Hopf Algebras and Combinatorics

Arjit Seth and M. Vinay

Manipal Institute of Technology, Manipal University

Context

These are notes created based on the lectures by Prof. Federico Ardila at San Francisco State University. Hopf algebras seem to be a difficult topic to introduce as a simple definition. They apparently have applications in quantum field theory, pertinent to Feynman diagrams.

Contents

1 Introduction

A Hopf algebra is a complicated mathematical structure with a definition involving lots of properties and scary commutative diagrams. It fundamentally invokes the tensor product, and it is difficult to construct as a complete definition in one sitting, so we must start with smaller definitions and build up the structure progressively.

Definition 1 (Hopf algebra)**.** A *Hopf algebra H* is a $\mathbb{K}$-vector space with five operations:

$$
\begin{aligned}
\text{Multiplication:} \quad & m\colon H \otimes H \to H \\
\text{Unit:} \quad & u\colon \mathbb{K} \to H \\
\text{Comultiplication:} \quad & \Delta\colon H \to H \otimes H \\
\text{Counit:} \quad & \epsilon\colon H \to \mathbb{K} \\
\text{Antipode:} \quad & S\colon H \to H
\end{aligned}
$$

These definitions invoke lots of commutative diagrams that are difficult to TeX, so they'll be shown later once a better understanding is developed. Now we must preliminarily define the tensor product.

Definition 2 (Tensor product)**.** Let $\mathbb{K}$ be a field, and let V and W be vector spaces over $\mathbb{K}$. The tensor product $V \otimes W$ is a vector space over $\mathbb{K}$ generated by vectors $v \otimes w$, $v \in V$, $w \in W$, and satisfying the following properties:

Distributivity over addition: $\quad (v + v') \otimes (w + w') = v \otimes w + v \otimes w' + v' \otimes w + v' \otimes w'$

Scalar multiplication independent of two arguments: $\quad \lambda(v \otimes w) = \lambda v \otimes w = v \otimes \lambda w$

The combination of these is a property called *bilinearity*. Note that this vector space is much larger than the product space $V \times W$:

$$\dim U \times V = \dim U + \dim V$$
$$\dim U \otimes V = \dim U \dim V$$

This is intuitively obvious because the tensor product defines an *actual* product between elements of V and W, instead of a restricted component structure induced by a Cartesian product, which constrains manipulations to V and W independently.

Example (Tensor product of bases)**.** If $\{v_i\}_{i \in I}$ and $\{w_j\}_{j \in J}$ are bases of V and W, then $\{v_i \otimes w_j \mid i \in I,\ j \in J\}$ is a basis for $V \otimes W$. Let $\{v_i\}$ and $\{w_j\}$ be the standard bases in two and three dimensions, represented as column and row vectors respectively. One element of the basis for the tensor product is:

$$v_1 \otimes w_3 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

It is clear that this construction develops the standard basis for $V_2 \otimes W_3$.

Example (Polynomial ring and matrices)**.** Let $V = \mathbb{R}[x]$ and $W = \mathrm{Mat}_{2 \times 2}(\mathbb{R})$. A formal expression of an element in $V \otimes W$ is, for example, $(2 + 2x) \otimes \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

Example (Dirac matrices)**.** The following anticommutation relations have a corresponding matrix representation called the Pauli matrices:

$$\sigma_i^2 = I_2,\quad \sigma_i \sigma_j + \sigma_j \sigma_i = 0,\ i \neq j$$
$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},\ \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix},\ \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

To develop the relativistic theory of the electron, Dirac constructed 4×4 matrices γ^μ, $\mu = 0, 1, 2, 3$, out of the Pauli matrices using tensor products. The details are not relevant, so we will just display the results:

$$\gamma := \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

$$\gamma^0 = \sigma_3 \otimes I_2 = \begin{bmatrix} I_2 & 0 \\ 0 & -I_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

$$\gamma^i = \gamma \otimes \sigma_i = \begin{bmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{bmatrix},\ i = 1, 2, 3$$

The Pauli and Dirac matrices form Clifford algebras respectively.

