Math 128

Problem set #4

Feb. 26, 2004, due March 4

The purpose of this problem set is to develop some facts about bialgebras, expecially Hopf algebras (see the definition below). We will go into more detail than we had time for in class. I will also give a more combinatorial description of the universal enveloping algebra of the free Lie algebra and this will lead to a beautiful generalization of the Baker-Campbell-Hausdorff theorem due to Chen in terms of "iterated integrals".

There is a lot of reading and 7 actual problems found on pages 5,6,12, and 22.

Contents

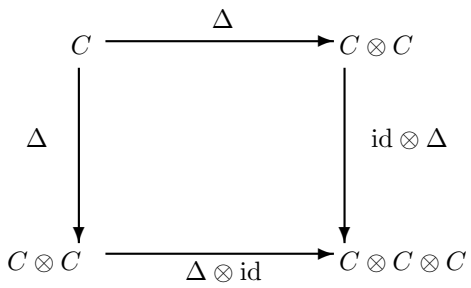


Figure 1:

1 Coalgebras.

We recall the definition: A vector space, C, with a linear map

$$\Delta : C \to C \otimes C$$

is called a **coalgebra** and the map Δ is called **comultiplication** because the dual diagram

$$m : A \otimes A \to A$$

is usually called a multiplication. The coalgebra (or the comultiplication) is called **coassociative** if

$$(\Delta \otimes id) \circ \Delta = (id \otimes \Delta) \circ \Delta \tag{1}$$

as maps from C to $C \otimes C \otimes C$. In other words the diagram in Figure 1. commutes.

1.1 Units and counits.

It is usual to assume that an associative algebra has a unit. Let $\mathbf{1}$ denote the unit element of the algebra A, so that

$$\mathbf{1}a = a\mathbf{1} = a$$

for all $a \in A$. Define the map

$$u : k \to A, \quad u(r) = r\mathbf{1}, \ r \in k.$$

Then the preceding equation can be written as

$$m(u(r) \otimes a) = m(a \otimes u(r)) = ra \tag{2}$$

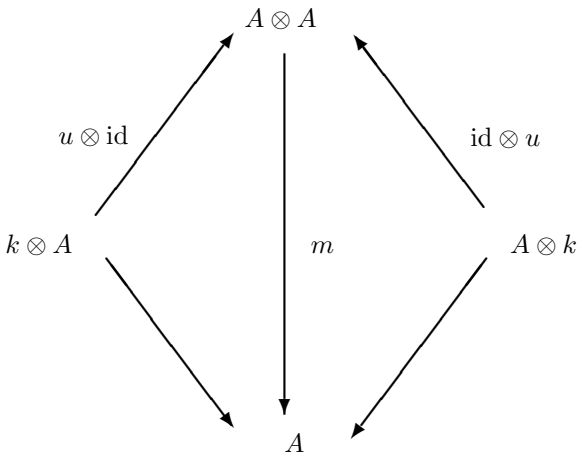


Figure 2:

for any $a \in A$.

Conversely, suppose that we start with a map $u : k \to A$ which satisfies (2). Then the element $\mathbf{1} = u(1)$ is a unit element of A, where 1 is the unit in k.

Thus we may *define* a unit in an algebra to be a map $u : k \to A$ such that (2) holds for any $r \in k$.

In terms of diagrams, we can express (2) as saying that the diagram in Figure 2 commutes: The two lower diagonal arrows are the natural isomorphisms of $k \otimes A$ and $A \otimes k$ with A, true for any vector space.

Dually, a *counit* for a coalgebra, C, over k, is defined to be a linear map

$$\epsilon : C \to k$$

satisfying

$$(\epsilon \otimes id) \circ \Delta = (id \otimes \epsilon) \circ \Delta = id \tag{3}$$

under the identification of $k \otimes C$ and $C \otimes k$ with C which holds for any vector space..

Let us write the comultiplication, Δ, in the "Sweedler notation":

$$\Delta(c) = \sum c_{(1)} \otimes c_{(2)} = \sum c_{i1} \otimes c_{i2}$$

where the middle expression is shorthand for the rightmost expression, see [5] . Then (3) becomes

$$\sum \epsilon(c_{(1)})c_{(2)} = \sum c_{(1)}\epsilon(c_{(2)}) = c. \tag{4}$$

In terms of diagrams, equation (3) says that the diagram in Figure 3 commutes.

The duality between the definitions of unit and counit become clear in terms of the diagrams: The diagram in Figure 3 is obtained from the diagram in Figure 2 by reversing all the arrows and by replacing u by ϵ and m by Δ.

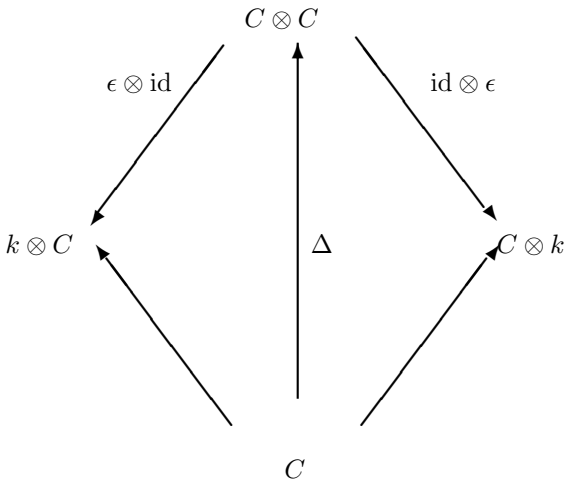


Figure 3:

Here are two important and instructive examples of coalgebras:

1) If X is any set, we will let $\mathcal{F}(X)$ denote the space of functions on this set with values in the ground field. If Y is another set, we can consider the tensor product $\mathcal{F}(X) \otimes \mathcal{F}(Y)$. A typical element of this space is an expression of the form $\sum_i f_i \otimes g_i$ the f_i are functions of $x \in X$ and the g_i are functions of $y \in Y$. This expression can be thought of as a function of "two variables"

$$\left(\sum_i f_i \otimes g_i \right)(x, y) = \sum_i f_i(x) g_i(y).$$

In this way we have defined a linear map

$$\mathcal{F}(X) \otimes \mathcal{F}(Y) \to \mathcal{F}(X \times Y).$$

It is easy to check that this map is injective. If X is finite, then $\mathcal{F}(X)$ is a finite dimensional vector space space whose dimension is equal to the number of elements in X. So if X and Y are finite sets, the above map is an isomorphism (by injectivity and dimension count).

