associative law for H.

5 Appendix: Algebraic Fondations

5.1 I-adic Topology and Completions

Suppose that we have an algebra A and an ideal $I \subset A$. We define the I-adic topology on A by taking as the open neighborhoods 0 as I^n (and then the

open neighborhoods of $a \in A$ as $a + I^n$). We can then form a complete algebra with respect to this topology by setting $\hat{A} = \lim_n A/I^n$ with the natural homomorphisms $\pi_{n,k} \colon A/I^n \to A/I^k$ for $k < n$. By the definition of (projective) limits there are homomorphisms $\hat{\pi}_n \colon \hat{A} \to A/I^n$ compatible with the $\pi_{n,k}$ in the sense that $\pi_{n,k} \circ \hat{\pi}_n = \hat{\pi}_k$. Furthermore, for any algebra B a system of compatible homomorphisms $B \to A/I^n$ is equivalent to a homomorphism $B \to \hat{A}$ The ideal I generates an ideal $\hat{I}$ of $\hat{A}$, namely all elements in the kernel of $\hat{\pi}_1 \colon \hat{A} \to A/I$.

The algebra $\hat{A}$ is *complete* with respect to the $\hat{I}$-adic topology in the sense that any sequence of elements x_n that is eventually constant modulo $\hat{I}^r$ for each $r > 0$ converges to a point of $\hat{A}$ in the $\hat{I}$-adic topology. If $\cap_{n>0} I^n = 0$, the the natural homomorphism $A \to \hat{A}$ is an injection with dense image in the I-adic topology. In this case $\hat{A}$ is the *completion* of A with respect to the I-adic topology.

Example (i). For a prime p, the p-adic integers $\hat{\mathbb{Z}}_p$ is the completion of $\mathbb{Z}$ with respect to the p-adic topology where the ideal is the prime ideal (p) generated by p. An element of this ring can be written uniquely as a power series

$$\sum_{n \geq 0} a_n p^n,$$

where the integers a_n range from 0 to $p - 1$. More generally, any series $\sum_{n \geq 0} a_n p^n$ for arbitrary integers a_n determines an element of $\hat{\mathbb{Z}}_p$.

Example (ii). Let $K[x]$ be a polynomial ring over a field and let $I = (x)$ be the ideal generated by x. Then the I-adic completion of $K[x]$ is the ring of formal power series in one variable $K[[x]]$.

Example (iii). Let $K[x_1, \ldots, x_n]$ be the polynomial ring in n variables over a field and let $I = (0)$ be the ideal of all polynomials vanishing at 0. This is the ideal generated by $(x_1, \ldots, x_n)$. The I-adic completion of $K[x_1, \ldots, x_n]$ is the formal power series ring on n variables $K[[x_1, \ldots, x_n]]$. This is thought of as the function field of a formal neighborhood of 0 in affine n-space.

Example (iv). Let $A = \oplus_{n=0}^{\infty} A^n$ be a graded ring or algebra, meaning that the multiplication is homogeneous with respect to the grading in the sense that $m \colon A^k \otimes A^\ell \to A^{k+\ell}$. Then let $I = \oplus_{n \geq 1} A^n$ be the ideal of element of positive degree. Clearly, I^k is the ideal of elements of degree at least k and $\cap_{k=1}^{\infty} I^k = 0$. The completion $\hat{A}$ with respect to the I-adic topology is $\prod_{n=0}^{\infty} A^n$ with the obvious multiplication. Elements in this ring are formal sums $\sum_{n=0}^{\infty} a_n$, with a_n of degree n and the multiplication is the natural one on these infinite series.

Example (v). There is a generalization of Example (iv). Instead of a

graded ring or algebra we consider a ring or algebra A with an increasing filtration

$$F_0(A) \subset F_1(A) \subset \cdots \subset F_n(A \subset \cdots$$

that is required to be multiplicative in the sense that $m\colon F_k(A) \otimes F_\ell(A) \to F_{k+\ell}(A)$. Then we can form the associated graded algebra

$$Gr^F(A) = \oplus_{n=0}^{\infty} F_n(A)/F_{n-1}(A)$$

with the induced graded multiplication. We can then form the completion of this graded ring as in Example (iv).