1.1 Motivating Examples for Hopf Algebras

Example 1 (Groups)**.** Let G be a *finite group* and $\mathbb{K}$ a *field*. To allow multiplication of scalars into the group, define the group ring:

$$H = \mathbb{K}G = \left\{ \sum_{i=1}^{n} \lambda_i g_i \;\middle|\; n \in \mathbb{N},\; \lambda_i \in \mathbb{K},\; g_i \in G \right\}$$

with the ring multiplication $g \cdot h = gh$, the group multiplication. Extending this linearly:

$$\left(\sum_{i=1}^{n} \lambda_i g_i \right) \left(\sum_{j=1}^{p} \mu_j h_j \right) = \sum_{i=1}^{n} \sum_{j=1}^{p} \lambda_i \mu_j (g_i h_j)$$

<u>Comultiplication:</u> $\Delta(g) = g \otimes g$. Extending this linearly:

$$\Delta\left(\sum_{i=1}^{n} \lambda_i g_i \right) = \sum_{i=1}^{n} \lambda_i (g_i \otimes g_i)$$

These definitions do give a Hopf algebra according to the lecturer.

Remark. The comultiplication $\Delta(g) = 1 \otimes g + g \otimes 1 + g \otimes g$ should also be valid, as the mapping *should* point to all possible information of where g can come from. As we will see later, this idea will be partially evident in future definitions of different structures.

Example 2 (Polynomial rings)**.** Let $H = \mathbb{K}[X]$, the *polynomial ring*.

<u>Multiplication:</u> $m\big(X^i \otimes X^j\big) = X^{i+j}$. Extending this linearly:

$$m\left(\sum_{i=0}^{n} \alpha_i X^i \otimes \sum_{j=0}^{p} \beta_j X^j \right) = \sum_{i=0}^{n} \sum_{j=0}^{p} \alpha_i \beta_j X^{i+j}, \quad \forall\, \alpha_i, \beta_j \in \mathbb{K},\; \forall\, n, p \in \mathbb{N}$$

Remark. This is a commutative multiplication because $X^{i+j} = X^{j+i}$.

<u>Comultiplication:</u> Extending the logic from groups:

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

Remark. The combination $(X \otimes X)$ is not included here probably because the addition of ring elements is analogous to multiplication of group elements, justifying $g \otimes g$ in the definition for groups.

This ring already has a multiplicative structure, which the coproduct should obey naturally:

$$\Delta\big(X^2\big) = [1 \otimes X + X \otimes 1] \cdot [1 \otimes X + X \otimes 1]$$
$$= 1 \otimes X^2 + 2(X \otimes X) + X^2 \otimes 1$$

This suggests the following extension:

$$\Delta(X^n) = \sum_i^n \binom{n}{i} \left(X^i \otimes X^{n-i} \right)$$

Notice the following, which shows the information about the coefficient of a comultiplication:

$$m(\Delta(X^n)) = \sum_{i=1}^n \binom{n}{i} X^n = 2^n X^n$$

Example 3 (Graphs). Let $[G]$ be the *isomorphism class of the graph* $G = (V, E)$. The collection of all isomorphic graphs is denoted by $L = \{\, [G] \mid G \in \mathcal{G}\,\}$, where $\mathcal{G}$ is the class of all graphs. Define the vector space $H = \mathbb{K}L$, in which the elements are linear combinations of graphs:

$$H = \left\{ \sum_i k_i [G_i] \,\middle|\, k \in \mathbb{K},\ G \in \mathcal{G} \right\}$$

Multiplication: Defined most naturally as a disjoint union of the arguments as components of the graph, which is commutative:

(1)

Comultiplication: Not entirely obvious, but a satisfactory definition can be deduced from the previous examples and observations. The comultiplication of an element is essentially the sum of all of its possible decompositions into two complementary subelements.

