Uniwersytet Wrocławski
Wydział Matematyki i Informatyki
Instytut Matematyczny
specjalność teoretyczna

Bartosz Sójka

Explanation ąćęłóśńźż of connection between Hopf algebras and Markov chains

Praca licencjacka
napisana pod kierunkiem
prof. dr. hab. Dariusza Buraczewskiego

Wrocław 2018

Contents

Abstract

In [DPR14] Persi Diaconis, Amy Pang and Arun Ram described how to use Hopf algebras to study Markov chains. As it involves ideas from quite different branches of mathematics, it could be hard to grasp a concept of it if someone is not familliar with them. The point of this paper is to describe some of their results in a more step-by-step, simplified way, so that they could be accessible to third year students after probability and abstract algebra courses. I will focus on the example of shuffling cards by inverse riffle shuffle method. Structure will be as follows: first there will be an introduction to both Hopf algebras and Markov chains, then it will be explained how to describe a Markov chain with a Hopf algebra, finally I will describe how to find left eigenbasis and right eigenbasis of Markov chain associated with riffle shuffling using Hopf algebras.

TO DO:

- finish

Chapter 1

Markov chains

Finite Markov chain is a random process on a finite set of states such that the probability of being in some state in the moment $n+1$ depends only on the state in which one was in the moment n. Now we will put this more formally.
Let $S = \{s_1, \ldots, s_k\}$. The sequence of random variables $(X_0, X_1, \ldots)$ with values in S is a Markov chain with state space S if for all $n \in \mathbb{N}$, for all $s_{i_0}, s_{i_1}, \ldots, s_{i_{n+1}} \in S$ such that

$$\mathbb{P}(X_0 = s_{i_0}, \ldots, X_n = s_{i_n}) > 0$$

following condition (called Markov property) holds:

$$\mathbb{P}(X_{n+1} = s_{i_{n+1}} \mid X_0 = s_{i_0}, \ldots, X_n = s_{i_n}) = \mathbb{P}(X_{n+1} = s_{i_{n+1}} \mid X_n = s_{i_n}). \tag{1.1}$$

It states that for all $s_i, s_j \in S$ the probability of moving from the state s_i to the state s_j is the same no matter what states $s_{i_0}, \ldots, s_{i_{n-1}}$ were visited before.
For the Markov chain $(X_0, X_1, \ldots)$ the $|S| \times |S|$ matrix $K_{i,j} = \mathbb{P}(X_{n+1} = s_j \mid X_n = s_i)$ is called the transition matrix. We will sometimes write $K(s_i, s_j)$ instead of $K_{i,j}$. Note that the sum of any row is equal to 1 since it is the sum of probabilities of moving somewhere frome s_i. Now $K^n_{i,j}$ is the chance of moving from s_i to s_j in n steps.
Markov chains can be also viewed as random walks on the directed, labeled graphs, where states are vertices and edge's label is the probability of moving from one vertex to another.
Card shuffling can be viewed as a Markov chain on all possible arrangements of the cards in the deck with $K(x, y)$ equal to probability of going from arrangement x to arrangement y in one shuffle.
More extensive indroduction can be found in [LPW17].

Cośtam cośtam stationary distirution.

1.1 Gilbert-Shannon-Reeds model of riffle shuffle

TO DO: BAM! Gilbert-Shannon-Reeds model of riffle shuffle with that specific forward

Chapter 2

Hopf algebras

Now I will give a full definition of a Hopf algebra. Although it is quite long and involves definition of ??? operations, I decided to put it in a consistent fragment, due to belief that thanks to that it will be a better reference.
If the reader feels lost in this section, it is recommended to read it in parallel to the section 2.3 where examples are provided or treat it just as a reference when formal definition will be needed. Another reason of arranging text like that (and possibility of treating this section just as a reference), is that for most of the time we will not be using full structure of a Hopf algebra. Nevertheless it is good to see the full shape of what we are dealing with.
So now comes the full definition but we will try to explain it piece by piece.

2.1 Preliminaries

2.1.1 Notational remarks

Remark. Let K be a field. In the following section k, if not stated otherwise, will denote an arbitrary elementof this field. If not stated otherwise, all vector spaces will be over K and all tensor products will be taken over K. Note that when we want to present a field multiplication from K as a linear map $K \otimes K \to K$ it will be denoted as ${}^K m$. As it is then an isomorphism let ${}^K\Delta := {}^K m^{-1}$. The 1 from K will be denoted as 1_K.

Remark. Let U, V, W, Z be vector spaces over field K. z5We will use notation $\varphi \otimes \psi : U \otimes V \to W \otimes Z$ which, for φ, ψ such that $\varphi : U \to W$, $\psi : V \to Z$, means a linear map that for all $u \in U$, $v \in V$ satisfies:

$$(\varphi \otimes \psi)(u \otimes v) = \varphi(u) \otimes \psi(v).$$

Because of linearity, for elements of shape $\sum_{i=1}^{n} u_i \otimes v_i$ it will take form:

$$(\varphi \otimes \psi)(\sum_{i=1}^{n} u \otimes v) = \sum_{i=1}^{n} \varphi(u) \otimes \psi(v).$$

I, if not stated otherwise, will be an identity in the adequate space.
T, if not stated otherwise, will be the twist map $T : V \otimes W \to W \otimes V$, which is linear map such that for any $v \otimes w \in V \otimes W$

$$T(v \otimes w) = w \otimes v.$$

For an n-tensor power $\overbrace{V \otimes \cdots \otimes V}^{n\ times}$ of a vector space V we will sometimes write $V^{\otimes n}$.
Throughout this paper, when there will be no risk of confusion, we will omit the "$\circ$" symbol of composition of functions and we will write $\varphi\psi(x)$ instead of $(\varphi \circ \psi)(x)$.

Dual spaces

We will use standard notation for dual spaces:
For a vector space V over a field K we will write V^* for a vector space dual to V - a vector space of all linear functions from V to K.

2.1.2 Tensor products

First we will introduce tensor product of the vector spaces. Let V, W be vector spaces over the field K. Let Z be a vector space with basis $V \times W$. Note that we are taking entire $V \times W$ as a basis of Z, not just a basis of $V \times W$. Consequently, every non-zero element of Z has unique representation in the form $\sum_{i=1}^{n} \alpha_i(v_i, w_i)$. Let $\simeq$ be the smallest equivalece relation on Z satisfying:
For all $v, v_1, v_2 \in V,\ w, w_1, w_2 \in W,\ k \in K$

$$\begin{aligned}(v, w_1) + (v, w_2) &\simeq (v, w_1 + w_2),\\ (v_1, w) + (v_2, w) &\simeq (v_1 + v_2, w),\\ k(v, w) &\simeq (kv, w),\\ k(v, w) &\simeq (v, kw).\end{aligned}$$

Since for all $z_1, z_2, z_3, z_4 \in Z$, all $k \in K$

$$z_1 \simeq z_2 \wedge z_3 \simeq z_4 \implies z_1 + z_3 \simeq z_2 + z_4 \text{ and}$$
$$z_1 \simeq z_2 \implies kz_1 \simeq kz_2,$$

we treat $Z/_{\simeq}$ as a vector space with operations

$$[z_1]_{\simeq} + [z_2]_{\simeq} \coloneqq [z_1 + z_2]_{\simeq},$$
$$k[z_1]_{\simeq} \coloneqq [kz_1]_{\simeq}.$$

We denote equivalence class $[(v,w)]_{\simeq}$ as $v \otimes w$. The tensor product $V \otimes W \coloneqq Z/_{\simeq}$. Note that in $V \otimes W$ there are vectors that cannot be written as $v \otimes w$ for any v, w. However, every $z \in V \otimes W$ can be written in as $z = \sum_{i=1}^{n} v_i \otimes w_i$ for some $v_1, \ldots, v_n \in V$, $w_1, \ldots, w_n \in W$. (More detailed explanation of this fact and the following example will come in the Observation 1..)
For example take $v_1, \ldots, v_n, w_1, \ldots, w_n$ such that they are linearly independent in corresponding spaces. Then take $\sum_{i=1}^{n} (v_i, w_i)$. There are no v, w such that $[(v,w)]_{\simeq} = \left[\sum_{i=1}^{n} (v_i, w_i)\right]_{\simeq}$. Thus, for the element $\left[\sum_{i=1}^{n} (v_i, w_i)\right]_{\simeq}$ of $V \otimes W$ there are no v, w such that $v \otimes w = \left[\sum_{i=1}^{n} (v_i, w_i)\right]_{\simeq}$. However, since $\left[\sum_{i=1}^{n} (v_i, w_i)\right]_{\simeq} = \sum_{i=1}^{n} [(v_i, w_i)]_{\simeq}$ it can be written as $\sum_{i=1}^{n} v_i \otimes w_i$.
Now we will make some further observations on how $V \otimes W$ looks like.