5.2 Bi-Algebras

Recall that if A and B are associatve unital algebras then so is $A \otimes B$. Its unit is the tensor product of the units of A and B and the multiplication is given by $(a \otimes b) \cdot (c \otimes d) = ac \otimes bd$.

Definition 5.1. Let A be an associative algebra over a field K with multiplication m with unit 1. A *bi-algebra structure* on A is in addition a *comultiplication*

$$c\colon A \to A \otimes A$$

that (i) is a (unital) algebra homomorphism and (ii) has a co-unit, which is a K-linear map $\epsilon A \to K$ satisfying

$$K \otimes A \xrightarrow[\mathrm{Id}_A \otimes \epsilon]{} A \otimes A \xrightarrow[m]{} A$$

is the natural identification of $K \otimes A \to A$, and analogously for $m \circ (\epsilon \otimes \mathrm{Id}_A)$.

Equivalently, we can suppose that $c\colon A \to A \otimes A$ is a co-algebra with co-unit and $m\colon A \otimes A \to A$ is an associative algebra with unit and a homomorphism of co-unital co-algebras.

We say that the bi-algebra is *co-commutative* if $T \circ c = c$ where $T\colon A \otimes A \to A \otimes A$ is the interchange of factors. We say that the bi-algebra is *co-associative* if

$$(1 \otimes c) \circ c = (c \otimes 1) \circ c.$$

Definition 5.2. In a co-algebra of a bi-algebra an element x is *primitive* if $c(x) = x \otimes 1 + 1 \otimes x$. In a co-algebra or bi-algebra an element is *group-like* if

$$c(x) = x \otimes x.$$

Example (i). Let $P(V)$ be the polynomial algebra on a finite dimensional vector space over a field of characteristic 0. The usual multiplication of polynomials makes this an associative algebra. Define $c\colon V \to V \otimes V$ by $c(v) = v \otimes 1 + 1 \otimes v$. Since V generates $P(V)$ as an algebra there is at most one algebra map $c_0\colon P(V) \to P(V) \otimes P(V)$ extending c. Since $P(V)$ is the free commutative and associative algebra generated by V and since $P(V) \otimes P(V)$ is also a commutative, associative algebra, there is a unique extension of c to a co-algebra map

$$c_0\colon P(V) \to P(V) \otimes P(V).$$

Its value on an n^{th} power is given by

$$c_0(v^n) = \sum_{k=0}^{n} \binom{n}{k} v^k \otimes v^{n-k}.$$

This is the *standard* co-multiplication on $P(V)$. It is an easy exercise to show it makes a co-commutative, co-associative bi-algebra.

It is an exercise to show (since K has characteristic 0) that every homogeneous polynomial of degree n is a sum of n^{th} powers.

Corollary 5.3. *For V a finite dimensional vector space over a field of characteristic the only primitive elements in $P(V)$ the usual co-multiplication are the homogeneous polynomials of degree* 1.

Proof. Since c_0 is homogeneous with respect to degree, if x is a primitive element then each of its homogeneous terms is. Thus, it suffices to assume that x is homogeneous, say of degree n. Polynomials of degree 0 are elements of K and since $c_0(1) = 1 \otimes 1$, there are no primitive elements of degree 0. Suppose that x is non-zero and homogeneous of degree $n \geq 1$. Then $x = \sum_i \lambda_i v_i^n$ and hence the terms of degree $(1, n-1)$ in $c(x)$ are $\sum_i n\lambda_i v_i \otimes v_i^{n-1}$ and the product of this term in $P(V)$ is nx. Thus, this term is non-zero. Hence, x is not primitive if $n \neq 1$. Lastly, if x is homogeneous of degree 1 then $x \in V$ and $c_0(x) = x \otimes 1 + 1 \otimes x$, so that x is primitive $\square$