This implies that if X, Y and Z are finite sets, and if we have a map

$$\phi : X \times Y \to Z,$$

then the "pull-back" map

$$\phi^* : \mathcal{F}(Z) \to \mathcal{F}(X \times Y), \quad h \mapsto \phi^* h, \ \phi^*(h)(x, y) := h(\phi(x, y))$$

gives a map

$$\mathcal{F}(Z) \to \mathcal{F}(X) \otimes \mathcal{F}(Y).$$

Now let $X = Y = Z = G$ be a finite group, and take ϕ to be the multiplication map $(x, y) \mapsto xy$.

1. Show that this makes $\mathcal{F}(G)$ into an associative coalgebra with counit. What is the counit?

If C is a finite dimensional coalgebra then its dual space C^* is an algebra.

2. Describe the algebra dual to the algebra $\mathcal{F}(G)$ of problem **1**. In more detail: For each $a \in G$, let δ_a denote the linear function on $\mathcal{F}(G)$ which assigns to every function f the value of f at a. So

$$\delta_a(f) := f(a).$$

The various δ_a as a ranges over G are linearly independent, and since there are $\#(G)$ of them and $\#(G)$ is the dimension of $\mathcal{F}(G)$ we see that they form a basis of $\mathcal{F}(G)^*$. The multiplication on $\mathcal{F}(G)^*$ is denoted by $\star$ and is called **convolution**. Show that

$$\delta_a \star \delta_b = \delta_{ab}.$$

Also express

$$\left(\sum_{a\in G} \phi(a)\delta_a\right) \star \left(\sum_{b\in G} \psi(b)\delta_b\right)$$

in the form

$$\sum c \in G \eta(c)\delta_c.$$

In other words, express η in terms of ϕ and ψ.

We can turn this around: If we start with an algebra A so we have a map $m : A \otimes A \to A$, we can let $C = A^*$ be the dual space of A and consider the transpose $M^* : A^* \to (A \otimes A)^*$. We always have an injection $A^* \otimes A^* \to (A \otimes A)^*$, and if A is finite dimensional this is an isomorphism (by dimension count). So our second class of examples is:

2) Start with a finite dimensional algebra A so we have the multiplication map $m : A \otimes A \to A$. Then $C = A^*$ with $\Delta := m^* : C \to C \otimes C$ is a co-algebra. If A is associative with unit then C is coassociative with co-unit.

For example, we may take $A = \mathrm{Mat}(n)$, the algebra of all $n \times n$ matrices. So C is the space of all linear functions on the space of $n \times n$ matrices. As a basis for C we may take the linear functions T_{ij} which assign to a matrix T its ijth entry. Explicitly, in terms of the basis $\{T_{mn}\}$ of C we have

$$\Delta T_{mn} = \Sigma T_{mj} \otimes T_{jn}. \quad \text{(matrix comultiplication)} \tag{5}$$

For example, in the two by two case where the matrix entries are a, b, c, d with $T_{11} = a, T_{12} = b$ etc., the comultiplication is

$$\begin{aligned} \Delta a &= a \otimes a + b \otimes c \\ \Delta b &= a \otimes b + b \otimes d \\ \Delta c &= c \otimes a + d \otimes c \\ \Delta d &= c \otimes b + d \otimes d. \end{aligned}$$

So we have a four dimensional coalgebra.

3. What is the co-unit for the coalgebra $C = \mathrm{Mat}(n)^*$? Verify that what you found is indeed the co-unit.

2 Representations, coalgebras, and addition laws for special functions.

Let us go back to the four dimensional coalgebra $\mathrm{Mat}(2)^*$ whose comulitplication table we listed just before the last problem:

$$\begin{aligned} \Delta a &= a \otimes a + b \otimes c \\ \Delta b &= a \otimes b + b \otimes d \\ \Delta c &= c \otimes a + d \otimes c \\ \Delta d &= c \otimes b + d \otimes d. \end{aligned}$$

Suppose we set $a = d = \cos$ and $c = \sin, b = -\sin$. We get

$$\Delta \cos = \cos \otimes \cos - \sin \otimes \sin, \qquad \Delta \sin = \sin \otimes \cos + \cos \otimes \sin .$$

These become the "addition laws" for the trigonometric functions:

$$\cos(x+y) = \cos x \cos y - \sin x \sin y, \qquad \sin(x+y) = \sin x \cos y + \cos x \sin y$$

if we think of x and y as "placeholders" with x indicating placement before the tensor product sign and y placement after the tensor product sign.

The general formulation is as follows: Recall that a finite dimensional representation of a group G is a map

$$\rho : G \to \mathrm{Mat}(n)$$

satisfying

$$\rho(ab) = \rho(a)\rho(b), \qquad \rho(e) = I$$

where e is the identity element of G and I is the identity matrix. The map ρ induces a "pull-back" map

$$\rho^* : \mathrm{Mat}(n)^* \to \mathcal{F}(G),$$

if L is a linear function on $\mathrm{Mat}(n)$ then the definition is

$$(\rho^*)L(a) := L(\rho(a)).$$

Let $\mathcal{F}_\rho(G)$ denote the subspace $\rho^*(\mathrm{Mat}(n)^*) \subset \mathcal{F}(G)$. Notice that $\mathcal{F}_\rho(G)$ is a finite dimensional subspace of $\mathcal{F}(G)$ and the fact that ρ is a representation translates into the assertion that the map

$$\Delta_G : \mathcal{F}(G) \to \mathcal{F}(G \times G), \quad (\Delta_G f)(a,b) = f(ab)$$

when restricted to $\mathcal{F}_\rho(G)$ maps $\mathcal{F}_\rho(G)$ to the image of $\mathcal{F}_\rho(G) \otimes \mathcal{F}_\rho(G)$ in $\mathcal{F}(G \times G)$. Indeed, if $L \in \mathrm{Mat}(n)^*$ and

$$\Delta L = \sum L_{(1)} \otimes L_{(2)}$$

in Sweedler notation, then

$$\Delta_G(\rho^* L)(a,b) = (\rho^* L)(ab) = L(\rho(a)\rho(b)) = \Delta L(\rho(a) \otimes \rho(b))$$

$$= \sum L_{(1)} \otimes L_{(2)}(\rho(a)) \otimes \rho(b)) = \sum L_{(1)}(\rho(a)) L_{((2)}\rho(b) = \sum (\rho^*(L_{(1)})(a)(\rho^* L_{(2)})(b)).$$

So even if the group G is not finite, we get a finite dimensional coalgebra from every finite dimensional representation of G.