Observation 1. *If $\{b_i\}_{i \in I}$, $\{c_j\}_{j \in J}$ are bases of, respectively, V and W, then $\{b_i \otimes c_j : i \in I, j \in J\}$ is the basis of $V \otimes W$.*

Proof. Let $z = \sum_{i=1}^{n} \alpha_i (v_i, w_i)$ be an arbitrary non-zero element of Z. We will show that $[z]_{\simeq}$ has representation as $\sum_{i=1}^{m} \beta_i [(b_i, c_i)]_{\simeq} \left(= \sum_{i=1}^{m} \beta_i (b_i \otimes c_i) \right)$.

$$
\begin{aligned}
[z]_{\simeq} = \left[\sum_{i=1}^{n} \alpha_i(v_i, w_i)\right]_{\simeq} &= \sum_{i=1}^{n} \alpha_i[(v_i, w_i)]_{\simeq} \\
&= \sum_{i=1}^{n} \alpha_i \left[\left(\sum_{j=1}^{l_1} \gamma_{i,j} b_{i,j}, \sum_{k=1}^{l_2} \gamma_{i,k} c_{i,k}\right)\right]_{\simeq} \\
&= \sum_{i=1}^{n} \alpha_i \left[\sum_{j=1}^{l_1} \gamma_{i,j} \left(b_{i,j}, \sum_{k=1}^{l_2} \gamma_{i,k} c_{i,k}\right)\right]_{\simeq} \\
&= \sum_{i=1}^{n} \alpha_i \left[\sum_{j=1}^{l_1} \gamma_{i,j} \left(\sum_{k=1}^{l_2} \gamma_{i,k} \left(b_{i,j}, c_{i,k}\right)\right)\right]_{\simeq} \\
&= \sum_{i=1}^{n} \alpha_i \left[\sum_{\substack{1 \leqslant j \leqslant l_1 \\ 1 \leqslant k \leqslant l_2}} \gamma_{i,j}\gamma_{i,k}(b_{i,j}, c_{i,k})\right]_{\simeq} \\
&= \sum_{i=1}^{n} \alpha_i \left(\sum_{\substack{1 \leqslant j \leqslant l_1 \\ 1 \leqslant k \leqslant l_2}} \gamma_{i,j}\gamma_{i,k}[(b_{i,j}, c_{i,k})]_{\simeq}\right) \\
&= \sum_{\substack{1 \leqslant i \leqslant n \\ 1 \leqslant j \leqslant l_1 \\ 1 \leqslant k \leqslant l_2}} \alpha_i \gamma_{i,j}\gamma_{i,k}[(b_{i,j}, c_{i,k})]_{\simeq}
\end{aligned}
$$

Thus $\{b_i \otimes c_j : i \in I, j \in J\}$ spans $V \otimes W$. To prove linear independence we can observe that if $\sum_{i=1}^{m} \alpha_i[(v_i, w_i)]_{\simeq} = 0$, then either $v_1, \ldots v_n$ or $w_1, \ldots, w_n$ have to be linearly dependent. It can't occur if $v_1, \ldots v_n$ and $w_1, \ldots, w_n$ are from the bases of V and W.
This observation also justifies recently cited fact and the example. □

Observation 2. *If V and W are finite dimensional and* $\dim(V) = n$, $\dim(W) = m$*, then* $\dim(V \otimes W) = nm$*.*

Proof. The proof is immediate from the Observation 1.. Since if $\{b_i\}_{i \in I}$, $\{c_j\}_{j \in J}$ are bases of, respectively, $V$and W and $\dim(V) = n$ and $\dim(W) = m$, then $| \{b_i \otimes c_j : i \in I, j \in J\} | = nm$ □

Observation 3. *$V \otimes W$ is a vector space of elements in the shape of $\sum_{i=1}^{n} v_i \otimes w_i$ with operations on them defined such that for all $v, v_1, v_2 \in V$, $w, w_1, w_2 \in$*

W, $k \in K$ there hold

$$v_1 \otimes w + v_2 \otimes w = (v_1 + v_2) \otimes w,$$
$$v \otimes w_1 + v \otimes w_2 = v \otimes (w_1 + w_2),$$
$$k(v \otimes w) = (kv) \otimes w = v \otimes (kw).$$

Proof. This observation is just a recollection of the definition. □

Observation 4. *For vector spaces U, V, W over the field K there is a natural isomorphism between $(U \otimes V) \otimes W$ and $U \otimes (V \otimes W)$ therefore there is no ambiguity in writing $U \otimes V \otimes W$ or a product of any greater number of vector spaces in that way. (We will also write "$u \otimes v \otimes w$" for some of their elements.) Form of elements, operations on them and structure of that vector spaces are fully analogous to the described above (with respect to all "coordinates" in terms like $u \otimes v \otimes w$ and so on). So the space $U \otimes V \otimes W$ has elements of shape $\sum_{i=1}^{n} u_i \otimes v_i \otimes w_i$ (each for some $u_1, \ldots, u_n \in U$, $v_1, \ldots v_n \in V$, $w_1, \ldots, w_n \in W$) and for all $u, u_1, u_2 \in U$, $v, v_1, v_2 \in V$, $w, w_1, w_2 \in W$, $k \in K$ there hold*

$$u_1 \otimes v \otimes w + u_2 \otimes v \otimes w = (u_1 + u_2) \otimes v \otimes w,$$
$$u \otimes v_1 \otimes w + u \otimes v_2 \otimes w = u \otimes (v_1 + v_2) \otimes w,$$
$$u \otimes v \otimes w_1 + u \otimes v \otimes w_2 = u \otimes v \otimes (w_1 + w_2),$$
$$k(u \otimes v \otimes w) = (ku) \otimes v \otimes w = u \otimes (kv) \otimes w = u \otimes v \otimes (kw).$$

Proof. Left to the reader. □

Observation 5. *If V is a vector space over K, then all elements of $K \otimes V$ ($V \otimes K$) can be expressed in form $1 \otimes v$ ($v \otimes 1$) and there are natural isomorphisms ${}^{L}m : K \otimes V \to V$, (${}^{R}m : V \otimes K \to V$) given by*

$${}^{L}m(k \otimes v) = kv,$$
$${}^{R}m(v \otimes k) = kv.$$

Proof. An arbitrary element of $K \otimes V$ has form $\sum_{i=1}^{n} k_i \otimes v_i$ but

$$\sum_{i=1}^{n} k_i \otimes v_i = \sum_{i=1}^{n} 1 \otimes k_i v_i = 1 \otimes \sum_{i=1}^{n} k_i v_i.$$

${}^{L}m$ is linear (left for the reader) and is bijective because for all $v, v_1, v_2 \in V$

$$\varphi(1 \otimes v) = v$$

and

$$1 \otimes v_1 = 1 \otimes v_2 \iff 1 \otimes v_1 - 1 \otimes v_2 = 0 \iff$$
$$1 \otimes (v_1 - v_2) = 0 \iff v_1 - v_2 = 0 \iff v_1 = v_2.$$

The proof for $V \otimes K$ and ${}^{R}m$ is analogous. In the later sections we will use notations of ${}^{L}m$ and ${}^{R}m$ for those isomorphism for any space. $\square$

Remark. In a special case when $V = W = K$ the natural isomorphisms described above take form of ${}^{K}m : K \otimes K \to K$ that for all $k_1, k_2 \in K$ ${}^{K}m(k_1 \otimes k_2) = k_1 k_2$. This isomorphism of $K \otimes K$ and K is just a field multiplication from K.

Remark. Thanks to Observation 3. there is no ambiguity in writing $kv \otimes w$. I hope that this third observation will also help us understand what the tensor product is and what it is not. It will be good to keep it in mind when we are intensively dealing with it in a combinatorical way in the following sections.

2.2 Algebras

Definition 1. *A* ***K-algebra*** *is a vector space $\mathcal{H}$ with additional associative, linear operation $m : \mathcal{H} \otimes \mathcal{H} \to \mathcal{H}$ called multiplication and a linear map $u : K \to \mathcal{H}$ called unit such that for all $a \in \mathcal{H}$*

$$m(u(1_K) \otimes a) = m(a \otimes u(1_K)) = a.$$

Explanation. Operation m defines on $\mathcal{H}$ a structure of a unitary ring by setting the ring multiplication (let it be denoted as "$\cdot$") as $a \cdot b = m(a \otimes b)$. The identity element of that ring multiplication is then $u(1)$. (We will be calling $u(1)$ also an identity element of multiplication m in K-algebra $\mathcal{H}$ or the 1 in the $\mathcal{H}$ and denote it as $1_{\mathcal{H}}$)
Proof. The fact that m is associative means that for all $a_1, a_2, a_3 \in \mathcal{H}$

$$m(m(a_1 \otimes a_2) \otimes a_3) = m(a_1 \otimes m(a_2 \otimes a_3)).$$

That implies that

$$(a \cdot b) \cdot c = m(m(a \otimes b) \otimes c) = m(a \otimes m(b \otimes c)) = a \cdot (b \cdot c).$$

So "$\cdot$" is proper ring multiplication. Recalling the definition of u, we can write that for all $a \in \mathcal{H}$

$$u(1_K) \cdot a = a \cdot u(1_K) = a$$

So indeed it is an identity element of that ring. As u is linear map it can be seen as natural insertion of a field K into an algebra $\mathcal{H}$ that maps 1_K to $1_{\mathcal{H}}$ (1 from the K to the identity element of multiplication in $\mathcal{H}$) and extends linearly. Given that we can observe that for all $a \in \mathcal{H}$, all $k \in K$, a multiplicated by $u(k)$ (no matter if from the left or right) is exactly the ka (an element of vector space $\mathcal{H}$). So we can think about $u[K]$ as a copy of K in $\mathcal{H}$ that acts on $\mathcal{H}$ just like K.

Because of associativity we can define $m^{[3]} : \mathcal{H}^{\otimes 3} \to \mathcal{H}$ as

$$m^{[3]} := m(m \otimes I)$$

and for all $a_1, a_2, a_3 \in \mathcal{H}$ write

$$m^{[3]}(a_1 \otimes a_2 \otimes a_3) = a_1 \cdot a_2 \cdot a_3$$

with no ambiguity. And futher:

Let A be an algebra with multiplication m and unit u. We will recurrently define a sequence of maps $(m^{[n]})_{n \geqslant 2}$, such that $m^{[n]} : \underbrace{A \otimes \cdots \otimes A}_{n\ times} \to A$ as follows:

$$\begin{aligned} m^{[2]} &:= m, \\ m^{[n]} &:= m^{[n-1]}(m \otimes \underbrace{I \otimes \cdots \otimes I}_{n-2 \text{ times}}) \end{aligned}$$

which is a multiplication of all factors together.
Because of that, for all $a_1, \ldots, a_n \in A$ we can write

$$m^{[n]}(a_1 \otimes \cdots \otimes a_n) = a_1 \cdot \ldots \cdot a_n.$$

Remark. An algebra A is said to be commutative iff for all $a_1, a_2 \in A$

$$m(a_1 \otimes a_2) = m(a_2 \otimes a_1).$$

Remark. Later in the text we will still be using "$\cdot$" as a symbol for an algebra multiplication in an algebra of our interest.

2.3 Coalgebras

Definition 2. *A **K-coalgebra** is a vector space $\mathcal{H}$ with additional coassociative, linear operation $\Delta : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$ called comultiplication and a linear map $\varepsilon : \mathcal{H} \to K$ called counit such that for all $a \in \mathcal{H}$*

$$\begin{aligned} &(\varepsilon \otimes I)\Delta(a) = 1 \otimes a \text{ and} \\ &(I \otimes \varepsilon)\Delta(a) = a \otimes 1. \end{aligned}$$

Note that properties of a unit from a K-algebra can also be written in that manner as:

$$m(u \otimes I)(1_K \otimes a) = a \text{ and}$$
$$m(I \otimes u)(a \otimes 1_K) = a$$

means exactly what was in the definition of u.

