Uniwersytet Wrocławski
Wydział Matematyki i Informatyki
Instytut Matematyczny
specjalność teoretyczna

Bartosz Sójka

Explanation of connection between Hopf algebras and Markov chains

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

Wrocław 2018

Abstract

In [DPR14] Persi Diaconis, Amy Pang and Arun Ram described how to use Hopf algebras for study Markov chains. As it involves ideas from quite different branches of mathematics it could be hard to grasp a concept 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 accesible for third year students who have passed probability and abstract algebra courses. I will focus on the example of shuffling cards by inverse riffle shuffle method. Structure will be as follows: firstly there will be introduction to both Hopf algebras and Markov chains, then there will be explanation 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.

1 Markov chains

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

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

following condition (called Markov property) holds:

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

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 $s_{i_0}, \dots, s_{i_{n-1}}$ were visited before.
For the Markov chain $(X_0, X_1, \dots)$ 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 the $K^n_{i,j}$ 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 arragments of the cards in the deck with $K(x, y)$ equal to probability of going from arragment x to arragment y in one shuffle.
More extensive indroduction can be found in [LPW17].
Cośtam cośtam stationary distirution.

2 Hopf algebras

2.1 Tensor products

First we will introduce tensor produkt 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 satisfaing:
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 $k \in K$

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

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

$$
\begin{aligned}
[z_1]_{/\simeq} + [z_2]_{/\simeq} &:= [z_1 + z_2]_{/\simeq}, \\
k[z_1]_{/\simeq} &:= [kz_1]_{/\simeq}.
\end{aligned}
$$

We denote equivalece class $[(v, w)]_{/\simeq}$ as $v \otimes w$. The tensor product $V \otimes W := Z/_{\simeq}$.
Note, that in $V \otimes W$ there are vectors that can not be writen as $v \otimes w$ for any v, w. However every $z \in V \otimes W$ can be writen 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 lineary 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 writen as $\sum_{i=1}^{n} v_i \otimes w_i$.
Now we will make some futher observations on how $V \otimes W$ looks like.

Observation 1. *If $\{b_i\}_{i \in I}$, $\{c_j\}_{j \in J}$ are basises 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 arbitraly 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 $v_1, \dots v_n$ or $w_1, \dots, w_n$ have to be lineary dependent. It can't occur if $v_1, \dots v_n$ and $w_1, \dots, w_n$ are from the basises of V and W.
This observation also justifies recently cited fact and the example. □

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

Proof. The proof is imediate from the Observation 1.. Since if $\{b_i\}_{i \in I}$, $\{c_j\}_{j \in J}$ are basises 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 recall of the definition. □

Observation 4. *For vector spaces U, V, W over the field K there is an 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. (Also we will 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 described above (in 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 isomophisms ${}^{L}m : K \otimes V \to V$, (${}^{R}m : V \otimes K \to V$) 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 bijection 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. □

Remark. In a special case when $V = W = K$ the natural isomorphisms descripted 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 in understanding what tensor product is and what is not. It will be good to keep it in mind when we will be intensively dealing with it in a combinatorical way in the following sections.

2.2 Hopf algebras

Now there will be 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 believe that thanks to that it will be a better reference.
If reader will feel lost in this section it is recommended to read it in parralell 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 refference), is that for most of the time we will not be using full structure of a Hopf algebra. Nethertheless it is good to see the full shape of what we are dealing with.
So now will come full definition nonetheless we will try to explain it piece by piece.

Remark. Let K be a field. In following section k, if not steted otherwise, will denotes an arbitrary element from 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 will 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 a vector spaces over field K. We 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 a n-tensor power $\overbrace{V \otimes \cdots \otimes V}^{n\ times}$ of a vector space V we will sometimes write $V^{\otimes n}$.

Throught 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 $(\varphi\circ\psi)(x)$.