Thus the map

$$x \mapsto \begin{pmatrix} \cos x & -\sin x \\ \sin x & \cos x \end{pmatrix}$$

gives a two dimensional real representation of the group $G = \mathbb{R}/2\pi\mathbb{Z}$ with $\mathcal{F}_\rho(G)$ the two dimensional space of functions spanned by cos and sin. The addition laws for these trigonometric functions is then just the comultiplication on the two dimensional coalgebra $\mathcal{F}_\rho(G)$.

I stress all of this because "addition laws" were discovered in the eighteenth and nineteenth centuries for various important functions long before the concept of a group or a representation or even of a matrix were formulated. Here are some examples. I will leave the group theoretical interpretation to a different course.

Weierstrass elliptic function.

$$\wp(x+y) = \frac{1}{4}\left\{\frac{\wp'(x) - \wp'(y)}{\wp(x) - \wp(y)}\right\}^2 - \wp(x) - \wp(y).$$

Jacobi elliptic functions. A typical addition formula (in Glaisher notation for the twelve basic functions) is:

$$cs(u+v) = \frac{cs(u)ns(v)ds(v) - cs(v)ns(u)ds(u)}{ns^2(v) - ns^2(u)}.$$

Bessel functions.

$$J_n(x+y) = \sum J_m(x) J_{n-m}(y).$$

Legendre polynomials. (Calc. Integ.(1815) pp.262-269.) This is difficult to understand in its original formulation. But here is a translation into semi-modern language: Express a rotation R in terms of its Euler angles

$$R = R(\phi, \theta, \psi).$$

Express the Euler angles of the product of two rotations in terms of their Euler angles,

$$\begin{aligned} R &= R_1 R_2 \\ R(\phi,\theta,\psi) &= R(\phi_1,\theta_1,\psi_1)R(\phi_2,\theta_2,\psi_2). \end{aligned}$$

The expression is messy:

$$\cos\theta = \cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2\cos(\phi_2+\psi_1) \qquad \textbf{(i)}$$

$$\begin{aligned} e^{i\phi} &= (e^{i\phi_1}/\sin\theta)[\sin\theta_1\cos\theta_2 \\ &\quad + \cos\theta_1\sin\theta_2\cos(\phi_2+\psi_1) + i\sin\theta_2\sin(\phi_2+\psi_1)] \end{aligned}$$

and an even more complicated expression for ψ. For each integer u there is a unique irreducible representation of dimension $2u+1$ of $SO(3)$. These correspond to the "integer spin" representations of $sl(2)$ that we studied in class. In terms of a suitable orthonormal basis, we will have a matrix expression

$$T^u_{mn}(R) = T^u_{mn}(R(\phi,\theta,\psi)).$$

This matrix entry is of the form

$$T^u_{mn}(R(\phi,\theta,\psi)) = c(u,m,n)e^{i(m\phi+n\psi)}\, P_u^{-n,m}(\cos\theta) \qquad \textbf{(ii)}$$

where c is a constant involving powers of i and quotients of factorials involving u, m and n. The fact that T^u is a representation implies that

$$T^u_{mn}(R_1R_2) = \sum T^u_{mj}(R_1)T^u_{jn}(R_2). \qquad \textbf{(iii)}$$

If we substitute (i) and (ii) into (iii) we get Legendre's addition formula.

3 The opposite algebra or coalgebra.

Given any algebra, we can define the "opposite" algebra, A^{op} where $A^{\mathrm{op}} = A$ as a vector space but we simply multiply the elements in reverse order: if $m(a\otimes b) = ab$ then $m^{\mathrm{op}}(a\otimes b) = ba$. Or, more abstractly,

$$m^{\mathrm{op}} = m\circ s$$

where $s : A\otimes A \to A\otimes A$ is the switching operator

$$s(a\otimes b) = b\otimes a. \qquad (6)$$

Clearly

$$m^{\mathrm{op}}(m^{\mathrm{op}}(a,b),c) = c(ba)$$

while

$$m^{\mathrm{op}}(a,m^{\mathrm{op}}(b,c)) = (cb)a.$$

So if A is associative, then so is $A^{\rm op}$. Dualizing, if we are given a coalgebra C, we can define the "opposite" coalgebra $C^{\rm op}$ where $C = C^{op}$ as a vector space, but where

$$\Delta^{op} = s \circ \Delta.$$

We can now *conclude* that if C is coassociative, then so is $C^{\rm op}$. For example, we can consider the opposite coalgebra of the algebra of matrix multiplication where

$$\Delta T_{mn} = \Sigma T_{in} \otimes T_{mi} \quad \text{(opmatrix comultiplication)}. \tag{7}$$

4 Bialgebras.

Suppose we consider the space, $\mathcal{F}(G)$, of all functions on a finite group G. It is a finite dimensional vector space with two operations; ordinary multiplication of functions which exists on the space of all functions on any set, and a comultiplication which exists because $\mathcal{F}(G)$ is a finite dimensional representation space of G (the "regular" representation). These two operations are consistent in the sense of the following definition which we gave in class:

Let B be a vector space which is both an algebra, with

$$\text{multiplication } m : B \otimes B \to B \text{ and unit } u : k \to B$$

and is also a coalgebra with

$$\text{comultiplication } \Delta : B \to B \otimes B \text{ and counit } \epsilon : B \to k.$$

We can give $B \otimes B$ the structure of an algebra where

$$(a \otimes b) \circ (c \otimes d) = ac \otimes bd.$$

More diagrammatically, we define

$$M : (B \otimes B) \otimes (B \otimes B) \to B \otimes B$$

as the composition

$$(B \otimes B) \otimes (B \otimes B) \stackrel{I \otimes s \otimes I}{\longrightarrow} (B \otimes B) \otimes (B \otimes B) \stackrel{m \otimes m}{\longrightarrow} B \otimes B$$

where s is the twist map

$$s(b \otimes c) = c \otimes b.$$

If $\mathbf{1} = u(1)$ is the unit element of B, then $\mathbf{1} \otimes \mathbf{1}$ is the unit of $B \otimes B$. In other words the unit map,

$$u_{B \otimes B} : k \to B \otimes B$$

is given by

$$u_{B \otimes B} = u_B \otimes u_B.$$

We can now require that the maps $\Delta : B \to B \otimes B$ and $\epsilon : B \to k$ be algebra maps. So we want

$$\Delta(xy) = \Delta(x)\Delta(y), \quad x, y \in B$$

and

$$\epsilon(xy) = \epsilon(x)\epsilon(y), \ x, y \in k.$$

These two equations say that Δ and ϵ carry multiplication into multiplication. The condition on the units is that

$$\Delta \circ u = (u \otimes u) \circ \delta$$

where $\delta : \ k \to k \otimes k$ is the standard identification and that

$$\epsilon \circ u = id$$

as maps from k to k. In terms of Sweedler notation with

$$\Delta(x) = \sum x_{(1)} \otimes x_{(2)}$$

we can write the first of these conditions as

$$\Delta(xy) = \sum x_{(1)}y_{(1)} \otimes x_{(2)}y_{(2)}. \tag{8}$$

Diagrammatically, the condition that Δ be an algebra map is that the diagrams (**1**) and (**2**) in Figure 4 commute, while the condition that ϵ be an algebra map is that diagrams (**3**) and (**4**) commute (unlabeled maps are the natural identifications):