Explanation. We will introduce a notation called Sweedler notation [Swe69] which will be very useful for writing coproducts. As for all $a \in \mathcal{H}$ we have $\Delta(a) = \sum_{i=1}^{n} a_{1,i} \otimes a_{2,i}$, we will write

$$\Delta(a) = \sum a_1 \otimes a_2.$$

This notation surpresses the index "i". Somewhere there can be also encountered an interjaced notation $\Delta(a) = \sum_{(a)} a_{(1)} \otimes a_{(2)}$.

In many cases comultiplication can be seen as a sum of possible decomposition of an element into elements "smaller" in some sense. For example, later it will come out that comultiplication is exactly the operation that models the process of cutting the deck of cards into pieces in riffle shuffle. In examples that we will work with (graded, connected Hopf algebras), comultiplication will represent some kind of natural decomposition in the more general way. What it means in the strict sense will be presented in Definition 8. when we will be introducing graded bialgebras.

Examples. **To do**

The coassociativity of Δ means that $(\Delta \otimes I)\Delta = (I \otimes \Delta)\Delta$. In Sweedler notation it can be written as

$$\forall_{a \in \mathcal{H}} \ \sum \Delta(a_1) \otimes a_2 = \sum a_1 \otimes \Delta(a_2)$$

or in a more expanded form as

$$\forall_{a \in \mathcal{H}} \ \sum \ a_{11} \otimes a_{12} \otimes a_2 = \sum a_1 \otimes a_{21} \otimes a_{22}. \tag{2.1}$$

Because of these equalities, terms from (2.1) can be written as $\sum a_1 \otimes a_2 \otimes a_3$ without ambiguity.

We can also define

$$\Delta^{[3]} := (\Delta \otimes I)\Delta$$

Now, for all $a \in \mathcal{H}$ there will be an equality

$$\Delta^{[3]}(a) = \sum a_1 \otimes a_2 \otimes a_3$$

which can be viewed as a sum of possible decompositions of a into three parts. In this point of view we can say that coassociativity of Δ means that Δ represents decomposition such that, when done twice, probabilities of possible outcomes are the same no matter which set of parts (a_1 or a_2) was been taken in the second iteration. It will be put more precise in the section 3. where we will present connection between Markov chains and Hopf algebras. Now we will take it a step futher:

Let C be a coalgebra with comultiplication Δ and counit ε. We will recurrently define a sequence of maps $(\Delta^{[n]})_{n\geqslant 2}$, such that $\Delta^{[n]} : C \to \underbrace{C \otimes \cdots \otimes C}_{n \text{ times}}$ as follows:

$$\begin{aligned}\Delta^{[2]} &:= \Delta,\\ \Delta^{[n]} &:= (\Delta \otimes \underbrace{I \otimes \cdots \otimes I}_{n-2 \text{ times}})\Delta_{n-1}.\end{aligned}$$

Which can be seen as composed iterations of Δ.
By induction it can be proved that for all $n \geqslant 3$, $i \in \{1, \ldots, n-2\}$, $m \in \{0, \ldots, n-i-1\}$ we have

$$\Delta^{[n]} = (\underbrace{I \otimes \cdots \otimes I}_{m \text{ times}} \otimes \Delta^{[i]} \otimes \underbrace{I \otimes \cdots \otimes I}_{n-i-1-m \text{ times}})\Delta^{[n-i]},$$

The proof can be found in [DNR00] (Proposition 1.1.7 and Lemma 1.1.10, sites 5-7). Note that the notation is slightly different there - it is $\Delta_1 := \Delta$ not $\Delta^{[2]} := \Delta$.
This formula is a generalization of coassociativity. It means that $\Delta^{[n]}$ is coproduct where Δ is applied $n-1$ times to any one tensor factor at each stage. Thanks to that we can write

$$\Delta^{[n]}(a) = \sum a_1 \otimes \cdots \otimes a_n$$

with no ambiguity.
Interpretation is an extension of that described in the previous paragraph for $n = 2$. Now we are just decomposing a into n parts and probabilities of outcomes do not depend on which factors we are applying Δ at each stage.

The counit property written in Sweedler notation takes form

$$\begin{aligned}\sum \varepsilon(a_1) \otimes a_2 &= 1 \otimes a,\\ \sum a_1 \otimes \varepsilon(a_2) &= a \otimes 1.\end{aligned}$$

Applying isomorphisms ${}^L m$ and ${}^R m$ from Observation 5. on both sides respectively we get

$$\sum \varepsilon(a_1)a_2 = a,$$
$$\sum a_1\varepsilon(a_2) = a.$$

Remark. A coalgebra C is said to be cocommutative iff for all $c \in C$

$$\sum c_1 \otimes c_2 = \sum c_2 \otimes c_1.$$

2.4 Bialgebras

Definition 3. *A **K-bialgebra** is vector space $\mathcal{H}$ with both an algebra structure $(\mathcal{H}, m, u)$ and a coalgebra structure $(\mathcal{H}, \Delta, \varepsilon)$ such that m, u are morphisms of coalgebras and Δ, ε are morphisms of algebras.*

Explanation. In fact, for a given vector space $\mathcal{H}$ with both an algebra structure $(\mathcal{H}, m, u)$ and a coalgebra structure $(\mathcal{H}, \Delta, \varepsilon)$, the fact that m and u are morphisms of coalgebras is equivalent to the fact that Δ and ε are morphisms of algebras and both are equivalent to conjuction of following conditions:

$$\begin{aligned}
\Delta m &= (m \otimes m)(I \otimes T \otimes I)(\Delta \otimes \Delta), \\
\varepsilon m &= {}^K m(\varepsilon \otimes \varepsilon), \\
\Delta u &= (u \otimes u)^K \Delta, \\
\varepsilon u &= I.
\end{aligned}$$

They can also be written as: for all $g, h \in \mathcal{H}$, all $k \in K$

$$\begin{aligned}
\sum (g \cdot h)_1 \otimes (g \cdot h)_2 &= \sum g_1 \cdot h_1 \otimes g_2 \cdot h_2, \\
\varepsilon(g \cdot h) &= \varepsilon(g)\varepsilon(h), \\
\sum (1_{\mathcal{H}})_1 \otimes (1_{\mathcal{H}})_2 &= 1_{\mathcal{H}} \otimes 1_{\mathcal{H}}, \\
\varepsilon(1_{\mathcal{H}}) &= 1_K.
\end{aligned}$$

or as: for all $g, h \in \mathcal{H}$, all $k \in K$

$$\begin{aligned}
\Delta(g \cdot h) &= \sum g_1 \cdot h_1 \otimes g_2 \cdot h_2, \\
\varepsilon(g \cdot h) &= \varepsilon(g)\varepsilon(h), \\
\Delta(1_{\mathcal{H}}) &= 1_{\mathcal{H}} \otimes 1_{\mathcal{H}}, \\
\varepsilon(1_{\mathcal{H}}) &= 1_K.
\end{aligned}$$

Remark. Note that for the condition $\Delta m = (m \otimes m)(I \otimes T \otimes I)(\Delta \otimes \Delta)$ we need the map $(I \otimes T \otimes I)$, because without it, the right side will be equal to $(m \otimes m)(\Delta \otimes \Delta)$ which, when applied to vector $g \otimes h$ yields $\sum g_1 \cdot g_2 \otimes h_1 \cdot h_2$ not $\sum g_1 \cdot h_1 \otimes g_2 \cdot h_2$ and we want comultiplication and multiplication to be done componentwise. Definition with one T is enough for all powers of m and Δ as stated in the following remark:
Remark. It can be proven by induction that for all ${}^1h, \dots, {}^nh \in \mathcal{H}$

$$\Delta^{[m]} m^{[n]}({}^1h \otimes \dots \otimes {}^nh) = \sum {}^1h_1 \cdot \ldots \cdot {}^nh_1 \otimes \dots \otimes {}^1h_m \cdot \ldots \cdot {}^nh_m. \quad (2.2)$$

Proof. Left to the reader. □

To simplify the notation, we will write a symbol for algebra multiplication also for componentwise multiplication,
so for all ${}^1h_1, \dots, {}^1h_m, \dots, {}^nh_1, \dots, {}^nh_m \in \mathcal{H}$:

$$({}^1h_1 \otimes \dots \otimes {}^1h_m) \cdot \ldots \cdot ({}^nh_1 \otimes \dots \otimes {}^nh_m) := {}^1h_1 \cdot \ldots \cdot {}^nh_1 \otimes \dots \otimes {}^1h_m \cdot \ldots \cdot {}^nh_m. \quad (2.3)$$

Definition 4. *Element b of a bialgebra $\mathcal{B}$ is said to be **primitive** iff*

$$\Delta(b) = 1_{\mathcal{B}} \otimes b + b \otimes 1_{\mathcal{B}}$$

Definition 5. *For a bialgebra $\mathcal{H}$ we define a **Hopf-square** map $\Psi^{[2]} : \mathcal{H} \to \mathcal{H}$ as $\Psi^{[2]} := m\Delta$.*

Comment. It will be a very important function in this paper. It will be this function, that will set a structure of a Markov chain on a Hopf algebra. In Hopf algebras that we will use for modelling Markov chains, the Hopf square map will preserve some of those algebras' (viewed as a vector space) finite dimensional subspaces. Bases of these preserved subspaces can be then treated as spaces of states (aces of spades, haha) of our associated Markov chains. Note that one Hopf algebra can set a structure of many Markov chains, each one having a basis of algebra's finite dimensional subspace preserved by $\Psi^{[2]}$ as its (chains) space of states. What's more, the matrix of $\Psi^{[2]}$ (viewed as a transformation of some fixed, finite-dimensional subspace of algebra) written in the base $\mathcal{B}$ of that subspace will be exactly a transition matrix $K_{i,j}$ of associated Markov chain on that basis. Finding eigenbasis of $K_{i,j}$ is then expressed as finding eigenvectors of $\Psi^{[2]}$. Later it will be put more carefully and precisely.
It will have a natural interpretation as "pulling apart" and then "putting pieces together", for example splitting the deck of cards and then shuffling it.

