

HOPF ALGEBRAS AND MARKOV CHAINS

(REVISED VERSION OF)

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF MATHEMATICS

AND THE COMMITTEE ON GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Chung Yin Amy Pang

Version of December 30, 2014

Because of time constraints, I did NOT submit this version to Stanford. This version differs from the submitted version in that the chapters are in a different order, and there are additional results. I prefer that you cite this version (<http://arxiv.org/abs/1412.8221>), or one of the related papers. Consult the table on my webpage for which papers contain which sections of this thesis.

Unless there is a major mathematical error, the version of this thesis on arXiv will not be updated. However, I aim to keep an updated version on my webpage, so please alert me to typos and confusing parts. Below is the list of major changes since the arXiv version; minor typographical corrections are not listed.

Abstract

This thesis introduces a way to build Markov chains out of Hopf algebras. The transition matrix of a *Hopf-power Markov chain* is (the transpose of) the matrix of the coproduct-then-product operator on a *combinatorial Hopf algebra* with respect to a suitable basis. These chains describe the breaking-then-recombining of the combinatorial objects in the Hopf algebra. The motivating example is the famous Gilbert-Shannon-Reeds model of riffle-shuffling of a deck of cards, which arises in this manner from the shuffle algebra.

The primary reason for constructing Hopf-power Markov chains, or for rephrasing familiar chains through this lens, is that much information about them comes simply from translating well-known facts on the underlying Hopf algebra. For example, there is an explicit formula for the stationary distribution (Theorem 4.5.1), and constructing quotient algebras show that certain statistics on a Hopf-power Markov chain are themselves Markov chains (Theorem 4.7.1). Perhaps the pinnacle is Theorem 2.5.1, a collection of algorithms for a full left and right eigenbasis in many common cases where the underlying Hopf algebra is commutative or cocommutative. This arises from a cocktail of the Poincare-Birkhoff-Witt theorem, the Cartier-Milnor-Moore theorem, Reutenauer's structure theory of the free Lie algebra, and Patras's Eulerian idempotent theory.

Since Hopf-power Markov chains can exhibit very different behaviour depending on the structure of the underlying Hopf algebra and its distinguished basis, one must restrict attention to certain styles of Hopf algebras in order to obtain stronger results. This thesis will focus respectively on a free-commutative basis, which produces "independent breaking" chains, and a cofree basis; there will be both general statements and in-depth examples.

Acknowledgement

First, I must thank my advisor Persi Diaconis for his patient guidance throughout my time at Stanford. You let me roam free on a landscape of algebra, probability and combinatorics in whichever direction I choose, yet are always ready with a host of ideas the moment I feel lost. Thanks in addition for putting me in touch with many algebraic combinatorialists.

The feedback from my thesis committee - Dan Bump, Tom Church, Eric Marberg - was invaluable to improving both the mathematics and the exposition in this thesis. Thanks also for a very enjoyable discussion at the defense - I doubt I'd have another chance to take an hour-long conversation solely on my work.

Thanks to my coauthor Arun Ram for your observations about card-shuffling and the shuffle algebra, from which grew the theory in this thesis. I also greatly appreciate your help in improving my mathematical writing while we worked on our paper together.

I'm grateful to Marcelo Aguiar for pointing me to Patras's work, which underlies a key part of this thesis, and for introducing me to many of the combinatorial Hopf algebras I describe here.

To the algebraic combinatorics community: thanks for taking me in as a part of your family at FPSAC and other conferences. Special mentions go to Sami Assaf and Aaron Lauve for spreading the word about my work; it makes a big difference to know that you are as enthusiastic as I am about my little project.

I'd like to thank my friends, both at Stanford and around the world, for mathematical insights, practical tips, and just cheerful banter. You've very much brightened my journey in the last five years. Thanks especially to Agnes and Janet for taking time from your busy schedules to hear me out during my rough times; because of you, I now have a more positive look on life.

Finally, my deepest thanks go to my parents. You are always ready to share my excitement at my latest result, and my frustrations at another failed attempt, even though I'm expressing it in way too much technical language. Thank you for always being there for me and supporting my every decision.

Contents

Abstract	iv
Acknowledgement	v
1 Introduction	1
1.1 Markov chains	1
1.2 Hopf algebras	6
1.3 Hopf-power Markov chains	7
2 Diagonalisation of the Hopf-power map	11
2.1 The Hopf-power Map	12
2.2 The Eulerian Idempotent	14
2.3 Eigenvectors of Higher Eigenvalue	16
2.4 Lyndon Words	18
2.5 Algorithms for a Full Eigenbasis	20
2.6 Basis for the Eigenspace of Largest Eigenvalue	27
3 Markov chains from linear operators	32
3.1 Construction	33
3.2 Diagonalisation	35
3.3 Stationarity and Reversibility	36
3.4 Projection	39

4	Construction and Basic Properties of Hopf-power Markov Chains	42
4.1	Combinatorial Hopf algebras	44
4.1.1	Species-with-Restrictions	48
4.1.2	Representation rings of Towers of Algebras	50
4.1.3	Subalgebras of Power Series	52
4.2	First Definition of a Hopf-power Markov Chain	57
4.3	General Definition of a Hopf-power Markov Chain	59
4.4	Description of a Hopf-power Markov chain	65
4.5	Stationary Distributions	68
4.6	Reversibility	71
4.7	Projection	72
5	Hopf-power Markov chains on Free-Commutative Bases	77
5.1	General Results	79
5.1.1	Independence	79
5.1.2	Unidirectionality	81
5.1.3	Probability Estimates from Eigenfunctions	84
5.1.4	Probability Estimates from Quasisymmetric Functions	96
5.2	Rock-Breaking	99
5.2.1	Constructing the Chain	100
5.2.2	Right Eigenfunctions	102
5.2.3	Left Eigenfunctions	105
5.2.4	Transition Matrix and Eigenfunctions when $n = 4$	107
5.3	Tree-Pruning	108
5.3.1	The Connes-Kreimer Hopf algebra	109
5.3.2	Constructing the Chain	113
5.3.3	Right Eigenfunctions	118
6	Hopf-power Markov Chains on Cofree Commutative Algebras	127
6.1	Riffle-Shuffling	127
6.1.1	Right Eigenfunctions	128
6.1.2	Left Eigenfunctions	135

6.1.3	Duality of Eigenfunctions	137
6.2	Descent Sets under Riffle-Shuffling	139
6.2.1	Notation regarding compositions and descents	142
6.2.2	Quasisymmetric Functions and Noncommutative Symmetric Functions	143
6.2.3	The Hopf-power Markov chain on $QSym$	146
6.2.4	Right Eigenfunctions Corresponding to Partitions	148
6.2.5	A full Basis of Right Eigenfunctions	152
6.2.6	Left Eigenfunctions Corresponding to Partitions	154
6.2.7	A full Basis of Left Eigenfunctions	157
6.2.8	Duality of Eigenfunctions	159
6.2.9	Transition Matrix and Eigenfunctions when $n = 4$	162

Bibliography		165
---------------------	--	------------

List of Tables

1.1 Corresponding sections of [DPR14] and the present thesis 10

List of Figures

4.1	An example coproduct calculation in \mathcal{G} , the Hopf algebra of graphs	49
5.1	The “two triangles with one common vertex” graph	87
5.2	The tree $[\bullet Q_3]$	110
5.3	Coproduct of $[\bullet Q_3]$	112
6.1	The trees $T_{(13245)}$ and $T_{(1122)}$	130
6.2	The Lyndon hedgerow $T_{(35142)}$	131

Chapter 1

Introduction

Sections 1.1 and 1.2 briefly summarise, respectively, the basics of the two worlds that this thesis bridges, namely Markov chains and Hopf algebras. Section 1.3 introduces the motivating example of riffle-shuffling of a deck of cards, and outlines the main themes in the thesis.

1.1 Markov chains

A friendly introduction to this topic is Part I of the textbook [LPW09].

A (discrete time) Markov chain is a simple model of the evolution of an object over time. The key assumption is that the state X_m of the object at time m only depends on X_{m-1} , its state one timestep prior, and not on earlier states. Writing $P\{A|B\}$ for the probability of the event A given the event B , this *Markov property* translates to

$$P\{X_m = x_m | X_0 = x_0, X_1 = x_1, \dots, X_{m-1} = x_{m-1}\} = P\{X_m = x_m | X_{m-1} = x_{m-1}\}.$$

Consequently,

$$\begin{aligned} & P\{X_0 = x_0, X_1 = x_1, \dots, X_m = x_m\} \\ &= P\{X_0 = x_0\}P\{X_1 = x_1 | X_0 = x_0\} \dots P\{X_m = x_m | X_{m-1} = x_{m-1}\}. \end{aligned}$$

The set of all possible values of the X_m is the *state space* - in this thesis, this will be a finite set, and will be denoted S or \mathcal{B} , as it will typically be the basis of a vector space.

All Markov chains in this thesis are time-invariant, so $P\{X_m = y | X_{m-1} = x\} = P\{X_1 = y | X_0 = x\}$. Thus a chain is completely specified by its *transition matrix*

$$K(x, y) := P\{X_1 = y | X_0 = x\}.$$

It is clear that $K(x, y) \geq 0$ for all $x, y \in S$, and $\sum_{y \in S} K(x, y) = 1$ for each $x \in S$. Conversely, any matrix K satisfying these two conditions defines a Markov chain. So this thesis will use the term “transition matrix” for any matrix with all entries non-negative and all row sums equal to 1. (A common equivalent term is *stochastic matrix*).

Note that

$$\begin{aligned} P\{X_2 = y | X_0 = x\} &= \sum_{z \in S} P\{X_2 = y | X_1 = z\} P\{X_1 = z | X_0 = x\} \\ &= \sum_{z \in S} K(z, y) K(x, z) = K^2(x, y); \end{aligned}$$

similarly, $K^m(x, y) = P\{X_m = y | X_0 = x\}$ - the powers of the transition matrix contain the transition probabilities after many steps.

Example 1.1.1. The process of card-shuffling is a Markov chain: the order of the cards after m shuffles depends only on their order just before the last shuffle, not on the orders prior to that. The state space is the $n!$ possible orderings of the deck, where n is the number of cards in the deck.

The most well-known model for card-shuffling, studied in numerous ways over the last 25 years, is due to Gilbert, Shannon and Reeds (GSR): first, cut the deck binomially (i.e. take i cards off the top of an n -card deck with probability $2^{-n} \binom{n}{i}$), then drop one by one the bottommost card from one of the two piles, chosen with probability proportional to the current pile size. Equivalently, all interleavings of the two piles which keep cards from the same pile in the same relative order are equally likely. This has been experimentally tested to be an accurate model of how the average person shuffles. Section 6.1 is devoted to this example, and contains references to the history and extensive literature.

After many shuffles, the deck is almost equally likely to be in any order. This is a common phenomenon for Markov chains: under mild conditions, the probability of being in state x after m steps tends to a limit $\pi(x)$ as $m \rightarrow \infty$. These limiting probabilities must satisfy $\sum_x \pi(x)K(x, y) = \pi(y)$, and any probability distribution satisfying this equation is known as a *stationary distribution*. With further mild assumptions (see [LPW09, Prop. 1.14]), $\pi(x)$ also describes the proportion of time the chain spends in state x .

The purpose of shuffling is to put the cards into a random order, in other words, to choose from all orderings of cards with equal probability. Similarly, Markov chains are often used as “random object generators”: thanks to the Markov property, running a Markov chain is a computationally efficient way to sample from π . Indeed, there are schemes such as Metropolis [LPW09, Chap. 3] for constructing Markov chains to converge to a desired stationary distribution. For these sampling applications, it is essential to know roughly how many steps to run the chain. The standard way to measure this rigorously is to equip the set of probability distributions on S with a metric, such as *total variation* or *separation distance*, and find a function $m(\varepsilon)$ for which $\|K^m(x_0, \cdot) - \pi(\cdot)\| < \varepsilon$. Such convergence rate bounds are outside the scope of this thesis, which simply views this as motivation for studying high powers of the transition matrix.

One way to investigate high powers of a matrix is through its spectral information.

Definition 1.1.2. Let $\{X_m\}$ be a Markov chain on the state space S with transition matrix K . Then

- A function $\mathbf{g} : S \rightarrow \mathbb{R}$ is a *left eigenfunction* of the chain $\{X_m\}$ of eigenvalue β if $\sum_{x \in S} \mathbf{g}(x)K(x, y) = \beta \mathbf{g}(y)$ for each $y \in S$.
- A function $\mathbf{f} : S \rightarrow \mathbb{R}$ is a *right eigenfunction* of the chain $\{X_m\}$ of eigenvalue β if $\sum_{y \in S} K(x, y)\mathbf{f}(y) = \beta \mathbf{f}(x)$ for each $x \in S$.

(It may be useful to think of \mathbf{g} as a row vector, and \mathbf{f} as a column vector.) Observe that a stationary distribution π is a left eigenfunction of eigenvalue 1. [DPR14, Sec. 2.1] lists many applications of both left and right eigenfunctions, of which two feature in this thesis. Chapter 5 and Section 6.1 employ their Use A: the expected value of a right eigenfunction

\mathbf{f} with eigenvalue β is

$$E\{\mathbf{f}(X_m)|X_0 = x_0\} := \sum_{s \in S} K^m(x_0, s)\mathbf{f}(s) = \beta^m \mathbf{f}(x_0).$$

The Proposition below records this, together with two simple corollaries.

Proposition 1.1.3 (Expectation estimates from right eigenfunctions). *Let $\{X_m\}$ be a Markov chain with state space S , and \mathbf{f}_i some right eigenfunctions with eigenvalue β_i .*

(i) *For each \mathbf{f}_i ,*

$$E\{\mathbf{f}_i(X_m)|X_0 = x_0\} = \beta_i^m \mathbf{f}_i(x_0).$$

(ii) *Suppose $\mathbf{f}: S \rightarrow \mathbb{R}$ is such that, for each $x \in S$,*

$$\sum_i \alpha_i \mathbf{f}_i(x) \leq \mathbf{f}(x) \leq \sum_i \alpha'_i \mathbf{f}_i(x)$$

for some non-negative constants α_i, α'_i . Then

$$\sum_i \alpha_i \beta_i^m \mathbf{f}_i(x_0) \leq E\{\mathbf{f}(X_m)|X_0 = x_0\} \leq \sum_i \alpha'_i \beta_i^m \mathbf{f}_i(x_0).$$

(iii) *Let S' be a subset of the state space S . Suppose the right eigenfunction \mathbf{f}_i is non-negative on S' and zero on $S \setminus S'$. Then*

$$\frac{\beta_i^m \mathbf{f}_i(x_0)}{\max_{s \in S'} \mathbf{f}_i(s)} \leq P\{X_m \in S' | X_0 = x_0\} \leq \frac{\beta_i^m \mathbf{f}_i(x_0)}{\min_{s \in S'} \mathbf{f}_i(s)}.$$

Proof. Part i is immediate from the definition of right eigenfunction. Part ii follows from the linearity of expectations. To see Part iii, specialise to $\mathbf{f} = \mathbf{1}_{S'}$, the indicator function of being in S' . Then it is true that

$$\frac{\mathbf{f}_i(x)}{\max_{s \in S'} \mathbf{f}_i(s)} \leq \mathbf{1}_{S'}(x) \leq \frac{\mathbf{f}_i(x)}{\min_{s \in S'} \mathbf{f}_i(s)}$$

and the expected value of an indicator function is the probability of the associated event. \square

A modification of [DPR14, Sec. 2.1, Use H] occurs in Corollary 6.2.18. Here is the basic, original version:

Proposition 1.1.4. *Let K be the transition matrix of a Markov chain $\{X_m\}$, and let $\{\mathbf{f}_i\}$, $\{\mathbf{g}_i\}$ be dual bases of right and left eigenfunctions for $\{X_m\}$ - that is, $\sum_j \mathbf{f}_i(j)\mathbf{g}_{i'}(j) = 0$ if $i \neq i'$, and $\sum_j \mathbf{f}_i(j)\mathbf{g}_i(j) = 1$. Write β_i for the common eigenvalue of \mathbf{f}_i and \mathbf{g}_i . Then*

$$P\{X_m = y | X_0 = x\} = K^m(x, y) = \sum_i \beta_i^m \mathbf{f}_i(x) \mathbf{g}_i(y).$$

Proof. Let D be the diagonal matrix of eigenvalues (so $D(i, i) = \beta_i$). Put the right eigenfunctions \mathbf{f}_j as columns into a matrix F (so $F(i, j) = \mathbf{f}_j(i)$), and the left eigenfunctions \mathbf{g}_i as rows into a matrix G (so $G(i, j) = \mathbf{g}_i(j)$). The duality means that $G = F^{-1}$. So, a simple change of coordinates gives $K = FDG$, hence $K^m = FD^mG$. Note that D^m is diagonal with $D^m(i, i) = \beta_i^m$. So

$$\begin{aligned} K^m(x, y) &= (FD^mG)(x, y) \\ &= \sum_{i, j} F(x, i) D^m(i, j) G(j, y) \\ &= \sum_i F(x, i) \beta_i^m G(i, y) \\ &= \sum_i \beta_i^m \mathbf{f}_i(x) \mathbf{g}_i(y). \end{aligned}$$

□

For general Markov chains, computing a full basis of eigenfunctions (a.k.a. “diagonalising” the chain) can be an intractable problem; this strategy is much more feasible when the chain has some underlying algebraic or geometric structure. For example, the eigenvalues of a random walk on a group come directly from the representation theory of the group [Dia88, Chap. 3E]. Similarly, there is a general formula for the eigenvalues and right eigenfunctions of a random walk on the chambers of a hyperplane arrangement [BHR99; Den12]. The purpose of this thesis is to carry out the equivalent analysis for Markov chains arising from Hopf algebras.

1.2 Hopf algebras

A graded, connected Hopf algebra is a graded vector space $\mathcal{H} = \bigoplus_{n=0}^{\infty} \mathcal{H}_n$ equipped with two linear maps: a product $m : \mathcal{H}_i \otimes \mathcal{H}_j \rightarrow \mathcal{H}_{i+j}$ and a coproduct $\Delta : \mathcal{H}_n \rightarrow \bigoplus_{j=0}^n \mathcal{H}_j \otimes \mathcal{H}_{n-j}$. The product is associative and has a unit which spans \mathcal{H}_0 . The corresponding requirements on the coproduct are *coassociativity*: $(\Delta \otimes \iota)\Delta = (\iota \otimes \Delta)\Delta$ (where ι denotes the identity map) and the counit axiom: $\Delta(x) - 1 \otimes x - x \otimes 1 \in \bigoplus_{j=1}^{n-1} \mathcal{H}_j \otimes \mathcal{H}_{n-j}$, for $x \in \mathcal{H}_n$. The product and coproduct satisfy the compatibility axiom $\Delta(wz) = \Delta(w)\Delta(z)$, where multiplication on $\mathcal{H} \otimes \mathcal{H}$ is componentwise. This condition may be more transparent in *Sweedler notation*: writing $\sum_{(x)} x_{(1)} \otimes x_{(2)}$ for $\Delta(x)$, the axiom reads $\Delta(wz) = \sum_{(w),(z)} w_{(1)}z_{(1)} \otimes w_{(2)}z_{(2)}$. This thesis will use Sweedler notation sparingly.

The definition of a general Hopf algebra, without the grading and connectedness assumptions, is slightly more complicated (it involves an extra *antipode* map, which is automatic in the graded case); the reader may consult [Swe69]. However, that reference (like many other introductions to Hopf algebras) concentrates on finite-dimensional Hopf algebras, which are useful in representation theory as generalisations of group algebras. These behave very differently from the infinite-dimensional Hopf algebras in this thesis.

Example 1.2.1 (Shuffle algebra). The shuffle algebra \mathcal{S} , as a vector space, has basis the set of all words in the letters $\{1, 2, \dots\}$. Write these words in parentheses to distinguish them from integers. The degree of a word is its number of letters, or *length*. The product of two words is the sum of all their interleavings (with multiplicity), and the coproduct is by deconcatenation; for example:

$$m((13) \otimes (52)) = (13)(52) = (1352) + (1532) + (1523) + (5132) + (5123) + (5213);$$

$$m((15) \otimes (52)) = (15)(52) = 2(1552) + (1525) + (5152) + (5125) + (5215);$$

$$\Delta((336)) = \emptyset \otimes (336) + (3) \otimes (36) + (33) \otimes (6) + (336) \otimes \emptyset.$$

(Here, \emptyset denotes the empty word, which is the unit of \mathcal{S} .)

More examples of Hopf algebras are in Section 4.1. This thesis will concentrate on Hopf algebras satisfying at least one of the following two symmetry conditions: \mathcal{H} is

commutative if $wz = zw$ for all $w, z \in \mathcal{H}$, and \mathcal{H} is *cocommutative* if $\sum_{(x)} x_{(1)} \otimes x_{(2)} = \sum_{(x)} x_{(2)} \otimes x_{(1)}$ for all $x \in \mathcal{H}$. In other words, if $\tau : \mathcal{H} \otimes \mathcal{H} \rightarrow \mathcal{H} \otimes \mathcal{H}$ is the linear map satisfying $\tau(w \otimes z) = z \otimes w$ for all $w, z \in \mathcal{H}$, then cocommutativity is the condition $\tau(\Delta(x)) = \Delta(x)$ for all x .

Hopf algebras first appeared in topology, where they describe the cohomology of a topological group or loop space. Cohomology is always an algebra under cup product, and the group product or the concatenation of loops induces the coproduct structure. Nowadays, the Hopf algebra is an indispensable tool in many parts of mathematics, partly due to structure theorems regarding abstract Hopf algebras. To give a flavour, a theorem of Hopf [Str11, Th. A49] states that any finite-dimensional, graded-commutative and graded-cocommutative Hopf algebra over a field of characteristic 0 is isomorphic as an algebra to a free exterior algebra with generators in odd degrees. More relevant to this thesis is the Cartier-Milnor-Moore theorem [Car07, Th. 3.8.1]: any cocommutative and conilpotent Hopf algebra \mathcal{H} over a field of characteristic zero is the universal enveloping algebra of its *primitive* subspace $\{x \in \mathcal{H} \mid \Delta(x) = 1 \otimes x + x \otimes 1\}$. That such a Hopf algebra is completely governed by its primitives will be important for Theorem 2.5.1.B, one of the algorithms diagonalising the Markov chains in this thesis.

1.3 Hopf-power Markov chains

To see the connection between the shuffle algebra and the GSR riffle-shuffle Markov chain, identify a deck of cards with the word whose i th letter denotes the value of the i th card, counting the cards from the top of the deck. So (316) describes a three-card deck with the card labelled 3 on top, card 1 in the middle, and card 6 at the bottom. Then, the probability that shuffling a deck x of n cards results in a deck y is

$$K(x, y) = \text{coefficient of } y \text{ in } 2^{-n} m\Delta(x).$$

In other words, the transition matrix of the riffle-shuffle Markov chain for decks of n cards is the transpose of the matrix of the linear map $2^{-n} m\Delta$ with respect to the basis of words of length n . Thus diagonalising the riffle-shuffle chain amounts to the completely algebraic

problem of finding an eigenbasis for $m\Delta$, the coproduct-then-product operator, on the shuffle algebra. Chapter 2 and Section 6.1 achieve this; although the resulting eigenfunctions are not dual in the sense of Proposition 1.1.4, this is the first time that full eigenbases for riffle-shuffling have been determined.

The subject of this thesis is to analogously model the breaking-then-recombining of other combinatorial objects. As described in Section 4.1, the literature contains numerous constructions of *combinatorial Hopf algebras*, which encode how to assemble and take apart combinatorial objects. For example, in the Hopf algebras of graphs (Example 4.1.3), the product is disjoint union, and the coproduct sends a graph to pairs of induced subgraphs on a subset of the vertices and on its complement. Then one can produce a “graph-breaking” model by defining the transition probabilities $K(x, y)$ to be the coefficient of y in $2^{-n}m\Delta(x)$, where n is the number of vertices of the graphs x and y . Then each step of the chain chooses a subset of the vertex set and severs all edges with exactly one endpoint in the chosen subset. Since this transition matrix is the matrix of the linear operator $2^{-n}m\Delta$, its eigenfunctions again come from the eigenvectors of $2^{-n}m\Delta$.

The obstacle to making the same definition on other Hopf algebras is that the coefficients of $2^{-n}m\Delta$ need not always sum to one. Fortunately, a clean workaround exists in the form of the *Doob transform*. Theorem 3.1.1 describes this very general method of building a transition matrix out of most non-negative linear operators, by rescaling the basis.

Since the transition matrix of such a *Hopf-power Markov chain* is the matrix of the coproduct-then-product operator $m\Delta$ (albeit with a rescaling of basis), many questions about these chains can be translated from probability into algebra. As previously mentioned, the eigenfunctions of the chain are the eigenvectors of $m\Delta$; this applies in particular to their stationary distributions. Reversibility of a Hopf-power Markov chain is equivalent to self-duality of the underlying Hopf algebra (Theorem 4.6.3), and the Projection Theorem (Theorem 4.7.1) explains how Markov statistics arise from certain maps between Hopf algebras. For example, Theorem 6.2.1 constructs a Hopf-morphism which sends a deck of distinct cards to its *descent set* (the positions where a card has greater value than its immediate successor). Consequently, tracking the descent set under riffle-shuffling of distinct cards is itself a Markov chain. In other words, the descent set after one shuffle only depends

on the current descent set, not on the precise ordering of the deck, an observation originally due to Stanley.

The Hopf-power Markov chain is a very general construction - it can exhibit various different behaviour depending on the structure of the underlying Hopf algebra, i.e. on the interplay of the breaking and combining rules. For example, the stationary distribution of the edge-removal chain is concentrated at the graph with no edges, whilst riffle-shuffling has a uniform distribution. In fact, for all combinatorial families with a “deterministic combining rule”, their chains are absorbing, and there is a standard procedure for approximating how close they are to absorption (Proposition 5.1.11).

The organisation of the thesis is as follows: Chapter 2 derives some results on the eigenvectors of $m\Delta$, which will be useful both in constructing and diagonalising Hopf-power Markov chains. It does not involve any probability. Chapter 3 is independent of Chapter 2 and describes the properties of the Doob transform under very general hypotheses, without reference to Hopf algebras. Chapter 4 is the centerpiece of the thesis - it contains the construction of Hopf-power Markov chains, and the theorems regarding their stationary distribution, reversibility, and Markov statistics. Chapter 5 opens with additional theory for chains with a “deterministic combining rule”, then illustrates this in detail on the examples of rock-breaking and tree-pruning. Chapter 6 is devoted to the initial example of riffle-shuffling - Section 6.1 derives a full left and right eigenbases and some associated probability estimates, and Section 6.2 interprets the left and right eigenbases of the descent set chain.

Remark. An earlier version of the Hopf-power Markov chain framework, restricted to free-commutative or free state space bases, appeared in [DPR14]. Table 1.1 pairs up the results and examples of that paper and their improvements in this thesis. (I plan to update this table on my website, as the theory advances and more examples are available.) In addition, a summary of Section 6.2, on the descent set Markov chain under riffle-shuffling, appeared in [Pan13].