We can read the commutativity of the diagrams in Figure 4 another way. We can make $B \otimes B$ into a coalgebra by defining the maps

$$B \otimes B \to B \otimes B \otimes B \otimes B$$

and

$$B \otimes B \to k$$

as the composite maps

$$B \otimes B \xrightarrow{\Delta \otimes \Delta} B \otimes B \otimes B \otimes B \xrightarrow{s_{23}} B \otimes B \otimes B \otimes B$$

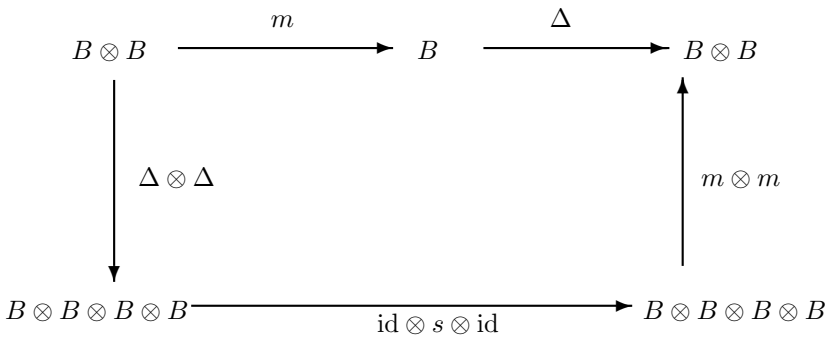
and

$$B \otimes B \xrightarrow{\epsilon \otimes \epsilon} k \otimes k \xrightarrow{\sim} k \ .$$

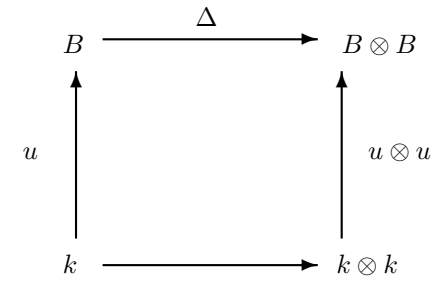
In the first of these formulas the symbol s_{23} means that we apply s to the second and third tensor positions, that is (in our example)

$$s_{23} = id \otimes s \otimes id.$$

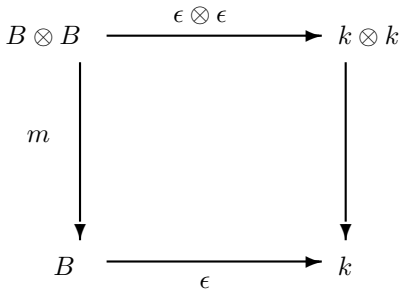
Then the commutativity of diagrams (**1**) and (**3**) in Figure 4 says that m is a coalgebra morphism while the commutativity of (**2**) and (**4**) says that u is a coalgebra morphism.



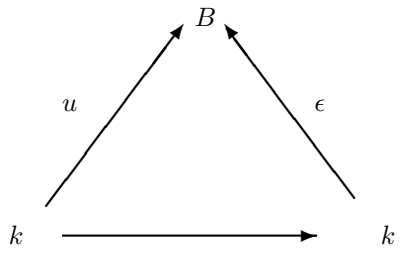
(**1**)



(**2**)



(**3**)



(**4**)

Figure 4:

Thus we make the following definition: a *bialgebra* is a vector space B which is both an algebra and a coalgebra and such that all four diagrams above commute. This is equivalent to requiring that (Δ, ϵ) be morphisms of the algebra structure or that (m, u) be morphisms of the coalgebra structure. More generally, we may want to replace the ground field, k, by a commutative ring, K and allow B to be a module over K. We will still use the word "bialgebra" in this more general situation. The example that we started with, where $B = \mathcal{F}(G)$ is the space of functions on a finite group, clearly satisfies the axioms, and so is a bialgebra. We will make use of this example to illustrate various constructs.

5 Antipodes and Hopf algebras.

Let us consider our model example, the bialgebra $B = \mathcal{F}(G)$ of all functions on a finite group, G. For any $f \in \mathcal{F}(G)$, let us define Tf by

$$(Tf)(a) = f(a^{-1}), \quad a \in G.$$

Then under the identification of $B \otimes B$ with $\mathcal{F}(G \times G)$ the comultiplication is given as

$$(\Delta f)(a, b) = f(ab).$$

Thus the element $h = (id \otimes T) \circ \Delta f$ is the function given by

$$h(a, b) = f(ab^{-1}).$$

The multiplication map $m : \mathcal{F}(G \times G) \to \mathcal{F}(G)$ is given by

$$m(u)(a) = u(a, a), \quad u \in \mathcal{F}(G \times G).$$

4. Verify this.

So

$$(mh)(a) = f(e)$$

if h is as above. In other words, the composite

$$m \circ (id \otimes T) \circ \Delta(f)$$

is just the constant function, $f(e) = \epsilon(f)$. But for any $r \in K$, the constant function $v \equiv r$ is the element $u(r) \in \mathcal{F}(G)$. This leads to the following general definition of *antipode* for a bialgebra, B over a ring, K: It is a map $T : B \to B$ such that the composite map

$$B \stackrel{\Delta}{\to} B \otimes B \stackrel{id \otimes T}{\longrightarrow} B \otimes B \stackrel{m}{\to} B$$

equals the composite map

$$B \stackrel{\epsilon}{\to} K \stackrel{u}{\to} B$$

where u is the unit and ϵ the counit.