We also define higher power maps for $n \geqslant 2$:

$$\Psi^{[n]} \coloneqq m^{[n]} \Delta^{[n]}.$$

Hopf-square in Sweedler notation looks like this:

$$\Psi^{[n]}(a) = \sum a_1 \cdot \ldots \cdot a_n.$$

Convolution

Definition 6. *Let (C, Δ, ε) be a coalgebra and (A, M, u) an algebra. On the set $Hom(C, A)$ we define an algebra structure in wich the multiplication, denoted by $*$, is given as follows: if $f, g \in Hom(C, A)$, then*

$$f * g \coloneqq m(f \otimes g)\Delta$$

*we call $*$ the **convolution** product.*

It can be also written as: for any $c \in C$, any $f, g \in Hom(C, A)$

$$(f * g)(c) = \sum f(c_1) \cdot g(c_2)$$

The multiplication defined above is associative, since for $f, g, h \in Hom(C, A)$ and $c \in C$ we have

$$\begin{aligned}((f * g) * h)(c) &= \sum (f * g)(c_1) \cdot h(c_2) \\ &= \sum f(c_1) \cdot g(c_2) \cdot h(c_3) \\ &= \sum f(c_1) \cdot (g * h)(c_2) \\ &= (f * (g * h))(c).\end{aligned}$$

The identity element of the algebra $Hom(C, A)$ is $u\varepsilon \in Hom(C, A)$ since

$$\begin{aligned}(f * u\varepsilon)(c) &= \sum f(c_1) \cdot u\varepsilon(c_2) \\ &= \sum f(c_1) \cdot \varepsilon(c_2) 1_A \\ &= \sum f(c_1)\varepsilon(c_2) \cdot 1_A \\ &= \left(\sum f(c_1)\varepsilon(c_2)\right) \cdot 1_A \\ &= f(c) \cdot 1_A = f(c)\end{aligned}$$

hence $f * u\varepsilon = f$. Similarly, $u\varepsilon * f = f$.
Let us note that if $A = K$, then $*$ is the convolution product defined on the dual algebra of the coalgebra C. This is why in the case A is an arbitrary

algebra we will also call $*$ the convolution product.

For a bialgebra $\mathcal{H}$ we denote $\mathcal{H}^A$, $\mathcal{H}^C$ as, respectively, the underlying algebra and coalgebra structure. We can define algebra structure on $Hom(\mathcal{H}^C, \mathcal{H}^A)$ as above. Note that the identity map $I : \mathcal{H} \to \mathcal{H}$ is an element of $Hom(\mathcal{H}^C, \mathcal{H}^A)$ but it is not the identity element of its algebra structure with convolution product. The $u\varepsilon$ is that identity element.

Definition 7. *Let $\mathcal{H}$ be a bialgebra. A linear map $S \in Hom(\mathcal{H}^C, \mathcal{H}^A)$ is called an* ***antipode*** *of the bialgebra $\mathcal{H}$ if S is the inverse of the identity map $I : \mathcal{H} \to \mathcal{H}$ with respect to the convolution product in $Hom(\mathcal{H}^C, \mathcal{H}^A)$*

The fact that $S \in Hom(\mathcal{H}^C, \mathcal{H}^A)$ is an antipode is written as

$$S * I = I * S = u\varepsilon.$$

and using Sweedler notation as:

$$\forall_{h\in\mathcal{H}} \sum S(h_1) \cdot h_2 = \sum h_1 \cdot S(h_2) = \varepsilon(h) 1_{\mathcal{H}}.$$

2.5 Hopf algebras

Definition 8. *A bialgebra having an antipode is called a* ***Hopf algebra****.*

Definition 9. *A* ***graded bialgebra*** *is a graded vector space $\mathcal{H} = \bigoplus_{i=0}^{\infty} \mathcal{H}_i$ with a bialgebra structure that is compatible with the grading.*

Explanation. A bialgebra structure is compatible with grading iff for all $i, j \in \mathbb{N}$:

$$m[\mathcal{H}_i \otimes \mathcal{H}_j] \subseteq \mathcal{H}_{i+j} \text{ and}$$
$$\Delta[H_n] \subseteq \bigoplus_{i=0}^{n} \mathcal{H}_i \otimes \mathcal{H}_{n-i}.$$

Now decomposition can be viewed as representing an element as the sum of pairs of lower-degree ("smaller") elements.

We can observe that

$$\begin{aligned}\Psi^{[2]}[\mathcal{H}_n] = m\Delta[\mathcal{H}_n] &\subseteq m[\bigoplus_{i=0}^{n} \mathcal{H}_i \otimes \mathcal{H}_{n-i}] \\ &= \bigoplus_{i=0}^{n} m[\mathcal{H}_i \otimes \mathcal{H}_{n-i}] \subseteq \bigoplus_{i=0}^{n} \mathcal{H}_n = \mathcal{H}_n,\end{aligned}$$

hence Hopf square $\Psi^{[2]}$ preserves grading (in the sence that $\Psi^{[2]}[\mathcal{H}_n] \subseteq \mathcal{H}_n$).

Definition 10. *A graded bialgebra* $\mathcal{H} = \bigoplus_{i=0}^{\infty} \mathcal{H}_i$ *is* ***connected*** *iff* $\mathcal{H}_0$ *is one-dimensional subspace spanned by* $1_{\mathcal{H}}$*.*

Explanation. Equivalently we can say that a graded bialgebra $\mathcal{H} = \bigoplus_{i=0}^{\infty} \mathcal{H}_i$ is connected iff $\mathcal{H}_0 = u[K]$ for u - unit in $\mathcal{H}$ treated as a K-algebra.

Theorem 1. *Any graded, connected bialgebra is a Hopf algebra with antipode:*

$$S = \sum_{k \geqslant 0} (u\varepsilon - I)^{*k}.$$

TO DO: MOŻE DAĆ JEDNAK TEN DOWÓD.

b This is the end of our algebraic definitions pfuuuu...

2.6 Examples

2.6.1 Graded, connected Hopf algebra of polinomials

Let P be a vector space of polynomials of one variable over the field K with natural grading by degree. Note that the standard polynomial multiplication is compatible with that grading as for polynomials with degrees i, j, their product has degree $i + j$. Connection comes from that the identity of multiplication is a polynomial of degree 0 ($1_p = X^0$).
P can be enriched with coalgebra structure with comultiplication Δ such that for all $n \in \mathbb{N}$:

$$\Delta(X^n) = \sum_{i=0}^{n} X^i \otimes X^{n-i}.$$

it extends linearly to the rest of P.
Counit is then 0 for all elements with positive degree (degree > 0). Here comes the proof:

Since for all $n \in \mathcal{N}$

$$
\begin{aligned}
(1_P \otimes \varepsilon)\Delta(X^n) &= X^n \otimes 1_K && \text{and} \\
(1_P \otimes \varepsilon)\Delta(X^n) &= \sum_{i=0}^{n} X^i \otimes \varepsilon(X^{n-i}) && \text{and} \\
\sum_{i=0}^{n} X^i \otimes \varepsilon(X^{n-i}) &= X^n \otimes \varepsilon(1_P) && + \sum_{i=0}^{n-1} X^i \otimes \varepsilon(X^{n-i}) \\
&= X^n \otimes 1_K && + \sum_{i=0}^{n-1} X^i \otimes \varepsilon(X^{n-i})
\end{aligned}
$$

we have that for all $n \in \mathbb{N}$

$$\sum_{i=0}^{n-1} X^i \otimes \varepsilon(X^{n-i}) = 0$$

but we also have that

$$\sum_{i=0}^{n-1} X^i \otimes \varepsilon(X^{n-i}) = \sum_{i=0}^{n-1} \varepsilon(X^{n-i})X^i \otimes 1_K = \left(\sum_{i=0}^{n-1} \varepsilon(X^{n-i})X^i\right) \otimes 1_K$$

Because $X_0, \ldots, X_{n-1}$ are linearly independent we have that $\forall_{0 \leqslant i \leqslant n-1}\ \varepsilon(X^{n-i}) = 0$. Keeping in mind that n was arbitrary, we have that for all $n \geqslant 1$ $\varepsilon(X^n) = 0$ and then by linearity of ε, that for every polynomial $p \in P$ with a positive degree we have that $\varepsilon(p) = 0$.
We can now check that P with that structure is a graded, connected Hopf algebra that is both commutative and cocommutative.
It is a bialebra, because:

$$a$$

TO DO: BAM! DO ROBOTY!

2.7 Graded, connected Hopf algebra of non-commuting variables

2.7.1 Free associative Hopf algebra

This is the main example of our interest. It will be used to describe inverse and forward riffle shuffling.

Let K be a field with characteristic 0. Let $\mathcal{X} = \{x_1, \ldots, x_N\}$ be a finite set. For every $n \in \mathbb{N}$ let $\mathcal{H}_n$ be a vector space having as a basis all words of length n made of elements of $\mathcal{X}$. (The basis of $\mathcal{H}_0$ is a singleton of an empty word). Let $\mathcal{H} \coloneqq \bigoplus_{i=0}^{\infty} \mathcal{H}_i$. Hence the basis of $\mathcal{H}$ is $\mathcal{X}^*$ - all finite words over an alphabet $\mathcal{X}$. Let $m : \mathcal{H} \otimes \mathcal{H} \to \mathcal{H}$ be the concatenation of words, that is, for all $s_1, s_2 \in \mathcal{X}^*$

$$m(s_1 \otimes s_2) \coloneqq s_1 s_2.$$

Let $\Delta : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$ be defined for all elements from $\mathcal{X}$ as

$$\Delta(x_i) = x_i \otimes 1_{\mathcal{H}} + 1_{\mathcal{H}} \otimes x_i.$$

and extends linearly and multiplically .
The unit is then $u : K \to \mathcal{H}$ such that

$$u(1_K) = \varepsilon$$

where ε is an empty word. And indeed $1_{\mathcal{H}} = \varepsilon$. **Lemma.** Then $\mathcal{H}$ is a graded, connected Hopf algebra that is cocommutative.