2.2.1 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 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 an 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 mutiplication. 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 lineary. Given that we can observe that for all $a \in \mathcal{H}$, all $k \in K$, a multiplicated by $u(k)$ (no matter if form 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 recurently define the sequence of maps $(m^{[n]})_{n\geqslant 2}$, such that $m^{[n]} : \underbrace{A\otimes\cdots\otimes A}_{n\ times} \to A$ as follows:

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

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

2.2.2 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 $a \in \mathcal{H}$*

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

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

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

means exactly what was in the definition of u.

Explanation. We will introduce a notation called Sweedler notation [Swe69] which will be very usefull 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 sens. For example, later it will come out that exactly the comultiplication will be the operation that will model the process of cutting the deck of cards into pieces in riffle shuffle. In examples that we will working with (graded, connected Hopf algebras) comultiplication will represent some kind of natural decomposition in the more general way. What does it mean in the strict sense will be presented in Definition 5. when we will be introducing graded Hopf Algebras.
Examples.

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

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

or in more expand form as

$$\forall_{a \in \mathcal{H}} \sum a_{11} \otimes a_{12} \otimes a_2 = \sum a_1 \otimes a_{21} \otimes a_{22}. \tag{2}$$

Because of these equalities, terms from (2) can be writen 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 represent decomposition such that, when did twice, probabilities of possible outcomes are the same no matter which set of parts (a_1 or a_2) have been tooked 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 recurently define the 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 $n \geqslant 3, i \in \{1, \dots, n-2\}, m \in \{0, \dots, n-i-1\}$ 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 there notatnion is slightly differnet - it is $\Delta_1 := \Delta$ not $\Delta^{[2]} := \Delta$. This formula is a generalization of a 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 to n parts and probabilieties of outcomes do not depend on which factors we are applying Δ at each stage.

The counit property writen 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 on both sides isomorphisms ${}^{L}m$ and ${}^{R}m$ from Observation 5. respectively we get

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

2.2.3 Bialgebras

Definition 3. *A **K-bialgebra** is a 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 morfisms of coalgebras and Δ, ε are morfisms 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 morfisms of coalgebras is equivalent to that Δ and ε are morfisms of algebras and both are equivalent to conjuction of following contditions:

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

They can be writen also 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 $k \in K$

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

Informal 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, right site will be equal to $(m \otimes m)(\Delta \otimes \Delta)$ which, when applied on 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 compultiplication and multiplication to be done componentwise. Definition with one T is enough for all powers of m and Δ as state in following remark:
Remark. It can be prooven by induction that for all ${}^{1}h, \ldots, {}^{n}h \in \mathcal{H}$

$$\Delta^{[m]}m^{[n]}({}^{1}h \otimes \cdots \otimes {}^{n}h) = \sum {}^{1}h_1 \cdot \ldots \cdot {}^{n}h_1 \otimes \cdots \otimes {}^{1}h_m \cdot \ldots \cdot {}^{n}h_m. \quad (3)$$

Proof. Left to the reader. □

Definition 4. *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 very important function in this paper. It will be it what will set a structure of a Markov chain on a Hopf algebra. In Hopf algebras that we will be using for modeling Markov chains the Hopf square map will preserve some of those algebras (viewed as a vector space) finall dimentional subspaces. Basises 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 algebras finite dimentional subspace preserved by $\Psi^{[2]}$ as its (chains) space of states. Whats more matrix of $\Psi^{[2]}$ (viewed as a trasformation of some fixed, finite-dimentional subspace of algebra) writen in a base $\mathcal{B}$ of that subspace will be exactly a transition matrix $K_{i,j}$ of associated Markov chain on that bases. 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 exaple split the deck of cards and then shuffling it.