The abstract idea behind this definition is as follows. Let $C,\ \Delta : C \to C \otimes C$ be a coalgebra and $A,\ m : A \otimes A \to A$ be an algebra, both over the same ring, K. Then $\mathrm{Hom}_K(C, A)$ has the structure of an algebra where the product, $f \star g$ of $f, g \in \mathrm{Hom}_k(C, A)$ is defined to be the composite map

$$C \stackrel{\Delta}{\to} C \otimes C \stackrel{f \otimes g}{\longrightarrow} A \otimes A \stackrel{m}{\to} A.$$

If the comultiplication on C and the multiplication on A are associative, then this multiplication is associative. The element $u \circ \epsilon : C \to A$ (where $\epsilon : C \to K$ is the comultiplication of C and $u : K \to A$ is the unit of A) is a unit for this multiplication. This because a bialgebra axiom says that $(\epsilon \otimes id) \circ \Delta$ is the natural isomorphism of C with $K \otimes C$ and an algebra axiom says that $m \circ (u \otimes id)$ is the natural isomorphism of $K \otimes A$ with A.

The definition given above for the antipode says that T is the inverse in the $\star$ multiplication of the identity map, $id : B \to B$.

A *Hopf algebra* is defined as a bialgebra which is associative and coassociative and which has an antipode.

6 The universal enveloping algebra as a Hopf algebra.

Let G be a Lie group with Lie algebra, $\mathcal{G}$. As we are no longer dealing with finite groups, the space $\mathcal{F}(G)$ will be infinite dimensional, and we no longer will have the identification of $\mathcal{F}(G \times G)$ with $\mathcal{F}(G) \otimes \mathcal{F}(G)$ and so $\mathcal{F}(G)$ will not be a coalgebra when G is not finite. Instead, we have the identification

$$\mathcal{F}(G \times G) \sim \mathcal{F}(G) \hat{\otimes} \mathcal{F}(G)$$

where $\hat{\otimes}$ denotes a completed tensor product: finite sums must be replaced by infinite series in the definition of Δ. This is already apparent in several of the examples the addition laws cited above. We could, of course, modify the definition of coalgebra to take this into account, and study topological coalgebras with completed tensor products replacing tensor products in the definition of comultiplication. But in the case of a connected Lie group we have an alternative: The "dual object" to the space of analytic functions on G is the universal enveloping algebra, $U(\mathcal{G})$, thought of as the space of distributions on G supported at the identity element, e. We can also think of $U(\mathcal{G})$ as the left invariant differential operators on G, and this gives the multiplication on $U(\mathcal{G})$. The comultiplication is dual to restriction to the diagonal on $\mathcal{F}(G \times G)$. For elements $X \in \mathcal{G}$ this translates into

$$\Delta(X) = X \otimes 1 + 1 \otimes X, \quad X \in \mathcal{G} \tag{9}$$

and extends by multiplication to $U(\mathcal{G})$. So the comultiplication is cocommutative, while the multiplication is non trivial and determined by the Lie algebra structure of $\mathcal{G}$. This is, of course, dual to the situation in $\mathcal{F}(G)$ where the

multiplication is commutative but the comultiplication carries the information relating to the group structure.

For a general Lie algebra, $\mathcal{G}$ over an arbitrary field,k, we gave in class an abstract definition of $U(\mathcal{G})$ as an associative algebra. It is determined as the solution to a "universal mapping problem": For any linear map ρ of $\mathcal{G}$ into an associative algebra A satisfying

$$\rho(X)\rho(Y) - \rho(Y)\rho(X) = \rho([X,Y]) \tag{10}$$

there is a unique algebra homomorphism, also denoted by ρ, of $U(\mathcal{G})$ into A. This determines $U(\mathcal{G})$ uniquely up to isomorphism, and $U(\mathcal{G})$ can be constructed as a quotient of the tensor algebra $T(\mathcal{G})$. If k is of characteristic zero, as we shall assume from now on, there is a natural embedding of $\mathcal{G}$ into $U(\mathcal{G})$, so we may identify $\mathcal{G}$ as a subspace of $U(\mathcal{G})$, and $U(\mathcal{G})$ is generated by $\mathcal{G}$ as an algebra. The map given by (9) clearly satisfies (10) and hence extends to an algebra homomorphism

$$\Delta : U(\mathcal{G}) \to U(\mathcal{G}) \otimes U(\mathcal{G}).$$

The map $\epsilon : \mathcal{G} \to k$ given by $\epsilon(X) = 0 \quad \forall X$ clearly satisfies (10) and hence extends to an algebra homomorphism, also denoted by ϵ. If follows immediately from (9) that

$$(\epsilon \otimes id)(\Delta a) = (id \otimes \epsilon)(\Delta a) = a$$

for $a = X \in \mathcal{G}$ and hence by multiplication for all $a \in U(\mathcal{G})$. So $U(\mathcal{G})$ is an associative, coassociative and cocommutative bialgebra.

The map

$$T : X \mapsto -X$$

satisfies (10) as a map from $\mathcal{G}$ to the opposite algebra, $U(\mathcal{G})^{op}$. Thus T extends as an algebra antihomomorphism

$$T : U(\mathcal{G}) \to U(\mathcal{G}), \quad T(ab) = T(b)T(a)$$

so

$$T(X_1 \cdots X_n) = (-1)^n X_n X_{n-1} \cdots X_1. \tag{11}$$

Clearly

$$[m \circ (T \otimes id) \circ \Delta](a) = [m \circ (id \otimes T) \circ \Delta](a) = 0$$

for $a = X \in \mathcal{G}$ and hence by multiplication for all products of X's. Hence T is an antipode for the bialgebra $U(\mathcal{G})$. We will refer to the Hopf algebra $(U(\mathcal{G}), m, \Delta, T)$ simply by $U(\mathcal{G})$. Clearly every element of $\mathcal{G}$ is primitive, and we proved in class that the converse is true

$$\mathrm{Prim}(U(\mathcal{G})) = \mathcal{G}. \tag{12}$$

7 Primitive elements and grouplike elements.

Recall that the primitive elements of an associative and coassociative bialgebra, B, are those elements which satisfy

$$\Delta(X) = X \otimes 1 + 1 \otimes X$$

and that the set, $\mathrm{Prim}(B)$, of primitive elements is a Lie algebra, i.e. is closed under commutator bracket. An element $b \in B$ is called *grouplike* if

$$\Delta(b) = b \otimes b. \tag{13}$$

Clearly the set of grouplike elements is closed under multiplication. Also,

$$[(\epsilon \otimes id) \circ \Delta](b) = \epsilon(b) \otimes b$$

if b is grouplike, but

$$(\epsilon \otimes id) \circ \Delta$$

is just the standard identification, $b \mapsto 1 \otimes b$ of B with $k \otimes B$ by the definition of the co-unit, see Figure 3. Thus