This idea can be implemented in this algebra as well by selecting a subset S of the vertex set of the graph $V(G)$ and taking the tensor product of the graphs induced by S and its complement $V(G) - S$ (a graph induced by a subset of the vertex set consists of vertices in that subset and all of the edges between these vertices that occur in the original graph). This is shown by (dropping the isomorphism class notation for brevity):

$$\Delta(G) = \sum_{S \subseteq V(G)} G\big|_S \otimes G\big|_{V(G)-S}$$

This is illustrated using the following example:

The number of tensor products is easily countable as the number of subsets of $V(G)$, which is 2^n, where n is the number of vertices of the graph G. For the above example, there are 5 vertices, so there are 32 tensor products to be summed over.

4

Remark. The product of the coproduct of a graph is:

$$m(\Delta(G)) \neq G$$

This indicates that the product of the coproduct of a graph does not bring back the original argument, so m and Δ are not inverse operations. (Note: This is also evident from the group and the polynomial ring examples.)

Example 4 (Permutations). Let P_n denote the set of all permutations of $\mathbb{N} \pmod{n} := \{1, 2, \ldots, n\}$, for a given non-negative integer n. Let P denote the set of all permutations of $\mathbb{N} \pmod{n}$, for all non-negative integers n — i.e., $P := \bigcup\limits_{n=0}^{\infty} P_n$. Here are two permutations in P:

$$p = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 3 & 4 & 5 \end{pmatrix}, \quad q = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

Remark. Although they might have the 'same' action, they are fundamentally different elements of P, as p is an element of P_5 and q that of P_2.

The vector space is defined as:

$$H = \mathbb{K}P := \left\{ \sum_{i=1}^{n} k_i p_i \ \middle|\ k_i \in \mathbb{K},\ p_i \in P,\ n \in \mathbb{N}_0 \right\}$$

Note that the coefficients do *not* denote the number of times of application of their corresponding permutations; they just depict a formal linear extension in the vector space.

<u>Multiplication</u>: An example is $(1\,2) \otimes (3\,2\,1)$ using one-line notation (*not* cycle notation):

$$
\begin{aligned}
m((1\,2) \otimes (3\,2\,1)) = &(1\,2\,5\,4\,3) + (1\,5\,2\,4\,3) + (1\,5\,4\,2\,3) + (1\,5\,4\,3\,2) + (5\,1\,2\,4\,3) + \\
&(5\,1\,4\,2\,3) + (5\,1\,4\,3\,2) + (5\,4\,1\,2\,3) + (5\,4\,1\,3\,2) + (5\,4\,3\,1\,2)
\end{aligned}
$$

This shows that it respects the relative order of the elements within each argument as in a riffle shuffle of cards and generates only *some* permutations in P_5 as a result. Also note how the elements $1, 2, 3$ in the second argument were relabelled $3, 4, 5$ respectively, to ensure that we get an element of P_5 after shuffling. This might be represented as given below, with u_i denoting the ith element in the permutation U:

$$
\begin{aligned}
m(U \otimes V) &= m((u_1 \ldots u_n) \otimes (v_1 \ldots v_p)) \\
&= \sum (w_1 \ldots w_{n+p})
\end{aligned}
$$

Where $(w_1 \ldots w_{n+p})$ stands for every possible permutation of the resultant set with the values of v_j right-shifted by n, under the rule that the relative orders of the elements of U and V respectively are not lost — i.e., for any two w_i, w_j, $1 \leq i < j \leq n + p$, one of the following holds:

1. $w_i = u_k$ and $w_j = v_r + n$

2. $w_i = v_r + n$ and $w_j = u_k$

3. $w_i = u_k$ and $w_j = u_l$, with $k < l$

4. $w_i = v_r + n$ and $w_j = v_s + n$, with $r < s$

where $1 \leq k, l \leq n$, $1 \leq r, s \leq p$.