Proof. Associativity of m and coassociativity of Δ are obvious. Actions fit together, because we define them so. Algebra is graded straight from definition and connected because an empty word is an identity element in respect of concatenation multiplication. Cocomutativity can be checked immediately. $\square$

Let $s = x_{i_0} \ldots x_{i_k} \in \mathcal{X}^*$. What is not so obvious is how $\Delta(x_{i_0} \ldots x_{i_k})$ looks like:

$$\begin{aligned}
\Delta(x_{i_0} \ldots x_{i_k}) &= \Delta m^{[k]}(x_{i_0} \otimes \cdots \otimes x_{i_k}) && (2.4)\\
&= (m^{[k]} \otimes m^{[k]}) \left(\sum (x_{i_0})_1 \otimes \cdots \otimes (x_{i_k})_1 \otimes (x_{i_0})_2 \otimes \cdots \otimes (x_{i_k})_2\right) && (2.5)\\
&= \sum (x_{i_0})_1 \ldots (x_{i_k})_1 \otimes (x_{i_0})_2 \ldots (x_{i_k})_2. && (2.6)
\end{aligned}$$

It may be unclear what this sum really is. It is taken over all possible combinations of all "possible values" of $(x_{i_j})_1$ and $(x_{i_j})_2$ for $0 \leqslant j \leqslant k$. We can recall that for all $x_i \in \mathcal{X}$ we have $\Delta(x_i) = x_i \otimes 1_{\mathcal{H}} + 1_{\mathcal{H}} \otimes x_i$. Writing that in Sweedler notation gives

$$\sum (x_i)_1 \otimes (x_i)_2 = x_i \otimes 1_{\mathcal{H}} + 1_{\mathcal{H}} \otimes x_i.$$

The sum we are discussing is then the sum over all possible partitions into two distinct subsequences of s, because for each component of that sum, for

each x_{i_j} we decide if we are taking it into the left subsequence (x_{i_j} as a "value" of $(x_{i_j})_1$ and $1_{\mathcal{H}}$ as a "value" of $(x_{i_j})_2$) or into the right subsequence ($1_{\mathcal{H}}$ as a "value" of $(x_{i_j})_1$ and x_{i_j} as a "value" of $(x_{i_j})_2$).
To denote it, let's write $s_1 \prec s$ for "s_1 is a subsequence of s" (a subsequence doesn't have to be a contiguous fragment) and for s_1, s such that $s_1 \prec s$, let $s_2 = s/s_1$ denote $s_2 \prec s$ created by removing s_1 from s. We can now write the sum from (2.5) as:

$$\sum (x_{i_0})_1 \dots (x_{i_k})_1 \otimes (x_{i_0})_2 \dots (x_{i_k})_2 = \sum_{\substack{s_1 \prec s \\ s_2 = s/s_1}} s_1 \otimes s_2.$$

Equivalently (and that expression can be found in [DPR14]) it can be written as

$$\sum_{S \subseteq \{i_0, \dots i_k\}} \prod_{j \in S} x_j \otimes \prod_{j \notin S} x_j.$$

where S is a multiset, because some of the $i_0, \dots, i_k$ can be the same.

This structure will describe the inverse riffle shuffling, as Δ will be an operation of randomly dividing a stack of cards into two stacks by putting each card with probability $\frac{1}{2}$ to the left or to the right and m will be an operation of deterministic putting the left stack on the top of the right stack. $\Psi^{[2]}$ will be then application of one iteration of inverse riffle shuffle.

2.7.2 Some futher remarks about structure

In paragraph 2.3 [DPR14] describes some aspects of the structure of free associative algebra. They will be important in the chapter about eigenbasises. Here we will present a shortened version for lookup.
GR89 shows that symmetrized sums of certain primitive elements form a basis of a free associative algebra. It will turn out that this will be the left eigenbasis of $m\Delta$. Here will be introduced concepts useful for construction of that basis. Explanation why this is an eigenbasis comes in Chapter 4.

Definition 11. *A word in an ordered alphabet is* ***Lyndon****, if it is strictly smaller (in lexicographical order) than all its cyclic rearrangements.*

Definition 12. *A* ***Lyndon factorization*** *of word w is a tuple of words $(l_1, l_2, \dots, l_k)$ such that $w = l_1 l_2 \dots l_k$, each l_i is a Lyndon word and $l_1 \geqslant l_2 \geqslant \dots \geqslant l_k$.*

Fact. **[Lot97, Th. 5.1.5]** Every word w has a unique Lyndon factorisation.

Definition 13. *For a Lyndon word l that has at least two letters a* ***standard factorisation*** *of l is a pair of words (l_1, l_2) such that $l = l_1 l_2$, both l_i are non-trivial (non-empty) Lyndon words and l_2 is the longest right Lyndon factor of l. A* ***standard factorisation*** *of a single letter word is that letter.*

Fact. Each Lyndon word l has a standard factorization.

Definition 14. *For a Lyndon word l a* ***standard bracketing $\boldsymbol{\lambda(l)}$*** *of l is defined recursively as $\lambda(a) := a$ for a letter and $\lambda(l) := [\lambda(l_1), \lambda(l_2)]$, where (l_1, l_2) is a standard factorisation of l. $[x, y] = x \cdot y - y \cdot x$ for every words x, y.*

Definition 15. *The* ***symmetrized product*** *of word w is*

$$\mathrm{sym}(w) = \sum_{\sigma \in S_k} \lambda(l_{\sigma(1)}) \cdot \ldots \cdot \lambda(l_{\sigma(k)}),$$

where $(l_1, \ldots, l_k)$ is a Lyndon factorization of w.

[GR89, Th. 5.2] shows that $\{\mathrm{sym}(w) : w \in \mathcal{X}^*\}$ form a basis for free associative algebra.
Let $|w|$ be the length of word w. For a word $w = a_1 \ldots a_{|w|}$ and permutation $\sigma \in S_{|w|}$ let $\sigma(w) := a_{\sigma(1)} \ldots a_{\sigma(|w|)}$.
Let $\simeq_{\mathrm{sym}}$ be a relation on $\mathcal{X}^* \times \mathcal{X}^*$ such that for all $w, v \in \mathcal{X}^*$

$$w \simeq_{\mathrm{sym}} v \iff \exists_{\sigma \in S_{|w|}} \, \sigma(w) = v$$

Observation. $\simeq_{\mathrm{sym}}$ is an equivalence relation on $\mathcal{X}^*$.

Proof. Obvious. ☐

Now we can provide a much finer grading.
With every $\nu \in \mathcal{X}^*$ we associate

$$\mathcal{H}_\nu := \mathrm{Lin}(\{w \in \mathcal{X}^* : w \simeq_{\mathrm{sym}} \nu\}). \tag{2.7}$$

So it is the subspace spanned by words that for each letter from $\mathcal{X}$ have the same number of instances of that letter as ν. (Of course for every $w, v \in \mathcal{X}^*$ such that $w \simeq_{\mathrm{sym}} v$ we have $\mathcal{H}_w = \mathcal{H}_v$.)
Now we can write $\mathcal{H}$ as

$$\mathcal{H} = \bigoplus_{[\nu]_{\simeq_{\mathrm{sym}}} \in \mathcal{X}^*_{/\simeq_{\mathrm{sym}}}} \mathcal{H}_\nu$$

Which is equivalent to

$$\mathcal{H} = \bigoplus_{S \in \mathcal{X}^*_{/\simeq_{\mathrm{sym}}}} \mathrm{Lin}(S)$$

This grading is also compatible with a bialgebra structure we have introduced in the sense that for all $\nu, \upsilon \in \mathcal{X}^*$

$$m[\mathcal{H}_\nu \otimes \mathcal{H}_\upsilon] \subseteq \mathcal{H}_{\nu\upsilon} \text{ and}$$
$$\Delta[\mathcal{H}_\nu] \subseteq \bigoplus_{\substack{s_1 \prec \nu \\ s_2 = \nu/s_1}} \mathcal{H}_{s_1} \otimes \mathcal{H}_{s_2}.$$

We can observe that

TO DO: pokrywa całość

This will be the grading we will be using for our probabilistic interpretation. It i

2.7.3 Alternative structure To do

Now we will describe an alternative graded and connected Hopf algebra structure on $\mathcal{H}$ - a vector space spanned by finite words over the fixed alphabet $\mathcal{X}$. It will describe the structure of forward riffle shuffle. We will denote that alternative Hopf algebra structure built on $\mathcal{H}$ as $\mathcal{H}^*$ and call it a graded dual of $\mathcal{H}^*$ (for reasons that will come later). (Note that $\mathcal{H}$ is isomorphic to $\mathcal{H}^*$ as a vector space and $\mathcal{H}^*$ in this sense is not the vector space dual to $\mathcal{H}$).
We define multiplication $\Delta^* : \mathcal{H}^* \otimes \mathcal{H}^* \to \mathcal{H}^*$ as for all $s_1, s_2 \in \mathcal{X}^*$:

$$\Delta^*(s_1 \otimes s_2) = \sum\{s : s_1 \prec s \text{ and } s_2 = s/s_1\}.$$

which is the sum of all possible interlaces of s_1 and s_2
and comultiplication $m^* : \mathcal{H}^* \to \mathcal{H}^* \otimes \mathcal{H}^*$ as for all $s \in \mathcal{X}^*$:

$$m^*(s) = \sum\{s_1 \otimes s_2 : s = s_1 s_2\}$$

which is the sum of all possible divisions of s into its prefix and suffix.
and both extended linearly.
Lemma. Then $\mathcal{H}^*$ is a graded, connected Hopf algebra that is commutative.

Proof. Associativity of Δ^* and coassociativity of m^* are obvious. Now we will prove that actions fits together which means that

$$m^*\Delta^* = (\Delta^* \otimes \Delta^*)(I \otimes T \otimes I)(m^* \otimes m^*)$$

Let $g, h \in \mathcal{X}^*$, $m^*\Delta^*(g \otimes h)$ and $(\Delta^* \otimes \Delta^*)(I \otimes T \otimes I)(m^* \otimes m^*)(g \otimes h)$ are sums of terms of shape $x \otimes y$. We make every letter in g and h different by giving them specific labels. We will show that then every term has a coefficient one in that sums, and then, that these terms are the same. Putting labels down will result in summation of some terms but as they are the same with the labels, they will be the same without them.