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

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

Hopf-square in sweedler notation looks like this:

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

Definition 5. *Let (C, Δ, ε) be a coalgebra and (A, M, u) an algebra. We define on the set $Hom(C, A)$ an algebra structure in with the multiplication, denoted by $*$ is given as follows: if $f, g \in Hom(C, A)$, then*

$$f * g := m(f \otimes g)\Delta$$

*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

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

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, respectivly, the underlying algebra and coalgebra structure. We can define as above algebra structure on $Hom(\mathcal{H}^C, \mathcal{H}^A)$. Note, that 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 6. *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:

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

2.2.4 Hopf algebras

Definition 7. *A bialgebra having an antipode is called a* ***Hopf algebra****.*

Definition 8. *A* ***graded Hopf algebra*** *is vector space $\mathcal{H} = \bigoplus_{i=0}^{\infty} \mathcal{H}_i$ with a Hopf algebra structure that is compatible with the grading.*

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

$$m[\mathcal{H}_i \otimes \mathcal{H}_j] \subseteq \mathcal{H}_{i+j} \text{ and}$$

$$\Delta[H_n] \subseteq \bigoplus_{i=0}^{n} \mathcal{H}_i \otimes \mathcal{H}_{n-i}.$$

Hopf square preserve grading

Definition 9. *A graded Hopf algebra $\mathcal{H} = \bigoplus_{i=0}^{\infty} \mathcal{H}_i$ is* ***connected*** *iff $\mathcal{H}_0$ is one-dimentional subspace spanned by $1_{\mathcal{H}}$.*

As an curiosity we will say that any graded, connected bialgebra is a Hopf algebra with antipode

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

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

2.3 Examples

2.3.1 Graded, connected Hopf algebra of polinomials

2.3.2 Graded, connected Hopf algebra of non-commuting variables

This is a main example of our interest.
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 H_n be a vector space having as a basis all words of length n made of elements of $\mathcal{X}$. Let $\mathcal{H} := \bigoplus_{i=0}^{\infty} \mathcal{H}_i$. Let $m : \mathcal{H} \otimes \mathcal{H} \to \mathcal{H}$ be concatenation of words, that is, for all $k, l \in \mathbb{N}$ for all $x_{i_0}, \ldots, x_{i_k}, x_{j_0}, \ldots, x_{j_l} \in \mathcal{X}$

$$m(x_{i_0} \ldots x_{i_k} \otimes x_{j_0} \ldots x_{j_l}) := x_{i_0} \ldots x_{i_k} x_{j_0} \ldots x_{j_l}.$$

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 lineary and multiplically
Lemma. Then $\mathcal{H}$ is the a graded, connected Hopf algebra that is cocomutative.

Proof. Associativity of m and coassociativity of Δ are obvious. Actions fit together, because we define them so. Algebra is graded and connnected because it is. $\square$

We can see $\Delta(x_{i_0} \ldots x_{i_k})$ how looks like:

$$\begin{aligned}
\Delta(x_{i_0} \ldots x_{i_k}) &= \Delta m^{[k]}(x_{i_0} \otimes \cdots \otimes x_{i_k}) \\
&= (m^{[k]} \otimes m^{[k]}) \left(\sum (x_{i_0})_1 \otimes \cdots \otimes (x_{i_k})_1 \otimes (x_{i_0})_2 \otimes \cdots \otimes (x_{i_k})_2 \right) \\
&= \sum (x_{i_0})_1 \ldots (x_{i_k})_1 \otimes (x_{i_0})_2 \ldots (x_{i_k})_2 \\
&= \sum_{S \subseteq \{i_0, \ldots i_k\}} \prod_{j \in S} x_j \otimes \prod_{j \notin S} x_j.
\end{aligned}$$

where S is a multiset, because some of the $i_0, \dot{i}_k$ can be the same. Form of that coproduct is like that because when for all $x_i \in \mathcal{X}$ we have $\Delta(x_i) = x_i \otimes 1_{\mathcal{H}} + 1_{\mathcal{H}} \otimes x_i$ then

3 Explanation of connection between Hopf algebras and Markov chains

Let take bla bla from eample ble ble.
Now we will show how free associative algebra from example 2.3.2 set the structure for inverse riffle shuffle.