Comultiplication: Similar to graphs, the coproduct of permutations is obtained by cutting a line between the numbers (in one-line notation, e.g. $(1\,2 \mid 3\,4\,5)$), and writing the outer product with each piece as a component, then summing over all complementary combinations.

$$\Delta(u_1 u_2 \ldots u_n) = \sum_{i=0}^{n} (u_1 \ldots u_i) \otimes (v_1 \ldots v_{n-i})$$

where $v_1, \ldots, v_{n-i}$ are obtained by relabelling $u_{i+1}, \ldots, u_n$ respectively in such a way that the resulting set $\{v_1, \ldots, v_{n-i}\} = \mathbb{N} \pmod{(n-i)}$, but preserving the relative order of elements. An example using a permutation of P_5:

$$\Delta(4\,2\,5\,3\,1) = (\varnothing \otimes 42531) + (1 \otimes 2431) + (21 \otimes 321) + \cdots$$

Remark. Nothing so far.

2 Algebras Over a Field/Ring

An *algebra over a field* is a structure that is simultaneously a *vector space* and a *ring*. The setup will be a field $\mathbb{K}$ and a ring A with a multiplicative identity 1, called a $\mathbb{K}$-algebra. Three definitions are introduced that will be shown to be equivalent.

Definition 3 ($\mathbb{K}$-algebra). *A forms a $\mathbb{K}$-algebra if $\mathbb{K} \subseteq Z(A)$ and $1_{\mathbb{K}} = 1_A$, where $Z(A)$ is the centre of the ring.*

Example (Polynomial rings). Let $A = \mathbb{K}[x]$. This forms a $\mathbb{K}$-algebra because the ring is commutative and the elements of the field in the ring are the constant polynomials, and $1_{\mathbb{K}}$ is the same as 1_A. This directly generalises to the multivariate polynomial ring $B = \mathbb{K}[x_1, \ldots, x_n]$, $n \in \mathbb{N}$.

An example of a ring that is a vector space, but does not form a $\mathbb{K}$-algebra as per Definition 3 is $A = \mathrm{Mat}_{n \times n}(\mathbb{K})$. This is because $\mathbb{K} \not\subseteq A$. However, a "copy" of $\mathbb{K}$ exists as the subfield of matrices of the form λI, where I is the identity matrix of corresponding dimensions, and $\lambda \in \mathbb{K}$. This structure should form a $\mathbb{K}$-algebra, which motivates the second definition.

Definition 4 ($\mathbb{K}$-algebra). *A is a $\mathbb{K}$-algebra if there is an embedding $u \colon \mathbb{K} \to A$ such that $u(\mathbb{K}) \subseteq Z(A)$ and $u(1_{\mathbb{K}}) = 1_A$.*

Remark. This allows $A = \mathrm{Mat}_{n \times n}(\mathbb{K})$ to form a $\mathbb{K}$-algebra if we define $u(\lambda) = \lambda I$, $\forall \lambda \in \mathbb{K}$. An observation is that any A now forms a $\mathbb{K}$-vector space because scalar multiplication is implemented as $\lambda \cdot a \coloneqq u(\lambda)a$, $\lambda \in \mathbb{K}$, $a \in A$. The construction of the polynomial ring $\mathbb{K}[x]$ as a $\mathbb{K}$-algebra also follows directly from this definition.

An *algebra over a ring*, by analogy, is a module and a ring with a multiplicative identity defined similarly:

Definition 5 (*R*-algebra)**.** M is an *R-algebra* if there is a ring homomorphism $u \colon R \to M$ such that $u(R) \subseteq Z(A)$ and $u(1_R) = 1_A$ with a left or right multiplication specified.