Each term in $m^*\Delta^*$ case corresponds to a pair of: possible interlace of g and h, and then, a possible division of its outcome. We want to show that for every term occurring in that sum, there is only one pair of interlace and division that leads to that term. Hence all terms will have a coefficient 1_K. Indeed - if the interlaces are different it means that at least two letters are in different order. After division they either will be in the different words (which points out the difference) or in one word in different order (which points out the difference too). If interlaces are the same divisions also must be the same to create a specific pair of words.

Each term in $(\Delta^* \otimes \Delta^*)(I \otimes T \otimes I)(m^* \otimes m^*)$ case corresponds to a pair of pairs: two divisions - one of g and one of h, and then, two interlaces - one interlace of created prefixes of g and h and one interlace of created suffixes of g and h. There is only one pair of pairs of divisions and interlaces that leads to a specific term. If at least one division is different it will lead to a words containing letters with different labels. If divisions are the same and at least one interlace is different it will lead to word with different order.

Now we will show that terms in $m^*\Delta^*(g \otimes h)$ and $(\Delta^* \otimes \Delta^*)(I \otimes T \otimes I)(m^* \otimes m^*)(g \otimes h)$ are the same.
We will do it by showing that for every pair of interlace and division (from $m^*\Delta^*$) there exist one pair of pairs of interlaces and divsions (from $(\Delta^* \otimes \Delta^*)(I \otimes T \otimes I)(m^* \otimes m^*)$) that leads to the same term and that for every pair of pairs of interlaces and divsions (from $(\Delta^* \otimes \Delta^*)(I \otimes T \otimes I)(m^* \otimes m^*)$) there exist one pair of interlace and division (from $m^*\Delta^*$) that leads to the same term.
The pair of interlace and division from $m^*\Delta^*$ generates divisions of g and h as "that letters that went to the prefix" and "that letters that went to the suffix" and interlaces of that prefixes and suffixes of g and h that are primal interlace restricted to a part of word.
Having pair of pairs of divisions and interlaces from $(\Delta^* \otimes \Delta^*)(I \otimes T \otimes I)(m^* \otimes m^*)$ we can construct a corresponding interlace of $g \otimes h$ by making that two interlaces at once. Then the division can be done such that restricted to word g and word h is the same as one of the original pair.
Other properties of bialgebra are easy to check. Algebra is connected because an empty word is still an identity element with respect to multiplication. Commutativity is clear. □

In the shuffling interpretation Δ^* will be the operation of dividing a stack of cards at some random point and putting the top pile on the left creating two stacks. m^* will be the operation of combining two stacks together with the same probability of every possible interlace of two stacks.

2.7.4 Graded dual

Now we will see that there is another method of introducing that structure. The structure of $\bigoplus_{i=0}^{\infty} \mathcal{H}_i^*$ (where for all $i \in \mathbb{N}$ $\mathcal{H}_i^*$ is a vector space dual to $\mathcal{H}$) with actions induced by actions from Hopf algebra $\mathcal{H}$ turns out to be one discrebed above. (Note that $\mathcal{H} = \bigoplus_{i=0}^{\infty} \mathcal{H}_i$ is isomorphic as a linear space to $\bigoplus_{i=0}^{\infty} \mathcal{H}_i^*$.)

Let $\mathcal{H}^{\mathrm{gd}*}$ denote $\bigoplus_{i=0}^{\infty} \mathcal{H}_i^*$. We define multiplication $\Delta^* : \mathcal{H}^{\mathrm{gd}*} \otimes \mathcal{H}^{\mathrm{gd}*} \to \mathcal{H}^{\mathrm{gd}*}$ and comultiplication $m^* : \mathcal{H}^{\mathrm{gd}*} \to \mathcal{H}^{\mathrm{gd}*} \otimes \mathcal{H}^{\mathrm{gd}*}$ as (for all $h_1^*, h_2^*, h^* \in \mathcal{H}^{\mathrm{gd}*}$):

$$\Delta^*(h_1^* \otimes h_2^*) = (h_1^* \otimes h_2^*)\Delta,$$
$$m^*(h^*) = h^* m.$$

TO DO: COŚTAM

Chapter 3

Connection

In the chapter one we said that with every Markov chain have corresponding matrix of transition probabilities. Thus with a Markov chain one can associate a certain linear transformation given by that matrix. The following theorem states dependency in opposite direction - that given linear transfomation with some features gives a Markov chain with the basis of transformed vector space as a state space.

Reminder. For a vector space V with basis $\mathcal{B}$, $b \in \mathcal{B}$ let b^* denote a linear functional such that $b^*(b) = 1$ and $\forall_{b' \in \mathcal{B} \setminus \{b\}}\ b^*(b') = 0$.

Theorem 2. *Let V be a linear space over field K that is $\mathbb{R}$ or $\mathbb{Q}$, with basis $\mathcal{B}$. Let $\psi : V \to V$ be a linear operation such that for all $b \in \mathcal{B}$ coefficients of vector $\psi(b)$ written in $\mathcal{B}$ are $\geqslant 0$ and their sum s is the same. ($\forall_{b_1, b_2 \in \mathbb{B}}\ b_1^* \psi(b_2) \geqslant 0 \wedge \exists_{s \in K}\ \forall_{b \in \mathcal{B}} \sum_{b_i \in \mathcal{B}} b_i^* \psi(b) = s$). Let $x_0 \in V \setminus \{0\}$ and $n_0 \in \mathbb{N}$, $\alpha_1, \ldots, \alpha_{n_0} \in K$, $b'_1, \ldots, b'_{n_0} \in \mathcal{B}$ be such that $x_0 = \sum_{i=0}^{n} \alpha_i b'_i$. Let $s_0 \coloneqq \sum_{i=1}^{n} \alpha_i$. ($s_0 = \sum_{b \in \mathcal{B}} b^*(x_0)$). Let $(\mathcal{B}^\omega, \mathbb{P})$ be a probabilistic space. For all $i \in \mathbb{N}$ let $X_i : \mathcal{B}^\omega \to \mathcal{B}$ be a projection on the i-th co-ordinate. Let $\mathbb{P}$ be such that for every $n \in \mathbb{N}$, every $b_0, \ldots, b_n \in \mathcal{B}$:*

$$\mathbb{P}(X_0 = b_0, X_1 = b_1, \ldots, X_n = b_n) = \frac{b_0^*(x_0)}{s_0} \prod_{i=0}^{n-1} \frac{b_{i+1}^* \psi(b_i)}{s}.$$

Then $(X_0, X_1, \ldots)$ is a Markov chain with state space $\mathcal{B}$ in which, for all $b_1, b_2 \in \mathcal{B}$, the probability of going from b_1 to b_2 is equal to the coefficient standing by b_2 in $\psi(b_1)$, written in $\mathcal{B}$, divided by s.

Proof. We have defined a measure $\mathbb{P}$ on basic open subsets of $\mathcal{B}^\omega$. Definition is valid, because for all $n \in \mathbb{N}$, all $b_0, \ldots, b_n \in \mathcal{B}$:

$$\sum_{b \in \mathcal{B}} \mathbb{P}(X_0 = b_0, X_1 = b_1, \ldots, X_n = b_n, X_{n+1} = b) = \sum_{b \in \mathcal{B}} \frac{b_0^*(x_0)}{s_0} \prod_{i=0}^{n-1} \frac{b_{i+1}^* \psi(b_i)}{s} \frac{b^* \psi(b_n)}{s} =$$

$$\frac{b_0^*(x_0)}{s_0} \prod_{i=0}^{n-1} \frac{b_{i+1}^* \psi(b_i)}{s} \sum \frac{b^* \psi(b_n)}{s} = \mathbb{P}(X_0 = b_0, X_1 = b_1, \ldots, X_n = b_n) \frac{\sum_{b \in \mathcal{B}} b^* \psi(b_n)}{s} =$$

$$\mathbb{P}(X_0 = b_0, X_1 = b_1, \ldots, X_n = b_n) \frac{s}{s} = \mathbb{P}(X_0 = b_0, X_1 = b_1, \ldots, X_n = b_n).$$

Now we will show that $(X_0, X_1, \ldots)$ is indeed a Markov. We will check the (1.1) property from Chapter 1.
Let $b_0, \ldots, b_m \in \mathbb{B}$ be a sequence such that

$$\mathbb{P}(X_0 = b_0, \ldots, X_m = b_m) > 0$$

We have:

$$\mathbb{P}(X_m = b_m \mid X_0 = b_0, \ldots, X_{m-1} = b_{m-1}) = \frac{\mathbb{P}(X_0 = b_0, X_1 = b_1, \ldots, X_{m-1} = b_{m-1}, X_m = b_m)}{\mathbb{P}(X_0 = b_0, \ldots, X_{m-1} = b_{m-1})} =$$

$$\frac{\frac{b_0^* \psi(x_0)}{s_0} \prod_{i=0}^{m-1} \frac{b_{i+1}^* \psi(b_i)}{s}}{\frac{b_0^* \psi(x_0)}{s_0} \prod_{i=0}^{m-2} \frac{b_{i+1}^* \psi(b_i)}{s}} = \frac{b_m^* \psi(b_{m-1})}{s}.$$