- dopisać comutative i cocomutative
- dopisać przykład polinomial i ciała do coalgebry
- dopisąć grupowy do Hopfa
- pokazać jak wygląda coproduct w noncommuting - ważne
- wprowadzić dualną do noncommuting - ważne
- w rozdziale 3 wyjaśnić
- non-commuting - Hopf square zachowuje skończone podprzestrzenie.
- do łańcuch ów markowa dopisacć dokładniejsy opis gilbert shannon reeds - jakie sa prawdopodobieństwa oraz że forward można rozumieć na dwa równoważne sposoby. To, że są do siebie odwrotne (dualne) będzie wyprowadzone przy uzyciu algebry).
- finer grading

We will be working of an examples of riffle shuffle and inverse riffle shuffle cards shuffling as our Markov chains.
How we will put them in the algebraic way?
First we will do this with inverse riffle shuffle, the forward riffle shuffle will then apear in a natural way.
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 denote it somehow. For that reason we spann a vector space $\mathcal{H}$ over some field K of characteristic 0 with basis $\mathcal{X}^*$ (finite words over $\mathcal{X}$, which means "all possible stacks of cards of types from $\mathcal{X}$ inlcluding an empty stack").
For all $s_1, \dots, s_n \in \mathcal{X}^*$, all $k_1, \dots, k_n \in K$ a non-zero vector $\sum_{i=1}^{n} k_i s_i$ is for all $i \in \{1, \dots, n\}$ interpreted as a state where we have a stack s_i with probability $\frac{k_i}{\sum_{i=1}^{n} k_i}$ or equvalently as a probabilistic measure on $\mathcal{X}^*$ with value $\frac{k_i}{\sum_{i=1}^{n} k_i}$ on s_i for every $i \in \{1, \dots, n\}$ and 0 elesewhere.
In that undestanding the "+" can be readed as "or".
We want also desribe a situation when we have multiple stacks of cards on a table, we percive it as a certain amount of stacks (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 of what is on our table.
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 shure 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 arragment looks like and we know a probability of eachone.
We will now focus on case when we have two decks on a table. At some point it will start to looks familiar.

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 arregements of two decks as a basis. Now we will think about how addition and scalar multiplication should look like there.
For $s_1, s_2 \in \mathcal{X}^*$ lets 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 arragement (s_1, s) with probability p and having arragement (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 annalogly to the second coordinate).
What is more note that 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) gives the same probabilty distribution on possible arragments of two decks as having s_1 on the left with probability p, having s_2 on the right for sure and with probability $1-p$ having z).
This leads us to conclusion, that we also want to $p(s_1, s_2) = (ps_1, s_2)$ (and annalogly to the second coordinaate).
At this point we can start to think that it is convinient to interprete a space of probabilistic measures on possible arragements of two stacks of cards as a $\mathcal{H} \otimes \mathcal{H}$.

int the Gilbert-Shannon-Reeds model of inverse riffle shuffling there are two steps. Firstly we are decomposing the deck by take cards from the top of deck deck - one after another and puting them to the left or to the right each with probability $frac12$. 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 we 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

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

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 proven 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 arregements of stack with coresponding probabilities. Ta- daaaam!

But where are that Markov chain? Where are these "subspaces preserved by Ψ"?
For an fixed deck of n cards $\nu = (\nu_1, \ldots, \nu_n) \in \mathcal{X}^n$ 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 rearagement of ν", more formally: spanned by S_ν, where:

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

($\sigma \in S_n$ is a permutation, S_n is a symmetric group of n (group of all permutations of n elements))
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 $\mathrm{Lin}(S_\nu)$ (which, as we can observe is finite-dimentional 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 and inverse riffle shuffle are the same shuffling method but once aplicated "forward" and once "backward". What we mean is that for all moving from

References

[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.