Example ($\mathbb{Z}$-algebra)**.** A quaternion $a + bi + cj + dk$ is a *Hurwitz quaternion* if a, b, c, and d are either all integers, or all half-integers — i.e., halves of odd integers. Let

$$H := \left\{ a + bi + cj + dk \,\middle|\, a, b, c, d \in \mathbb{Z} \text{ or } a, b, c, d \in \mathbb{Z} + \frac{1}{2} \right\}$$

be the set of all Hurwitz quaternions. Then H is a $\mathbb{Z}$-algebra, since its centre consists of all Hurwitz quaternions with imaginary parts zero, which contains a copy of $\mathbb{Z}$ as a subring.

2.1 Algebra Homomorphism

Definition 6. Let A_1, A_2 be $\mathbb{K}$-algebras. An *algebra homomorphism* is $\phi \colon A_1 \to A_2$ which is a ring homomorphism and linear map simultaneously.

Remark. Observations:

- $B \subseteq A$ is a subalgebra if B is both a subspace and subring.

- A structure that quotients the algebra must quotient the vector space and the ring simultaneously. A subspace quotients the vector space, and an ideal quotients the ring. All ideals form subspaces, therefore ideals quotient the algebra.

- If $\phi \colon A \to B$ is a $\mathbb{K}$-algebra homomorphism, then $\operatorname{Im} \phi \cong A / \ker \phi$, where $\ker \phi$ forms an ideal I. This is analogous to the first isomorphism theorem for groups, rings, vector spaces, etc.

- The direct product of two $\mathbb{K}$-algebras is also a $\mathbb{K}$-algebra, as the vector space and ring structures are independently inherited. The tensor product also forms a $\mathbb{K}$-algebra.

Definition 7 ($\mathbb{K}$-algebra)**.** A $\mathbb{K}$-algebra is a vector space A over a field $\mathbb{K}$ equipped with linear maps $m \colon A \otimes A \to A$, $u \colon \mathbb{K} \to A$ such that the following diagrams commute:

$$
\begin{array}{ccc}
A \otimes A \otimes A & \xrightarrow{m \otimes \mathrm{id}} & A \otimes A \\
{\scriptstyle \mathrm{id} \otimes m} \downarrow & & \downarrow {\scriptstyle m} \\
A \otimes A & \xrightarrow{\quad m \quad} & A
\end{array}
\qquad
\begin{array}{ccccc}
\mathbb{K} \otimes A & \xrightarrow{u \otimes \mathrm{id}} & A \otimes A & \xleftarrow{\mathrm{id} \otimes u} & A \otimes \mathbb{K} \\
& {\scriptstyle f} \searrow & \downarrow {\scriptstyle m} & \swarrow {\scriptstyle g} & \\
& & A & &
\end{array}
$$

where f and g are the natural isomorphisms $\mathbb{K} \otimes A \to A$, $\lambda \otimes x \mapsto \lambda x$ and $A \otimes \mathbb{K} \to A$, $x \otimes \lambda \mapsto \lambda x$ respectively.

Remark. For vectors x and y, we write $m(x, y)$ as xy.

For the first diagram: $(m \otimes \mathrm{id})(x \otimes y \otimes z) = xy \otimes z$, $m(xy \otimes z) = (xy)z$ and $(\mathrm{id} \otimes m)(x \otimes y \otimes z) = x \otimes yz$, $m(x \otimes yz) = x(yz)$, which shows that the multiplication m is *associative*:

$$m(m(x \otimes y) \otimes z) = m(x \otimes m(y \otimes z))$$

For the second diagram: $(u \otimes \mathrm{id})(\lambda \otimes x) = u(\lambda) \otimes x$, and $m(u(\lambda) \otimes x) = u(\lambda)x$; also, $f(\lambda \otimes x) = \lambda x$. Since the diagram commutes, $u(\lambda)x = \lambda x$. This shows that left-multiplying a vector x by the vector $u(\lambda)$ is equivalent to multiplying it by the scalar λ. The same holds for right multiplication. If $1_A := u(1_{\mathbb{K}})$, then $1_A x = 1_{\mathbb{K}} x = x$ and $x1_A = x1_{\mathbb{K}} = x$, which shows that the multiplication m is *unitary*, with unity 1_A.