	[DPR14]	thesis
construction	3.2	4.2,4.3
stationary distribution	3.7.1	4.5
reversibility		4.6
projection		4.7
diagonalisation		
general	3.5	2
algorithm for free-commutative basis	Th. 3.15	Th. 2.5.1.A
algorithm for basis of primitives		Th. 2.5.1.B
algorithm for shuffle basis		Th. 2.5.1.A'
algorithm for free basis	Th. 3.16	Th. 2.5.1.B'
unidirectionality for free-commutative basis	3.3	5.1.2
right eigenfunctions for free-commutative basis	3.6	5.1.3
link to terminality of $QSym$	3.7.2	5.1.4
examples		
rock-breaking	4	5.2
tree-pruning		5.3
riffle-shuffling ¹	5	6.1
descent sets under riffle-shuffling		6.2

Table 1.1: Corresponding sections of [DPR14] and the present thesis

Chapter 2

Diagonalisation of the Hopf-power map

This chapter collects together some results on the eigenvectors of the Hopf-power map; these will be useful in subsequent chapters for constructing and diagonalising Hopf-power Markov chains. These results do not require any probability, and may be of interest independently of Hopf-power Markov chains.

Section 2.1 introduces the Hopf-power map and its dual. The next three sections build towards Theorem 2.5.1, a set of four explicit algorithms for full eigenbases of the Hopf-power map Ψ^a on a commutative or cocommutative (graded connected) Hopf algebra. These allow explicit computations of left and right eigenbases of the associated Markov chains. Each algorithm follows the same general two-step principle: first, produce the eigenvectors of smallest eigenvalue, using the Eulerian idempotent (Section 2.2), then, combine these into eigenvectors of higher eigenvalue, following Section 2.3. Section 2.4 explains the Lyndon word terminology necessary to implement Theorems 2.5.1.A' and 2.5.1.B'; these extended algorithms are useful when the information required for Theorems 2.5.1.A and 2.5.1.B are not readily available. Section 2.5 contains all four algorithms and their proofs.

Section 2.6 drops the assumptions of commutativity or cocommutativity, and proves that the eigenbases algorithms still hold, in some sense, for the highest eigenvalue. This last result encodes the stationary distributions for any Hopf-power Markov chain (Theorem 4.5.1), and offers some explanation as to why certain bases cannot produce Markov chains through the Doob transform (end of Section 4.3).

2.1 The Hopf-power Map

The Markov chains in this thesis are built from the *ath* Hopf-power map $\Psi^a : \mathcal{H} \rightarrow \mathcal{H}$, defined to be the *a-fold coproduct* followed by the *a-fold product*: $\Psi^a := m^{[a]}\Delta^{[a]}$. Here $\Delta^{[a]} : \mathcal{H} \rightarrow \mathcal{H}^{\otimes a}$ is defined inductively by $\Delta^{[a]} := (\iota \otimes \cdots \otimes \iota \otimes \Delta)\Delta^{[a-1]}$, $\Delta^{[1]} = \iota$ (recall ι denotes the identity map), and $m^{[a]} : \mathcal{H}^{\otimes a} \rightarrow \mathcal{H}$ by $m^{[a]} := m(m^{[a-1]} \otimes \iota)$, $m^{[1]} = \iota$. So the Hopf-square is coproduct followed by product: $\Psi^2 := m\Delta$. Observe that, on a graded Hopf algebra, the Hopf-powers preserve degree: $\Psi^a : \mathcal{H}_n \rightarrow \mathcal{H}_n$.

The Hopf-power map first appeared in [TO70] in the study of group schemes. The notation Ψ^a comes from [Pat93]; [Kas00] writes $[a]$, and [AL13] writes ι^{*a} , since it is the *ath* convolution power of the identity map. [LMS06] denotes $\Psi^a(x)$ by $x^{[a]}$; they study this operator on finite-dimensional Hopf algebras as a generalisation of group algebras. The nomenclature ‘‘Hopf-power’’ comes from the fact that these operators exponentiate the basis elements of a group algebra; in this special case, $\Psi^a(g) = g^a$. Since this thesis deals with graded, connected Hopf algebras, there will be no elements satisfying $\Psi^a(g) = g^a$, other than multiples of the unit. However, the view of Ψ^a as a power map is still helpful: on commutative or cocommutative Hopf algebras, the *power rule* $\Psi^a\Psi^{a'} = \Psi^{aa'}$ holds. Here is a simple proof [Kas00, Lem. 4.1.1], employing Sweedler notation:

$$\begin{aligned}
& \Psi^{a'}\Psi^a(x) \\
&= \sum_{(x)} \Psi^{a'}(x_{(1)} \cdots x_{(a)}) \\
&= \sum [(x_{(1)})_{(1)}(x_{(2)})_{(1)} \cdots (x_{(a)})_{(1)}] [(x_{(1)})_{(2)} \cdots (x_{(a)})_{(2)}] \cdots [(x_{(1)})_{(a')} \cdots (x_{(a)})_{(a')}] \\
&= \sum [x_{(1)}x_{(a'+1)} \cdots x_{(a'(a-1)+1)}] [x_{(2)} \cdots x_{(a'(a-1)+2)}] \cdots [x_{(a')} \cdots x_{(aa')}] \\
&= \sum x_{(1)}x_{(2)} \cdots x_{(aa')} = \Psi^{aa'}(x).
\end{aligned}$$

(The third equality uses coassociativity, and the fourth uses commutativity or cocommutativity.)

The Hopf-power Markov chains of this thesis arise from applying the Doob transform to the Hopf-power map $\Psi^a : \mathcal{H}_n \rightarrow \mathcal{H}_n$. As Theorem 3.1.1 will explain, the Doob transform

requires a special eigenvector of the dual map to Ψ^a . This dual map is in fact also a Hopf-power map, but on the dual Hopf algebra, as defined below.

Definition 2.1.1. Let $\mathcal{H} = \bigoplus_{n \geq 0} \mathcal{H}_n$ be a graded, connected Hopf algebra over \mathbb{R} with basis $\mathcal{B} = \coprod_n \mathcal{B}_n$. The (graded) dual of \mathcal{H} is $\mathcal{H}^* := \bigoplus_{n \geq 0} \mathcal{H}_n^*$, where \mathcal{H}_n^* is the set of linear functionals on \mathcal{H}_n . (This is the dual of \mathcal{H}_n in the sense of vector spaces, as described at the start of Chapter 3.) The product and coproduct on \mathcal{H}^* are defined by

$$m(f \otimes g)(x) := (f \otimes g)(\Delta x); \quad \Delta(f)(w \otimes z) = f(wz)$$

for $x, z, w \in \mathcal{H}$ and $f, g \in \mathcal{H}^*$. (Here, $(f \otimes g)(a \otimes b) = f(a)g(b)$.)

The symmetry of the Hopf axioms ensures that \mathcal{H}^* is also a (graded, connected) Hopf algebra. Note that, for $x \in \mathcal{H}$ and $f \in \mathcal{H}^*$,

$$(m^{[a]} \Delta^{[a]} f)(x) = (\Delta^{[a]} f)(\Delta^{[a]} x) = f(m^{[a]} \Delta^{[a]} x)$$

so the a th Hopf-power of \mathcal{H}_n^* is the dual map (in the linear algebraic sense) to the a th Hopf-power on \mathcal{H}_n .

Example 2.1.2. The dual of the shuffle algebra \mathcal{S} is the free associative algebra \mathcal{S}^* , whose basis is also indexed by words in the letters $\{1, 2, \dots\}$. The product in \mathcal{S}^* is concatenation, for example:

$$m((12) \otimes (231)) = (12231)$$

and the coproduct is “deshuffling”:

$$\Delta(w_1 \dots w_n) = \sum_{S \subseteq \{1, 2, \dots, n\}} \prod_{i \in S} w_i \otimes \prod_{i \notin S} w_i.$$

For example,

$$\begin{aligned} \Delta((316)) &= \emptyset \otimes (316) + (3) \otimes (16) + (1) \otimes (36) + (6) \otimes (31) \\ &\quad + (31) \otimes (6) + (36) \otimes (1) + (16) \otimes (3) + (316) \otimes \emptyset. \end{aligned}$$

Observe that the free associative algebra is noncommutative and cocommutative. In general, the dual of a commutative algebra is cocommutative, and vice versa.

2.2 The Eulerian Idempotent

The first step in building an eigenbasis for the Hopf-power map Ψ^a is to use the Eulerian idempotent map to produce eigenvectors of smallest eigenvalue. Defining this map requires the notion of the *reduced coproduct* $\bar{\Delta}(x) := \Delta(x) - 1 \otimes x - x \otimes 1$. It follows from the counit axiom that $\bar{\Delta}(x)$ consists precisely of the terms of $\Delta(x)$ where both tensor-factors have strictly positive degree. Define inductively the a -fold reduced coproduct: $\bar{\Delta}^{[1]} := \iota$, and $\bar{\Delta}^{[a]} := (\iota \otimes \cdots \otimes \iota \otimes \bar{\Delta})\bar{\Delta}^{[a-1]}$, which picks out the terms in $\Delta^{[a]}(x)$ with all a tensor-factors having strictly positive degree. This captures the notion of breaking into a non-trivial pieces. Note that $\bar{\Delta}^{[2]} = \bar{\Delta}$.

Definition 2.2.1 (Eulerian idempotent). [Pat93, Def. 2.2] Let \mathcal{H} be a Hopf algebra over a field of characteristic zero which is *conilpotent* (i.e. for each x , there is some a with $\bar{\Delta}^{[a]}x = 0$). Then the (first) Eulerian idempotent map $e : \mathcal{H} \rightarrow \mathcal{H}$ is given by

$$e(x) = \sum_{r \geq 1} \frac{(-1)^{r-1}}{r} m^{[r]} \bar{\Delta}^{[r]}(x).$$

(Conilpotence ensures this sum is finite).

Clearly, graded Hopf algebras are conilpotent: if $x \in \mathcal{H}_n$, then $\bar{\Delta}^{[r]}(x) = 0$ whenever $r > n$.

Patras proved that, if \mathcal{H} is commutative or cocommutative, then the image of e is the eigenspace for Ψ^a of eigenvalue a . Furthermore, if \mathcal{H} is cocommutative, [Sch94, Th. 9.4] shows that this image is the subspace of *primitive* elements of \mathcal{H} , defined to be $\{x \in \mathcal{H} \mid \Delta(x) = 1 \otimes x + x \otimes 1\}$. Note that this subspace is precisely the kernel of the reduced coproduct map $\bar{\Delta}$. A brief explanation of these properties of $\text{im}(e)$ is at the end of this section, after an example of calculating $e(x)$.

Example 2.2.2. Work in the shuffle algebra \mathcal{S} of Example 1.2.1, where the product is interleaving and the coproduct is deconcatenation.

$$\begin{aligned}
 e((12)) &= (12) - \frac{1}{2}m\bar{\Delta}(12) \\
 &= (12) - \frac{1}{2}(1)(2) \\
 &= (12) - \frac{1}{2}[(12) + (21)] \\
 &= \frac{1}{2}[(12) - (21)].
 \end{aligned}$$

Observe that

$$\bar{\Delta}\left(\frac{1}{2}[(12) - (21)]\right) = \frac{1}{2}[(1) \otimes (2) - (2) \otimes (1)],$$

so, by commutativity, $m\bar{\Delta}e((12)) = 0$, but $\bar{\Delta}e((12)) \neq 0$. Thus $e((12))$ is an eigenvector for Ψ^a of eigenvalue a , but is not primitive.

Here is one more demonstration of the Eulerian idempotent:

$$\begin{aligned}
 e((123)) &= (123) - \frac{1}{2}m\bar{\Delta}(123) + \frac{1}{3}m^{[3]}\bar{\Delta}^{[3]}(123) \\
 &= (123) - \frac{1}{2}[(12)(3) + (1)(23)] + \frac{1}{3}(1)(2)(3) \\
 &= (123) - \frac{1}{2}[2(123) + (132) + (312) + (213) + (231)] \\
 &\quad + \frac{1}{3}[(123) + (132) + (312) + (213) + (231) + (321)] \\
 &= \frac{1}{6}[2(123) - (132) - (312) - (213) - (231) + 2(321)].
 \end{aligned}$$

The idea of the Eulerian idempotent came independently from Reutenauer and from Patras: Reutenauer analysed it on the free associative algebra \mathcal{S}^* (see Example 2.1.2), and Patras derived the same properties for a general commutative or cocommutative conilpotent algebra. They both define the Eulerian idempotent as the logarithm of the identity map in the algebra (under convolution product) of endomorphisms of \mathcal{H} . To obtain the explicit formula of Definition 2.2.1 above, use the Taylor expansion of $\log(1+x)$ with x being $\iota - 1$, where 1 is projection to \mathcal{H}_0 (or, more generally, the counit followed by unit). From

the familiar identity

$$y^a = e^{a \log y} = \sum_{i=0}^{\infty} \frac{a^i}{i!} (\log y)^i$$

applied to the identity map, Patras concludes in his Proposition 3.2 that $\Psi^a = \sum_{i=0}^{\infty} a^i e_i$ where the e_i are his *higher Eulerian idempotents*, the i th convolution power of e scaled by $i!$:

$$e_i := \frac{1}{i!} m^{[i]}(e \otimes \cdots \otimes e) \Delta^{[i]}.$$

Hence the usual Eulerian idempotent e is e_1 . Recall from Section 2.1 that, if \mathcal{H} is commutative or cocommutative, then the power law holds: $\Psi^a \Psi^{a'} = \Psi^{aa'}$ (the left hand side is the composition of two Hopf-powers). In terms of Eulerian idempotents, this says

$$\sum_{i,j=0}^{\infty} a^i e_i a'^j e_j = \sum_{k=0}^{\infty} (aa')^k e_k.$$

Equating coefficients of aa' then shows that the e_i are orthogonal idempotents under composition: $e_i e_i = e_i$ and $e_i e_j = 0$ for $i \neq j$. Combining this knowledge with the expansion $\Psi^a = \sum_{i=0}^{\infty} a^i e_i$ concludes that e_i is the orthogonal projection of \mathcal{H} onto the a^i -eigenspace of Ψ^a .

2.3 Eigenvectors of Higher Eigenvalue

As just discussed, on a commutative or cocommutative graded Hopf algebra, Patras's higher Eulerian idempotent maps e_k are projections to the a^k -eigenspaces for the a th Hopf-power. However, this thesis chooses instead to build the a^k -eigenspace out of k -tuples of eigenvectors of eigenvalue a .

First, consider the case where \mathcal{H} is commutative. Then, as noted in [Pat93], the power-map Ψ^a is an algebra homomorphism:

$$\begin{aligned} \Psi^a(xy) &= m^{[a]} \sum_{(x),(y)} x_{(1)} y_{(1)} \otimes \cdots \otimes x_{(a)} y_{(a)} \\ &= \sum_{(x),(y)} x_{(1)} y_{(1)} \cdots x_{(a)} y_{(a)} = \sum_{(x),(y)} x_{(1)} \cdots x_{(a)} y_{(1)} \cdots y_{(a)} = \Psi^a(x) \Psi^a(y). \end{aligned}$$

Then it follows easily that:

Theorem 2.3.1. *Work in a commutative Hopf algebra. Suppose x_1, x_2, \dots, x_k are eigenvectors of Ψ^a of eigenvalue a . Then $x_1 x_2 \dots x_k$ is an eigenvector of Ψ^a with eigenvalue a^k .* \square

If \mathcal{H} is not commutative, then a strikingly similar construction holds, if one restricts the x_i to be primitive rather than simply eigenvectors of eigenvalue a . The reasoning is completely different:

Theorem 2.3.2 (Symmetrisation Lemma). *Let x_1, x_2, \dots, x_k be primitive elements of any Hopf algebra, then $\sum_{\sigma \in S_k} x_{\sigma(1)} x_{\sigma(2)} \dots x_{\sigma(k)}$ is an eigenvector of Ψ^a with eigenvalue a^k .*

Proof. The proof is essentially a calculation. For concreteness, take $a = 2$. Then

$$\begin{aligned}
& m\Delta \left(\sum_{\sigma \in S_k} x_{\sigma(1)} x_{\sigma(2)} \dots x_{\sigma(k)} \right) \\
&= m \left(\sum_{\sigma \in S_k} (\Delta x_{\sigma(1)}) (\Delta x_{\sigma(2)}) \dots (\Delta x_{\sigma(k)}) \right) \\
&= m \left(\sum_{\sigma \in S_k} (x_{\sigma(1)} \otimes 1 + 1 \otimes x_{\sigma(1)}) \dots (x_{\sigma(k)} \otimes 1 + 1 \otimes x_{\sigma(k)}) \right) \\
&= m \left(\sum_{A_1 \amalg A_2 = \{1, 2, \dots, k\}} \sum_{\sigma \in S_k} \prod_{i \in A_1} x_{\sigma(i)} \otimes \prod_{j \in A_2} x_{\sigma(j)} \right) \\
&= |\{(A_1, A_2) \mid A_1 \amalg A_2 = \{1, 2, \dots, k\}\}| \sum_{\sigma \in S_k} x_{\sigma(1)} \dots x_{\sigma(k)} \\
&= 2^k \sum_{\sigma \in S_k} x_{\sigma(1)} \dots x_{\sigma(k)}.
\end{aligned}$$

For higher a , the same argument shows that

$$\begin{aligned}
& \Psi^a \left(\sum_{\sigma \in \mathcal{S}_k} x_{\sigma(1)} x_{\sigma(2)} \cdots x_{\sigma(k)} \right) \\
&= m^{[a]} \left(\sum_{A_1 \amalg \cdots \amalg A_a = \{1, 2, \dots, k\}} \sum_{\sigma \in \mathcal{S}_k} \left(\prod_{i \in A_1} x_{\sigma(i)} \right) \otimes \cdots \otimes \left(\prod_{i \in A_a} x_{\sigma(i)} \right) \right) \\
&= a^k \sum_{\sigma \in \mathcal{S}_k} x_{\sigma(1)} \cdots x_{\sigma(k)}.
\end{aligned}$$

□

2.4 Lyndon Words

The previous two sections show that, for commutative or cocommutative \mathcal{H} , (symmetrised) products of images under the Eulerian idempotent map are eigenvectors of the Hopf-power maps Ψ^a . A natural question follows: to which elements of \mathcal{H} should one apply the Eulerian idempotent map in order for this process to output a basis? One possible answer is “the generators of \mathcal{H} ”, in a sense which Theorems 2.5.1.A and 2.5.1.B will make precise. Such generators can sometimes be conveniently determined, but in many cases it is easier to first relate the combinatorial Hopf algebra to the shuffle algebra or the free associative algebra, and then use the structure theory of these two famous algebras to pick out the required generators. This is the main idea of Theorems 2.5.1.A' and 2.5.1.B' respectively, and this section explains, following [Lot97, Sec. 5.1], the Lyndon word terminology necessary for this latter step.

Definition 2.4.1 (Lyndon word). A word is *Lyndon* if it is lexicographically strictly smaller than its cyclic rearrangements.

For example, (11212) is Lyndon, as it is lexicographically strictly smaller than (12121), (21211), (12112) and (21121). The word (1212) is not Lyndon as it is equal to one of its cyclic rearrangements. (31421) is also not Lyndon - for example, it does not begin with its minimal letter.

Definition 2.4.2 (Lyndon factorisation). The *Lyndon factorisation* $u_1 \cdots u_k$ of w is obtained by taking u_k to be the lexicographically smallest tail of w , then u_{k-1} is the lexicographically smallest tail of w with u_k removed, and so on. Throughout this thesis, $k(w)$ will always denote the number of Lyndon factors in w .

Observe that w is the concatenation of its Lyndon factors, not the product of these factors in the sense of the shuffle algebra. Indeed, all this terminology is independent of the product on the shuffle algebra.

[Lot97, Th. 5.1.5, Prop. 5.1.6] asserts that such u_i are each Lyndon - indeed, this is the only way to deconcatenate w into Lyndon words with $u_1 \geq u_2 \geq \cdots \geq u_k$ in lexicographic order. It follows from this uniqueness that each unordered k -tuple of Lyndon words (possibly with repeats) is the Lyndon factorisation of precisely one word, namely their concatenation in decreasing lexicographic order.

Example 2.4.3. Let $w = (31421)$. The tails of w are (1), (21), (421), (1421) and (31421), and the lexicographically smallest of these is (1). The lexicographically smallest tail of (3142) is (142). So $k(w) = 3$ and the Lyndon factors of w are $u_1 = (3)$, $u_2 = (142)$ and $u_3 = (1)$.

Definition 2.4.4 (Standard factorisation). A Lyndon word u of length greater than 1 has *standard factorisation* $u_1 \cdot u_2$, where u_2 is the longest Lyndon tail of u that is not u itself, and u_1 is the corresponding head. By [Lot97, Prop. 5.1.3], the head u_1 is also Lyndon.

Example 2.4.5. The Lyndon word $u = (1323)$ has two tails which are Lyndon (and are not u itself): (3) and (23). The longer Lyndon tail is (23), so the standard factorisation of u is $(1323) = (13 \cdot 23)$

When using Theorems 2.5.1.A' and 2.5.1.B' below, it will be more convenient to work with an alphabet of combinatorial objects rather than the positive integers - all the above notions are well-defined for “words” whose letters are drawn from any totally-ordered set. In addition, if this set is graded, then one can assign the degree of a word to be the sum of the degree of its letters.

Example 2.4.6. If \bullet comes before x in an alphabet, then the word $\bullet x \bullet$ has Lyndon factorisation $\bullet x \bullet$, and the Lyndon word $\bullet \bullet x$ has standard factorisation $\bullet \bullet x$. If $\deg(x) = 2$, then

both $\bullet x \bullet$ and $\bullet \bullet x$ have degree 4. Example 2.5.3 below will demonstrate the eigenvector corresponding to $\bullet x \bullet$.

2.5 Algorithms for a Full Eigenbasis

Theorem 2.5.1 below collects together four algorithms for a full eigenbasis of the Hopf-power Ψ^a . Immediately following are calculations illustrating Parts A' and B', before the proofs of all four algorithms. These algorithms will be useful in Chapters 5 and 6 to compute eigenfunctions of Hopf-power Markov chains.

One more ingredient is necessary to state Part A of Theorem 2.5.1: the dual Cartier-Milnor-Moore theorem [Car07, Th. 3.8.3] states that every graded connected commutative Hopf algebra \mathcal{H} (over a field \mathbb{F} of characteristic 0) is a polynomial algebra, i.e. $\mathcal{H} = \mathbb{F}[c_1, c_2, \dots]$ for homogeneous elements c_i . $\{c_1, c_2, \dots\}$ is then called a *free generating set* for \mathcal{H} . (The usual Cartier-Milnor-Moore theorem, for cocommutative Hopf algebras, also plays a role in the eigenbasis algorithms; see the proof of Part B).

Theorem 2.5.1 (Eigenbasis algorithms). *In all four parts below, $\mathcal{H} = \bigoplus_{n \geq 0} \mathcal{H}_n$ is a graded connected Hopf algebra over \mathbb{R} with each \mathcal{H}_n finite-dimensional.*

- (A) *Suppose \mathcal{H} is commutative, and let \mathcal{C} be a free generating set for \mathcal{H} . Then $\{e(c_1) \dots e(c_k) \mid k \in \mathbb{N}, \{c_1, \dots, c_k\} \text{ a multiset in } \mathcal{C}\}$ is an eigenbasis for Ψ^a on \mathcal{H} , and the eigenvector $e(c_1) \dots e(c_k)$ has eigenvalue a^k . So the multiplicity of the eigenvalue a^k in \mathcal{H}_n is the coefficient of $x^n y^k$ in $\prod_{c \in \mathcal{C}} (1 - yx^{\deg c})^{-1}$.*
- (B) *Suppose \mathcal{H} is cocommutative, and let \mathcal{P} be a basis of its primitive subspace. Then $\{\frac{1}{k!} \sum_{\sigma \in \mathfrak{S}_k} p_{\sigma(1)} \dots p_{\sigma(k)} \mid k \in \mathbb{N}, \{p_1, \dots, p_k\} \text{ a multiset in } \mathcal{P}\}$ is an eigenbasis for Ψ^a on \mathcal{H} , and the eigenvector $\frac{1}{k!} \sum_{\sigma \in \mathfrak{S}_k} p_{\sigma(1)} \dots p_{\sigma(k)}$ has eigenvalue a^k . So the multiplicity of the eigenvalue a^k in \mathcal{H}_n is the coefficient of $x^n y^k$ in $\prod_{p \in \mathcal{P}} (1 - yx^{\deg p})^{-1}$.*
- (A') *Suppose \mathcal{H} is isomorphic, as a non-graded algebra only, to the shuffle algebra, and write P_w for the image in \mathcal{H} of the word w under this isomorphism. (So $\{P_w\}$ is a basis of \mathcal{H} indexed by words such that $P_w P_{w'} = \sum_v P_v$, summing over all interleavings*

v of w and w' with multiplicity.) For each word w , define $g_w \in \mathcal{H}$ recursively to be:

$$\begin{aligned} g_w &:= e(P_w) && \text{if } w \text{ is a Lyndon word;} \\ g_w &:= g_{u_1} \cdots g_{u_k} && \text{if } w \text{ has Lyndon factorisation } w = u_1 \cdots u_k. \end{aligned}$$

Then $\{g_w\}$ is an eigenbasis for Ψ^a on \mathcal{H} , and the eigenvector g_w has eigenvalue $a^{k(w)}$, where $k(w)$ is the number of factors in the Lyndon factorisation of w . So the multiplicity of the eigenvalue a^k in \mathcal{H}_n is the coefficient of $x^n y^k$ in $\prod_{w \text{ Lyndon}} (1 - yx^{\deg P_w})^{-1}$.

(B') Suppose \mathcal{H} is cocommutative, and is isomorphic, as a non-graded algebra only, to the free associative algebra $\mathbb{R}\langle S_1, S_2, \dots \rangle$. For each word $w = w_1 \dots w_l$, where each w_i is a letter, write S_w for $S_{w_1} \dots S_{w_l}$, so $\{S_w\}$ is a free basis with concatenation product. For each word w , define $g_w \in \mathcal{H}$ recursively by:

$$\begin{aligned} g_w &:= e(S_w) && \text{if } w \text{ is a single letter;} \\ g_w &:= [g_{u_1}, g_{u_2}] := g_{u_1} g_{u_2} - g_{u_2} g_{u_1} && \text{if } w \text{ is Lyndon with standard factorisation } w = u_1 u_2; \\ g_w &:= \frac{1}{k!} \sum_{\sigma \in \mathfrak{S}_k} g_{u_{\sigma(1)}} \cdots g_{u_{\sigma(k)}} && \text{if } w \text{ has Lyndon factorisation } w = u_1 \cdots u_k. \end{aligned}$$

Then $\{g_w\}$ is an eigenbasis for Ψ^a on \mathcal{H} , and the eigenvector g_w has eigenvalue $a^{k(w)}$, where $k(w)$ is the number of factors in the Lyndon factorisation of w . So the multiplicity of the eigenvalue a^k in \mathcal{H}_n is the coefficient of $x^n y^k$ in $\prod_{w \text{ Lyndon}} (1 - yx^{\deg S_w})^{-1}$.

Remarks.