$$\epsilon(b) = 1 \tag{14}$$

if b is grouplike. If B is a Hopf algebra with antipode T then

$$\begin{aligned} T(b)b &= m(T \otimes id)\Delta(b) \\ &= \epsilon(b) \\ &= 1 \end{aligned}$$

if b is grouplike. So in a Hopf algebra the grouplike elements form a group with

$$b^{-1} = T(b).$$

We will denote the set of grouplike elements of B by $\mathrm{Group}(B)$. We would like to be able to say that

$$\exp : \mathrm{Prim}(B) \to \mathrm{Group}(B)$$

where the exponential map is defined by its usual series:

$$\exp(X) = 1 + X + \frac{1}{2!}X^2 + \frac{1}{3!}X^3 + \cdots .$$

The problem, of course, is that the series need not converge. For example, if $B = U(\mathcal{G})$ is the universal enveloping algebra of a Lie algebra, $\mathcal{G}$, such an infinite sum makes no sense. But suppose we were in a situation where series such as the exponential series converge. Then we could argue as follows: For any polynomial with coefficients in k, the consistency of multiplication and comultiplication implies that $\Delta(P(X)) = P(\Delta(X))$ for any $X \in B$. Assuming

that we can also apply this to the convergent power series for the exponential, we have, for X primitive,

$$\begin{aligned}\Delta(\exp(X)) &= \exp(\Delta(X)) \\ &= \exp(X\otimes 1 + 1\otimes X) \\ &= \exp(X\otimes 1)\cdot\exp(1\otimes X) \\ &= (\exp X\otimes 1)\cdot(1\otimes\exp X) \\ &= \exp X\otimes\exp X\end{aligned}$$

so

$$\exp X \in \mathrm{Grou}(B).$$

In passing from the second line to the third we have use the fact that $X\otimes 1$ commutes with $1\otimes X$ and then that the exponential of the sum of two commuting elements is the product of their exponentials.

A similar type argument shows that the logarithm of a group like element is primitive, where $\log(1+x)$ is defined by its power series. One uses the fact that the logarithm of the product of two commuting elements is the sum of their logarithms.

In the next section I will describe a situation where there is no trouble with convergence. It gives another way of looking at what we did in class.

8 The shuffle algebra.

In this section and the next, we want to describe some ideas of Chen that appear in an important series of articles [3] and the references listed there. The results of this section are based on an important paper of Ree,[4].

Let $\mathbf{I}$ be a collection of symbols. In what follows I, J etc. will denote finite ordered sequences of elements of $\mathbf{I}$. When we write $I\subset J$ we mean that I appears as an ordered subset of J . If $\#J$ denotes the number of elements of J and $1\le k\le \#J$ then J_k denotes the sequence of length k consisting of the first k elements of J and J^k denotes the sequence of length $\#J-k$ consisting of the last $\#J-k$ elements of J . Thus if I consists of the positive integers then

$$(4,3,2,2,5,6)_4 = (4,3,2,2) \quad\text{and}\quad (4,3,2,2,5,6)^4 = (5,6).$$

If J and K are sequences we let (J,K) denote the sequence of length $\#J+\#K$ whose first $\#J$ elements constitute J and whose last $\#K$ elements consist of K . We consider the k module $C = C(\mathbf{I})$ generated by 1 and elements Y_I where I ranges over all finite ordered sequences of elements of $\mathbf{I}$. We define a comultiplication $\Delta : C\to C\otimes C$ by $\Delta 1 = 1\otimes 1$ and

$$\Delta Y_I = \sum_{(J,K)=I} Y_J\otimes Y_K = \sum_{k=0}^{\#I} Y_{I_k}\otimes Y^{I^k}. \tag{15}$$

Here $Y_\emptyset$ corresponding to subscript 0 or superscript $\#I$ is taken as 1. Thus

$$\Delta Y_{(1,2,3)} = Y_{(1,2,3)} \otimes 1 + Y_{(1,2)} \otimes Y_{(3)} + Y_{(1)} \otimes Y_{(2,3)} + 1 \otimes Y_{(1,2,3)}$$

for example. Notice that Δ is coassociative as

$$(\Delta \otimes id) \circ \Delta(Y_I) = (id \otimes \Delta) \circ \Delta(Y_I) = \sum_{(J,K,L)=I} Y_J \otimes Y_K \otimes Y_L.$$

The counit $\epsilon : C \to k$ is defined by $\epsilon(1) = 1$ and $\epsilon(Y_I) = 0, I \neq \emptyset$. It clearly satisfies the axioms for a counit, so C is an associative coalgebra.

8.1 The dual space to C is the algebra of formal power series in non-commuting variables.

Before defining the multiplication and the antipode on C and showing that it is a Hopf algebra, we study some immediate consequences of the coalgebra structure. Let F denote the dual space to C. So F consists of all *infinite* series of the form

$$\sum a_I X^I$$

where the X^I are dual to the Y_I . Then F inherits a multiplication given by

$$\langle X^J \cdot X^K, Y_I \rangle = \langle X^J \otimes X^K, \Delta Y_I \rangle.$$

This last expression is 0 unless $(J, K) = I$ and is 1 if $(J, K) = I$ as follows from the definition of the comultiplication on C. Thus

$$X^J \cdot X^K = X^{(J,K)}.$$

So if $J = (p, q, r, ..., s)$ then $X^J = X^p \cdot X^q \cdot X^r \cdots X^s$ where we write X^p instead of $X^{(p)}$ etc.. In other words, F is the ring of noncommutative formal power series in the variables indexed by I.