And on the other hand:

$$\mathbb{P}(X_m = b_m \mid X_{m-1} = b_{m-1}) = \frac{\mathbb{P}(X_{m-1} = b_{m-1}, X_m = b_m)}{\mathbb{P}(X_{m-1} = b_{m-1})} =$$

$$\frac{\sum\limits_{(c_0,\ldots,c_{m-2})\in\mathcal{B}^{m-1}} \mathbb{P}(X_0 = c_0, \ldots, X_{m-2} = c_{m-2}, X_{m-1} = b_{m-1}, X_m = b_m)}{\sum\limits_{(c_0,\ldots,c_{m-2})\in\mathcal{B}^{m-1}} \mathbb{P}(X_0 = c_0, \ldots, X_{m-2} = c_{m-2}, X_{m-1} = c_{m-1})} =$$

$$\frac{\sum\limits_{(c_0,\ldots,c_{m-2})\in\mathcal{B}^{m-1}} \frac{b_0^*\psi(x_0)}{s_0} \left(\prod\limits_{i=0}^{m-3} \frac{c_{i+1}^*\psi(c_i)}{s}\right) \frac{b_{m-1}^*\psi(c_{m-2})}{s} \frac{b_m^*\psi(b_{m-1})}{s}}{\sum\limits_{(c_0,\ldots,c_{m-2})\in\mathcal{B}^{m-1}} \left(\prod\limits_{i=0}^{m-3} \frac{c_{i+1}^*\psi(c_i)}{s}\right) \frac{b_{m-1}^*\psi(c_{m-2})}{s}} =$$

$$\frac{\frac{b_m^*\psi(b_{m-1})}{s} \sum\limits_{(c_0,\ldots,c_{m-2})\in\mathcal{B}^{m-1}} \frac{b_0^*\psi(x_0)}{s_0} \left(\prod\limits_{i=0}^{m-3} \frac{c_{i+1}^*\psi(c_i)}{s}\right) \frac{b_{m-1}^*\psi(c_{m-2})}{s}}{\sum\limits_{(c_0,\ldots,c_{m-2})\in\mathcal{B}^{m-1}} \left(\prod\limits_{i=0}^{m-3} \frac{c_{i+1}^*\psi(c_i)}{s}\right) \frac{b_{m-1}^*\psi(c_{m-2})}{s}} = \frac{b_m^*\psi(b_{m-1})}{s}$$

and the chain indeed has claimed transition probability, because for all $b_1, b_2 \in \mathcal{B}$, $\frac{b_2^*\psi(b_1)}{s}$ is a coefficient standing by b_2 in $\psi(b_1)$, written in $\mathcal{B}$, divided by s. □

To do Cocommutative Hopf algebra of non-commuting variables is a model for inverse riffle shuffling and commutative Hopf algebra of non-cocommutative variables is a model for forward riffle shuffling. It goes as follows. Let $\mathcal{X}$ be the finite set of all possible types of cards. Let ν be a tuple of elements from $\mathcal{X}$ that represents our actual deck of cards (the same type of card can occur multiple times, the order in ν should be the order in which we think cards are ordered). Now we can take a look at a subspace $\mathcal{H}_\nu$ of vector space $\mathcal{H}$ built over $\mathcal{X}$ as described in 2.6.2. $\mathcal{H}_\nu$ will be the subspaces spanned by words that for every type of cards consists of exactly the same number of cards of that type as ν. So the basis of $\mathcal{H}_\nu$ will be the set of words for every arrangement of our deck of cards. Let's name this basis $\mathcal{B}_\nu$. Note that then $\mathcal{H}_\nu$ is finite dimensional. As was previously described, there are two ways of equipping $\mathcal{H}$ with Hopf algebra structure. One will correspond to inverse version of riffle shuffle and another one to the forward one. With given arrangement of cards $s \in \mathcal{B}_\nu$ applying $m\Delta$ to s yields the sum of possible outcomes after one inverse riffle shuffle while applying $\Delta^* m^*$ yields the same for forward riffle shuffle. In both cases coefficients (after normalization) are probabilities of corresponding outcomes.

TO DO: How does it set?

Let's take a non-commuting variables algebra from its example. Let's take $\mathcal{H}_\nu$ for some $\nu \in \mathcal{X}^*$. Then $\Psi^{[2]}$ sets the Markov chain of inverse riffle shuffle the deck of cards containing cards labeled by xs appearing in ν. Chains state space is then the basis of $\mathcal{H}_\nu$. Let's call it $\mathcal{B}_\nu$ Chains transition matrix is equal to transition matrix of $\Psi^{[2]}$ written in $\mathcal{B}_\nu$.

Here will come an intuition why tensor products and Hopf algebras suits for describing probabilistic issues.

Let $\mathcal{X} = \{x_1, \dots x_N\}$ be our set of all possible types of cards.
We will denote a stack of k cards containing (from top to bottom) $x_{i_1}, \dots, x_{i_k}$ simply as $x_{i_1} \dots x_{i_k}$.
Imagine, that you have stack of cards $x_{i_1} \dots x_{i_k}$. After shuffling it you can get one of finitely many stack of cards each with certain probability. We want to have some representation of it in our structure. For that reason we span a vector space $\mathcal{H}$, over $\mathbb{Q}$ (but can be $\mathbb{R}$ if someone likes), with basis $\mathcal{X}^*$ (finite words over $\mathcal{X}$, which means "all possible stacks of cards of types from $\mathcal{X}$ including an empty stack").
For all $s_1, \dots, s_n \in \mathcal{X}^*$, all $0 \leqslant q_1, \dots, q_n \in K$ a non-zero vector $\sum_{i=1}^{n} q_i s_i$ is for all $i \in \{1, \dots, n\}$ interpreted as a state where we have a stack s_i with probability $\frac{q_i}{\sum_{i=1}^n q_i}$ or equivalently as a probabilistic measure on $\mathcal{X}^*$ with value $\frac{q_i}{\sum_{i=1}^n q_i}$ on s_i for every $i \in \{1, \dots, n\}$ and 0 elsewhere.
In that understanding the "+" can be read as "or".
We want also desribe a situation when: we have multiple stacks of cards on a table (some of them maybe empty), there are only finitely many options how these stacks can exactly look like and we know a probability of every option.
It is very natural situation during shuffling as when we for example split a stack of cards at some random point (with known probabilities of where the split can be) we for sure have two stacks of cards (as soon as we agree that one of them can be empty), there are only finetely many options how exactly arrangement looks like and we know a probability of each one.

We will now focus on case when we have two decks on a table.
We want to deal with that matter in similar way as we done for setting "probabilistic options" to one deck of cards. We will span a vector space with all possible arrangements of two decks as a basis. That vector space will be $\mathcal{H} \otimes \mathcal{H}$. Now we will try to give some explanation why in fact this is quite intuitive.
For $s_1, s_2 \in \mathcal{X}^*$ let's denote (s_1, s_2) as having s_1 on the left stack and s_2 on the right stack.
Let's make an observation that for all $s, s_1, s_2 \in \mathcal{X}^*$ situation of having arrangement (s_1, s) with probability p and having arrangement (s_2, s) with probability $1-p$ is the same situation as having s_1 with probability p or having s_2 with probability $1-p$ on the left stack and for sure having s on the right stack. Making connection with our previously introduced notation so we want to $p(s_1, s) + (1-p)(s_2, s) = (ps_1 + (1-p)s_2, s)$ (and analogously to the second coordinate).
What is more, having for sure s_1 on the left and s_2 on the right with probability p (and with probability $(1-p)$ some else arragement, let's call it (z_1, z_2)) gives the same probability distribution on possible arrangments of two decks as, having s_1 on the left and having s_2 on the right, with probability p and with probability $1-p$ having (z_1, z_2)).
This leads us to conclusion, that we also want to $p(s_1, s_2) = (ps_1, s_2)$ (and analogously to the second coordinaate).
In the Gilbert-Shannon-Reeds model of inverse riffle shuffling there are two steps. First we are decomposing the deck by take cards from the top of deck - one after another and putting them to the left or to the right each with probability $\frac{1}{2}$. Secondly putting left stack on the right stack.
That pulling apart causes a split into two stacks, each of them can be any subset of original stack (with preservation of order) with equal probability of each option.
For $s_1, s \in \mathcal{X}^*$ let denote that s_1 is subsequence of s (a subset with preservation of order) as $s_1 \prec s_2$. Let's denote a stack arisen from removing form s its subsequence s_1 as s/s_1.
Let denote that pulling apart as a $\Delta : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$, then for all $s \in \mathcal{X}^*$ it will give

$$\Delta(s) = \sum_{\substack{s_1 \prec s \\ \wedge s_2 = s/s_1}} s_1 \otimes s_2.$$

For putting two piles back together by placing left on the top let us write a linear map $m : \mathcal{H} \otimes \mathcal{H} \to \mathcal{H}$ that is concatenation, which means, that for all $s_1, s_2 \in \mathcal{X}^*$

$$m(s_1 \otimes s_2) = s_1 s_2.$$

What we just define here is exactly an algebra of non-commuting variables from example 2.3.2.
Facts about its algebraic nature are presented in that section.
We can observe now that Hopf-square map $\Psi^{[2]} = m\Delta$ for Δ, m defined as above describes one iteration of the inverse riffle shuffle. For every $s \in \mathcal{X}^*$, $\Psi^{[2]}(s)$ is a sum of possible arrangements of stack with corresponding probabilities (without normalisation)).
Ta- daaaam!

But where are that Markov chain? Where are these "subspaces preserved by Ψ"?
For a fixed deck of n cards $\nu = (\nu_1, \dots, \nu_n) \in \mathcal{X}^n$ the Markov chain of shuffling that deck is set by $\Psi^{[2]}$ restricted to the subspace spanned by $S_\nu =$ "all $s \in \mathcal{X}^*$ that are some rearangement of ν", more formally: spanned by S_ν, where:

$$S_\nu = \{s = x_{i_1} \dots x_{i_n} \in \mathcal{X}^* \mid \exists_{\sigma \in S_n} x_{\sigma(i_1)} \dots x_{\sigma(i_n)} = \nu_1 \dots \nu_n.\}.$$

($\sigma \in S_n$ is a permutation, S_n is a symmetric group of n (group of all permutations of n elements)). Its equivalent to that $S_\nu = [\nu]_{\simeq_{\mathrm{sym}}}$.
Then the state space of that chain is S_ν. The transition matrix of that chain is exactly a matrix of $\Psi^{[2]}$ truncated to $\mathcal{H}_\nu := \mathrm{Lin}(S_\nu)$ (which, as we can observe is finite-dimensional and preserved by $\Psi^{[2]}$).