1. The notation P and S for the bases in Parts A' and B' are intentionally suggestive of dual power sums and complete noncommutative symmetric functions respectively, see Section 6.2.
2. Part A does not imply that the map $x_i \rightarrow e(c_i)$ is a Hopf-isomorphism from the polynomial algebra $\mathbb{R}[x_1, x_2, \dots]$ to any graded connected commutative Hopf algebra, as

the $e(c_i)$ need not be primitive. This map is only a Hopf-isomorphism if the Hopf algebra in question is cocommutative in addition to being commutative. See Section 5.3 on the tree-pruning process for a counterexample. Similarly, Part A' does not imply that the shuffle algebra is Hopf-isomorphic to any Hopf algebra with a shuffle product structure via the map $w \rightarrow e(P_w)$ for Lyndon w ; even if all the $e(P_i)$ were primitive, $\bar{\Delta}(e(P_{12}))$ might not be $e(P_1) \otimes e(P_2)$. In short, the presence of a shuffle product structure is not sufficiently restrictive on the coproduct structure to uniquely determine the Hopf algebra.

3. In contrast, the map $i \rightarrow e(S_i)$ in Part B' does construct a (non-graded) Hopf-isomorphism from the free associative algebra \mathcal{S}^* to any cocommutative Hopf algebra with a free basis. This is because the image under e of a cocommutative Hopf algebra is primitive. In fact, the eigenvectors g_w are simply the images of an eigenbasis for the free associative algebra \mathcal{S}^* under this isomorphism. Hence the approach of this thesis is as follows: Section 6.1.1 uses Part B' above to generate an eigenbasis for \mathcal{S}^* , and writes these, up to scaling, as

$$\sum_{w' \in \mathcal{S}_{\deg(w)}} \mathbf{f}_w^{\mathcal{S}}(w')w'.$$

(The notation $\mathbf{f}_w^{\mathcal{S}}$ comes from these being the right eigenfunctions of riffle-shuffling.) It explains a method to calculate them in terms of decreasing Lyndon hedgerows. Thereafter, the thesis will ignore Part B' and simply use

$$g_w = \sum_{w'} \mathbf{f}_w^{\mathcal{S}}(w')e(S_{w'_1}) \dots e(S_{w'_l})$$

to obtain the necessary eigenvectors in Section 6.2.5, taking advantage of the graphical way to calculate $\mathbf{f}_w^{\mathcal{S}}$. Here the sum runs over all w' containing the same letters as w , and w'_i denotes the i th letter of w' . This alternative expression differs from the g_w in Part B' above by a scaling factor, but for the probability applications in this thesis, this alternative scaling is actually more convenient.

4. Each part of the Theorem closes with the generating function for the multiplicities of each eigenvalue on subspaces of each degree. These are simple generalisations of the generating function for partitions, since each eigenvector of eigenvalue a^k corresponds to a k -tuple (unordered, possibly with repeats) of generators (Part A), primitives (Part B), or Lyndon words (Parts A' and B'). See [Wil94, Th. 3.14.1]. All four generating functions hold for Hopf algebras that are multigraded - simply replace all x s, n s and degrees by tuples, and read the formula as multi-index notation. For example, for a bigraded commutative algebra \mathcal{H} with free generating set \mathcal{C} (so Part A applies), the multiplicity of the a^k -eigenspace in $\mathcal{H}_{m,n}$ is the coefficient of $x_1^m x_2^n y^k$ in $\prod_{c \in \mathcal{C}} \left(1 - y x_1^{\deg_1 c} x_2^{\deg_2 c}\right)^{-1}$, where $\deg(c) = (\deg_1(c), \deg_2(c))$. This idea will be useful in Section 6.1 for the study of riffle-shuffling.
5. To analyse Markov chains, one ideally wants expressions for left and right eigenfunctions of the transition matrix that are “dual”, in the sense of Proposition 1.1.4. For Hopf-power Markov chains, Proposition 3.2.1 below translates this goal to an eigenbasis for the Hopf-power Ψ^a on \mathcal{H} and the dual eigenbasis for Ψ^a on \mathcal{H}^* . Thus it would be best to apply the above algorithms to \mathcal{H} and \mathcal{H}^* in such a way that the results interact nicely. Theorem 5.1.9 achieves this when a free-commutative basis of \mathcal{H} is explicit, using Part A on \mathcal{H} and Part B on \mathcal{H}^* .

Example 2.5.2. Theorem 2.5.1.A' applies to the shuffle algebra, with $P_w = w$ for each word w . Take $w = (3141)$, which has Lyndon factorisation $(3 \cdot 14 \cdot 1)$. Then the associated eigenvector g_w , which has eigenvalue a^3 , is

$$\begin{aligned}
& e((3))e((14))e((1)) \\
&= (3) \left[(14) - \frac{1}{2}(1)(4) \right] (1) \\
&= (3) \left[\frac{1}{2}(14) - \frac{1}{2}(41) \right] (1) \\
&= (3) \frac{1}{2} [(141) + 2(114) - 2(411) - (141)] \\
&= (3114) + (1314) + (1134) + (1143) - (3411) - (4311) - (4131) - (4113).
\end{aligned}$$

Example 2.5.3. Consider applying Theorem 2.5.1.B' to a Hopf algebra with a free basis to find the eigenvector corresponding to the word $\bullet x \bullet$, where \bullet and x are letters with $\deg(\bullet) = 1$, $\deg(x) = 2$, and \bullet coming before x in “alphabetical order”. (This would, for example, construct a right eigenfunction for the Markov chain of the descent set under riffle-shuffling corresponding to the composition $(1, 2, 1)$, if x were $S^{(2)}$. See Example 6.2.8.) As noted in Example 2.4.6, the Lyndon factorisation of $\bullet x \bullet$ is $\bullet x \bullet$, so, according to Theorem 2.5.1.B'

$$g_{\bullet x \bullet} = \frac{1}{2!} (g_{\bullet x} g_{\bullet} + g_{\bullet} g_{x \bullet}).$$

The first Lyndon factor $\bullet x$ has standard factorisation $\bullet \cdot x$, so

$$g_{\bullet x} = g_{\bullet} g_x - g_x g_{\bullet} = e(\bullet) e(x) - e(x) e(\bullet).$$

As $\deg(\bullet) = 1$, it follows that $e(\bullet) = \bullet$. Hence

$$\begin{aligned} g_{\bullet x \bullet} &= \frac{1}{2!} ((\bullet e(x) - e(x) \bullet) \bullet + \bullet (\bullet e(x) - e(x) \bullet)) \\ &= \frac{1}{2} (\bullet \bullet e(x) - e(x) \bullet \bullet). \end{aligned}$$

Alternatively, use the formulation in Remark 3 above,

$$g_{\bullet x \bullet} = \sum_{w'} \mathbf{f}_{\bullet x \bullet}^{\mathcal{L}}(w') e(S_{w'_1}) \dots e(S_{w'_l}).$$

summing over all words w' whose letters are \bullet, x, \bullet . Thus

$$\begin{aligned} g_{\bullet x \bullet} &= \left[\mathbf{f}_{\bullet x \bullet}^{\mathcal{L}}(\bullet \bullet x) \right] e(\bullet) e(\bullet) e(x) + \left[\mathbf{f}_{\bullet x \bullet}^{\mathcal{L}}(\bullet x \bullet) \right] e(\bullet) e(x) e(\bullet) + \left[\mathbf{f}_{\bullet x \bullet}^{\mathcal{L}}(x \bullet \bullet) \right] e(\bullet) e(\bullet) e(x) \\ &= \left[\mathbf{f}_{\bullet x \bullet}^{\mathcal{L}}(\bullet \bullet x) \right] \bullet \bullet e(x) + \left[\mathbf{f}_{\bullet x \bullet}^{\mathcal{L}}(\bullet x \bullet) \right] \bullet e(x) \bullet + \left[\mathbf{f}_{\bullet x \bullet}^{\mathcal{L}}(x \bullet \bullet) \right] e(x) \bullet \bullet. \end{aligned}$$

The graphical calculation of $\mathbf{f}_{\bullet x \bullet}^{\mathcal{L}}$ then shows that $\mathbf{f}_{\bullet x \bullet}^{\mathcal{L}}(\bullet \bullet x) = 1$, $\mathbf{f}_{\bullet x \bullet}^{\mathcal{L}}(\bullet x \bullet) = 0$ and $\mathbf{f}_{\bullet x \bullet}^{\mathcal{L}}(x \bullet \bullet) = 1$, so this gives twice the eigenvector found before. As $\bullet x \bullet$ has two Lyndon factors, the eigenvector $g_{\bullet x \bullet}$ has eigenvalue a^2 .

Proof of Theorem 2.5.1.A. As explained in Sections 2.2 and 2.3 respectively, $e(c_i)$ is an eigenvector of Ψ^a with eigenvalue a , and the product of eigenvectors is again an eigenvector, with the product eigenvalue. Hence $e(c_1)\dots e(c_k)$ is an eigenvector of eigenvalue a^k .

To deduce that $\{e(c_1)\dots e(c_k) \mid k \in \mathbb{N}, \{c_1, \dots, c_k\} \text{ a multiset in } \mathcal{C}\}$ is a basis, it suffices to show that the matrix changing $\{e(c_1)\dots e(c_k) \mid k \in \mathbb{N}, \{c_1, \dots, c_k\} \text{ a multiset in } \mathcal{C}\}$ to $\{c_1 \dots c_k \mid k \in \mathbb{N}, \{c_1, \dots, c_k\} \text{ a multiset in } \mathcal{C}\}$ is uni-triangular, under any ordering which refines the length k . By definition of the Eulerian idempotent map, $e(c_i) = c_i + \text{products}$. So

$$e(c_1)\dots e(c_k) = c_1 \dots c_k + \text{products of at least } k+1 \text{ factors.}$$

Expanding these products in terms of the free generating set \mathcal{C} requires at least $k+1$ c 's in each summand. \square

Proof of Theorem 2.5.1.B. The Symmetrisation Lemma (Theorem 2.3.2) asserts that, if x_1, \dots, x_k are all primitive, then $\sum_{\sigma \in \mathfrak{S}_k} x_{\sigma(1)} \dots x_{\sigma(k)}$ is an eigenvector of Ψ^a of eigenvalue a^k . That these symmetrised products give a basis follows directly from the following two well-known theorems on the structure of Hopf algebras (recall from Section 2.2 that a graded Hopf algebra is conilpotent because $\bar{\Delta}^{[\deg x+1]}(x) = 0$):

Theorem (Cartier-Milnor-Moore). [Car07, Th. 3.8.1] *A connected, conilpotent and co-commutative Hopf algebra \mathcal{H} (over a field of characteristic 0) is isomorphic to $\mathcal{U}(\mathfrak{g})$, the universal enveloping algebra of a Lie algebra \mathfrak{g} , where \mathfrak{g} is the Lie algebra of primitive elements of \mathcal{H} .*

Theorem (Poincare-Birkhoff-Witt, symmetrised version). [Kna02, Prop. 3.23] *If $\{x_1, x_2, \dots\}$ is a basis for a Lie algebra \mathfrak{g} , then the symmetrised products $\sum_{\sigma \in \mathfrak{S}_k} x_{i_{\sigma(1)}} x_{i_{\sigma(2)}} \dots x_{i_{\sigma(k)}}$, for $1 \leq i_1 \leq i_2 \leq \dots \leq i_k$, form a basis for $\mathcal{U}(\mathfrak{g})$.*

\square

Proof of Theorem 2.5.1.A'. Apply Theorem 2.5.1.A, the eigenbasis algorithm for commutative Hopf algebras, with $\{P_w \mid w \text{ Lyndon}\}$ as the free generating set \mathcal{C} , since [Reu93, Th. 6.1.i] asserts that the Lyndon words generate the shuffle algebra freely as a commutative algebra. \square

Proof of Theorem 2.5.1.B'. [ABT13, Prop. 22] shows that $\{g_i | i \text{ a single letter}\}$ generates \mathcal{H} freely. Since each $g_i = e(S_i)$ is primitive, the map $i \rightarrow g_i$ is a Hopf-isomorphism from the free associative algebra to \mathcal{H} . Now, by [Lot97, Th. 5.3.1], the “standard bracketing” of Lyndon words is a basis for the primitive subspace of the free associative algebra, and its image under this Hopf-isomorphism is precisely $\{g_w | w \text{ Lyndon}\}$. So applying Theorem 2.5.1.B to $\mathcal{P} = \{g_w | w \text{ Lyndon}\}$ gives the result.

Here is a second proof employing length-triangularity arguments similar to those in the proof of Theorem 2.5.1.A. First observe that, if x, y are primitive, then so is $[x, y] = xy - yx$:

$$\begin{aligned}
\Delta(xy - yx) &= \Delta(x)\Delta(y) - \Delta(y)\Delta(x) \\
&= (1 \otimes x + x \otimes 1)(1 \otimes y + y \otimes 1) - (1 \otimes y + y \otimes 1)(1 \otimes x + x \otimes 1) \\
&= 1 \otimes xy + y \otimes x + x \otimes y + xy \otimes 1 - (1 \otimes yx + x \otimes y + y \otimes x + yx \otimes 1) \\
&= 1 \otimes xy + xy \otimes 1 - 1 \otimes yx - yx \otimes 1 \\
&= 1 \otimes (xy - yx) + (xy - yx) \otimes 1.
\end{aligned}$$

Applying this argument recursively shows that, for Lyndon w , the vector g_w as defined in the Theorem is indeed primitive. So, by the Symmetrisation Lemma (Theorem 2.3.2), the g_w for general w , which are the symmetrised products of the primitive g_w , are indeed eigenvectors of Ψ^a .

To deduce that these give a basis for \mathcal{H} , it suffices to show that the matrix changing $\{g_w\}$ to the basis $\{s[w]\}$ of [GR89, Th. 5.2] is uni-triangular, under any ordering which refines the length $l(w)$. (Recall that the length $l(w)$ is the number of letters in w). The $\{s[w]\}$ basis is defined recursively as follows:

$$\begin{aligned}
s[w] &:= S_w && \text{if } w \text{ is a single letter;} \\
s[w] &:= s[u_1]s[u_2] - s[u_2]s[u_1] && \text{if } w \text{ is Lyndon with standard factorisation } w = u_1u_2; \\
s[w] &:= \frac{1}{k!} \sum_{\sigma \in \mathfrak{S}_k} s[u_{\sigma(1)}] \dots s[u_{\sigma(k)}] && \text{if } w \text{ has Lyndon factorisation } w = u_1 \dots u_k.
\end{aligned}$$

(For a Lyndon word w , the expression $s[w]$ is known as its *standard bracketing*.) For single-letter words w , $g_w = e(w) = S_w + \text{products}$, by definition of the Eulerian idempotent map.

The recursive definition of both g_w and $s[w]$ show that

$$g_w = s[w] + \text{products of at least } l(w) + 1 \text{ factors.}$$

As in the proof of Theorem 2.5.1.A, expressing these products in the basis $\{S_w\}$ involves words of length at least $l(w) + 1$. It is clear from the definition of $s[w]$ that all S_u appearing in the S -expansion of $s[v]$ have $l(u) = l(v)$, so all $s[v]$ in the s -expansion of these products have $l(v) \geq l(w) + 1$. \square

2.6 Basis for the Eigenspace of Largest Eigenvalue

What are the eigenvectors and eigenvalues of the Hopf-power map Ψ^a on a Hopf algebra that is neither commutative nor cocommutative? The power rule need not hold in this case, so the Eulerian idempotent map may not produce eigenvectors. By the Symmetrisation Lemma (Theorem 2.3.2), the symmetrised products of k primitives are eigenvectors of eigenvalue a^k . Appealing to the Poincare-Birkhoff-Witt theorem on the universal enveloping algebra of the primitives, these symmetrised products can be made linearly independent, but, without cocommutativity, these will in general not span the eigenspace.

Recently [AL13] found the eigenvalues of Ψ^a and their algebraic multiplicities (i.e. the exponents of the factors in the characteristic polynomial) by passing to $\text{gr}(\mathcal{H})$, the associated graded Hopf algebra of \mathcal{H} with respect to the coradical filtration. The key to their argument is a simple linear algebra observation: the eigenvalues and algebraic multiplicities of Ψ^a are the same for \mathcal{H} as for $\text{gr}(\mathcal{H})$. By [AS05a, Prop. 1.6], $\text{gr}(\mathcal{H})$ is commutative, so the eigenbasis algorithm in Theorem 2.5.1.A above applies. So the last assertion of the algorithm gives the following formula:

Theorem 2.6.1. [AL13, Th. 4 and remark in same section] *Let $\mathcal{H} = \bigoplus_{n \geq 0} \mathcal{H}_n$ be a graded connected Hopf algebra over \mathbb{R} , and write b_i for the number of degree i elements in a free generating set of $\text{gr}(\mathcal{H})$. In other words, b_i are the numbers satisfying $\prod_i (1 - x^i)^{-b_i} = \sum_n \dim \text{gr}(\mathcal{H})_n x^n = \sum_n \dim \mathcal{H}_n x^n$. Then the algebraic multiplicity of the eigenvalue a^k for $\Psi^a : \mathcal{H}_n \rightarrow \mathcal{H}_n$ is the coefficient of $x^n y^k$ in $\prod_i (1 - yx^i)^{-b_i}$. Equivalently, this multiplicity is*

the number of ways to choose k elements, unordered and possibly with repetition, out of b_i elements in degree i , subject to the condition that their degrees sum to n . \square

Remarks.

1. The proof in [AL13] applies the Poincare-Birkhoff-Witt theorem to the dual of $\text{gr}(\mathcal{H})$, instead of appealing to the eigenbasis algorithm on commutative Hopf algebras.
2. Explicit calculations on $FQSym$, the Malvenuto-Reutenauer Hopf algebra of permutations [MR95, Th. 3.3; AS05b] show that Ψ^a need not be diagonalisable on a non-commutative, noncocommutative Hopf algebra - in other words, there are non-trivial Jordan blocks.

Happily, in the special case $k = n$ (corresponding to the largest eigenvalue), this multiplicity formula implies that the Symmetrisation Lemma indeed builds all eigenvectors of eigenvalue a^n , provided $\mathcal{H}_1 \neq \emptyset$:

Theorem 2.6.2. *Let $\mathcal{H} = \bigoplus_{n \geq 0} \mathcal{H}_n$ be a graded connected Hopf algebra over \mathbb{R} . Suppose $\mathcal{H}_1 \neq \emptyset$, and let \mathcal{B}_1 be a basis of \mathcal{H}_1 . Then a^n is the largest eigenvalue of the Hopf-power map Ψ^a on \mathcal{H}_n , and the corresponding eigenspace has basis*

$$\mathcal{E} := \left\{ \sum_{\sigma \in \mathfrak{S}_n} c_{\sigma(1)} \cdots c_{\sigma(n)} \mid \{c_1, \dots, c_n\} \text{ a multiset in } \mathcal{B}_1 \right\}.$$

As Theorem 4.5.1 below shows, this identifies all stationary distributions of a Hopf-power Markov chain.

Proof. For each monomial $x^n y^k$ in the generating function $\prod_i (1 - yx^i)^{-b_i}$ of Theorem 2.6.1, it must be that $k \leq n$. Hence all eigenvalues a^k of Ψ^a on \mathcal{H}_n necessarily have $k \leq n$, and thus a^n is the largest possible eigenvalue.

Next observe that, since the c_i each have degree 1, they are necessarily primitive. So $\sum_{\sigma \in \mathfrak{S}_n} c_{\sigma(1)} \cdots c_{\sigma(n)}$ is a symmetrised product of n primitives, which the Symmetrisation Lemma (Theorem 2.3.2) asserts is an eigenvector of Ψ^a of eigenvalue a^n . Working in

the universal enveloping algebra of \mathcal{H}_1 , the Poincare-Birkhoff-Witt theorem gives linear independence of $\{\sum_{\sigma \in \mathfrak{S}_n} c_{\sigma(1)} \cdots c_{\sigma(n)}\}$ across all choices of multisets $\{c_1, \dots, c_n\} \subseteq \mathcal{B}_1$.

To conclude that the set \mathcal{E} of symmetrised products span the a^n -eigenspace, it suffices to show that $|\mathcal{E}|$ is equal to the algebraic multiplicity of the eigenvalue a^n as specified by Theorem 2.6.1. Clearly $|\mathcal{E}| = \binom{|\mathcal{B}_1|+n-1}{n}$, the number of ways to choose n unordered elements, allowing repetition, from \mathcal{B}_1 . On the other hand, the algebraic multiplicity is $\binom{b_1+n-1}{n}$, since choosing n elements whose degrees sum to n constrains each element to have degree 1. By equating the coefficient of x in the equality $\prod_i (1-x^i)^{-b_i} = \sum_n \dim \mathcal{H}_n x^n$, it is clear that $b_1 = \dim \mathcal{H}_1 = |\mathcal{B}_1|$. So $|\mathcal{E}|$ is indeed the algebraic multiplicity of the eigenvalue a^n . \square

The condition $\mathcal{H}_1 \neq \emptyset$ is satisfied for the vast majority of combinatorial Hopf algebras, so this thesis will not require the analogous, clumsier, result for general \mathcal{H} , though I include it below for completeness. To determine the highest eigenvalue, first define the sets $\mathcal{D} := \{d > 0 \mid \mathcal{H}_d \neq \emptyset\}$, and $\mathcal{D}' = \{d \in \mathcal{D} \mid d \neq d_1 + d_2 \text{ with } d_1, d_2 \in \mathcal{D}\}$. In the familiar case where $\mathcal{D} = \{1, 2, 3, \dots\}$, the set \mathcal{D}' is $\{1\}$. It is possible to build Hopf algebras with \mathcal{D} being any additively-closed set - for example, take a free associative algebra with a generator in degree d for each $d \in \mathcal{D}$, and let all these generators be primitive. The reason for considering \mathcal{D}' is that $\bigoplus_{d \in \mathcal{D}'} \mathcal{H}_d$ consists solely of primitives: for $x \in \mathcal{H}_d$, the counit axiom mandates that $\bar{\Delta}(x) \in \bigoplus_{d_1+d_2=d} \mathcal{H}_{d_1} \otimes \mathcal{H}_{d_2}$, and this direct sum is empty if $d \in \mathcal{D}'$. However, there may well be primitives in higher degrees.

For a fixed degree $n \in \mathcal{D}$, define a \mathcal{D}' -partition of n to be an unordered tuple $\lambda := (\lambda_1, \dots, \lambda_{l(\lambda)})$ such that each $\lambda_i \in \mathcal{D}'$ and $\lambda_1 + \dots + \lambda_{l(\lambda)} = n$. The parts λ_i need not be distinct. Then $l(\lambda)$ is the length of λ . (The analogous notion of a \mathcal{D} -partition will be useful in the proof of Theorem 2.6.4.)

Example 2.6.3. Suppose $\mathcal{D} = \{5, 6, 7, 9, 10, 11, \dots\} = \mathbb{N} \setminus \{1, 2, 3, 4, 8\}$, so $\mathcal{D}' = \{5, 6, 7, 9\}$. There are four \mathcal{D}' -partitions of 23: $(6, 6, 6, 5)$, $(7, 6, 5, 5)$, $(9, 7, 7)$, $(9, 9, 5)$. These have length 4, 4, 3, 3 respectively.

Theorem 2.6.4. Let $\mathcal{H} = \bigoplus_{n \in \mathcal{D}} \mathcal{H}_n$ be a graded connected Hopf algebra over \mathbb{R} . Then the highest eigenvalue of the Hopf-power map Ψ^a on \mathcal{H}_n is $a^{K(n)}$, where $K(n)$ denotes the

maximal length of a \mathcal{D}' -partition of n . A basis for the corresponding eigenspace is

$$\mathcal{E} := \left\{ \sum_{\sigma \in \mathfrak{S}_K} c_{\sigma(1)} \cdots c_{\sigma(K)} \mid \{c_1, \dots, c_K\} \text{ a multiset in } \mathcal{B} \text{ with } \deg c_1 + \cdots + \deg c_K = n \right\}.$$

More explicitly, for each \mathcal{D}' -partition λ of n of the maximal length K , set

$$\mathcal{E}_\lambda := \left\{ \sum_{\sigma \in \mathfrak{S}_K} c_{\sigma(1)} \cdots c_{\sigma(K)} \mid \begin{array}{l} \{c_1, \dots, c_{m_1}\} \text{ a multiset in } \mathcal{B}_1, \\ \{c_{m_1+1}, \dots, c_{m_1+m_2}\} \text{ a multiset in } \mathcal{B}_2, \dots \end{array} \right\},$$

where m_i is the number of parts of size i in λ . Then $\mathcal{E} = \coprod \mathcal{E}_\lambda$, over all \mathcal{D}' -partitions λ of n having length K .

Example 2.6.5. Continue from Example 2.6.3. In degree 23, the highest eigenvalue of Ψ^a is a^4 , and its corresponding eigenspace has basis $\mathcal{E}_{(6,6,6,5)} \amalg \mathcal{E}_{(7,6,5,5)}$, where

$$\mathcal{E}_{(6,6,6,5)} := \left\{ \sum_{\sigma \in \mathfrak{S}_4} c_{\sigma(1)} c_{\sigma(2)} c_{\sigma(3)} c_{\sigma(4)} \mid \begin{array}{l} c_1 \in \mathcal{B}_5, \\ \{c_2, c_3, c_4\} \text{ a multiset in } \mathcal{B}_6 \end{array} \right\},$$

$$\mathcal{E}_{(7,6,5,5)} := \left\{ \sum_{\sigma \in \mathfrak{S}_4} c_{\sigma(1)} c_{\sigma(2)} c_{\sigma(3)} c_{\sigma(4)} \mid \begin{array}{l} \{c_1, c_2\} \text{ a multiset in } \mathcal{B}_5, \\ c_3 \in \mathcal{B}_6, c_4 \in \mathcal{B}_7 \end{array} \right\}.$$

Proof. The argument below is essentially a more careful version of the proof of Theorem 2.6.2.

By Theorem 2.6.1, a^k is an eigenvalue of $\Psi^a : \mathcal{H}_n \rightarrow \mathcal{H}_n$ if and only if there are k elements in \mathcal{H} whose degrees sum to n . In other words, a^k is an eigenvalue precisely when there is a \mathcal{D} -partition of n of length k . Note that a \mathcal{D} -partition of n with maximal length must be a \mathcal{D}' -partition: if a part λ_i of λ is not in \mathcal{D}' , then $\lambda_i = d_1 + d_2$ with $d_1, d_2 \in \mathcal{D}$, and replacing λ_i with two parts d_1, d_2 in λ creates a longer partition. Hence the largest eigenvalue of $\Psi^a : \mathcal{H}_n \rightarrow \mathcal{H}_n$ corresponds to the maximal length of a \mathcal{D}' -partition of n .

As observed earlier, every element of $\bigoplus_{d \in \mathcal{D}'} \mathcal{H}_d$ is primitive, by degree considerations. So each element in \mathcal{E} is a symmetrised product of K primitives; by the Symmetrisation Lemma (Theorem 2.3.2), they are eigenvectors of Ψ^a of eigenvalue a^K . As before, applying

the Poincare-Birkhoff-Witt theorem to the universal enveloping algebra of $\bigoplus_{d \in \mathcal{D}'} \mathcal{H}_d$ gives linear independence of \mathcal{E} .