Let $A = A(\mathbf{I})$ be the ring of non-commutative polynomials in the variable $X^p,\ p \in \mathbf{I}$, so that F is the completion of A in theformal power series topology. We can also think of A as the tensor algebra $A = T(V)$ where V is the vector space with basis $\{X^p\}$. It follows from the defining properties of the universal enveloping algebra, as we showed in class, that $A = U(\mathcal{G})$ where $\mathcal{G}$ if the free Lie algebra generated by the X^p. Hence A is a Hopf algebra. Since F is the completion of A it inherits a Hopf algebra structure. In particular, if we let $\mathcal{L}$ denote the completion of $\mathcal{G}$ in F, then the Lie algebra $\mathcal{L} = \mathrm{Prim}(F)$. In F there is no problem about convergence of power series such as $\exp a$ or $\ln(1 + a)$ so long as $a \in m$, the maximal ideal of F. In particular the argument we gave at the beginning of this section is legitimate and we have the result that we proved in class:

Proposition 1 *The map* $\exp : Prim(F) \to Group(F)$ *is a bijection.*

The comultiplication on F induces a multiplication on C by the formula

$$\langle X^K, \mu(Y_I \otimes Y_J)\rangle = \langle \Delta X^K, Y_I \otimes Y_J\rangle. \tag{16}$$

To see what this is more explicitly, recall that the comultiplication on F is generated multiplicatively by

$$\Delta X^p = X^p \otimes 1 + 1 \otimes X^p.$$

For example, if $K = (i, j, k)$ then

$$\begin{aligned}
\Delta(X^K) &= \Delta X^i \Delta X^j \Delta X^k \\
&= (X^i \otimes 1 + 1 \otimes X^i)(X^j \otimes 1 + 1 \otimes X^j)(X^k \otimes 1 + 1 \otimes X^k) \\
&= X^i X^j X^k \otimes 1 + X^i X^j \otimes X^k + X^i X^k \otimes X^j + X^j X^k \otimes X^i \\
&\quad X^i \otimes X^j X^k + X^j \otimes X^i X^k + X^k \otimes X^i X^j + 1 \otimes X^i X^j X^k.
\end{aligned}$$

If we look at

$$I = (p), \quad J = (q, r)$$

then we get a contribution of 1 to (16) only if K is such that p is one of the three letters i, j or k and the remaining two are q and r in the correct order. Otherwise the contribution will be zero. In other words, we get the non-zero contribution if and only if K is a riffle shuffle of (I, J). (A riffle shuffle with a cut at k is a permutation $\sigma \in S_n$ such that $\sigma(x) < \sigma(y)$ if $x, y \leq k$ or $k < x < y \leq n$.) So we define

$$\mu : C \otimes C \to C$$

by

$$\mu(Y_I \otimes Y_J) = \sum_{\sigma(I,J)=K} Y_K \tag{17}$$

where the sum is over all riffle shuffles. From general principles we know that Δ and μ are compatible since they are the duals of the compatible multiplication and comultiplication on F.

It is interesting to study the composition $\mu \circ \Delta : C \to C$. If we think of Y_I as a stack of cards labeled by the elements of I, then Δ can be thought of as the direct sum of all possible cuts of the deck. Applying μ then adds up all possible shuffles. Thus the symmetric element

$$\sum_{\sigma \in S_{\#I}} Y_{\sigma I}$$

is clearly an eigenvector of $\mu \circ \Delta$, since every term in the expansion of $\mu \circ \Delta$ into a sum of riffle shuffles leaves this completely symmetric element invariant. The eigenvalue will be the number of summands and this equals $2^{\#I}$ since the position, k, of the cut (in the expression for $\Delta = \Delta_C$) ranges over $0 \leq k \leq I$ and there are

$$\left(\begin{array}{c} \#I \\ k \end{array} \right)$$

choices of riffle shuffle with cut at k.

In fact, the operator $\mu \circ \Delta$ is completely reducible and all its eigenvalues of are powers of two. To see this, observe that the transpose of $\mu \circ \Delta$ is

$$m \circ \Delta, \quad \Delta = \Delta_F : F \to F \otimes F, \quad m : F \otimes F \to F,$$

where, with apologies for the notation, $\Delta = \mu^*$ is the comultiplication in F and $m = \Delta_C^*$ is the multiplication. The assertion about the complete reducibility of $\mu \circ \Delta_C$ is an immediate consequence of the following proposition, valid for any Lie algebra.

Proposition 2 *Let $\mathcal{G}$ be a Lie algebra over a field of characteristic zero. Let*

$$W_n : S^n(\mathcal{G}) \to U(\mathcal{G})$$

denote the completely symmetric (Weyl) embedding, so that we have the vector space identification

$$U(\mathcal{G}) = \bigoplus W_n(S^n(\mathcal{G})). \tag{18}$$

Then (18) is a decomposition of $U(\mathcal{G})$ into eigenspaces of $m \circ \Delta$ with eigenvalues 2^n on $W_n(S^n(\mathcal{G}))$.

Indeed, in terms of a Poincaré Birkhoff Witt basis, $\{X^K\}$, the space $W_n(S^n(\mathcal{G}))$ is spanned by

$$\sum_{\sigma \in S_n} X^{\sigma K}, x, \quad \#K = n.$$

But

$$(m \circ \Delta) X^K = \sum_{\tau \text{ a shuffle}} X^{\tau K}.$$

Applying the corresponding formula to each of the $X^{\sigma K}$ and using the fact that there are 2^n shuffles of a deck of size n proves the proposition. We will need this proposition in Chapter 11.

8.2 Repeated shuffling converges to randomness.

As an aside, we observe that if the elements of I are distinct, then on the space spanned by the

$$\{Y_{\sigma I}\}, \quad \sigma \in S_{\#I}$$

the totally symmetric element is the unique eigenvector with eigenvalue $2^{\#I}$, the space being spanned by eigenvectors, with remaining eigenvalues being $2^r, 1 \leq r < \#I$. Indeed, the totally symmetric element, $\sum Y_{\sigma I}$ is characterized (up to scalar multiple) as being orthogonal to $W_k(S^k(\mathcal{L})), k \neq n$. To illustrate, consider the case where $I = (1, 2, 3)$. In F, we need only take into account

expressions which involve sums of products of X^1, X^2 and X^3 in various orders. This six dimensional space is spanned by

$$\begin{aligned} [X_1,[X_2,X_3]],[[X_1,X_2].X_3] &\in W_1(S^1(\mathcal{L})) \\ X^1[X^2,X^3]+[X^2,X^3]X^1, & \\ X^2[X^3,X^1]+[X^3,X^1]X^2, & \\ X^3[X^1,X^2]+[X^1,X^2]X^3 &\in W_2(S^2(\mathcal{L})) \\ \sum_{\sigma\in S_3} X^{\sigma I} &\in W_3(S^3(\mathcal{L})). \end{aligned}$$

Clearly $\sum Y_{\sigma I}$ is orthogonal to all these elements except the last.

The operator

$$\frac{1}{2^n}\mu\circ\Delta, \quad n=\#I$$

acts on the space spanned by the Y_σ as an element of the group algebra of S_n with positive coefficients supported on the shuffles, and whose coefficients add up to one. In other words, it is a probability measure concentrated on the shuffles. As an operator, its maximal eigenvalue = 1, and all other eigenvalues are positive and strictly less than 1. The assertion above about the uniqueness of the maximal eigenvector implies that repeated shuffling, i.e. applying powers of the operator, converges to projection onto one dimensional space of totally symmetric elements. Expressed in terms of probability measures on S_n, it says that repeated convolution converges to the uniform measure, i.e. to complete randomness. For a full discussion of more realistic shuffling paradigms see [2] and the references given there.