For forward riffle shuffle we will be working with the same space $\mathcal{H}$ (as we still are dealing with the same set of types of cards) but with differnt actions (as operations of "pulling apart" and "putting together" look now different).
We will proove that indeed forward riffle shuffle $(F_i)_{i \geqslant 0}$ and inversed riffle shuffle $(I_i)_{i \geqslant 0}$ are the same shuffling method but once aplicated "forward" and once "backward". What we mean is that for fixed deck of n cards $\nu = (\nu_1, \dots, \nu_n) \in \mathcal{X}^n$, for all $s_1, s_2 \in \mathcal{H}_\nu$, all $n \geqslant 0$:

$$\mathbb{P}\{F_{n+1} = s_2 \mid F_n = s_1\} = \mathbb{P}\{I_{n+1} = s_1 \mid I_n = s_2\}.$$

Which means that probability of going from state s_1 to state s_2 in one step in forward riffle shuffle is qual to probability of going form s_2 to s_1 in one step in inverse riffle shuffle.
As remarked in Section 1. forward riffle shuffle can be defined as cutting the deck at some point with uniform distribution on "where" ($n + 1$ options for a deck of size n) and then putting back two piles together in the way that *everyone-had-seen-at-some-point-in-the-life* (*trrrrrrrr*) with the same probability of every possible "*trrrrrrrr*".
Let us denote $\mathcal{H}$ as a dual to $\mathcal{H}$ what we want to do is to see how induced

multiplication and comultiplication look like.
"here it come".
It is forward fiffle sfufle, we can check it that corresponds to that , that to that fold product is exactly that and that, so the coeficient matches and that is ok.
And then it is exactly an riffle shuffle as it is bla bla.
Let $\Delta_F : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$ be an linear map for decomposition of the deck for forward riffle shuffle, then for all $s \in \mathcal{X}^*$

$$\Delta_F(s) = \sum_{s_1 s_2 = s} s_1 \otimes s_2.$$

Let $m_F : \mathcal{H} \otimes \mathcal{H} \to \mathcal{H}$ be a linear map coresponding to *trrrrrrrr*. Then for all $s_1, s_2 \in \mathcal{X}^*$:

$$m_F(s_1 \otimes s_2) = \sum_{\substack{s_1 \prec s \\ s_2 = s/s_1}} s$$

which is sum of all possible entanglements of s_1 and s_2.

Now let us consider a vector space that is dual to $\mathcal{H}$. It is vector space of linear functions on $\mathcal{H}$ whith basis bla bla $1s^* : \mathcal{H} \to K$ such that for all $s \in \mathcal{X}$ We can define multiplication and comultiplication of $\mathcal{H}^*$ in the natural way. bla bla ble ble

We ckech, and what? That are exactly m_F and Δ_F matrixes of $(F_i)_{i \geqslant 0}$ and $(I_i)_{i \geqslant 0}$ are transpositions of each other.

TO DO: Do poprawki, bo żal

Chapter 4

Left and right eigenbasises

The reasons we bother with finding the eigenbasis are desribed in 1.?.?. For a given deck ν we will find left and right eigenbasises of $\mathcal{H}_\nu$ for inverse and forward riffle shuffling. Note that because $m\Delta$ and $\Delta^* m^*$ are dual to each other left eigenbasis for inverse riffle-shuffle is right eigenbasis for forward riffle-shuffle and right eigenbasis for inverse is left eigenbasis for forward.

4.1 Left eigenbasis

Construction of left eigenbasis begins with a general observation that for every Hopf algebra, primitive elements are eigenvectors of $\Psi^{[a]}$ for every $a \in \mathbb{N}$.

Observation 6. *For every primitive element $h \in \mathcal{H}$, for every $a \in \mathbb{N}$ it holds that h is an eigenvector of $\Psi^{[a]}$ with an eigenvalue a.*

Proof. Let $h \in \mathcal{H}$ be a primitive element.

$$\Psi^{[a]}(h) =$$
$$m^{[a]}\Delta^{[a]}(h) =$$
$$m^{[a]}(\sum h_1 \otimes \cdots \otimes h_i) =$$
$$m^{[a]}(\underbrace{\overbrace{h \otimes 1_{\mathcal{H}} \otimes \cdots \otimes 1_{\mathcal{H}}}^{i \text{ factors}} + \overbrace{1_{\mathcal{H}} \otimes h \otimes \cdots \otimes 1_{\mathcal{H}}}^{a \text{ factors}} + \cdots + \overbrace{1_{\mathcal{H}} \otimes 1_{\mathcal{H}} \otimes \cdots \otimes h}^{a \text{ factors}}}_{a \text{ summands}}) =$$
$$\underbrace{m^{[a]}(\overbrace{h \otimes 1_{\mathcal{H}} \otimes \cdots \otimes 1_{\mathcal{H}}}^{a \text{ factors}}) + m^{[a]}(\overbrace{1_{\mathcal{H}} \otimes h \otimes \cdots \otimes 1_{\mathcal{H}}}^{a \text{ factors}}) + \cdots + m^{[a]}(\overbrace{1_{\mathcal{H}} \otimes 1_{\mathcal{H}} \otimes \cdots \otimes h}^{a \text{ factors}})}_{a \text{ summands}} =$$
$$\underbrace{h + \cdots + h}_{a \text{ summands}} =$$
$$ah$$

□

Next observation is presented in [DPR14] and it is called the symmeterization lemma.

Theorem 3. *(Symmetrization lemma). Let $x_1, \ldots, x_n$ be primitive elements of any Hopf alebra $\mathcal{H}$, then $\sum_{\sigma \in S_k} x_{\sigma(1)} \cdot \ldots \cdot x_{\sigma(k)}$ is an eigenvector of $\Psi^{[a]}$ with eigenvalue a^k.*

Now for the basis introduced at the end of **2.6.2** to be eigenbasis we only need to check that for every Lyndon word l, element $\lambda(l)$ is primitive.

Lemma 1. *For every x, y that are primitive, $[x, y]$ is primitive.*

Proof. Let x, y be primitive elements of a bialgebra $\mathcal{H}$.

$$\begin{gathered}
\Delta([x,y]) = \\
\Delta(x \cdot y - y \cdot x) = \\
\Delta(x \cdot y) - \Delta(y \cdot x) = \\
\Delta m(x \otimes y) - \Delta m(y \otimes x) = \\
\sum (x_1 \otimes x_2) \cdot (y_1 \otimes y_2) - \sum (y_1 \otimes y_2) \cdot (x_1 \otimes x_2) = \\
(x \otimes 1_{\mathcal{H}}) \cdot (y \otimes 1_{\mathcal{H}}) + (x \otimes 1_{\mathcal{H}}) \cdot (1_{\mathcal{H}} \otimes y) + (1_{\mathcal{H}} \otimes x) \cdot (y \otimes 1_{\mathcal{H}}) + (1_{\mathcal{H}} \otimes x) \cdot (1_{\mathcal{H}} \otimes y) + \\
-\Big((y \otimes 1_{\mathcal{H}}) \cdot (x \otimes 1_{\mathcal{H}}) + (y \otimes 1_{\mathcal{H}}) \cdot (1_{\mathcal{H}} \otimes x) + (1_{\mathcal{H}} \otimes y) \cdot (x \otimes 1_{\mathcal{H}}) + (1_{\mathcal{H}} \otimes y) \cdot (1_{\mathcal{H}} \otimes x) \Big) = \\
x \cdot y \otimes 1_{\mathcal{H}} + x \otimes y + y \otimes x + 1_{\mathcal{H}} \otimes x \cdot y - \Big(y \cdot x \otimes 1_{\mathcal{H}} + y \otimes x + x \otimes y + 1_{\mathcal{H}} \otimes y \cdot x \Big) = \\
(x \cdot y - y \cdot x) \otimes 1_{\mathcal{H}} + 1_{\mathcal{H}} \otimes (x \cdot y - y \cdot x) = \\
[x,y] \otimes 1_{\mathcal{H}} + 1_{\mathcal{H}} \otimes [x,y].
\end{gathered}$$

□

Because of Lemma 1. and the fact that single-letter word is a primitive element we have that $\lambda(l)$ is primitive for every Lyndon word l. Hence

$$\mathcal{E} \coloneqq \{\mathrm{sym}(w) : w \in \mathcal{X}^*\}$$

is a left eigenbasis of $\mathcal{H}$ with respect to $\Psi^{[a]}$ for every a. What is more:

Theorem 4. *Let $\mathcal{H}$ be a free associative Hopf algebra of words over the alphabet $\mathcal{X}$. Let $\nu \in \mathcal{X}^*$. Then $\mathcal{E}_\nu \coloneqq \{\mathrm{sym}(w) : w \in \mathcal{B}_\nu\}$ is an eigenbasis of $\mathcal{H}_\nu$ with respect to $\Psi^{[a]}$ for every a.*

To prove it, we will need a following lemma:
Lemma. $w \in \mathcal{H}_\nu$ iff $\mathrm{sym}(w) \in \mathcal{H}_\nu$

Proof. We can observe that for every $x, y \in \mathcal{X}^*$ there holds $xy \simeq_{\mathrm{sym}} yx$. Because of that for every Lyndon word l we have that $\lambda(l) \in \mathcal{H}_l$. So we have that for every word w with Lyndon factorisation $(l_1, \ldots, l_k)$ there holds $\lambda(l_1) \cdot \ldots \cdot \lambda(l_k) \in \mathcal{H}_w$. From that, for every $\sigma \in S_k$, we have that $\lambda(l_{\sigma(1)}) \cdot \ldots \cdot \lambda(l_{\sigma(k)}) \in \mathcal{H}_w$.

□

Proof. (of the theorem)
$\mathcal{E}_\nu \subseteq \mathcal{E}$, so $\mathcal{E}_\nu$ is linearly independent. $w \in \mathcal{H}_\nu$ iff $\mathrm{sym}(w) \in \mathcal{H}_\nu$ so $\mathcal{E} \cap \mathcal{H}_\nu = \mathcal{E}_\nu$, so $\mathcal{E}_\nu$ spans the $\mathcal{H}_\nu$. □

4.2 Right eigenbasis

The right eigenbasis can be obtained as stated in the following theorm:

Theorem 5. *a*

4.3 Reference to [DPR14]

Now we will recall two theorems from [DPR14] describing eigenbases.

Theorem 6. *bla bla eigenbasismn*

Theorem 7. *bla bla dual eigen basis*

To prove them we will need a simmetrization lemma
bla bla

TO DO: Remark of no primitive generators in other algebras

Now we can make some further observations about shuffling.
ble ble

Chapter 5

Summation

Bibliography

[DNR00] Sorin Dascalescu, Constantin Nastasescu, and Serban Raianu. *Hopf algebras: an introduction*. Pure and Applied Mathematics. CRC Press, 1 edition, 2000.

[DPR14] Persi Diaconis, C. Y. Amy Pang, and Arun Ram. Hopf algebras and markov chains: two examples and a theory. *Journal of Algebraic Combinatorics*, 39(3):527–585, May 2014.

[LPW17] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov Chains and Mixing Times*. AMS MBK #107. American Mathematical Society, 2017.