It remains to show that \mathcal{E} spans the a^K -eigenspace. The dimension-counting argument which closes the proof of Theorem 2.6.2 will function, so long as $b_i = |\mathcal{B}_i|$ for each $i \in \mathcal{D}'$. Recall that b_i is defined by $\prod_i (1 - x^i)^{-b_i} = \sum_{d \in \mathcal{D}} |\mathcal{B}_d| x^d$. Equating coefficients of x^d for $d \notin \mathcal{D}$ shows that $b_d = 0$ for $d \notin \mathcal{D}$, so the left hand side is $\prod_{i \in \mathcal{D}} (1 - x^i)^{-b_i}$. Now, for each $i \in \mathcal{D}'$, there is no $d_1, d_2 \in \mathcal{D}$ with $i = d_1 + d_2$, so the coefficient of x^i in $\prod_{i \in \mathcal{D}} (1 - x^i)^{-b_i}$ is b_i . \square

Chapter 3

Markov chains from linear operators

As outlined previously in Section 1.3, one advantage of relating riffle-shuffling to the Hopf-square map on the shuffle algebra is that Hopf algebra theory supplies the eigenvalues and eigenvectors of the transition matrix. Such a philosophy applies whenever the transition matrix is the matrix of a linear operator. Although this thesis treats solely the case where this operator is the Hopf-power, some arguments are cleaner in the more general setting, as presented in this chapter. The majority of these results have appeared in the literature under various guises.

Section 3.1 explains how the Doob transform normalises a linear operator to obtain a transition matrix. Then Sections 3.2, 3.3, 3.4 connect the eigenbasis, stationary distribution and time-reversal, and projection of this class of chains respectively to properties of its originating linear map.

A few pieces of notation: in this chapter, all vector spaces are finite-dimensional over \mathbb{R} . For a linear map $\theta : V \rightarrow W$, and bases $\mathcal{B}, \mathcal{B}'$ of V, W respectively, $[\theta]_{\mathcal{B}, \mathcal{B}'}$ will denote the matrix of θ with respect to \mathcal{B} and \mathcal{B}' . In other words, the entries of $[\theta]_{\mathcal{B}, \mathcal{B}'}$ satisfy

$$\theta(v) = \sum_{w \in \mathcal{B}'} [\theta]_{\mathcal{B}, \mathcal{B}'}(w, v)w$$

for each $v \in \mathcal{B}$. When $V = W$ and $\mathcal{B} = \mathcal{B}'$, shorten this to $[\theta]_{\mathcal{B}}$. The *transpose* of a matrix A is given by $A^T(x, y) := A(y, x)$. The *dual vector space* to V , written V^* , is the set of linear functions from V to \mathbb{R} . If \mathcal{B} is a basis for V , then the natural basis to use for

V^* is $\mathcal{B}^* := \{x^* | x \in \mathcal{B}\}$, where x^* satisfies $x^*(x) = 1$, $x^*(y) = 0$ for all $y \in \mathcal{B}$, $y \neq x$. In other words, x^* is the linear extension of the indicator function on x . When elements of V are expressed as column vectors, it is often convenient to view these functions as row vectors, so that evaluation on an element of V is given by matrix multiplication. The *dual map* to $\theta : V \rightarrow W$ is the linear map $\theta^* : W^* \rightarrow V^*$ satisfying $(\theta^* f)(v) = f(\theta v)$. Note that $[\theta^*]_{\mathcal{B}^*, \mathcal{B}^*} = [\theta]_{\mathcal{B}, \mathcal{B}'}^T$.

3.1 Construction

The starting point is as follows: V is a vector space with basis \mathcal{B} , and $\Psi : V \rightarrow V$ is a linear map. Suppose the candidate transition matrix $K := [\Psi]_{\mathcal{B}}^T$ has all entries non-negative, but its rows do not necessarily sum to 1.

One common way to resolve this is to divide each entry of K by the sum of the entries in its row. This is not ideal for the present situation since the outcome is no longer a matrix for Ψ . For example, an eigenbasis of Ψ will not give the eigenfunctions of the resulting matrix.

A better solution comes in the form of Doob's h -transform. This is usually applied to a transition matrix with the row and column corresponding to an absorbing state removed, to obtain the transition matrix of the chain conditioned on non-absorption. Hence some of the references listed in Theorem 3.1.1 below assume that K is *sub-Markovian* (i.e. $\sum_y K(x, y) < 1$), but, as the calculation in the proof shows, that is unnecessary.

The Doob transform works in great generality, for continuous-time Markov chains on general state spaces. In the present discrete case, it relies on an eigenvector η of the dual map Ψ^* , that takes only positive values on the basis \mathcal{B} . Without imposing additional constraints on Ψ (which will somewhat undesirably limit the scope of this theory), the existence of such an eigenvector η is not guaranteed. Even when η exists, it may not be unique in any reasonable sense, and different choices of η will in general lead to different Markov chains. However, when Ψ is a Hopf-power map, there is a preferred choice of η , given by Definition 4.3.1. Hence this thesis will suppress the dependence of this construction on the eigenvector η .

Theorem 3.1.1 (Doob h -transform for non-negative linear maps). [Gan59, Sec. XIII.6.1; KSK66, Def. 8.11, 8.12; Zho08, Lemma 4.4.1.1; LPW09, Sec.17.6.1; Swa12, Lem. 2.7] Let V be a vector space with basis \mathcal{B} , and $\Psi : V \rightarrow V$ be a non-zero linear map for which $K := [\Psi]_{\mathcal{B}}^T$ has all entries non-negative. Suppose there is an eigenvector η of the dual map Ψ^* taking only positive values on \mathcal{B} , and let β be the corresponding eigenvalue. Then

$$\check{K}(x, y) := \frac{1}{\beta} K(x, y) \frac{\eta(y)}{\eta(x)}$$

defines a transition matrix. Equivalently, $\check{K} := \left[\frac{\Psi}{\beta} \right]_{\check{\mathcal{B}}}^T$, where $\check{\mathcal{B}} := \left\{ \check{x} := \frac{x}{\eta(x)} \mid x \in \mathcal{B} \right\}$.

Call the resulting chain a Ψ -Markov chain on \mathcal{B} (neglecting the dependence on its rescaling function η as discussed previously). See Example 4.3.6 for a numerical illustration of this construction.

Proof. First note that $K := [\Psi^*]_{\mathcal{B}^*}$, so $\Psi^* \eta = \beta \eta$ translates to $\sum_y K(x, y) \eta(y) = \beta \eta(x)$. (Functions satisfying this latter condition are called *harmonic*, hence the name h -transform.) Since $\eta(y) > 0$ for all y , $K(x, y) \geq 0$ for all x, y and $K(x, y) > 0$ for some x, y , the eigenvalue β must be positive. So $\check{K}(x, y) \geq 0$. It remains to show that the rows of \check{K} sum to 1:

$$\sum_y \check{K}(x, y) = \frac{\sum_y K(x, y) \eta(y)}{\beta \eta(x)} = \frac{\beta \eta(x)}{\beta \eta(x)} = 1.$$

□

Remarks.

1. β , the eigenvalue of η , is necessarily the largest eigenvalue of Ψ . Here's the reason: by the Perron-Frobenius theorem for non-negative matrices [Gan59, Ch. XIII Th. 3], there is an eigenvector ξ of Ψ , with largest eigenvalue β_{\max} , whose components are all non-negative. As η has all components positive, the matrix product $\eta^T \xi$ results in a positive number. But $\beta \eta^T \xi = (\Psi^* \eta)^T \xi = \eta^T (\Psi \xi) = \beta_{\max} \eta^T \xi$, so $\beta = \beta_{\max}$.

2. Rescaling the basis \mathcal{B} does not change the chain: suppose $\mathcal{B}' = \{x' := \alpha_x x \mid x \in \mathcal{B}\}$ for some non-zero constants α_x . Then, since η is a linear function,

$$\check{x}' := \frac{x'}{\eta(x')} = \frac{\alpha_x x}{\alpha_x \eta(x)} = \check{x}.$$

Hence the transition matrix for both chains is the transpose of the matrix of Ψ with respect to the same basis. This is used in Theorem 3.3.3 to give a condition under which the chain is reversible.

3. In the same vein, if η' is a multiple of η , then both eigenvectors η' and η give rise to the same Ψ -Markov chain, since the transition matrix depends only on the ratio $\frac{\eta(y)}{\eta(x)}$.

3.2 Diagonalisation

Recall that the main reason for defining the transition matrix \check{K} to be the transpose of a matrix for some linear operator Ψ is that it reduces the diagonalisation of the Markov chain to identifying the eigenvectors of Ψ and its dual Ψ^* . Proposition 3.2.1 below records precisely the relationship between the left and right eigenfunctions of the Markov chain and these eigenvectors; it is immediate from the definition of \check{K} above.

Proposition 3.2.1 (Eigenfunctions of Ψ -Markov chains). [*Zho08, Lemma 4.4.1.4; Swa12, Lem. 2.11*] *Let V be a vector space with basis \mathcal{B} , and $\Psi : V \rightarrow V$ be a linear operator allowing the construction of a Ψ -Markov chain (whose transition matrix is $\check{K} := \left[\frac{\Psi}{\beta} \right]_{\check{\mathcal{B}}}^T$, where $\check{\mathcal{B}} := \left\{ \check{x} := \frac{x}{\eta(x)} \mid x \in \mathcal{B} \right\}$). Then:*

(L) *Given a function $\mathbf{g} : \mathcal{B} \rightarrow \mathbb{R}$, define a vector $g \in V$ by*

$$g := \sum_{x \in \mathcal{B}} \frac{\mathbf{g}(x)}{\eta(x)} x.$$

Then \mathbf{g} is a left eigenfunction, of eigenvalue β' , for this Ψ -Markov chain if and only if g is an eigenvector, of eigenvalue $\beta\beta'$, of Ψ . Consequently, given a basis $\{g_i\}$ of V

with $\Psi g_i = \beta_i g_i$, the set of functions

$$\{\mathbf{g}_i(x) := \text{coefficient of } x \text{ in } \eta(x)g_i\}$$

is a basis of left eigenfunctions for the Ψ -Markov chain, with $\sum_x \check{K}(x, y)\mathbf{g}(x) = \frac{\beta_i}{\beta} \mathbf{g}(y)$ for all y .

(R) Given a function $\mathbf{f}: \mathcal{B} \rightarrow \mathbb{R}$, define a vector f in the dual space V^* by

$$f := \sum_{x \in \mathcal{B}} \mathbf{f}(x)\eta(x)x^*.$$

Then \mathbf{f} is a right eigenfunction, of eigenvalue β' , for this Ψ -Markov chain if and only if f is an eigenvector, of eigenvalue $\beta\beta'$, of the dual map Ψ^* . Consequently, given a basis $\{f_i\}$ of V^* with $\Psi^* f_i = \beta_i f_i$, the set of functions

$$\left\{ \mathbf{f}_i(x) := \frac{1}{\eta(x)} f_i(x) \right\}$$

is a basis of right eigenfunctions for the Ψ -Markov chain, with $\sum_x \check{K}(x, y)\mathbf{f}(y) = \frac{\beta_i}{\beta} \mathbf{f}(x)$ for all x . \square

Remark. In the Markov chain literature, the term “left eigenvector” is often used interchangeably with “left eigenfunction”, but this thesis will be careful to make a distinction between the eigenfunction $\mathbf{g}: \mathcal{B} \rightarrow \mathbb{R}$ and the corresponding eigenvector $g \in V$ (and similarly for right eigenfunctions).

3.3 Stationarity and Reversibility

Recall from Section 1.1 that, for a Markov chain with transition matrix K , a *stationary distribution* $\pi(x)$ is one which satisfies $\sum_x \pi(x)K(x, y) = \pi(y)$, or, if written as a row vector, $\pi K = \pi$. So it is a left eigenfunction of eigenvalue 1. These are of interest as they include all possible limiting distributions of the chain. The following Proposition is essentially a specialisation of Proposition 3.2.1.L to the case $\beta' = 1$:

Proposition 3.3.1 (Stationary Distributions of Ψ -Markov chains). [Zho08, Lemma 4.4.1.2; Swa12, Lem. 2.16] *Work in the setup of Theorem 3.1.1. The stationary distributions π of a Ψ -Markov chain are in bijection with the eigenvectors $\xi = \sum_{x \in \mathcal{B}} \xi_x x$ of the linear map Ψ of eigenvalue β , which have $\xi_x \geq 0$ for all $x \in \mathcal{B}$, and are scaled so $\eta(\xi) = \sum_x \eta(x) \xi_x = 1$. The bijection is given by $\pi(x) = \eta(x) \xi_x$. \square*

Observe that a stationary distribution always exists: as remarked after Theorem 3.1.1, β is the largest eigenvalue of Ψ , and the Perron-Frobenius theorem guarantees a corresponding eigenvector with all entries non-negative. Rescaling this then gives a ξ satisfying the conditions of the Proposition.

For the rest of this section, assume that β has multiplicity 1 as an eigenvalue of Ψ , so there is a unique stationary distribution π and corresponding eigenvector ξ of the linear map Ψ . (Indeed, Proposition 3.3.1 above asserts that β having multiplicity 1 is also the necessary condition.) Assume in addition that $\pi(x) > 0$ for all $x \in \mathcal{B}$. Then, there is a well-defined notion of the Markov chain run backwards; that is, one can construct a stochastic process $\{X_m^*\}$ for which

$$P\{X_0^* = x_i, X_1^* = x_{i-1}, \dots, X_i^* = x_0\} = P\{X_0 = x_0, X_1 = x_1, \dots, X_i = x_i\}$$

for every i . As [LPW09, Sec. 1.6] explains, if the original Markov chain *started in stationarity* (i.e. $P(X_0 = x) = \pi(x)$), then this reversed process is also a Markov chain - the formal *time-reversal* chain - with transition matrix

$$K^*(x, y) = \frac{\pi(y)}{\pi(x)} K(y, x).$$

Theorem 3.3.2 below shows that, if the forward chain is built from a linear map via the Doob transform, then its time-reversal corresponds to the dual map.

Theorem 3.3.2 (Time-reversal of a Ψ -Markov chain). *Work in the framework of Theorem 3.1.1. If the time-reversal of a Ψ -Markov chain is defined, then it arises from applying the Doob transform to the linear-algebraic-dual map $\Psi^* : V^* \rightarrow V^*$ with respect to the dual basis \mathcal{B}^* .*

Proof. Let K^* denote the transpose of the matrix of Ψ^* with respect to the basis \mathcal{B}^* . Then $K^*(x^*, y^*) = K(y, x)$. By definition, the transition matrix of a Ψ^* -Markov chain is

$$\check{K}^*(x^*, y^*) = \frac{K^*(x^*, y^*) \eta^*(y^*)}{\beta^* \eta^*(x^*)},$$

where η^* is an eigenvector of the dual map to Ψ^* with $\eta^*(x^*) > 0$ for all $x^* \in \mathcal{B}^*$, and β^* is its eigenvalue. Identify the dual map to Ψ^* with Ψ ; then ξ is such an eigenvector, since the condition $\pi(x) > 0$ for the existence of a time-reversal is equivalent to $\xi(x^*) = \xi_x > 0$. Then $\beta^* = \beta$, so

$$\begin{aligned} \check{K}^*(x^*, y^*) &= \frac{K^*(x^*, y^*) \xi_y}{\beta \xi_x} \\ &= \frac{K(y, x) \xi_y \eta(y) \eta(x)}{\beta \xi_x \eta(x) \eta(y)} \\ &= \frac{\pi(y) K(y, x) \eta(x)}{\pi(x) \beta \eta(y)} \\ &= \frac{\pi(y)}{\pi(x)} \check{K}(y, x). \end{aligned}$$

□

Remark. This time-reversed chain is in fact the only possible Ψ^* -Markov chain on \mathcal{B}^* ; all possible rescaling functions η^* give rise to the same chain. Here is the reason: as remarked after Theorem 3.1.1, a consequence of the Perron-Frobenius theorem is that all eigenvectors with all coefficients positive must correspond to the largest eigenvalue. Here, the existence of a time-reversal constrains this eigenvalue to have multiplicity 1, so any other choice of η^* must be a multiple of ξ , hence defining the same Ψ^* -Markov chain on \mathcal{B}^* .

Markov chains that are *reversible*, that is, equal to their own time-reversal, are particularly appealing as they admit more tools of analysis. It is immediate from the definition of the time-reversal that the necessary and sufficient conditions for reversibility are $\pi(x) > 0$ for all x in the state space, and the *detailed balance equation* $\pi(x)K(x, y) = \pi(y)K(y, x)$. Thanks to Theorem 3.3.2, a Ψ -Markov chain is reversible if and only if $[\Psi]_{\mathcal{B}} = [\Psi^*]_{\mathcal{B}^*}$. As the right hand side is $[\Psi]_{\mathcal{B}}^T$, this equality is equivalent to $[\Psi]_{\mathcal{B}}$ being a symmetric matrix. A less coordinate-dependent rephrasing is that Ψ is self-adjoint with respect to some inner

product where the basis \mathcal{B} is orthonormal. Actually, it suffices to require that the vectors in \mathcal{B} are pairwise orthogonal; the length of the vectors are unimportant since, as remarked after Theorem 3.1.1, all rescalings of a basis define the same chain. To summarise:

Theorem 3.3.3. *Let V be a vector space with an inner product, and \mathcal{B} a basis of V consisting of pairwise orthogonal vectors. Suppose $\Psi : V \rightarrow V$ is a self-adjoint linear map admitting the construction of a Ψ -Markov chain on \mathcal{B} , and that this chain has a unique stationary distribution, which happens to take only positive values. Then this chain is reversible.* \square

3.4 Projection

Sometimes, one is interested only in one particular feature of a Markov chain. A classic example from [ADS11] is shuffling cards for a game of Black-Jack, where the suits of the cards are irrelevant. In the same paper, they also study the position of the ace of spades. In situations like these, it makes sense to study the *projected* process $\{\theta(X_m)\}$ for some function θ on the state space, rather than the original chain $\{X_m\}$. Since θ effectively merges several states into one, the process $\{\theta(X_m)\}$ is also known as the *lumping* of $\{X_m\}$ under θ .

Since the projection $\{\theta(X_m)\}$ is entirely governed by $\{X_m\}$, information about $\{\theta(X_m)\}$ can shed some light on the behaviour of $\{X_m\}$. For example, the convergence rate of $\{\theta(X_m)\}$ is a lower bound for the convergence rate of $\{X_m\}$. So, when $\{X_m\}$ is too complicated to analyse, one may hope that some $\{\theta(X_m)\}$ is more tractable - after all, its state space is smaller. For chains on algebraic structures, quotient structures often provide good examples of projections. For instance, if $\{X_m\}$ is a random walk on a group, then θ can be a group homomorphism. Section 4.7 will show that the same applies to Hopf-power Markov chains.

In the ideal scenario, the projection $\{\theta(X_m)\}$ is itself a Markov chain also. As explained in [KS60, Sec. 6.3], $\{\theta(X_m)\}$ is a Markov chain for any starting distribution if and only if the sum of probabilities $\sum_{y:\theta(y)=\bar{y}} K(x, y)$ depends only on $\theta(x)$, not on x . This condition is commonly known as *Dynkin's criterion*. (Weaker conditions suffice if one desires $\{\theta(X_m)\}$

to be Markov only for particular starting distributions, see [KS60, Sec. 6.4].) Writing \bar{x} for $\theta(x)$, the chain $\{\theta(X_m)\}$ then has transition matrix

$$\bar{K}(\bar{x}, \bar{y}) = \sum_{y:\theta(y)=\bar{y}} K(x, y) \quad \text{for any } x \text{ with } \theta(x) = \bar{x}.$$

Equivalently, as noted in [KS60, Th. 6.3.4], if R is the matrix with 1 in positions $x, \theta(x)$ for all x , and 0 elsewhere, then $KR = R\bar{K}$.

To apply this to chains from linear maps, take $\theta : V \rightarrow \bar{V}$ to be a linear map and suppose θ sends the basis \mathcal{B} of V to a basis $\bar{\mathcal{B}}$ of \bar{V} . (θ must be surjective, but need not be injective - several elements of \mathcal{B} may have the same image in \bar{V} , as long as the distinct images are linearly independent.) Then the matrix R above is $[\theta]_{\mathcal{B}, \bar{\mathcal{B}}}^T$. Recall that $K = [\Psi]_{\mathcal{B}}^T$, and let $\bar{K} = [\bar{\Psi}]_{\bar{\mathcal{B}}}^T$ for some linear map $\bar{\Psi} : \bar{V} \rightarrow \bar{V}$. Then the condition $KR = R\bar{K}$ is precisely $[\theta\Psi]_{\mathcal{B}, \bar{\mathcal{B}}}^T = [\bar{\Psi}\theta]_{\mathcal{B}, \bar{\mathcal{B}}}^T$. A θ satisfying this type of relation is commonly known as an *intertwining map*. So, if K, \bar{K} are transition matrices, then $\theta\Psi = \bar{\Psi}\theta$ guarantees that the chain built from Ψ lumps to the chain built from $\bar{\Psi}$.

When K is not a transition matrix, so the Doob transform is non-trivial, an extra hypothesis is necessary:

Theorem 3.4.1. *Let V, \bar{V} be vector spaces with bases $\mathcal{B}, \bar{\mathcal{B}}$, and let $\Psi : V \rightarrow V, \bar{\Psi} : \bar{V} \rightarrow \bar{V}$ be linear maps allowing the Markov chain construction of Theorem 3.1.1, using dual eigenvectors $\eta, \bar{\eta}$ respectively. Let $\theta : V \rightarrow \bar{V}$ be a linear map with $\theta(\mathcal{B}) = \bar{\mathcal{B}}$ and $\theta\Psi = \bar{\Psi}\theta$. Suppose in addition that at least one of the following holds:*

- (i) *all entries of $[\Psi]_{\mathcal{B}}$ are positive;*
- (ii) *the largest eigenvalue of Ψ has multiplicity 1;*
- (iii) *for all $x \in \mathcal{B}$, $\bar{\eta}(\theta(x)) = \alpha\eta(x)$ for some constant $\alpha \neq 0$*

Then θ defines a projection of the Ψ -Markov chain to the $\bar{\Psi}$ -Markov chain.

Remark. Condition iii is the weakest of the three hypotheses, and the only one relevant to the rest of the thesis, as there is an easy way to verify it on Hopf-power Markov chains. This then leads to Theorem 4.7.1, the Projection Theorem of Hopf-power Markov chains. Hypotheses i and ii are potentially useful when there is no simple expression for $\eta(x)$.

Proof. Let $\beta, \bar{\beta}$ be the largest eigenvalues of $\Psi, \bar{\Psi}$ respectively. The equality $[\theta\Psi]_{\check{\mathcal{B}}, \check{\mathcal{B}}}^T = [\bar{\Psi}\theta]_{\check{\mathcal{B}}, \check{\mathcal{B}}}^T$ gives $(\beta\check{K})\check{R} = \check{R}(\bar{\beta}\check{K})$, where $\check{R} = [\theta]_{\check{\mathcal{B}}, \check{\mathcal{B}}}^T$. The goal is to recover $\check{K}R = R\check{K}$ from this: first, show that $\beta = \bar{\beta}$, then, show that $\check{R} = \alpha R$.

To establish that the top eigenvalues are equal, appeal to [Pik13, Thms. 1.3.1.2, 1.3.1.3], which in the present linear-algebraic notation reads: (the asterisks denote taking the linear-algebraic dual map)

Proposition 3.4.2.

- (i) If \bar{f} is an eigenvector of $\bar{\Psi}^*$ with eigenvalue β' , then $f := \theta^* \bar{f}$ (i.e. $f(x) = \bar{f}(\bar{x})$), if non-zero, is an eigenvector of Ψ^* with eigenvalue β' .
- (ii) If g is an eigenvector of Ψ with eigenvalue β'' , then $\bar{g} := \theta g$ (i.e. $\bar{g}_{\bar{x}} = \sum_{x|\theta(x)=\bar{x}} g_x$), if non-zero, is an eigenvector of $\bar{\Psi}$ with eigenvalue β'' . □

So it suffices to show that $\theta^* \bar{f} \neq 0$ for at least one eigenvector \bar{f} of $\bar{\Psi}^*$ with eigenvalue $\bar{\beta}$, and $\theta g \neq 0$ for at least one eigenvector g of Ψ with eigenvalue β . Since \bar{f} is non-zero, it is clear that $f(x) = \bar{f}(\bar{x}) \neq 0$ for some x . As for g , the Perron-Frobenius theorem guarantees that each component of g is non-negative, and since some component of g is non-zero, $\bar{g}_{\bar{x}} = \sum_{x|\theta(x)=\bar{x}} g_x$ is non-zero for some \bar{x} .

Now show $\check{R} = \alpha R$. Recall that $\check{R} = [\theta]_{\check{\mathcal{B}}, \check{\mathcal{B}}}^T$, so its $x, \theta(x)$ entry is $\frac{\bar{\eta}(\bar{x})}{\eta(x)}$. The corresponding entries of R are all 1, and all other entries of both \check{R} and R are zero. So hypothesis iii exactly ensures that $\check{R} = \alpha R$. Hypothesis i clearly implies hypothesis ii via the Perron-Frobenius theorem. To see that hypothesis ii implies hypothesis iii, use Proposition 3.4.2.i in the above paragraph: the composite function $\bar{\eta}\theta$, sending x to $\bar{\eta}(\bar{x})$, is a non-zero eigenvector of Ψ^* with eigenvalue $\bar{\beta} = \beta$; as this eigenvalue has multiplicity 1, it must be some multiple of η . □

Chapter 4

Construction and Basic Properties of Hopf-power Markov Chains

This chapter covers all theory of Hopf-power Markov chains that do not involve diagonalisation, and does not require commutativity or cocommutativity. The goal is the following routine for initial analysis of a Hopf-power Markov chain:

- (Definition 4.3.3) discern whether the given Hopf algebra \mathcal{H} and basis \mathcal{B} are suitable for building a Hopf-power Markov chain (whether \mathcal{B} satisfies the conditions of a state space basis);
- (Definition 4.3.4) build the Hopf-power Markov chain;
- (Definition 4.3.1) calculate the rescaling function η ;
- (Theorem 4.4.1) describe the chain combinatorially without using the Hopf algebra structure;
- (Theorem 4.5.1) obtain its stationary distributions;
- (Theorem 4.6.1) describe the time-reversal of this process.

Two examples will be revisited throughout Sections 4.3-4.6 to illustrate the main theorems, building the following two blurbs step by step.

Example (Riffle-shuffling). The shuffle algebra \mathcal{S} has basis \mathcal{B} consisting of words. The product of two words is the sum of their interleavings, and the coproduct is deconcatenation (Example 4.1.1). The rescaling function is the constant function 1; in other words, no rescaling is necessary to create the associated Markov chain (Example 4.3.5). The a th Hopf-power Markov chain is the Bayer-Diaconis a -handed generalisation of the GSR riffle-shuffle (Example 4.4.2):

1. Cut the deck multinomially into a piles.
2. Interleave the a piles with uniform probability.

Its stationary distribution is the uniform distribution (Example 4.5.3). Its time-reversal is inverse-shuffling (Example 4.6.2):

1. With uniform probability, assign each card to one of a piles, keeping the cards in the same relative order.
2. Place the first pile on top of the second pile, then this combined pile on top of the third pile, etc.

Example (Restriction-then-induction). Let \mathcal{H} be the vector space spanned by representations of the symmetric groups \mathfrak{S}_n , over all $n \in \mathbb{N}$. Let \mathcal{B} be the basis of irreducible representations. The product of representations of \mathfrak{S}_n and \mathfrak{S}_m is the induction of their external product to \mathfrak{S}_{n+m} , and the coproduct of a representation of \mathfrak{S}_n is the sum of its restrictions to $\mathfrak{S}_i \times \mathfrak{S}_{n-i}$ for $0 \leq i \leq n$ (Example 4.1.4). For any irreducible representation x , the rescaling function $\eta(x)$ evaluates to its dimension $\dim x$ (Example 4.3.2). One step of the a th Hopf-power Markov chain, starting from an irreducible representation x of \mathfrak{S}_n , is the following two-fold process (Example 4.4.3):

1. Choose a Young subgroup $\mathfrak{S}_{i_1} \times \cdots \times \mathfrak{S}_{i_a}$ multinomially.
2. Restrict the starting state x to the chosen subgroup, induce it back up to \mathfrak{S}_n , then pick an irreducible constituent with probability proportional to the dimension of its isotypic component.