8.3 The antipode of the shuffle algebra.

The antipode, S, on C, is defined by $S1=1$ and

$$SY_I=(-1)^{\#I}Y_{I_{\text{rev}}}$$

where I_{rev} is I with the order of its elements reversed. So

$$SY_{(1,2,3,4)}=Y_{(4,3,2,1)},$$

for example. This is the dual of the antipode, T, on F given by

$$T(X^I)=(-1)^{\#I}X^{I_{\text{rev}}}.$$

the is the standard antipode for the universal enveloping algebra, generated from $TX^p=-X^p$. Since $S=T^*$ it follows that S is an antipode for C. The fact that S is an antipode, i.e. that

$$[\mu\circ(S\otimes id)\circ\Delta]Y_I=0, \quad I\neq\emptyset.$$

contains nontrivial combinatorial information. Here is a sample of how a direct computation of the cancelation involved in the above equation goes, where we have dropped the commas and parentheses on the subscripts to save space:

$$\Delta Y_{1234} = Y_{1234} \otimes 1 + Y_{123} \otimes Y_4 + Y_{12} \otimes Y_{34} + Y_1 \otimes Y_{234} + 1 \otimes Y_{1234}$$

so

$$(S \otimes id) \circ \Delta Y_{1234} = Y_{4321} \otimes 1 - Y_{321} \otimes Y_4 + Y_{21} \otimes Y_{34} - Y_1 \otimes Y_{234} + 1 \otimes Y_{1234}.$$

Hence $[\mu \circ (S \otimes id) \circ \Delta](Y_{1234})$ is given by

$$\begin{array}{llllll} Y_{4321} & + & & & & \\ -Y_{4321} & -Y_{3421} & -Y_{3241} & -Y_{3214} & & \\ +Y_{2134} & +Y_{3421} & +Y_{3241} & +Y_{3214}+ & Y_{2341}+ & Y_{2314} \\ -Y_{2134} & -Y_{1234} & & & -Y_{2341} & -Y_{2314} \\ & & +Y_{1234} & & & \\ & & = 0. & & & \end{array}$$

When $\mathbf{I}$ consists of a single element, i.e. when $\#\mathbf{I} = 1$, then the Hopf algebra C is called the algebra of *divided powers*.

9 Chen's iterated integrals.

Chen, in his series of remarkable papers, associates to piecewise differentiable paths, α in $\mathbf{R}^n$ an element $\theta(\alpha) \in \text{Group}F$ where $\mathbf{I} = \{1, 2, \ldots, n\}$

This assignment carries composition of paths into multiplication in the group, Group(F). Here is how it works: Let $f_i, i \in \mathbf{I}$ be a collection of (real or complex) piecewise continuous functions defined on some interval containing the points a, b, s etc. Define the iterated integral

$$\int_a^b J(f)$$

inductively as follows: If $\#J = 1$ and J consists of the single index, j, then

$$\int_a^b J(f)dt = \int_a^b f_j(t)dt,$$

the usual integral. If the iterated integral has been defined for all sequences of length n, and $\#J = n + 1$ with $J = (K, j)$ and $\#K = n$ (so that the iterated integral for K is defined) then define

$$\int_a^b J(f)dt = \int_a^b \int_a^\tau K(f)dt f_j(\tau)d\tau. \tag{19}$$

We then have Chen's lemma which asserts that

$$\int_a^s I(f)dt = \sum_{(J,K)=I} \int_a^b J(f)dt \int_b^s K(f)dt \tag{20}$$

it being understood that, in the sum on the right, the integrals corresponding to empty K or J are just 1 . So if $\#I = 1$ this equation says that the ordinary integral from a to b plus the ordinary integral from b to s equals that ordinary integral from a to s. So (20) is clearly true for $\#I = 1$.

5. Prove Chen's lemma. [Hint: Use induction and differentiation.]

In particular, consider a piecewise differentiable path, α, in $\mathbb{R}^n$, and let

$$f_i(t) = x_i'(t)$$

where the x_i are the coordinates. Then iterated integration assigns a number, call it $Y_I(\alpha)$, to each piecewise differentiable path. If α and β are two paths with the end point of α coinciding with the initial point of β, let $\gamma = \alpha\beta$ denote the combined path. Let $\theta(\alpha)$ be the linear function on C which assigns the value $Y_I(\alpha)$ to each Y_I . Thus

$$\theta(\alpha) \in F$$

and Chen's lemma, equation (20), implies that

$$\theta(\alpha\beta) = \theta(\alpha)\theta(\beta). \tag{21}$$

where the multiplication on the right is in F.

6. Let $(a_1, \ldots, a_n)$ be a point of $\mathbb{R}^n$ and let $\alpha : [0, 1] \to \mathbb{R}^n$ be the curve

$$\alpha(t) = (ta_1, \ldots, ta_n) \qquad 0 \leq t \leq 1$$

so that α is the straight line segment joining the origin to $(a_1, \ldots, a_n)$. Show that

$$\theta(\alpha) = \exp(a_1X_1 + \cdots a_nX_n).$$

From this and Proposition 1 it follows that $\theta(\alpha) \in \text{Group}(F)$ for this particular curve. Since θ is defined in terms of the derivatives of the path, it is invariant under translation.

7. Conclude Chen's theorem which says that $\theta(\alpha) \in$ Group F for any piecewise differentiable path and hence θ defines a homomorphism from the space of piecewise differentiable paths under composition to into the group like elements of F.

This is Chen's generalization of the Campbell-Baker-Hausdorff theorem.

References

[1] E.Abe, *Hopf algebras*, Cambridge University Press, Cambridge-New York, 1980. [MR83a:16010]

[2] D.Bayer and P.Diaconis, *Trailing the dovetail shuffle to its lair*, Ann. Appl. Probab. **2** (1992), 294–313. [MR93d:60014]

[3] K.T.Chen, *Pullback path fibrations, homotopies and iterated integrals*, Bull. Inst. Math. Acad. Sinica **8** (1980), no. 2-3, part 1, 263–275. [MR82e:58004]

[4] R.Ree, *Lie elements and an algebra associated with shuffles*, Ann. Math. **68** (1958), 210–220. [MR20#6447]

[5] M.E.Sweedler, *Hopf algebras*, Benjamin, New York, 1969. [MR40#5705]