This definition is equivalent to the previous one, which can be verified by checking if the structure shares the remaining property of a ring, *distributivity*, which is satisfied because of the linearity of m:

$$m(u \otimes (v + w)) = m(u \otimes v) + m(u \otimes w) = uv + uw$$

Remark. To show that the previous definition implies the current one, the only non-trivial construction required is m, which is defined as a multiplication under the previous definition: $m \colon A \otimes A \to A$ such that $m(a \otimes b) := ab$; this remark is a little troubling because we do not have a clear definition of the tensor product and its relations with other algebraic structures, which will be investigated now.

3   Tensor Products

Definition 8 (Free vector space). Let $F(V \times W) = \mathbb{K}\text{-span}\{\, (v, w) \mid v \in V, w \in W \,\}$, called the *free vector space* on $V \times W$.

Definition 9 (Tensor product). The tensor product space is the quotient:

$$V \otimes W = F(V \times W)/I$$

where I is generated by linear relations of v and w:

$$(v_1 + v_2, w) \sim (v_1, w) + (v_2, w)$$
$$(v, w_1 + w_2) \sim (v, w_1) + (v, w_2)$$
$$(cv, w) \sim (v, cw) \sim c(v, w)$$

Remark. The tensor $a \otimes b = \overline{(a, b)}$, an equivalence class, and $a' \otimes b'$ can be equal to $a \otimes b$, $a' \neq a$, $b' \neq b$.

To define a linear $m \colon V \otimes W \to W$, we would have to define an $n \colon F(V \times W) \to W$ such that $n(I) = 0$, which implies that n is a bilinear map.

3.1 Universal Property

Theorem 1 (Universal property of tensor products). *Let $\phi\colon V \times W \to V \otimes W$ be a bilinear map. If $f\colon V \times W \to L$ is a bilinear map, then there exists a unique linear map $\tilde{f}\colon V \otimes W \to L$ such that $f = \tilde{f} \cdot \phi$, which satisfied the commutative diagram:*

$$
\begin{array}{ccc}
V \times W & \xrightarrow{\ \phi\ } & V \otimes W \\
& \searrow{\scriptstyle f} & \downarrow{\scriptstyle \tilde{f}} \\
& & L
\end{array}
$$

Proof. Let $\phi(v, w) := (v, w) + I = v \otimes w$. The following are evident from their respective properties:

$$
\begin{aligned}
f(\lambda v, w) &= \lambda f(v, w) && \text{(bilinearity)} \\
\phi(\lambda v, w) = \phi(v, \lambda w) &= \lambda(v, w) + I = \lambda(v \otimes w) && \text{(bilinearity)} \\
\tilde{f}(\lambda(v \otimes w)) = \lambda \tilde{f}(v \otimes w) &:= \lambda f(v, w) && \text{(linearity)}
\end{aligned}
$$

Extending $\tilde{f}$ linearly:

$$
\tilde{f}\left(\sum_{i=1}^{m} \lambda_i v_i \otimes w_i\right) := \sum_{i=1}^{n} \lambda_i f(v_i, w_i)
$$

If $\sum_{i=1}^{m} \lambda_i(v_i, w_i) = \sum_{j=1}^{n} \mu_j(x_j, y_j)$, then for the map to be well-defined, we must have:

$$
f\left(\sum_{i=1}^{m} \lambda_i(v_i, w_i)\right) = f\left(\sum_{j=1}^{n} \mu_j(x_j, y_j)\right)
$$

Because $\sum_{i=1}^{m} \lambda_i(v_i, w_i) - \sum_{j=1}^{n} \mu_j(x_j, y_j) \in I$. $\qquad\square$