The stationary distribution of this chain is the famous Plancherel measure (Example 4.5.3). This chain is reversible (Example 4.6.4).

Section 4.1 reviews the literature on combinatorial Hopf algebras. Section 4.2 gives a rudimentary construction of Hopf-power Markov chains, which is improved in Section 4.3, using the Doob transform of Section 3.1. Section 4.4 derives an interpretation of these chains as a breaking step followed by a combining step. Section 4.5 gives a complete description of the stationary distributions. Sections 4.6 and 4.7 employ the theory of Sections 3.3 and 3.4 respectively to deduce that the time-reversal of a Hopf-power chain is that associated to its dual algebra, and that the projection of a Hopf-power chain under a Hopf-morphism is the Hopf-power chain on the target algebra.

4.1 Combinatorial Hopf algebras

Recall from Section 1.2 the definition of a graded connected Hopf algebra: it is a vector space $\mathcal{H} = \bigoplus_{n=0}^{\infty} \mathcal{H}_n$ with a product map $m : \mathcal{H}_i \otimes \mathcal{H}_j \rightarrow \mathcal{H}_{i+j}$ and a coproduct map $\Delta : \mathcal{H}_n \rightarrow \bigoplus_{j=0}^n \mathcal{H}_j \otimes \mathcal{H}_{n-j}$ satisfying $\Delta(wz) = \Delta(w)\Delta(z)$ and some other axioms. To construct the Markov chains in this thesis, the natural Hopf algebras to use are *combinatorial Hopf algebras*, where the product and coproduct respectively encode how to combine and split combinatorial objects. These easily satisfy the non-negativity conditions required to define the associated Markov chain, which then has a natural interpretation in terms of breaking an object and then reassembling the pieces. A motivating example of a combinatorial Hopf algebra is:

Example 4.1.1 (Shuffle algebra). The shuffle algebra $\mathcal{S}(N)$, as defined in [Ree58], has as its basis the set of all words in the letters $\{1, 2, \dots, N\}$. The number of letters N is usually unimportant, so we write this algebra simply as \mathcal{S} . These words are notated in parantheses to distinguish them from integers.

The product of two words is the sum of all their interleavings, with multiplicity. For example,

$$m((13) \otimes (52)) = (13)(52) = (1352) + (1532) + (1523) + (5132) + (5123) + (5213),$$

$(12)(231) = 2(12231) + (12321) + (12312) + (21231) + (21321) + (21312) + (23121) + 2(23112)$. [Reu93, Sec. 1.5] shows that deconcatenation is a compatible coproduct. For example,

$$\Delta((316)) = \emptyset \otimes (316) + (3) \otimes (16) + (31) \otimes (6) + (316) \otimes \emptyset.$$

(Here, \emptyset denotes the empty word, which is the unit of \mathcal{S} .)

The associated Markov chain is the GSR riffle-shuffle of Example 1.1.1; below Example 4.4.2 will deduce this connection from Theorem 4.4.1.

The idea of using Hopf algebras to study combinatorial structures was originally due to Joni and Rota [JR79]. The concept enjoyed increased popularity in the late 1990s, when [Kre98] linked a combinatorial Hopf algebra on trees (see Section 5.3 below) to renormalisation in theoretical physics. Today, an abundance of combinatorial Hopf algebras exists; see the introduction of [Foi12] for a list of references to many examples. An instructive and entertaining overview of the basics and the history of the subject is in [Zab10]. [LR10] gives structure theorems for these algebras analogous to the Poincare-Birkhoff-Witt theorem (see Section 2.5 above) for cocommutative Hopf algebras.

A particular triumph of this algebrisation of combinatorics is [ABS06, Th. 4.1], which claims that $QSym$, the algebra of quasisymmetric functions (Example 4.1.6 below) is the terminal object in the category of combinatorial Hopf algebras with a multiplicative linear functional called a *character*. Their explicit map from any such algebra to $QSym$ unifies many ways of assigning polynomial invariants to combinatorial objects, such as the chromatic polynomial of graphs and Ehrenboug’s quasisymmetric function of a ranked poset. Section 5.1.4 makes the connection between these invariants and the probability of absorption of the associated Hopf-power Markov chains.

There is no universal definition of a combinatorial Hopf algebra in the literature; each author considers Hopf algebras with slightly different axioms. What they do agree on is that it should have a distinguished basis \mathcal{B} indexed by “combinatorial objects”, such as permutations, set partitions, or trees, and it should be graded by the “size” of these objects. The Hopf algebra is connected since the empty object is the only object of size 0.

For $x, y, z_1, \dots, z_a \in \mathcal{B}$, define *structure constants* $\xi_{z_1, \dots, z_a}^y, \eta_x^{z_1, \dots, z_a}$ by

$$z_1 \cdots z_a = \sum_{y \in \mathcal{B}} \xi_{z_1, \dots, z_a}^y y, \quad \Delta^{[a]}(x) = \sum_{z_1, \dots, z_a \in \mathcal{B}} \eta_x^{z_1, \dots, z_a} z_1 \otimes \cdots \otimes z_a.$$

Note that, by the inductive definitions of $m^{[a]}$ and $\Delta^{[a]}$, all structure constants are determined by $\xi_{w,z}^y$ and $\eta_x^{w,z}$ (see the proof of Lemma 4.2.1). Shorten these to ξ_{wz}^y and η_x^{wz} , without the comma in between w and z . In a combinatorial Hopf algebra, these two numbers should have interpretations respectively as the (possibly weighted) number of ways to combine w, z and obtain y , and the (possibly weighted) number of ways to break x into w, z . Then, the compatibility axiom $\Delta(wz) = \Delta(w)\Delta(z)$ translates roughly into the following: suppose y is one possible outcome when combining w and z ; then every way of breaking y comes (bijectively) from a way of breaking w and z separately. The axioms $\deg(wz) = \deg(w) + \deg(z)$ and $\Delta(x) \in \bigoplus_{i=0}^{\deg(x)} \mathcal{H}_i \otimes \mathcal{H}_{\deg(x)-i}$ simply say that the “total size” of an object is conserved under breaking and combining.

These are the minimal conditions for a combinatorial Hopf algebra, and will be sufficient for this thesis. For interest, a common additional hypothesis is the existence of an internal product $\mathcal{H}_n \otimes \mathcal{H}_n \rightarrow \mathcal{H}_n$, and perhaps also an internal coproduct. Note that commutativity of a combinatorial Hopf algebra indicates a symmetric assembling rule, and a symmetric breaking rule induces a cocommutative Hopf algebra.

Many families of combinatorial objects have a single member of size 1, so \mathcal{H}_1 is often one-dimensional. For example, there is only one graph on one vertex, and only one partition of total size 1. In such cases, \bullet will denote this sole object of size 1, so $\mathcal{B}_1 = \{\bullet\}$. A larger \mathcal{B}_1 may be the sign of a disconnected state space. That is, the associated Markov chain may separate into two (or more) chains running on disjoint subsets of the state space. For example, the usual grading on the shuffle algebra is by the length of the words. Then \mathcal{S}_3 contains both permutations of $\{1, 2, 3\}$ and permutations of $\{1, 1, 2\}$, but clearly no amount of shuffling will convert from one set to the other. To study these two Markov chains separately, refine the degree of a word w to be a vector whose i th component is the number of occurrences of i in w . (Trailing 0s in this vector are usually omitted.) So summing the components of this multidegree gives the old notion of degree. Now $\mathcal{S}_{(1,1,1)}$ contains the permutations of $\{1, 2, 3\}$, whilst $\mathcal{S}_{(2,1)}$ contains the permutations of $\{1, 1, 2\}$.

As Proposition 4.1.2 below will show, there is often an analogous multigrading on any combinatorial Hopf algebra with $|\mathcal{B}_1| > 1$. The catch is that elements of the basis \mathcal{B} may not be homogeneous in this multigrading, that is, \mathcal{B} might not be the disjoint union of bases \mathcal{B}_ν for each degree ν subspace \mathcal{H}_ν . (Currently, I do not know of any examples of such non-homogeneous bases.) In the case where $\mathcal{B} = \coprod_\nu \mathcal{B}_\nu$, Theorem 4.5.1.ii shows that the stationary distribution of the associated Markov chains (on each subspace \mathcal{H}_ν) is unique.

Proposition 4.1.2. *Let $\mathcal{H} = \bigoplus_{n \geq 0} \mathcal{H}_n$ be a graded connected Hopf algebra over \mathbb{R} . Suppose $\mathcal{B}_1 := \{\bullet_1, \bullet_2, \dots, \bullet_{|\mathcal{B}_1|}\}$ is a basis of \mathcal{H}_1 . For each $\nu = (\nu_1, \dots, \nu_{|\mathcal{B}_1|}) \in \mathbb{N}^{|\mathcal{B}_1|}$, set $c_1 = c_2 = \dots = c_{\nu_1} = \bullet_1$, $c_{\nu_1+1} = \dots = c_{\nu_1+\nu_2} = \bullet_2$, etc., and define*

$$\mathcal{H}_\nu := \{x \in \mathcal{H} \mid \bar{\Delta}^{[|\nu|]}(x) \in \text{span}\{c_{\sigma(1)} \otimes \dots \otimes c_{\sigma(|\nu|)} \mid \sigma \in \mathfrak{S}_{|\nu|}\}.$$

If $\mathcal{H}_n = \bigoplus_{|\nu|=n} \mathcal{H}_\nu$, then this gives a multigrading on \mathcal{H} refining the \mathbb{N} -grading. This is the unique multigrading satisfying $\deg(\bullet_1) = (1, 0, \dots, 0)$, $\deg(\bullet_2) = (0, 1, 0, \dots, 0)$, \dots , $\deg(\bullet_{|\mathcal{B}_1|}) = (0, \dots, 0, 1)$.

Proof. Comultiplication respects this notion of degree as coassociativity implies $\Delta^{[i+j]}(x) = (\Delta^{[i]} \otimes \Delta^{[j]})(\Delta x)$.

It is trickier to see the product respecting the degree. Take $z \in \mathcal{H}_i, w \in \mathcal{H}_j$. Then $\Delta^{[i+j]}(zw) = \Delta^{[i+j]}(z)\Delta^{[i+j]}(w)$. Since $\deg(z) = i$, at least j tensor-factors in each term of $\Delta^{[i+j]}(z)$ are in \mathcal{H}_0 , and the same is true for at least i tensor-factors in each term of $\Delta^{[i+j]}(w)$. Hence a term in $\bar{\Delta}^{[i+j]}(zw)$ must arise from terms in $\Delta^{[i+j]}(z), \Delta^{[i+j]}(w)$ which have exactly j and i tensor-factors respectively in \mathcal{H}_0 , in complementary positions. A term of $\Delta^{[i+j]}(z)$ with j tensor-factors in degree 0 must have the remaining i tensor-factors in degree 1, hence it corresponds to a term in $\bar{\Delta}^{[i]}(z)$, and similarly for w . So there is a bijection

$$\left\{ \begin{array}{c} \text{terms in} \\ \bar{\Delta}^{[i]}(z) \end{array} \right\} \times \left\{ \begin{array}{c} \text{terms in} \\ \bar{\Delta}^{[j]}(w) \end{array} \right\} \times \left\{ \begin{array}{c} \text{subsets of} \\ \{1, 2, \dots, i+j\} \\ \text{of size } i \end{array} \right\} \leftrightarrow \left\{ \begin{array}{c} \text{terms in} \\ \bar{\Delta}^{[i+j]}(zw) \end{array} \right\}$$

$c_1 \otimes \dots \otimes c_i, \quad c'_1 \otimes \dots \otimes c'_j, \quad k_1 < \dots < k_i \rightarrow k_r \text{th tensor-factor is } c_r,$

c'_1, \dots, c'_j in other tensor-factors.

And so the multidegree $\deg(zw)$ is $\deg(z) + \deg(w)$.

As for uniqueness: suppose $\bar{\Delta}^{[|v|]}(x) \in \text{span}\{c_{\sigma(1)} \otimes \cdots \otimes c_{\sigma(|v|)} \mid \sigma \in \mathfrak{S}_{|v|}\}$. Then, since the coproduct respects the multigrading, it must be that $\deg(x) = \deg c_1 + \cdots + \deg c_{|v|} = v$. \square

The rest of this section is a whistle-stop tour of three sources of combinatorial Hopf algebras. A fourth important source is operads [Hol04], but that theory is too technical to cover in detail here.

4.1.1 Species-with-Restrictions

This class of examples is especially of interest in this thesis, as the associated Markov chains have two nice properties. Firstly, constructing these chains does not require the Doob transform (Definition 4.2.2). Secondly, the natural bases of these Hopf algebras are *free-commutative* in the sense of Chapter 5, so additional tools are available to study the associated Markov chains. For instance, these chains are absorbing, and Section 5.1.3 provides bounds for the probability of being “far from absorption”.

The theory of species originated in [Joy81], as an abstraction of common manipulations of generating functions. Loosely speaking, a species is a type of combinatorial structure which one can build on sets of “vertices”. Important examples include (labelled) graphs, trees and permutations. The formal definition of a species is as a functor from the category of sets with bijections to the same category. In this categorical language, the species of graphs maps a set V to the set of all graphs whose vertices are indexed by V . There are operations on species which correspond to the multiplication, composition and differentiation of their associated generating functions; these are not so relevant to the present Markov chain construction, so the reader is referred to [BLL98] for further details.

Schmitt [Sch93] first makes the connection between species and Hopf algebras. He defines a species-with-restrictions, or R -species, to be a functor from sets with coinjections to the category of functions. (A *coinjection* is a partially-defined function whose restriction to where it’s defined is a bijection; an example is $f : \{1, 2, 3, 4\} \rightarrow \{7, 8\}$ with $f(1) = 8$, $f(3) = 7$ and $f(2), f(4)$ undefined.) Intuitively, these are combinatorial structures with a notion of restriction to a subset of their vertex set; for example, one can restrict a graph to a

$$\Delta(\frown) = \frown \otimes 1 + \diagup \otimes \bullet + \diagdown \otimes \bullet + \bullet \otimes \bullet + \bullet \otimes \diagup + \bullet \otimes \diagdown + \bullet \otimes \bullet + 1 \otimes \frown$$

Figure 4.1: An example coproduct calculation in \mathcal{G} , the Hopf algebra of graphs

subset of its vertices by considering only the edges connected to this subset (usually known as the *induced subgraph*). Schmitt fashions from each such species a Hopf algebra which is both commutative and cocommutative; Example 4.1.3 below explains his construction via the species of graphs.

Example 4.1.3 (The Hopf algebra of graphs). [Sch94, Sec. 12; Fis10, Sec. 3.2] Let \mathcal{G} be the vector space with basis the set of simple graphs (no loops or multiple edges). The vertices of such graphs are unlabelled, so these may be considered the isomorphism classes of graphs. Define the degree of a graph to be its number of vertices. The product of two graphs is their disjoint union, and the coproduct is

$$\Delta(G) = \sum G_S \otimes G_{S^c}$$

where the sum is over all subsets S of vertices of G , and G_S, G_{S^c} denote the subgraphs that G induces on the vertex set S and its complement. As an example, Figure 4.1 calculates the coproduct of P_3 , the path of length 3. Writing P_2 for the path of length 2, and \bullet for the unique graph on one vertex, this calculation shows that

$$\Delta(P_3) = P_3 \otimes 1 + 2P_2 \otimes \bullet + \bullet^2 \otimes \bullet + 2\bullet \otimes P_2 + \bullet \otimes \bullet^2 + 1 \otimes P_3.$$

As mentioned above, this Hopf algebra, and analogous constructions from other species-with-restrictions, are both commutative and cocommutative.

As Example 4.2.3 will describe, the Hopf-power Markov chain on \mathcal{G} models the removal of edges: at each step, colour each vertex independently and uniformly in one of a colours, and disconnect edges between vertices of different colours. This chain will act as the running example in Section 5.1, to illustrate general results concerning a Hopf-power

Markov chain on a free-commutative basis. However, because the concept of graph is so general, it is hard to say anything specific or interesting without restricting to graphs of a particular structure. For example, restricting to unions of complete graphs gives the rock-breaking chain of Section 5.2. I aim to produce more such examples in the near future.

Recently, Aguiar and Mahajan [AM10] extended vastly this construction to the concept of a *Hopf monoid in species*, which is a finer structure than a Hopf algebra. Their Chapter 15 gives two major pathways from a species to a Hopf algebra: the Bosonic Fock functor, which is essentially Schmitt’s original idea, and the Full Fock functor. (Since the product and coproduct in the latter involves “shifting” and “standardisation” of labels, the resulting Hopf algebras lead to rather contrived Markov chains, so this thesis will not explore the Full Fock functor in detail.) In addition there are decorated and coloured variants of these two constructions, which allow the input of parameters. Many popular combinatorial Hopf algebras, including all examples in this thesis, arise from Hopf monoids; perhaps this is an indication that the Hopf monoid is the “correct” setting to work in. The more rigid set of axioms of a Hopf monoid potentially leads to stronger theorems.

In his masters’ thesis, Pineda [Pin14] transfers some of the Hopf-power Markov chain technology of this thesis to the world of Hopf monoids, building a Markov chain on faces of a permutohedra. His chain has many absorbing states, a phenomenon not seen in any of the chains in this thesis. This suggests that a theory of Markov chains from Hopf monoids may lead to a richer collection of examples.

4.1.2 Representation rings of Towers of Algebras

The ideas of this construction date back to Zelevinsky [Zel81, Sec. 6], which the lecture notes [GR14, Sec. 4] retell in modern notation. The archetype is as follows:

Example 4.1.4 (Representations of symmetric groups). Let \mathcal{B}_n be the irreducible representations of the symmetric group \mathfrak{S}_n , so \mathcal{H}_n is the vector space spanned by all representations of \mathfrak{S}_n . The product of representations w, z of $\mathfrak{S}_n, \mathfrak{S}_m$ respectively is defined using induction:

$$m(w \otimes z) = \text{Ind}_{\mathfrak{S}_n \times \mathfrak{S}_m}^{\mathfrak{S}_{n+m}} w \times z,$$

and the coproduct of x , a representation of \mathfrak{S}_n , is the sum of its restrictions:

$$\Delta(x) = \bigoplus_{i=0}^n \text{Res}_{\mathfrak{S}_i \times \mathfrak{S}_{n-i}}^{\mathfrak{S}_n} x.$$

Mackey theory ensures these operations satisfy $\Delta(wz) = \Delta(w)\Delta(z)$. This Hopf algebra is both commutative and cocommutative, as $\mathfrak{S}_n \times \mathfrak{S}_m$ and $\mathfrak{S}_m \times \mathfrak{S}_n$ are conjugate in \mathfrak{S}_{n+m} ; however, the general construction need not have either symmetry. The associated Markov chain describes the restriction then induction of representations, see Example 4.4.3.

It's natural to attempt this construction with, instead of $\{\mathfrak{S}_n\}$, any series of algebras $\{A_n\}$ where an injection $A_n \otimes A_m \subseteq A_{n+m}$ allows this outer product of its modules. For the result to be a Hopf algebra, one needs some additional hypotheses on the algebras $\{A_n\}$; this leads to the definition of a *tower of algebras* in [BL09]. In general, two Hopf algebras can be built this way: one using the finitely-generated modules of each A_n , and one from the finitely-generated projective modules of each A_n . (For the above example of symmetric groups, these coincide, as all representations are semisimple.) These are graded duals in the sense of Section 2.1. For example, [KT97, Sec. 5] takes A_n to be the 0-Hecke algebra, then the Hopf algebra of finitely-generated modules is $QSym$, the Hopf algebra of quasisymmetric functions. Example 4.1.6 below will present $QSym$ in a different guise that does not require knowledge of Hecke algebras. The Hopf algebra of finitely-generated projective modules of the 0-Hecke algebras is **Sym**, the algebra of noncommutative symmetric functions of Section 6.2.2. Further developments regarding Hopf structures from representations of towers of algebras are in [BLL12].

It will follow from Definition 4.3.4 of a Hopf-power Markov chain that, as long as every irreducible representation of A_n has a non-zero restriction to some proper subalgebra $A_i \otimes A_{n-i}$ ($1 \leq i \leq n$), one can build a Markov chain on the irreducible representations of the tower of algebras $\{A_n\}$. (Unfortunately, when A_n is the group algebra of GL_n over a finite field, the cuspidal representations violate this hypothesis.) These chains should be some variant of restriction-then-induction. It is highly possible that the precise description of the chain is exactly as in Example 4.4.3: starting at an irreducible representation of A_n , pick $i \in [0, n]$ binomially, restrict to $A_i \otimes A_{n-i}$, then induce back to A_n and pick an irreducible representation with probability proportional to the dimension of the isotypic component.

Interestingly, it is sometimes possible to tell a similar story with the basis \mathcal{B}_n being a set of reducible representations, possibly with slight tweaks to the definitions of product and coproduct. In [Agu+12; BV13; ABT13; And14], \mathcal{B}_n is a *supercharacter* theory of various matrix groups over finite fields. This means that the matrix group can be partitioned into *superclasses*, which are each a union of conjugacy classes, such that each supercharacter (the characters of the representations in \mathcal{B}_n) is constant on each superclass, and each irreducible character of the matrix group is a constituent of exactly one supercharacter. [DI08] gives a unified method to build a supercharacter theory on many matrix groups; this is useful as the irreducible representations of these groups are extremely complicated.

4.1.3 Subalgebras of Power Series

The starting point for this approach is the algebra of symmetric functions, widely considered as the first combinatorial Hopf algebra in history, and possibly the most extensively studied. Thorough textbook introductions to its algebra structure and its various bases are [Mac95, Chap. 1] and [Sta99, Chap. 7].

Example 4.1.5 (Symmetric functions). Work in the algebra $\mathbb{R}[[x_1, x_2, \dots]]$ of power series in infinitely-many commuting variables x_i , graded so $\deg(x_i) = 1$ for all i . The *algebra of symmetric functions* Λ is the subalgebra of power series of finite degree invariant under the action of the infinite symmetric group \mathfrak{S}_∞ permuting the variables. (These elements are often called “polynomials” due to their finite degree, even though they contain infinitely-many monomial terms.)

An obvious basis of Λ is the sum of monomials in each \mathfrak{S}_∞ orbit; these are the *monomial symmetric functions*:

$$m_\lambda := \sum_{\substack{(i_1, \dots, i_l) \\ i_j \text{ distinct}}} x_{i_1}^{\lambda_1} \dots x_{i_l}^{\lambda_l}.$$

Here, λ is a partition of $\deg(m_\lambda)$: $\lambda_1 + \cdots + \lambda_{l(\lambda)} = \deg(m_\lambda)$ with $\lambda_1 \geq \cdots \geq \lambda_{l(\lambda)}$. For example, the three monomial symmetric functions of degree three are:

$$\begin{aligned} m_{(3)} &= x_1^3 + x_2^3 + \cdots; \\ m_{(2,1)} &= x_1^2 x_2 + x_1^2 x_3 + \cdots + x_2^2 x_1 + x_2^2 x_3 + x_2^2 x_4 + \cdots; \\ m_{(1,1,1)} &= x_1 x_2 x_3 + x_1 x_2 x_4 + \cdots + x_1 x_3 x_4 + x_1 x_3 x_5 + \cdots + x_2 x_3 x_4 + \cdots \end{aligned}$$

It turns out [Sta99, Th. 7.4.4, Cor. 7.6.2] that Λ is isomorphic to a polynomial ring in infinitely-many variables: $\Lambda = \mathbb{R}[h_{(1)}, h_{(2)}, \dots]$, where

$$h_{(n)} := \sum_{i_1 \leq \cdots \leq i_n} x_{i_1} \cdots x_{i_n}.$$

(This is often denoted h_n , as it is standard to write the integer n for the partition (n) of single part.) For example,

$$h_{(2)} = x_1^2 + x_1 x_2 + x_1 x_3 + \cdots + x_2^2 + x_2 x_3 + \cdots$$

So, setting $h_\lambda := h_{(\lambda_1)} \cdots h_{(\lambda_{l(\lambda)})}$ over all partitions λ gives another basis of Λ , the *complete symmetric functions*.

Two more bases are important: the *power sums* are $p_{(n)} := \sum_i x_i^n$, $p_\lambda := p_{(\lambda_1)} \cdots p_{(\lambda_{l(\lambda)})}$; and the *Schur functions* $\{s_\lambda\}$ are the image of the irreducible representations under the *Frobenius characteristic isomorphism* from the representation rings of the symmetric groups (Example 4.1.4) to Λ [Sta99, Sec. 7.18]. This map is defined by sending the indicator function of an n -cycle of \mathfrak{S}_n to the scaled power sum $\frac{p_{(n)}}{n}$. (I am omitting the elementary basis $\{e_\lambda\}$, as it has similar behaviour as $\{h_\lambda\}$.)

The coproduct on Λ comes from the ‘‘alphabet doubling trick’’. This relies on the isomorphism between the power series algebras $\mathbb{R}[[x_1, x_2, \dots, y_1, y_2, \dots]]$ and $\mathbb{R}[[x_1, x_2, \dots]] \otimes \mathbb{R}[[y_1, y_2, \dots]]$, which simply rewrites the monomial $x_{i_1} \cdots x_{i_k} y_{j_1} \cdots y_{j_l}$ as $x_{i_1} \cdots x_{i_k} \otimes y_{j_1} \cdots y_{j_l}$. To calculate the coproduct of a symmetric function f , first regard f as a power series in two sets of variables $x_1, x_2, \dots, y_1, y_2, \dots$; then $\Delta(f)$ is the image of

$f(x_1, x_2, \dots, y_1, y_2, \dots)$ in $\mathbb{R}[[x_1, x_2, \dots]] \otimes \mathbb{R}[[y_1, y_2, \dots]]$ under the above isomorphism. Because f is a symmetric function, the power series $f(x_1, x_2, \dots, y_1, y_2, \dots)$ is invariant under the permutation of the x_i s and y_i s separately, so $\Delta(f)$ is in fact in $\Lambda \otimes \Lambda$. For example,

$$\begin{aligned} h_{(2)}(x_1, x_2, \dots, y_1, y_2, \dots) &= x_1^2 + x_1x_2 + x_1x_3 + \dots + x_1y_1 + x_1y_2 + \dots \\ &\quad + x_2^2 + x_2x_3 + \dots + x_2y_1 + x_2y_2 + \dots \\ &\quad + \dots \\ &\quad + y_1^2 + y_1y_2 + y_1y_3 + \dots \\ &\quad + y_2^2 + y_2y_3 + \dots \\ &\quad + \dots \\ &= h_{(2)}(x_1, x_2, \dots) + h_{(1)}(x_1, x_2, \dots)h_{(1)}(y_1, y_2, \dots) + h_{(2)}(y_1, y_2, \dots), \end{aligned}$$

so $\Delta(h_{(2)}) = h_{(2)} \otimes 1 + h_{(1)} \otimes h_{(1)} + 1 \otimes h_{(2)}$. In general, $\Delta(h_{(n)}) = \sum_{i=0}^n h_{(i)} \otimes h_{(n-i)}$, with the convention $h_{(0)} = 1$. (This is Geissenger's original definition of the coproduct [Gei77].) Note that $\Delta(p_{(n)}) = 1 \otimes p_{(n)} + p_{(n)} \otimes 1$; this property is the main reason for working with the power sum basis.

The Hopf-power Markov chain on $\{h_\lambda\}$ describes an independent multinomial rock-breaking process, see Section 5.2.

The generalisation of Λ is easier to see if the \mathfrak{S}_∞ action is rephrased in terms of a function to a fundamental domain. Observe that each orbit of the monomials, under the action of the infinite symmetric group permuting the variables, contains precisely one term of the form $x_1^{\lambda_1} \dots x_l^{\lambda_l}$ for some partition λ . Hence the set $\mathcal{D} := \left\{ x_1^{\lambda_1} \dots x_l^{\lambda_l} \mid l, \lambda_i \in \mathbb{N}, \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_l > 0 \right\}$ is a fundamental domain for this \mathfrak{S}_∞ action. Define a function f sending a monomial to the element of \mathcal{D} in its orbit; explicitly,

$$f\left(x_{j_1}^{i_1} \dots x_{j_l}^{i_l}\right) = x_1^{i_{\sigma(1)}} \dots x_l^{i_{\sigma(l)}},$$

where $\sigma \in \mathfrak{S}_l$ is such that $i_{\sigma(1)} \geq \dots \geq i_{\sigma(l)}$. For example, $f(x_1x_3^2x_4) = x_1^2x_2x_3$. It is clear that the monomial symmetric function m_λ , previously defined to be the sum over \mathfrak{S}_∞ orbits,

is the sum over preimages of f :

$$m_\lambda := \sum_{f(x)=x^\lambda} x,$$

where x^λ is shorthand for $x_1^{\lambda_1} \dots x_l^{\lambda_l}$. Summing over preimages of other functions can give bases of other Hopf algebras. Again, the product is that of power series, and the coproduct comes from alphabet doubling. Example 4.1.6, essentially a simplified, commutative, version of [NT06, Sec. 2], builds the algebra of quasisymmetric functions using this recipe. This algebra is originally due to Gessel [Ges84], who defines it in terms of P -partitions.

Example 4.1.6 (Quasisymmetric functions). Start again with $\mathbb{R}[[x_1, x_2, \dots]]$, the algebra of power series in infinitely-many commuting variables x_i . Let pack be the function sending a monomial $x_{j_1}^{i_1} \dots x_{j_l}^{i_l}$ (assuming $j_1 < \dots < j_l$) to its *packing* $x_1^{i_1} \dots x_l^{i_l}$. For example, $\text{pack}(x_1 x_3^2 x_4) = x_1 x_2^2 x_3$. A monomial is *packed* if it is its own packing, in other words, its constituent variables are consecutive starting from x_1 . Let \mathcal{D} be the set of packed monomials, so $\mathcal{D} := \{x_1^{i_1} \dots x_l^{i_l} \mid l, i_j \in \mathbb{N}\}$. Writing I for the *composition* (i_1, \dots, i_l) and x^I for $x_1^{i_1} \dots x_l^{i_l}$, define the *monomial quasisymmetric functions* to be:

$$M_I := \sum_{\text{pack}(x)=x^I} x = \sum_{j_1 < \dots < j_{l(I)}} x_{j_1}^{i_1} \dots x_{j_{l(I)}}^{i_{l(I)}}.$$

For example, the four monomial quasisymmetric functions of degree three are:

$$\begin{aligned} M_{(3)} &= x_1^3 + x_2^3 + \dots; \\ M_{(2,1)} &= x_1^2 x_2 + x_1^2 x_3 + \dots + x_2^2 x_3 + x_2^2 x_4 + \dots + x_3^2 x_4 + \dots; \\ M_{(1,2)} &= x_1 x_2^2 + x_1 x_3^2 + \dots + x_2 x_3^2 + x_2 x_4^2 + \dots + x_3 x_4^2 + \dots; \\ M_{(1,1,1)} &= x_1 x_2 x_3 + x_1 x_2 x_4 + \dots + x_1 x_3 x_4 + x_1 x_3 x_5 + \dots + x_2 x_3 x_4 + \dots \end{aligned}$$

$QSym$, the *algebra of quasisymmetric functions*, is then the subalgebra of $\mathbb{R}[[x_1, x_2, \dots]]$ spanned by the M_I .

Note that the monomial symmetric function $m_{(2,1)}$ is $M_{(2,1)} + M_{(1,2)}$; in general, $m_\lambda = \sum M_I$ over all compositions I whose parts, when ordered decreasingly, are equal to λ . Thus Λ is a subalgebra of $QSym$.

The basis of $QSym$ with representation-theoretic significance, analogous to the Schur functions of Λ , are the fundamental quasisymmetric functions:

$$F_I = \sum_{J \geq I} M_J$$

where the sum runs over all compositions J refining I (i.e. I can be obtained by gluing together some adjacent parts of J). For example,

$$F_{(2,1)} = M_{(2,1)} + M_{(1,1,1)} = \sum_{j_1 \leq j_2 < j_3} x_{j_1} x_{j_2} x_{j_3}.$$

The fundamental quasisymmetric functions are sometimes denoted L_I or Q_I in the literature. They correspond to the irreducible modules of the 0-Hecke algebra [KT97, Sec. 5]. The analogue of power sums are more complex (as they naturally live in the dual Hopf algebra to $QSym$), see Section 6.2.2 for a full definition.

The Hopf-power Markov chain on the basis of fundamental quasisymmetric functions $\{F_I\}$ is the change in descent set under riffle-shuffling, which Section 6.2 analyses in detail.

In the last decade, a community in Paris have dedicated themselves [DHT02; NT06; FNT11] to recasting familiar combinatorial Hopf algebras in this manner, a process they call *polynomial realisation*. They usually start with power series in noncommuting variables, so the resulting Hopf algebra is not constrained to be commutative. The least technical exposition is probably [Thi12], which also provides a list of examples. The simplest of these is **Sym**, a noncommutative analogue of the symmetric functions; its construction is explained in Section 6.2.2 below. For a more interesting example, take M_T to be the sum of all noncommutative monomials with Q -tableau equal to T under the Robinson-Schensted-Knuth algorithm [Sta99, Sec. 7.11]; then their span is **FSym**, the Poirier-Reutenauer Hopf algebra of tableaux [PR95]. [Hiv07, Th. 31] and [Pri13, Th. 1] give sufficient conditions on the functions for this construction to produce a Hopf algebra. One motivation for this program is to bring to light various bases that are free (like h_λ), interact well with the coproduct (like p_λ) or are connected to representation theory (like s_λ), and to carry over some of the vast amount of machinery developed for the symmetric functions to analyse these combinatorial objects in new ways. Indeed, Joni and Rota anticipated in their original paper

[JR79] that “many an interesting combinatorial problem can be formulated algebraically as that of transforming this basis into another basis with more desirable properties”.

4.2 First Definition of a Hopf-power Markov Chain

Recall from Section 1.1 the GSR riffle-shuffle of a deck of cards: cut the deck into two piles according to a symmetric binomial distribution, then drop the cards one by one from the bottom of the piles, chosen with probability proportional to the current pile size. As mentioned in Section 1.3, a direct calculation shows that, for words x, y of length n in the shuffle algebra of Example 4.1.1, the coefficient of y in $2^{-n}m\Delta(x)$ is the probability of obtaining a deck of cards in order y after applying a GSR riffle-shuffle to a deck in order x :

$$2^{-n}m\Delta(x) = \sum_y K(x, y)y. \quad (4.1)$$

(Here, identify the word $x_1x_2\dots x_n$ in the shuffle algebra with the deck whose top card has value x_1 , second card has value x_2 , and so on, so x_n is the value of the bottommost card.) In other words, the matrix of the linear operator $2^{-n}m\Delta$ on \mathcal{H}_n , with respect to the basis of words, is the transpose of the transition matrix of the GSR shuffle. Furthermore, the matrix of the a th Hopf-power map $a^{-n}\Psi^a := a^{-n}m^{[a]}\Delta^{[a]}$ on \mathcal{H}_n (with respect to the basis of words) is the transpose of the transition matrix of an a -handed shuffle of [BD92]; this will follow from Theorem 4.4.1 below. An a -handed shuffle is a straightforward generalisation of the GSR shuffle: cut the deck into a piles according to the symmetric multinomial distribution, then drop the cards one by one from the bottom of the pile, where the probability of dropping from any particular pile is proportional to the number of cards currently in that pile. This second step is equivalent to all interleavings of the a piles being equally likely; more equivalent views are in [BD92, Chap. 3].

This relationship between a -handed shuffles and the a th Hopf-power map on the shuffle algebra motivates the question: for which graded Hopf algebras \mathcal{H} and bases \mathcal{B} does Equation 4.1 (and its analogue for $a > 2$) define a Markov chain? In other words, what conditions on \mathcal{H} and \mathcal{B} guarantee that the coefficients of $a^{-n}\Psi^a(x)$ are non-negative and

sum to 1? Achieving a sum of 1 is the subject of the next section; as for non-negativity, one solution is to mandate that the product and coproduct structure constants are non-negative:

Lemma 4.2.1. *Let \mathcal{H} be a Hopf algebra over \mathbb{R} with basis \mathcal{B} such that:*

- (i) *for all $w, z \in \mathcal{B}$, $wz = \sum_{y \in \mathcal{B}} \xi_{wz}^y y$ with $\xi_{wz}^y \geq 0$ (non-negative product structure constants);*
- (ii) *for all $x \in \mathcal{B}$, $\Delta(x) = \sum_{w, z \in \mathcal{B}} \eta_x^{wz} w \otimes z$ with $\eta_x^{wz} \geq 0$ (non-negative coproduct structure constants).*

Then, for all $x, y \in \mathcal{B}$, the coefficient of y in $\Psi^a(x)$ is non-negative, for all a .

Proof. In the notation for structure constants at the start of Section 4.1, the coefficient of y in $\Psi^a(x)$ is $\sum_{z_1, \dots, z_n} \xi_{z_1, \dots, z_n}^y \eta_x^{z_1, \dots, z_n}$. By definition of a -fold multiplication and comultiplication,

$$\xi_{z_1, \dots, z_a}^y = \sum_z \xi_{zz_a}^y \xi_{z_1, \dots, z_{a-1}}^z, \quad \eta_x^{z_1, \dots, z_a} = \sum_z \eta_x^{z a} \eta_z^{z_1, \dots, z_{a-1}},$$

so, by induction on a (the base case of $a = 2$ being the hypothesis), both ξ_{z_1, \dots, z_a}^y and $\eta_x^{z_1, \dots, z_a}$ are non-negative. \square

So the following indeed specifies a Markov chain:

Definition 4.2.2 (First definition of Hopf-power Markov chain). Let $\mathcal{H} = \bigoplus_{n \geq 0} \mathcal{H}_n$ be a graded connected Hopf algebra over \mathbb{R} , with each \mathcal{H}_n finite-dimensional. Let $\mathcal{B} = \coprod_{n \geq 0} \mathcal{B}_n$ be a basis of \mathcal{H} with non-negative structure constants (i.e. satisfying conditions i, ii of Lemma 4.2.1 above). Assume in addition that, for all $x \in \mathcal{B}_n$, the coefficients (with respect to \mathcal{B}_n) of $a^{-n} \Psi^a(x)$ sum to 1. Then the a th Hopf-power Markov chain on \mathcal{B}_n has transition matrix $K_{a,n} := [a^{-n} \Psi^a]_{\mathcal{B}_n}^T$, the transpose of the matrix of $a^{-n} \Psi^a$ with respect to the basis \mathcal{B}_n .

Observe that, if \mathcal{H} comes from a species-with-restrictions in the method of Section 4.1.1, then the coefficients of $a^{-n} \Psi^a(x)$ sum to 1, for all a and all n . This is because the terms in $\Delta^{[a]}(x)$ correspond to the a^n ways of partitioning the underlying set into a (possibly trivial) subsets (the order of the subsets matter), and each such term gives only a single term under $m^{[a]}$.

Example 4.2.3. Take $\mathcal{H} = \overline{\mathcal{G}}$, the algebra of graphs of Example 4.1.3. Recall that the product of two graphs is their disjoint union, and the coproduct gives the induced subgraphs on two complementary subsets of the vertex set. Thus one step of the associated a th Hopf-power Markov chain is the following: independently assign to each vertex one of a colours, each with an equal probability of $\frac{1}{a}$. Then remove all edges between vertices of different colours. As an example, take $a = 2$ and start at P_3 , the path of length 3. Write P_2 for the two-vertex graph with a single edge. By Figure 4.1,

$$\Delta(P_3) = P_3 \otimes 1 + 2P_2 \otimes \bullet + \bullet^2 \otimes \bullet + 2\bullet \otimes P_2 + \bullet \otimes \bullet^2 + 1 \otimes P_3.$$

Hence $\Psi^2(P_3) = 2P_3 + 4P_2\bullet + 2\bullet^3$. So, starting at P_3 , the chain stays at P_3 with probability $\frac{2}{2^3} = \frac{1}{4}$, or moves to $P_2\bullet$ with probability $\frac{4}{2^3} = \frac{1}{2}$, or moves to the disconnected graph with probability $\frac{2}{2^3} = \frac{1}{4}$.

4.3 General Definition of a Hopf-power Markov Chain

One would like to remove from Definition 4.2.2 above the restrictive condition that the sum of the coefficients of $a^{-n}\Psi^a(x)$ is 1. In other words, it would be good to build a Markov chain out of Ψ^a even when the matrix $K_{a,n} := [a^{-n}\Psi^a]_{\mathcal{B}_n}^T$ does not have every row summing to 1. Lemma 3.1.1, the Doob h -transform for linear maps, gives one possible answer: instead of \mathcal{B}_n , work with the basis $\check{\mathcal{B}}_n := \left\{ \check{x} := \frac{x}{\eta_n(x)} \mid x \in \mathcal{B}_n \right\}$, where $\eta_n \in \mathcal{H}_n^*$ is a “positive” eigenvector for the map dual to Ψ^a . Recall from Section 2.1 that this dual map is again a Hopf-power map Ψ^a , but on the (graded) dual Hopf algebra \mathcal{H}^* . On a combinatorial Hopf algebra, one choice of η_n has a remarkably simple description as “the number of ways to break into singletons”, and is usually a well-investigated number. The first two definitions of η_n below are more intuitive, as they avoid direct reference to \mathcal{H}^* , whilst the third streamlines the proofs.

Definition 4.3.1. Three equivalent definitions of the *rescaling functions* $\eta_n : \mathcal{B}_n \rightarrow \mathbb{R}$ are:

- (i) $\eta_n(x)$ is the sum of coproduct structure constants (over all ordered n -tuples, possibly with repetition of the c_i):

$$\eta_n(x) := \sum_{c_1, c_2, \dots, c_n \in \mathcal{B}_1} \eta_x^{c_1, \dots, c_n};$$

- (ii) $\eta_n(x)$ is the sum of the coefficients of $\bar{\Delta}^{[n]}(x)$, the n -fold reduced coproduct of x , when expanded in the basis $\mathcal{B}^{\otimes n}$. (Recall from Section 2.2 that $\bar{\Delta}(x) := \bar{\Delta}^{[2]}(x) := \Delta(x) - 1 \otimes x - x \otimes 1$, and $\bar{\Delta}^{[n]} := (\iota \otimes \dots \otimes \iota \otimes \bar{\Delta})\bar{\Delta}^{[n-1]}$, so $\bar{\Delta}^{[n]} \in \mathcal{H}_1^{\otimes n}$.)
- (iii) Let $\bullet^* \in \mathcal{H}_1^*$ be the linear function on \mathcal{H} taking value 1 on each element of \mathcal{B}_1 and 0 on all other basis elements. (In the dual basis notation from the start of Chapter 3, $\bullet^* := \sum_{c \in \mathcal{B}_1} c^*$; in particular, if $\mathcal{B}_1 = \{\bullet\}$ then this agrees with the dual basis notation.) Then set $\eta_n := (\bullet^*)^n$. In other words, $\eta_n(x) := (\bullet^* \otimes \dots \otimes \bullet^*)\Delta^{[n]}(x)$.

Since, for each $n \in \mathbb{N}$, the rescaling function η_n has a different domain (namely \mathcal{H}_n), no confusion arises from abbreviating $\eta_{\deg x}(x)$ by $\eta(x)$. Observe though that such a function η is not an element of the (graded) dual \mathcal{H}^* , as it is an infinite sum of linear functions on the subspaces \mathcal{H}_n . However, the variant $\frac{\eta_{\deg x}(x)}{\deg x!}$ is a *character* in the sense of [ABS06], as it is multiplicative; see Lemma 5.1.2.

Example 4.3.2. Recall from Example 4.1.4 the Hopf algebra of representations of the symmetric groups, with product arising from induction and coproduct from restriction. Its distinguished basis \mathcal{B} is the set of irreducible representations. So \mathcal{B}_1 consists only of the trivial representation \bullet , thus, by the first of the equivalent definitions above, $\eta(x) = \eta_x^{\bullet, \dots, \bullet}$. For an irreducible representation x of \mathfrak{S}_n , $\text{Res}_{\mathfrak{S}_1 \times \dots \times \mathfrak{S}_1}^{\mathfrak{S}_n} x = \dim x(\bullet \otimes \dots \otimes \bullet)$, so $\eta(x) = \dim x$.

A simple application of the Symmetrisation Lemma (Theorem 2.3.2) shows that η_n is an eigenvector of $\Psi^a : \mathcal{H}_n^* \rightarrow \mathcal{H}_n^*$ of eigenvalue a^n , since \bullet^* has degree 1 and is hence primitive. In order to use η_n in the Doob transform, we must ensure that $\eta_n(x) > 0$ for all $x \in \mathcal{B}_n$. (It suffices to force $\eta_n(x) \neq 0$ for all $x \in \mathcal{B}_n$, since, as a sum of coproduct structure constants, η_n takes non-negative values on \mathcal{B}_n .) This is the purpose of condition iii in Definition 4.3.3 below. This requirement essentially translates to “every object of

size greater than 1 breaks non-trivially”; the intuition is that repeatedly applying such non-trivial breaks to the pieces provides a way to reduce x to singletons. Theorem 4.3.7 below rigorises this heuristic, and explains why it is necessary to forbid primitive basis elements of degree greater than one in order to apply the Doob transform to the Hopf-power map, for all choices of rescaling functions.

Definition 4.3.3 (State space basis). Let $\mathcal{H} = \bigoplus_{n \geq 0} \mathcal{H}_n$ be a graded connected Hopf algebra over \mathbb{R} , with each \mathcal{H}_n finite-dimensional. A basis $\mathcal{B} = \coprod_{n \geq 0} \mathcal{B}_n$ of \mathcal{H} is a *state space basis* if:

- (i) for all $w, z \in \mathcal{B}$, $wz = \sum_{y \in \mathcal{B}} \xi_{wz}^y y$ with $\xi_{wz}^y \geq 0$ (non-negative product structure constants);
- (ii) for all $x \in \mathcal{B}$, $\Delta(x) = \sum_{w, z \in \mathcal{B}} \eta_x^{wz} w \otimes z$ with $\eta_x^{wz} \geq 0$ (non-negative coproduct structure constants);
- (iii) for all $x \in \mathcal{B}$ with $\deg(x) > 1$, it holds that $\Delta(x) \neq 1 \otimes x + x \otimes 1$ (no primitive elements in \mathcal{B} of degree greater than 1).

Note that \mathcal{H} may contain primitive elements of any degree, so long as those of degree greater than one are not in the basis \mathcal{B} . Applying the Doob transform to $\Psi^a : \mathcal{H}_n \rightarrow \mathcal{H}_n$ (with the rescaling function η) then creates the family of Markov chains defined below.

Definition 4.3.4 (General definition of Hopf-power Markov chain). Let $\mathcal{H} = \bigoplus_{n \geq 0} \mathcal{H}_n$ be a graded connected Hopf algebra over \mathbb{R} , with each \mathcal{H}_n finite-dimensional, and with state space basis \mathcal{B} . Take η_n according to Definition 4.3.1. Then the *ath Hopf-power Markov chain* on \mathcal{B}_n has transition matrix $\check{K}_{a,n} := [a^{-n} \Psi^a]^T_{\check{\mathcal{B}}_n}$, where $\check{\mathcal{B}}_n := \left\{ \check{x} := \frac{x}{\eta_n(x)} \mid x \in \mathcal{B}_n \right\}$. In other words,

$$a^{-n} \Psi^a(\check{x}) = \sum_{y \in \mathcal{B}_n} \check{K}_{a,n}(x, y) \check{y},$$

or, equivalently,

$$a^{-n} \Psi^a(x) = \sum_{y \in \mathcal{B}_n} \frac{\eta_n(x)}{\eta_n(y)} \check{K}_{a,n}(x, y) y.$$

Recall that, if \mathcal{H} is commutative or cocommutative, then the power law $\Psi^a \Psi^{a'} = \Psi^{aa'}$ holds. Thus long term behaviour of Hopf-power Markov chains may be deduced from

increasing the power further and further: taking m steps of the a th Hopf-power chain is equivalent to a single step of the a^m th Hopf-power chain. This will be relevant in Section 5.1.4, on approximations of absorbing probabilities using quasisymmetric functions.

Example 4.3.5. In the shuffle algebra of Example 4.1.1, for any word x , and any $c_1, \dots, c_n \in \mathcal{B}_1$, the coproduct structure constant $\eta_x^{c_1, \dots, c_n} = 0$ unless x is the concatenation of c_1, c_2, \dots, c_n in that order, in which case $\eta_x^{c_1, \dots, c_n} = 1$. So $\eta(x) = 1$ for all $x \in \mathcal{B}$, thus no rescaling of the basis is necessary to define the Hopf-power Markov chain. (No rescaling is necessary whenever η is a constant function on each \mathcal{B}_n - this constant may depend on n .)

Example 4.3.6. Take \mathcal{H} to be the Hopf algebra of representations of the symmetric groups, as in Example 4.1.4. \mathcal{B}_3 is the set of irreducible representations of \mathfrak{S}_3 , comprising the trivial representation, the sign representation and the two-dimensional irreducible representation. From explicit computation of $m\Delta = \bigoplus_{i=0}^3 \text{Ind}_{\mathfrak{S}_i \times \mathfrak{S}_{3-i}}^{\mathfrak{S}_3} \text{Res}_{\mathfrak{S}_i \times \mathfrak{S}_{3-i}}^{\mathfrak{S}_3}$ for these three representations, it follows that

$$K_{2,3} := [2^{-3}m\Delta]_{\mathcal{B}_3}^T = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{3}{4} \end{bmatrix}.$$

Observe that $(1, 1, 2)$, the vector of dimensions of these representations, is a (right) eigenvector of $K_{2,3}$ of eigenvalue 1, as predicted by Example 4.3.2. So applying the Doob transform to $K_{2,3}$ is to divide the third row by two and multiply the third column by 2, giving

$$\check{K}_{2,3} = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{8} & \frac{1}{8} & \frac{3}{4} \end{bmatrix}.$$

This is a transition matrix as its rows sum to 1. Example 4.4.3 below interprets this Markov chain as restriction-then-induction.

As promised, here is a check that η indeed takes positive values on a state space basis, and that, assuming $\mathcal{H}_1 \neq \emptyset$, there is no suitable rescaling function for bases which are not

state space bases (i.e. there are primitive basis elements of degree greater than one.) In this sense, η is an optimal rescaling function. Example 4.3.8 gives a numerical illustration of this second fact.

Theorem 4.3.7. *Suppose $\mathcal{H} = \bigoplus_{n \geq 0} \mathcal{H}_n$ is a graded connected Hopf algebra over \mathbb{R} with non-negative coproduct structure constants in the basis $\mathcal{B} = \coprod_{n \geq 0} \mathcal{B}_n$. Assume also that $\mathcal{H}_1 \neq \emptyset$.*

- (i) *If $\Delta(x) \neq 1 \otimes x + x \otimes 1$ for all $x \in \mathcal{B}$ with $\deg(x) > 1$, then the functions η_n of Definition 4.3.1 satisfy $\eta_{\deg x}(x) > 0$ for all $x \in \mathcal{B}$.*
- (ii) *If $\Delta(x) = 1 \otimes x + x \otimes 1$ for some $x \in \mathcal{B}_n$ with $n > 1$, then $\eta'_n(x) = 0$ for all eigenvectors η'_n of $\Psi^a : \mathcal{H}_n^* \rightarrow \mathcal{H}_n^*$ of highest eigenvalue.*

Proof. Recall that the intuition behind Part i is that “repeatedly breaking x non-trivially gives a way to reduce it to singletons”. So proceed by induction on $\deg x$. If $\deg x = 1$, then $\eta_1(x) = 1$ by definition. Otherwise, by hypothesis, $\bar{\Delta}(x) \neq 0$. Take a term $w \otimes z$ in $\bar{\Delta}(x)$, so $\eta_x^{wz} > 0$. Then the counit axiom forces $\deg w, \deg z < \deg x$. Consequently

$$\begin{aligned}
\eta_{\deg x}(x) &= (\bullet^*)^{\deg x}(x) \\
&= (\bullet^*)^{\deg w} (\bullet^*)^{\deg z}(x) \\
&= \left[(\bullet^*)^{\deg w} \otimes (\bullet^*)^{\deg z} \right] (\Delta x) \\
&= \left[(\bullet^*)^{\deg w} \otimes (\bullet^*)^{\deg z} \right] \left(\sum_{w', z' \in \mathcal{B}} \eta_x^{w'z'} w' \otimes z' \right) \\
&= \sum \eta_x^{w'z'} \eta_{\deg w}(w') \eta_{\deg z}(z')
\end{aligned}$$

where the last sum is over all $w' \in \mathcal{B}_{\deg w}, z' \in \mathcal{B}_{\deg z}$, because on all other summands, $(\bullet^*)^{\deg w} \otimes (\bullet^*)^{\deg z}$ evaluates to 0. The coproduct structure constants $\eta_x^{w'z'}$ are non-negative, and, by inductive hypothesis, $\eta_{\deg w}(w'), \eta_{\deg z}(z') > 0$. So all summands above are non-negative and the summand $\eta_x^{wz} \eta_{\deg w}(w') \eta_{\deg z}(z')$ is positive, so the sum is positive.

To see Part ii, it suffices to show that $\eta'_n(x) = 0$ for η' belonging to the basis in Theorem 2.6.2 of the eigenspace of $\Psi^a : \mathcal{H}_n^* \rightarrow \mathcal{H}_n^*$ of highest eigenvalue. Such basis eigenvectors

have the form $\eta' = \sum_{\sigma \in \mathfrak{S}_n} c_{\sigma(1)}^* \cdots c_{\sigma(n)}^*$ for some $c_1^*, \dots, c_n^* \in \mathcal{H}_1^*$. Now, because multiplication in \mathcal{H}^* is dual to comultiplication in \mathcal{H} ,

$$\begin{aligned}
 \eta'(x) &= \left(\sum_{\sigma \in \mathfrak{S}_n} c_{\sigma(1)}^* \cdots c_{\sigma(n)}^* \right) (x) \\
 &= \sum_{\sigma \in \mathfrak{S}_n} \left(c_{\sigma(1)}^* \otimes \cdots \otimes c_{\sigma(n)}^* \right) (\Delta^{[n]}x) \\
 &= \sum_{\sigma \in \mathfrak{S}_n} c_{\sigma(1)}^*(x) \otimes c_{\sigma(2)}^*(1) \otimes \cdots \otimes c_{\sigma(n)}^*(1) \\
 &\quad + c_{\sigma(1)}^*(1) \otimes c_{\sigma(2)}^*(x) \otimes c_{\sigma(3)}^*(1) \otimes \cdots \otimes c_{\sigma(n)}^*(1) + \dots \\
 &\quad + c_{\sigma(1)}^*(1) \otimes \cdots \otimes c_{\sigma(n-1)}^*(1) \otimes c_{\sigma(n)}^*(x) \\
 &= 0,
 \end{aligned}$$

since $c_{\sigma(i)}^*(x)$, $c_{\sigma(i)}^*(1)$ are all zero by degree considerations. (The third equality used that x is primitive.) \square

Example 4.3.8. Work in the algebra Λ of symmetric functions, and take \mathcal{B} to be the power sums, as described in Example 4.1.5. So $\mathcal{B}_3 = \{p_1^3, p_1 p_2, p_3\}$ and $\Delta(p_n) = 1 \otimes p_n + p_n \otimes 1$ for each n . By explicit computation,

$$K_{2,3} := [2^{-3} m\Delta]_{\mathcal{B}_3}^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{4} \end{bmatrix}.$$

Simply rescaling the basis \mathcal{B}_3 cannot make the rows of this matrix sum to 1, as rescaling the basis does not change the diagonal entries, and can only change non-zero non-diagonal entries.

It is easy to see how this problem generalises: for any primitive element $x \in \mathcal{B}_n$, it happens that $m\Delta(x) = 2x$, so the row corresponding to x in $K_{2,n}$ is 2^{-n+1} in the main diagonal and zeroes elsewhere. Then this row sum cannot change under basis rescaling.

To end this section, here is a brief word on how to modify the above notions for the case where $\mathcal{H}_1 = \emptyset$. As in Section 2.6, set $\mathcal{D} := \{d > 0 \mid \mathcal{H}_d \neq \emptyset\}$, $\mathcal{D}' = \{d \in \mathcal{D} \mid d \neq d_1 + d_2 \text{ with } d_1, d_2 \in \mathcal{D}\}$, so $\bigoplus_{d \in \mathcal{D}'} \mathcal{H}_d$ consists solely of primitive elements. Then define \mathcal{B} to be a state space basis if it contains no primitive elements outside of $\bigoplus_{d \in \mathcal{D}'} \mathcal{H}_d$. For each $n \in \mathcal{D}$, let $K(n)$ denote the maximal length of a \mathcal{D}' -partition of n , so, by Theorem 2.6.4, $a^{K(n)}$ is the largest eigenvalue of $\Psi^a : \mathcal{H}_n \rightarrow \mathcal{H}_n$. Then the value of the rescaling function $\eta_n(x)$ should be the sum of the coefficients of $\bar{\Delta}^{[K(n)]}(x)$, and the transition matrix of the Hopf-power Markov chain is $\check{K}_{a,n} := \left[a^{-K(n)} \Psi^a \right]_{\check{\mathcal{B}}_n}^T$, where $\check{\mathcal{B}}_n := \left\{ \check{x} := \frac{x}{\eta_n(x)} \mid x \in \mathcal{B}_n \right\}$.

4.4 Description of a Hopf-power Markov chain

Definition 4.3.4 gives the exact transition probabilities of a Hopf-power Markov chain, but this is not very enlightening without an intuitive description of the chain. Such descriptions can be very specific to the underlying Hopf algebra (see Theorem 5.3.8 regarding tree-pruning). The starting point to finding these interpretations is Theorem 4.4.1, which separates each timestep of the chain into breaking (steps 1 and 2) and recombining (step 3). The probabilities involved in both stages are expressed in terms of the structure constants of \mathcal{H} and the rescaling function η .

Theorem 4.4.1 (Three-step description for Hopf-power Markov chains). *A single step of the a th Hopf-power Markov chain, starting at $x \in \mathcal{B}_n$, is equivalent to the following three-step process:*

1. Choose a composition (i_1, \dots, i_a) of n (that is, non-negative integers with $i_1 + \dots + i_a = n$) according to the multinomial distribution with parameter $1/a$. In other words, choose (i_1, \dots, i_a) with probability $a^{-n} \binom{n}{i_1 \dots i_a}$.
2. Choose $z_1 \in \mathcal{B}_{i_1}, z_2 \in \mathcal{B}_{i_2}, \dots, z_a \in \mathcal{B}_{i_a}$ with probability $\frac{1}{\eta(x)} \eta_x^{z_1, \dots, z_a} \eta(z_1) \dots \eta(z_a)$.
3. Choose $y \in \mathcal{B}_n$ with probability $\left(\binom{n}{\deg_{z_1} \dots \deg_{z_a}} \eta(z_1) \dots \eta(z_a) \right)^{-1} \xi_{z_1, \dots, z_a}^y \eta(y)$.

Example 4.4.2. Applying Theorem 4.4.1 to the shuffle algebra \mathcal{S} recovers the description of the a -handed shuffle at the start of Section 4.1. Since the coproduct on \mathcal{S} is deconcatenation, the coproduct structure constant $\eta_x^{z_1, \dots, z_a} = 0$ unless x is the concatenation of

z_1, z_2, \dots, z_a in that order, so there is no choice at step 2. Hence steps 1 and 2 combined correspond to a multinomially-distributed cut of the deck. As for step 3: $\eta(y) = 1$ for all y , so y is chosen with probability proportional to ξ_{z_1, \dots, z_a}^y , the number of ways to interleave z_1, \dots, z_a to obtain y . Hence all interleavings are equally likely.

Example 4.4.3. How does Theorem 4.4.1 interpret the chain on the irreducible representations of the symmetric groups? Recall from Example 4.1.4 that the product is external induction and the coproduct is restriction. For simplicity, first take $a = 2$. Then, starting at a representation x of \mathfrak{S}_n , the first step is to binomially choose an integer i between 0 and n . It turns out that a cleaner description emerges if steps 2 and 3 above are combined. This merged step is to choose an irreducible representation y with probability proportional to $\sum \eta_x^{z_1 z_2} \xi_{z_1 z_2}^y \eta(y)$, where the sum is over all irreducible representations z_1 of \mathfrak{S}_i , and z_2 of \mathfrak{S}_{n-i} . Now $\sum \eta_x^{z_1 z_2} \xi_{z_1 z_2}^y$ is the coefficient or the multiplicity of the representation y in $\text{Ind}_{\mathfrak{S}_i \times \mathfrak{S}_{n-i}}^{\mathfrak{S}_n} \text{Res}_{\mathfrak{S}_i \times \mathfrak{S}_{n-i}}^{\mathfrak{S}_n}(x)$, and Example 4.3.2 showed that $\eta(y) = \dim y$. So the product of these two numbers have a neat interpretation as the dimension of the y isotypic component.

So, for general a , the chain on irreducible representations of the symmetric group has the following description:

1. Choose a Young subgroup $\mathfrak{S}_{i_1} \times \dots \times \mathfrak{S}_{i_a}$ according to a symmetric multinomial distribution.
2. Restrict the starting state x to the chosen subgroup, induce it back up to \mathfrak{S}_n , then pick an irreducible constituent with probability proportional to the dimension of its isotypic component.

A similar interpretation holds for other Hopf-power Markov chains on Hopf algebras of representations of other towers of algebras. For this particular case with the symmetric groups, this representation Hopf algebra is isomorphic to the cohomology of the infinite Grassmannian: the product is cup product, and the coproduct comes from a product on the infinite Grassmannian, which is taking direct sums of the subspaces. This isomorphism sends the basis \mathcal{B} of irreducible representations to the Schubert classes. So perhaps the restriction-then-induction chain on irreducible representations has an alternative interpretation in terms of decomposing a Schubert variety in terms of smaller Grassmannians, then taking the intersection.

A variant of this restriction-then-induction chain, where the choice of Young subgroup is fixed instead of random, appears in [Ful05]. There, it generates central limit theorems for character ratios, via Stein's method.

Proof of Theorem 4.4.1, the three-step description. First check that the probabilities in step 2 do sum to 1:

$$\begin{aligned}
& \sum_{z_1 \in \mathcal{B}_{i_1}, \dots, z_a \in \mathcal{B}_{i_a}} \eta_x^{z_1, \dots, z_a} \eta(z_1) \dots \eta(z_a) \\
&= \left((\bullet^*)^{i_1} \otimes \dots \otimes (\bullet^*)^{i_a} \right) \left(\sum_{z_1 \in \mathcal{B}_{i_1}, \dots, z_a \in \mathcal{B}_{i_a}} \eta_x^{z_1, \dots, z_a} z_1 \otimes \dots \otimes z_a \right) \\
&= \left((\bullet^*)^{i_1} \dots (\bullet^*)^{i_a} \right) (\Delta^a(x)) \\
&= (\bullet^*)^n(x) \\
&= \eta(x)
\end{aligned}$$

where the first equality uses Definition 4.3.1.iii of the rescaling function η_x , the second equality is because $(\bullet^*)^i(x_j) = 0$ if $\deg(x_j) \neq i$, and the third equality is by definition of the product of \mathcal{H}^* . And similarly for the probabilities in step 3, the combining step:

$$\begin{aligned}
\sum_{y \in \mathcal{B}_n} \xi_{z_1, \dots, z_a}^y \eta(y) &= (\bullet^*)^n \left(\sum_{y \in \mathcal{B}_n} \xi_{z_1, \dots, z_a}^y y \right) \\
&= (\bullet^*)^n(z_1 \dots z_a) \\
&= \Delta^a((\bullet^*)^n)(z_1 \otimes \dots \otimes z_a) \\
&= \left(\sum_{i_1, \dots, i_n} \binom{n}{i_1 \dots i_n} (\bullet^*)^{i_1} \otimes \dots \otimes (\bullet^*)^{i_n} \right) (z_1 \otimes \dots \otimes z_n) \\
&= \binom{n}{\deg z_1 \dots \deg z_n} \eta(z_1) \dots \eta(z_n).
\end{aligned}$$

Finally, the probability of moving from x to y under the three-step process is

$$\begin{aligned}
& \sum_{z_1 \dots z_a} a^{-n} \binom{n}{\deg z_1 \dots \deg z_a} \frac{\eta_x^{z_1, \dots, z_a} \eta(z_1) \dots \eta(z_a)}{\eta(x)} \frac{\xi_{z_1, \dots, z_a}^y \eta(y)}{\binom{n}{\deg z_1 \dots \deg z_a} \eta(z_1) \dots \eta(z_a)} \\
&= a^{-n} \sum_{z_1, \dots, z_a} \frac{\eta(y)}{\eta(x)} \xi_{z_1, \dots, z_a}^x \eta_y^{z_1, \dots, z_a} \\
&= \check{K}_{a,n}(x, y).
\end{aligned}$$

□

4.5 Stationary Distributions

The theorem below classifies all stationary distributions of a Hopf-power Markov chain; they have a simple expression in terms of the product structure constants and the rescaling function η of Definition 4.3.1.

Theorem 4.5.1 (Stationary distribution of Hopf-power Markov chains). *Follow the notation of Definition 4.3.4. Then, for each multiset $\{c_1, \dots, c_n\}$ in \mathcal{B}_1 , the function*

$$\pi_{c_1, \dots, c_n}(x) := \frac{\eta(x)}{n!^2} \sum_{\sigma \in \mathfrak{S}_n} \xi_{c_{\sigma(1)}, \dots, c_{\sigma(n)}}^x$$

is a stationary distribution for the a th Hopf-power Markov chain on \mathcal{B}_n , and any stationary distribution of this chain can be uniquely written as a linear combination of these π_{c_1, \dots, c_n} .

In particular,

(i) *if $\mathcal{B}_1 = \{\bullet\}$, then*

$$\pi_n(x) := \frac{\eta(x)}{n!} \xi_{\bullet, \dots, \bullet}^x$$

is the unique stationary distribution of the chain on \mathcal{B}_n ;

(ii) if \mathcal{H} is multigraded ($\mathcal{H} = \bigoplus_{\mathbf{v}} \mathcal{H}_{\mathbf{v}}$, $\mathcal{B} = \prod_{\mathbf{v}} \mathcal{B}_{\mathbf{v}}$) and $\mathcal{B}_{(1,0,\dots,0)} = \{\bullet_1\}$, $\mathcal{B}_{(0,1,0,\dots,0)} = \{\bullet_2\}$ and so on, then

$$\pi_{\mathbf{v}}(x) := \frac{\eta(x)}{n!^2} \sum_{\sigma \in \mathfrak{S}_n} \xi_{c_{\sigma(1)}, \dots, c_{\sigma(n)}}^x$$

with $c_1 = c_2 = \dots = c_{v_1} = \bullet_1$, $c_{v_1+1} = \dots = c_{v_1+v_2} = \bullet_2$, etc. is the unique stationary distribution of the chain on $\mathcal{B}_{\mathbf{v}}$;

and these are also necessary conditions.

Intuitively, the sum of product structure constants $\sum_{\sigma \in \mathfrak{S}_n} \xi_{c_{\sigma(1)}, \dots, c_{\sigma(n)}}^x$ counts the ways that x can be assembled from c_1, \dots, c_n in any order. So $\pi_{c_1, \dots, c_n}(x)$ is proportional to the number of ways to assemble x from c_1, \dots, c_n , and then repeatedly break it down into objects of size 1.

Proof. First, show that π_{c_1, \dots, c_n} is a probability distribution. As remarked in the proof of Lemma 4.2.1, $\xi_{c_{\sigma(1)}, \dots, c_{\sigma(n)}}^x \geq 0$, so π_{c_1, \dots, c_n} is a non-negative function. To see that $\sum_{x \in \mathcal{B}_n} \pi_{c_1, \dots, c_n}(x) = 1$, appeal to the second displayed equation of the proof of Theorem 4.4.1. Taking $a = n$, it shows that, for each $\sigma \in \mathfrak{S}_n$,

$$\sum_{x \in \mathcal{B}_n} \xi_{c_{\sigma(1)}, \dots, c_{\sigma(n)}}^x \eta(x) = \binom{n}{\deg c_{\sigma(1)} \dots \deg c_{\sigma(n)}} \eta(c_1) \dots \eta(c_n) = n! \cdot 1 \dots 1.$$

Next, recall that the stationary distributions are the left eigenfunctions of the transition matrix of eigenvalue 1. So, by Proposition 3.2.1.L, it suffices to show that $\sum_{x \in \mathcal{B}_n} \pi_{c_1, \dots, c_n}(x) \frac{x}{\eta(x)} = \sum_{\sigma \in \mathfrak{S}_n} c_{\sigma(1)} \dots c_{\sigma(n)}$ is a basis for the a^n -eigenspace of Ψ^a . This is precisely the assertion of 2.6.2.

Finally, the two uniqueness results are immediate by taking the sole choice of c_i s. \square

The first example below describes the typical behaviour when $\mathcal{B}_1 = \{\bullet\}$ and \mathcal{B} is a free-commutative basis: the unique stationary distribution is concentrated at a single state. Such a chain is said to be absorbing, and Sections 5.1.3 and 5.1.4 give some methods for estimating the probability of absorption after a given time.

Example 4.5.2. Continue with the edge-removal chain of Example 4.2.3, which arises from the Hopf algebra \mathcal{G} of graphs. Here, the only element of \mathcal{B}_1 is the graph \bullet with a single vertex. So Part i of Theorem 4.5.1 applies, and the unique stationary distribution is

$$\pi_n(x) = \frac{\eta(x)}{n!} \xi_{\bullet, \dots, \bullet}^x,$$

which is the point mass at the graph with no edges. This is because the structure constant $\xi_{\bullet, \dots, \bullet}^x$ is 0 for all other graphs x . Indeed, one would expect after many steps of this chain, that all edges would be removed.

Example 4.5.3. Continuing from Example 4.1.4, take \mathcal{H} to be the representation rings of the symmetric groups. The only irreducible representation of \mathfrak{S}_1 is the trivial representation, so again Theorem 4.5.1.i above applies. Now \bullet^n is the induced representation from $\mathfrak{S}_1 \times \dots \times \mathfrak{S}_1$ to \mathfrak{S}_n of the trivial representation, which gives the regular representation. So $\xi_{\bullet, \dots, \bullet}^x$ is the multiplicity of the irreducible representation x in the regular representation, which is $\dim x$. Recall from Example 4.3.2 that the rescaling constant $\eta(x)$ is also $\dim x$. Thus the unique stationary distribution of this restriction-then-induction chain is $\pi_n(x) = \frac{1}{n!^2} (\dim x)^2$. This is the well-studied Plancherel measure. It appears as the distribution of partitions growing one cell at a time under the Plancherel growth process [Ker99]. [Oko00] identifies its limit as $n \rightarrow \infty$, suitably rescaled, with the distribution of eigenvalues of a Gaussian random Hermitian matrix; the proof involves some combinatorially flavoured topology and illuminates a connection to the intersection theory on moduli spaces of curves.

Example 4.5.4. Consider $\mathcal{S}_{(1,1,\dots,1)}$, the degree $(1, 1, \dots, 1)$ subspace of the shuffle algebra. This corresponds to riffle-shuffling a distinct deck of cards. Use Theorem 4.5.1.ii with $c_i = (i)$. It is clear that, for each word x in $\mathcal{S}_{(1,\dots,1)}$, there is a unique way to interleave $(1), (2), \dots, (n)$ to obtain x . So $\xi_{(1), \dots, (n)}^x = 1$, and by commutativity, $\xi_{(\sigma(1)), \dots, (\sigma(n))}^x = 1$ for all permutations σ . Recall also that $\eta(x) = 1$ for all words. So the unique stationary distribution for riffle-shuffling is the uniform distribution $\pi(x) \equiv \frac{1}{n!}$.

All the chains appearing in this thesis have unique stationary distributions. For an example of a Hopf-power Markov chain with several absorbing states, see Pineda's example on the Hopf monoid of permutohedra [Pin14].

4.6 Reversibility

Recall from Section 3.3 that the time-reversal of a Markov chain from a linear map is given by the dual map. As observed in Section 2.1, the dual map to $\Psi^a : \mathcal{H}_n \rightarrow \mathcal{H}_n$ is the Hopf-power map on the dual Hopf algebra, $\Psi^a : \mathcal{H}_n^* \rightarrow \mathcal{H}_n^*$. Thus Theorem 3.3.2 specialises to the following for Hopf-power chains:

Theorem 4.6.1 (Time-reversal of Hopf-power Markov chains). *Let \mathcal{H} be a graded, connected Hopf algebra over \mathbb{R} with state space basis \mathcal{B} satisfying $\mathcal{B} = \{\bullet\}$ (or \mathcal{H} is multi-graded and $\mathcal{B}_{(1,0,\dots,0)} = \{\bullet_1\}$, $\mathcal{B}_{(0,1,0,\dots,0)} = \{\bullet_2\}$ and so on). Suppose in addition that, for all $y \in \mathcal{B}$ with $\deg(y) > 1$, there is some $w, z \in \mathcal{B}$ with $\deg(w), \deg(z) > 0$ such that $\xi_{wz}^y \neq 0$. Then the time-reversal of the a th Hopf-power Markov chain on \mathcal{B}_n (or \mathcal{B}_v) is the a th Hopf-power Markov chain on the dual basis \mathcal{B}_n^* (or \mathcal{B}_v^*) of the (graded) dual Hopf algebra \mathcal{H}^* . \square*

Note the the condition $\xi_{wz}^y \neq 0$ is equivalent to \mathcal{B}^* being a state space basis, since dualising the Hopf algebra simply exchanges the product and coproduct structure constants:

$$\xi_{w^*z^*}^{x^*} = \eta_x^{wz}; \quad \eta_y^{w^*z^*} = \xi_{wz}^y.$$

Then, applying Theorem 4.3.7.i to \mathcal{H}^* implies $\xi_{\bullet,\dots,\bullet}^y > 0$ for all $y \in \mathcal{B}$. So the stationary distribution of the Hopf-power chain on \mathcal{H} is nowhere zero, and the time-reversal chain is indeed defined.

Example 4.6.2. Recall from Example 2.1.2 that the dual of the shuffle algebra \mathcal{S} is the free associative algebra \mathcal{S}^* , with concatenation product and deshuffling coproduct. Its associated Hopf-square Markov chain has this interpretation in terms of decks of cards: uniformly and independently assign each card to the left or right pile, keeping cards which land in the same pile in the same relative order, then put the left pile on top of the right pile. This agrees with the description of inverse shuffling of [BD92, Sec. 3].

The final result of Section 3.3 states that, if Ψ is self-adjoint with respect to an inner product where the state space basis is orthogonal, and if a Ψ -Markov chain has a well-defined time-reversal, then this chain is reversible. The condition that the Hopf-power be

self-adjoint is a little odd; a stronger but more natural hypothesis is that the product and coproduct are adjoint, in the manner described below.

Theorem 4.6.3 (Reversibility of Hopf-power Markov chains). *Let \mathcal{H} be a graded, connected Hopf algebra over \mathbb{R} equipped with an inner product $\langle \cdot, \cdot \rangle$ adjoining product and coproduct, that is, $\langle wz, x \rangle = \langle w \otimes z, \Delta(x) \rangle$. (Here, $\langle w \otimes z, a \otimes b \rangle = \langle w, a \rangle \langle z, b \rangle$.) Let \mathcal{B} be a state space basis of \mathcal{H} which is orthogonal under this inner product, with $\mathcal{B}_1 = \{\bullet\}$ (or \mathcal{H} is multigraded and $\mathcal{B}_{(1,0,\dots,0)} = \{\bullet_1\}$, $\mathcal{B}_{(0,1,0,\dots,0)} = \{\bullet_2\}$ and so on). Assume in addition that, for all $y \in \mathcal{B}$ with $\deg(y) > 1$, there is some $w, z \in \mathcal{B}$ with $\deg(w), \deg(z) > 0$ such that $\xi_{wz}^y \neq 0$. Then the a th Hopf-power Markov chain on \mathcal{B}_n (or \mathcal{B}_v) is reversible. \square*

Zelevinsky's classification [Zel81, Th. 2.2, 3.1] of *positive self-dual* Hopf algebras says that, if one restricts to Hopf algebras with integral structure constants, then the example below is essentially the only chain satisfying the hypothesis of Theorem 4.6.3 above.

Example 4.6.4. Equip the representation rings of the symmetric group with the usual inner product where the irreducible representations are orthonormal. (This is equivalent to the Hall inner product of symmetric functions, see [Sta99, Sec. 7.9].) That this inner product adjoins the multiplication and comultiplication is simply Frobenius reciprocity:

$$\langle \text{Ind}_{\mathfrak{S}_i \times \mathfrak{S}_j}^{\mathfrak{S}_{i+j}} w \times z, x \rangle = \langle w \otimes z, \text{Res}_{\mathfrak{S}_i \times \mathfrak{S}_j}^{\mathfrak{S}_{i+j}} x \rangle.$$

(Note that, if w, z are representations of $\mathfrak{S}_i, \mathfrak{S}_j$ respectively, then $\langle w \otimes z, \text{Res}_{\mathfrak{S}_k \times \mathfrak{S}_{i+j-k}}^{\mathfrak{S}_{i+j}} x \rangle = 0$ unless $k = i$.) As calculated in Example 4.5.3, the associated restriction-then-induction chain has a unique stationary distribution given by the Plancherel measure $\pi(x) = \frac{\dim x^2}{n!} > 0$. So this chain is reversible.

4.7 Projection

Recall the mantra of Section 3.4: intertwining maps give rise to projections of Markov chains. For Hopf-power Markov chains, the natural maps to use are Hopf-morphisms. A linear map $\theta : \mathcal{H} \rightarrow \mathcal{H}$ is a *Hopf-morphism* if $\theta(1) = 1$, $\deg(\theta(x)) = \deg(x)$, $\theta(w)\theta(z) =$

$\theta(wz)$ and $\Delta(\theta(x)) = (\theta \otimes \theta)(\Delta(x))$ for all $x, w, z \in \mathcal{H}$. Then

$$\theta(m\Delta(x)) = m(\theta \otimes \theta)(\Delta(x)) = m\Delta(\theta(x)),$$

so θ intertwines the Hopf-square maps on \mathcal{H} and on $\bar{\mathcal{H}}$. Indeed, a simple (co)associativity argument shows that $\theta m^{[a]} = m^{[a]}\theta^{\otimes a}$ and $\theta^{\otimes a}\Delta^{[a]} = \Delta^{[a]}\theta$ for all a , so $\theta\Psi_{\mathcal{H}}^a = \Psi_{\bar{\mathcal{H}}}^a\theta$. (Note that Ψ^a is not a Hopf-morphism in general.)

Specialising Theorem 3.4.1, concerning projections of chains from linear maps, to the Hopf-power map, gives the following:

Theorem 4.7.1 (Projection Theorem for Hopf-power Markov Chains). *Let \mathcal{H} , $\bar{\mathcal{H}}$ be graded, connected Hopf algebras over \mathbb{R} with bases \mathcal{B} , $\bar{\mathcal{B}}$ respectively. Suppose in addition that \mathcal{B} is a state space basis. If $\theta : \mathcal{H} \rightarrow \bar{\mathcal{H}}$ is a Hopf-morphism such that $\theta(\mathcal{B}_n) = \bar{\mathcal{B}}_n$ for some n , and $\theta(\mathcal{B}_1) \subseteq \bar{\mathcal{B}}_1$, then the Hopf-power Markov chain on $\bar{\mathcal{B}}_n$ is the projection via θ of the Hopf-power Markov chain on \mathcal{B}_n .*

Remarks.

1. As in the more general Theorem 3.4.1, the condition $\theta(\mathcal{B}_n) = \bar{\mathcal{B}}_n$ does not mean that the restriction $\theta : \mathcal{H}_n \rightarrow \bar{\mathcal{H}}_n$ is an isomorphism. Although θ must be surjective onto $\bar{\mathcal{B}}_n$, it need not be injective - the requirement is simply that distinct images of \mathcal{B}_n under θ are linearly independent.
2. The theorem does not require $\theta(\mathcal{B}_n) = \bar{\mathcal{B}}_n$ to hold for all n . Section 6.2.3, regarding the descent sets under riffle-shuffling, is an important example where the domain \mathcal{H} is multigraded, and $\theta(\mathcal{B}_\nu) = \bar{\mathcal{B}}_{|\nu|}$ for only certain values of ν .
3. The proof will show that the weaker assumption $\theta(\mathcal{B}_1) \subseteq \alpha\bar{\mathcal{B}}_1 := \{\alpha\bar{c} \mid \bar{c} \in \bar{\mathcal{B}}_1\}$ is sufficient. (Here, α can be any non-zero constant.)

Proof. As discussed before the statement of the theorem, $\theta\Psi_{\mathcal{H}}^a = \Psi_{\bar{\mathcal{H}}}^a\theta$. So it suffices to show that the condition $\theta(\mathcal{B}_1) \subseteq \bar{\mathcal{B}}_1$ guarantees $\eta(x) = \eta(\theta(x))$ for all $x \in \mathcal{B}_n$. Then Theorem 3.4.1, concerning projections of chains from linear maps, applies to give the desired result.

Let $n = \deg x = \deg(\theta(x))$. Recall that the rescaling function $\eta(x)$ is the sum of the coefficients of $\bar{\Delta}^{[n]}(x)$ when expanded in the basis $\mathcal{B}_1^{\otimes n}$:

$$\eta(x) = \sum_{c_1, \dots, c_n \in \mathcal{B}_1} \eta_x^{c_1, \dots, c_n},$$

so

$$\eta(\theta(x)) = \sum_{c_1, \dots, c_n \in \bar{\mathcal{B}}_1} \eta_{\theta(x)}^{c_1, \dots, c_n}.$$

Now expanding the equality $\bar{\Delta}^{[n]}(\theta(x)) = \theta^{\otimes n}(\bar{\Delta}^{[n]}(x))$ in the basis $\mathcal{B}_1^{\otimes n}$ gives:

$$\begin{aligned} \sum_{\bar{c}_1, \dots, \bar{c}_n \in \bar{\mathcal{B}}_1} \eta_{\theta(x)}^{\bar{c}_1, \dots, \bar{c}_n} \bar{c}_1 \otimes \dots \otimes \bar{c}_n &= \theta^{\otimes n} \left(\sum_{c_1, \dots, c_n \in \mathcal{B}_1} \eta_x^{c_1, \dots, c_n} c_1 \otimes \dots \otimes c_n \right) \\ &= \sum_{c_1, \dots, c_n \in \mathcal{B}_1} \eta_x^{c_1, \dots, c_n} \theta(c_1) \otimes \dots \otimes \theta(c_n) \\ &= \sum_{\bar{c}_1, \dots, \bar{c}_n \in \bar{\mathcal{B}}_1} \left(\sum_{c_1, \dots, c_n, \theta(c_i) = \bar{c}_i} \eta_x^{c_1, \dots, c_n} \right) \bar{c}_1 \otimes \dots \otimes \bar{c}_n, \end{aligned}$$

where the last equality uses the assumption $\theta(c_i) \in \mathcal{B}_1$. So the coefficient sums of the left and right hand sides are equal, and these are $\eta(\theta(x))$ and $\eta(x)$ respectively. \square

Example 4.7.2. Work in \mathcal{S}^* , the free associative algebra introduced in Example 2.1.2, where the product of two words is their concatenation, and the coproduct is deshuffle. As seen in Example 4.6.2, the associated Hopf-power Markov chain describes inverse riffle-shuffling: randomly place each card on the left or right pile, then place the left pile on top of the right. Let $\bar{\mathcal{S}}^*$ be the quotient of \mathcal{S}^* , as an algebra, by the relations $\{ij = ji \mid |i - j| > 1\}$. Then $\bar{\mathcal{S}}^*$ is one example of a free partially commutative algebra of [Sch90], based on the free partially commutative monoids of [CF69]. The technical Lemmas 4.7.4 and 4.7.5 below prove respectively that the quotient map $\mathcal{S}^* \rightarrow \bar{\mathcal{S}}^*$ is a map of Hopf algebras, and that this map sends the basis of words of \mathcal{S}^* to a basis of $\bar{\mathcal{S}}^*$. Thus this quotient map shows that inverse riffle-shuffling while forgetting the orders of cards with nonconsecutive values is a Markov chain. For example, this would identify (231124) with (213412). When all cards in the deck are distinct, this amounts to keeping track only of whether card 1 is

above or below card 2, whether card 2 is above or below card 3, etc. This statistic is known as the *idescent set* (or recoil):

$$\text{ides}(w) = \{i \mid i+1 \text{ occurs before } i \text{ in } w\}$$

as it is the descent set of the inverse of w , when regarding w as a permutation in one-line notation. The projection of inverse riffle-shuffling by idescent set is studied in [AD10, Ex. 5.12.ii].

The same construction goes through for other sets of commutation relations. Specifically, let G be a graph with vertex set $\{1, 2, \dots\}$ and finitely-many edges, and set $\tilde{\mathcal{S}}^*_G$ to be the quotient of \mathcal{S}^* , as an algebra, by the relations $\{ij = ji \mid (i, j) \text{ not an edge of } G\}$. Thus the edges of G indicate noncommuting pairs of letters in $\tilde{\mathcal{S}}^*_G$. The example above, where only nonconsecutive values commute, corresponds to a path. The Lemmas below show that, for any graph G , the quotient map $\theta_G : \mathcal{S} \rightarrow \tilde{\mathcal{S}}^*_G$ satisfies the conditions of the Projection Theorem, so these maps all give Markov statistics for inverse shuffling. To interpret these statistics, appeal to [KMLR82, Prop. 2]. For a word w , let w_{ij} denote the subword of w obtained by deleting all letters not equal to i or j . Thus $(231124)_{12} = (2112)$, $(231124)_{23} = (232)$. Then their proposition asserts that $\theta_G(w)$ is recoverable from the set of w_{ij} over all edges (i, j) of G . To summarise:

Theorem 4.7.3. *Let G be a graph with vertex set $\{1, 2, \dots\}$ and finitely-many edges. For a deck of cards w , let w_{ij} be the subdeck obtained by throwing out all cards not labelled i or j . Then the set of all w_{ij} over all edges (i, j) of G is a Markov statistic under inverse shuffling. \square*

Below are the promised technical Lemmas necessary to establish this result.

Lemma 4.7.4. *Let G be a graph with vertex set $\{1, 2, \dots\}$ and finitely-many edges. Denote by I_G the ideal in the free associative algebra \mathcal{S}^* generated by $\{ij - ji \mid (i, j) \text{ not an edge of } G\}$. Then I_G is also a coideal, (i.e. $\Delta(I_G) \subseteq \mathcal{S}^* \otimes I_G + I_G \otimes \mathcal{S}^*$), so the quotient $\tilde{\mathcal{S}}^*_G := \mathcal{S}^*/I_G$ is a Hopf algebra.*

Proof. Since Δ is linear and $\Delta(xy) = \Delta(x)\Delta(y)$, it suffices to check the coideal condition only on the generators of I_G , that is, that $\Delta(ij - ji) \subseteq \mathcal{S}^* \otimes I_G + I_G \otimes \mathcal{S}^*$ whenever (i, j)

is not an edge of G . Now

$$\begin{aligned}
\Delta(ij - ji) &= \Delta(i)\Delta(j) - \Delta(j)\Delta(i) \\
&= (1 \otimes i + i \otimes 1)(1 \otimes j + j \otimes 1) - (1 \otimes j + j \otimes 1)(1 \otimes i + i \otimes 1) \\
&= 1 \otimes ij + j \otimes i + i \otimes j + ij \otimes 1 - (1 \otimes ji + i \otimes j + j \otimes i + ji \otimes 1) \\
&= 1 \otimes ij + ij \otimes 1 - 1 \otimes ji - ji \otimes 1 \\
&= 1 \otimes (ij - ji) + (ij - ji) \otimes 1 \\
&\subseteq \mathcal{S}^* \otimes I_G + I_G \otimes \mathcal{S}^*.
\end{aligned}$$

□

Lemma 4.7.5. *Let $\theta_G : \mathcal{S}^* \rightarrow \bar{\mathcal{S}}^*_G$ be the quotient map, by the ideal I_G in Lemma 4.7.4 above. Write \mathcal{B} the basis of words in the free associative algebra \mathcal{S}^* . Then $\bar{\mathcal{B}} := \theta_G(\mathcal{B})$ is a basis of $\bar{\mathcal{S}}^*_G$.*

Proof. (The main idea of this proof arose from a discussion with Zeb Brady.) Clearly $\bar{\mathcal{B}}$ spans $\bar{\mathcal{S}}^*_G$, so the only issue is linear independence. This will follow from

$$I_G = J := \left\{ a_1 b_1 + \cdots + a_m b_m \mid b_i \in \mathcal{B}, \sum_{i: \theta_G(b_i) = \bar{b}} a_i = 0 \text{ for each } \bar{b} \in \bar{\mathcal{B}} \right\}.$$

The quotient map θ_G clearly sends each element of J to 0, so $J \subseteq \ker \theta_G = I_G$. To see $I_G \subseteq J$, it suffices to show that J is an ideal containing the generators $ij - ji$ of I_G . First, J is clearly closed under addition. J is closed under multiplication by elements of \mathcal{S}^* because, for any letter c (i.e. any generator of \mathcal{S}^*), $c(a_1 b_1 + \cdots + a_m b_m) = a_1 (cb_1) + \cdots + a_m (cb_m)$ with each $cb_i \in \mathcal{B}$, and, if $\theta_G(b_i) = \theta_G(b_j)$, then $\theta_G(cb_i) = \theta_G(cb_j)$. Lastly, if (i, j) is not an edge of G , then $\theta_G(ij) = \theta_G(ji)$, so $ij - ji \in J$. □

Chapter 5

Hopf-power Markov chains on Free-Commutative Bases

This chapter concentrates on a class of Hopf-power Markov chains whose behaviour is “simple” in two ways, thanks to the additional hypothesis that the state space basis is free-commutative, as defined below.

Definition (Free generating set, free-commutative basis). Let \mathcal{H} be a graded connected commutative Hopf algebra over \mathbb{R} . Then the dual Cartier-Milnor-Moore theorem [Car07, Th. 3.8.3] states that \mathcal{H} is isomorphic as an algebra to the polynomial algebra $\mathbb{R}[c_1, c_2, \dots]$ for some elements c_i , which may have any degree. (In fact, it suffices that the base field be of characteristic 0.) The set $\mathcal{C} := \{c_1, c_2, \dots\}$ is a *free generating set* for \mathcal{H} , and the basis $\mathcal{B} = \{c_1 \dots c_l \mid l \in \mathbb{N}, \{c_1, \dots, c_l\} \text{ a multiset in } \mathcal{C}\}$, consisting of all products of the c_i s, is a *free-commutative basis*.

One can think of a free-commutative basis as the basis of monomials in the c_i , but this thesis prefers to reserve the terminology “monomial” for analogues of the monomial symmetric functions, which are cofree.

An archetypal chain on a free-commutative basis is the edge-removal of graphs (or indeed the analogous construction for any species-with-restrictions, as discussed in Section 4.1.1). Specialising to disjoint unions of complete graphs gives the independent multinomial breaking of rocks, as discussed in Section 5.2.

Example (Edge-removal of graphs). Recall from Examples 4.1.3 and 4.2.3 the Hopf algebra \mathcal{G} of graphs: the degree $\deg(G)$ of a graph G is its number of vertices, the product of two graphs is their disjoint union, and the coproduct is

$$\Delta(G) = \sum G_S \otimes G_{S^c}$$

where the sum is over all subsets S of the vertex set of G , and G_S, G_{S^c} denote the subgraphs that G induces on the vertex set S and its complement. The set \mathcal{B} of all graphs is a free-commutative basis, and the free generating set \mathcal{C} consists of the connected graphs.

The a th Hopf-power Markov chain describes edge-removal: at each step, assign uniformly and independently one of a colours to each vertex, and remove the edges connecting vertices of different colours. There is no need to rescale the state space basis \mathcal{B} to define this chain: for all graphs G with n vertices, the rescaling function $\eta(G)$ counts the ways to break G into n (ordered) singletons, of which there are $n!$, irrespective of G . An easy application of Theorem 4.5.1 shows that its unique stationary distribution takes value 1 on the graph with no edges and 0 on all other states, so the chain is *absorbing*.

The first “simplicity” feature of this edge-removal chain is that each connected component behaves independently. Section 5.1.1 explains the analogous behaviour for all chains on a free-commutative basis as a consequence of the Hopf-power map Ψ^a being an algebra homomorphism, since the underlying Hopf algebra is commutative. The second aspect of interest is that the edge-removal chain never returns to a state it has left. Indeed, at each step the chain either stays at the same graph or the number of connected components increases. Section 5.1.2 will show that a Hopf-power Markov chain on a free-commutative state space basis always has a triangular transition matrix; then, applying Perron-Frobenius to each minor gives right eigenfunctions that are non-negative in the first few coordinates and zero in the last coordinates. Section 5.1.3 identifies these as the output of Theorem 2.5.1.B, and outlines how they give upper bounds for the probability of being “far from absorbed”. Section 5.1.4 then repackages the exact probabilities in terms of a “generalised chromatic quasisymmetric function” constructed in [ABS06], though this is a theoretical discussion only as I have no effective way to compute or bound such functions. This appears to require weaker hypotheses than a free-commutative state space basis, but it is

unclear whether there are non-free-commutative state space bases that satisfy the weaker hypotheses, nor what the conclusions mean in this more general setup.

Sections 5.2 and 5.3 apply these techniques to a rock-breaking and tree-pruning process respectively, arising from the algebra of symmetric functions and the Connes-Kreimer algebra of rooted forests.

5.1 General Results

5.1.1 Independence

The following theorem converts the fact that Ψ^a is an algebra homomorphism into “independent breaking” of the Hopf-power Markov chain if the starting state is a product. For example, in the Hopf algebra \mathcal{G} of graphs, a graph is the product of its connected components, so the associated edge-removal Markov chain behaves independently on each connected component. As a result, to understand a Hopf-power Markov chain on a free-commutative basis, it suffices to describe one step of the chain starting only from the generators, i.e. to apply Theorem 4.4.1, the three-step interpretation, only to states which are not products.

Theorem 5.1.1. *Let $x_1, x_2 \in \mathcal{B}$, a free-commutative state space basis. Then one step of the a th Hopf-power Markov chain on \mathcal{B} starting at $x := x_1 x_2$ is equivalent to the following: take one step of the a th Hopf-power Markov chain from x_1 and from x_2 , and move to the product of the results.*

Proof. Let n, n_1, n_2 be the degrees of x, x_1, x_2 respectively. By definition, the probability of moving from x to y in the a th Hopf-power Markov chain is

$$\check{K}_{a,n}(x,y) = y^*(a^{-n}\Psi^a(x))\frac{\eta(y)}{\eta(x)}.$$

So the probability of moving from x to y under the composite process described in the theorem is

$$\begin{aligned} & \sum_{y_1 y_2 = y} \check{K}_{a,n_1}(x_1, y_1) \check{K}_{a,n_2}(x_2, y_2) \\ &= \sum_{y_1 y_2 = y} y_1^*(a^{-n_1} \Psi^a(x_1)) \frac{\eta(y_1)}{\eta(x_1)} y_2^*(a^{-n_2} \Psi^a(x_2)) \frac{\eta(y_2)}{\eta(x_2)}. \end{aligned}$$

Since \mathcal{B} is a free-commutative basis, the structure constant $\xi_{y_1 y_2}^y$ is 1 if $y_1 y_2 = y$, and 0 otherwise. So the above probability is

$$\begin{aligned} & \sum_{y_1 \in \mathcal{B}_{n_1}, y_2 \in \mathcal{B}_{n_2}} \xi_{y_1 y_2}^y (y_1^* \otimes y_2^*) (a^{-n} \Psi^a(x_1) \otimes \Psi^a(x_2)) \frac{\eta(y_1)}{\eta(x_1)} \frac{\eta(y_2)}{\eta(x_2)} \\ &= \Delta^*(y) (a^{-n} \Psi^a(x_1) \otimes \Psi^a(x_2)) \frac{\eta(y_1)}{\eta(x_1)} \frac{\eta(y_2)}{\eta(x_2)} \\ &= y^*(a^{-n} \Psi^a(x_1) \Psi^a(x_2)) \frac{\eta(y_1)}{\eta(x_1)} \frac{\eta(y_2)}{\eta(x_2)} \\ &= y^*(a^{-n} \Psi^a(x_1 x_2)) \frac{\eta(y_1)}{\eta(x_1)} \frac{\eta(y_2)}{\eta(x_2)}. \end{aligned}$$

The last step uses that Ψ^a is an algebra homomorphism since the Hopf algebra is commutative. Lemma 5.1.2 below shows that $\frac{\eta(y_1)}{\eta(x_1)} \frac{\eta(y_2)}{\eta(x_2)} = \frac{\eta(y_1 y_2)}{\eta(x_1 x_2)} = \frac{\eta(y)}{\eta(x)}$, so this probability is indeed $\check{K}_{a,n}(x, y)$. \square

Lemma 5.1.2. *The rescaling function η satisfies*

$$\eta(x_1 x_2) = \binom{\deg(x_1 x_2)}{\deg(x_1)} \eta(x_1) \eta(x_2).$$

In other words, $\frac{\eta(x)}{(\deg x)!}$ is multiplicative.

Proof. There is a short proof via $\eta(x) = (\bullet^*)^{\deg x}$, but the enumerative argument here is more transparent and more versatile - similar lines of reasoning lie behind Proposition 5.1.13 and (to a lesser extent) Theorems 5.3.6 and 5.3.10.

Write n, n_1, n_2 for the degrees of x, x_1, x_2 respectively. $\eta(x_1x_2)$ is the sum of the coefficients of $\bar{\Delta}^{[n]}(x_1x_2)$. The Hopf axiom $\Delta(x_1x_2) = \Delta(x_1)\Delta(x_2)$ gives the following bijection:

$$\left\{ \begin{array}{l} \text{terms in} \\ \bar{\Delta}^{[n]}(x_1x_2) \end{array} \right\} \leftrightarrow \left\{ \begin{array}{l} \text{terms in} \\ \bar{\Delta}^{[n_1]}(x_1) \end{array} \right\} \times \left\{ \begin{array}{l} \text{terms in} \\ \bar{\Delta}^{[n_2]}(x_2) \end{array} \right\} \times \left\{ \begin{array}{l} \text{choices of } n_1 \text{ tensor-factors} \\ \text{amongst } n \text{ to place} \\ \text{the term from } \bar{\Delta}^{[n]}(x_1) \end{array} \right\}.$$

Taking coefficients of both sides recovers the lemma. \square

5.1.2 Unidirectionality

Call a Markov chain *unidirectional* if it cannot return to any state it has left. (The term “unidirectional” is a suggestion from John Pike, since “monotone” and “acyclic” already have technical meanings in Markov chain theory.) An equivalent phrasing is that the state space is a poset under the relation “is accessible from”. Yet another characterisation of a unidirectional chain is that its transition matrix is triangular for some suitable ordering of the states.

The edge-removal chain at the start of this chapter is unidirectional as the chain either stays at the current graph, or the number of connected components increases. Corollary 5.1.5 below shows that this phenomenon occurs for all Hopf-power Markov chains on a free-commutative basis. The generalisation of “number of connected components” is the *length*: for $x \in \mathcal{B}$, its length $l(x)$ is the number of factors in the factorisation of x into generators. Lemma 5.1.3 below explains the way the length changes under product and coproduct. It requires one more piece of notation: define $x \rightarrow x'$ for $x, x' \in \mathcal{B}$ if x' appears in $\Psi^a(x)$ (when expanded in the basis \mathcal{B}) for some a . This is precisely the relation “is accessible from” discussed in the previous paragraph.

Lemma 5.1.3. *Let $x, y, x_i, x_{(i)}$ be elements of a free-commutative basis. Then*

- (i) $l(x_1 \dots x_a) = l(x_1) + \dots + l(x_a)$;
- (ii) *For any summand $x_{(1)} \otimes \dots \otimes x_{(a)}$ in $\Delta^{[a]}(x)$, $l(x_{(1)}) + \dots + l(x_{(a)}) \geq l(x)$;*
- (iii) *if $x \rightarrow y$, then $l(y) \geq l(x)$.*

Proof. (i) is clear from the definition of length.

Prove (ii) by induction on $l(x)$. Note that the claim is vacuously true if x is a generator, as each $l(x_{(i)}) \geq 0$, and not all $l(x_{(i)})$ may be zero. If x factorises non-trivially as $x = st$, then, as $\Delta^{[a]}(x) = \Delta^{[a]}(s)\Delta^{[a]}(t)$, it must be the case that $x_{(i)} = s_{(i)}t_{(i)}$, for some $s_{(1)} \otimes \cdots \otimes s_{(a)}$ in $\Delta^{[a]}(s)$, $t_{(1)} \otimes \cdots \otimes t_{(a)}$ in $\Delta^{[a]}(t)$. So $l(x_{(1)}) + \cdots + l(x_{(a)}) = l(s_{(1)}) + \cdots + l(s_{(a)}) + l(t_{(1)}) + \cdots + l(t_{(a)})$ by (i), and by inductive hypothesis, this is at least $l(s) + l(t) = l(x)$.

(iii) follows trivially from (i) and (ii): if $x \rightarrow y$, then $y = x_{(1)} \cdots x_{(a)}$ for a term $x_{(1)} \otimes \cdots \otimes x_{(a)}$ in $\Delta^{[a]}(x)$. So $l(y) = l(x_{(1)}) + \cdots + l(x_{(a)}) \geq l(x)$. \square

Here is the algebraic fact which causes unidirectionality; the proof is four paragraphs below.

Proposition 5.1.4. *Let \mathcal{H} be a Hopf algebra with free-commutative basis \mathcal{B} , where all coproduct structure constants η_x^{wz} are non-negative. Then the relation \rightarrow defines a partial order on \mathcal{B} , and the partial-ordering by length refines this partial-order: if $x \rightarrow y$ and $x \neq y$, then $l(x) < l(y)$. Furthermore, for any integer a and any $x \in \mathcal{B}$,*

$$\Psi^a(x) = a^{l(x)}x + \sum_{l(y) > l(x)} \alpha_{xy}y$$

for some $\alpha_{xy} \geq 0$.

The probability consequence is immediate from Definition 4.3.4 of a Hopf-power Markov chain:

Corollary 5.1.5. *Let $\{X_m\}$ be the a th Hopf-power Markov chain on a free-commutative basis \mathcal{B}_n . Then*

$$P\{X_{m+1} = x | X_m = x\} = a^{l(x)-n},$$

and $P\{X_{m+1} = y | X_m = x\}$ is non-negative only if $l(y) \geq l(x)$. \square

In other words, if the states are totally ordered to refine the partial-ordering by length, then the transition matrices are upper-triangular with a^{l-n} on the main diagonal. In particular, states with length n are absorbing - which also follows from the stationary distribution expressions in Theorem 4.5.1. These states are precisely the products of elements of \mathcal{B}_1 .

Proof of Proposition 5.1.4. It is easier to first prove the expression for $\Psi^a(x)$. Suppose x has factorisation into generators $x = c_1 c_2 \dots c_{l(x)}$. As \mathcal{H} is commutative, Ψ^a is an algebra homomorphism, so $\Psi^a(x) = \Psi^a(c_1) \dots \Psi^a(c_{l(x)})$. Recall from Section 2.2 that $\bar{\Delta}(c) = \Delta(c) - 1 \otimes c - c \otimes 1 \in \bigoplus_{i=1}^{\deg(c)-1} \mathcal{H}_i \otimes \mathcal{H}_{\deg(c)-i}$, in other words, $1 \otimes c$ and $c \otimes 1$ are the only terms in $\Delta(c)$ which have a tensor-factor of degree 0. As $\Delta^{[3]} = (\iota \otimes \Delta)\Delta$, the only terms in $\Delta^{[3]}(c)$ with two tensor-factors of degree 0 are $1 \otimes 1 \otimes c$, $1 \otimes c \otimes 1$ and $c \otimes 1 \otimes 1$. Inductively, we see that the only terms in $\Delta^{[a]}(c)$ with all but one tensor-factor having degree 0 are $1 \otimes \dots \otimes 1 \otimes c$, $1 \otimes \dots \otimes 1 \otimes c \otimes 1, \dots, c \otimes 1 \otimes \dots \otimes 1$. So $\Psi^a(c) = ac + \sum_{l(y) > 1} \alpha_{cy} y$ for generators c , and $\alpha_{cy} \geq 0$ by the hypothesis that all coproduct structure constants are non-negative. As $\Psi^a(x) = \Psi^a(c_1) \dots \Psi^a(c_l)$, and length is multiplicative (Lemma 5.1.3.i), the expression for $\Psi^a(x)$ follows.

It is then clear that \rightarrow is reflexive and antisymmetric. Transitivity follows from the power rule: if $x \rightarrow y$ and $y \rightarrow y'$, then y appears in $\Psi^a(x)$ for some a and y' appears in $\Psi^{a'}(y)$ for some a' . So y' appears in $\Psi^{a'} \Psi^a(x) = \Psi^{a+a'}(x)$. (The non-negativity of coproduct structure constants ensures that the y' term in $\Psi^{a+a'}(x)$ cannot cancel out due to contributions from an intermediary different from x' .) \square

Remark. It is possible to adapt the above arguments to Hopf algebras with a (noncommutative) free basis $\mathcal{B} = \{S_1 S_2 \dots S_k \mid k \in \mathbb{N}, S_i \in \mathcal{C}\}$ (see Theorem 2.5.1.B). This shows that, for $x \in \mathcal{B}$, all terms in $\Psi^a(x)$ are either a permutation of the factors of x , or have length greater than that of x . In particular, for the associated Markov chain, the probability of going from x to some permutation of its factors (as opposed to a state of greater length, from which there is no return to x) is $a^{l(x) - \deg(x)}$. However, it is easier to deduce such information by working in the *abelianisation* of the underlying Hopf algebra; that is, quotient it by commutators $xy - yx$, which would send the free basis \mathcal{B} to a free-commutative basis. By Theorem 4.7.1, such quotienting corresponds to a projection of the Markov chain.

Here are two more technical results in this spirit, which will be helpful in Section 5.1.3 for deducing a triangularity feature of the eigenfunctions.

Lemma 5.1.6. *Let x, x_i, y_i be elements of a free-commutative basis, with respect to which all coproduct structure constants are non-negative. If $x = x_1 \dots x_k$ and $x_i \rightarrow y_i$ for each i , then $x \rightarrow y_1 \dots y_k$.*

Proof. For readability, take $k = 2$ and write $x = st$, $s \rightarrow s'$, $t \rightarrow t'$. By definition of the relation \rightarrow , it must be that $s' = s_{(1)} \dots s_{(a)}$ for some summand $s_{(1)} \otimes \dots \otimes s_{(a)}$ of $\bar{\Delta}^{[a]}(s)$. Likewise $t' = t_{(1)} \dots t_{(a')}$ for some a' . Suppose $a > a'$. Coassociativity implies that $\Delta^{[a]}(t) = (t \otimes \dots \otimes t \otimes \Delta^{[a-a']})\Delta^{[a']}(t)$, and $t_{(a')} \otimes 1 \otimes \dots \otimes 1$ is certainly a summand of $\Delta^{[a-a']}(t_{(a')})$, so $t_{(1)} \otimes \dots \otimes t_{(a')} \otimes 1 \otimes \dots \otimes 1$ occurs in $\Delta^{[a]}(t)$. So, taking $t_{(a'+1)} = \dots = t_{(a)} = 1$, we can assume $a = a'$. Then $\Delta^{[a]}(x) = \Delta^{[a]}(s)\Delta^{[a]}(t)$ contains the term $s_{(1)}t_{(1)} \otimes \dots \otimes s_{(a)}t_{(a)}$. Hence $\Psi^a(x)$ contains the term $s_{(1)}t_{(1)} \dots s_{(a)}t_{(a)}$, and this product is $s't'$ by commutativity. (Again, this instance of $s't'$ in $\Psi^a(x)$ cannot cancel out with another term in $\Psi^a(x)$ because the coproduct structure constants are non-negative.) \square

Lemma 5.1.7. *Let x, y be elements of a free-commutative basis, with respect to which all coproduct structure constants are non-negative. Suppose y has factorisation into generators $y = c_1 \dots c_l$. If $x \rightarrow y$ then a coproduct structure constant of the form $\eta_x^{c_{\sigma(1)}, \dots, c_{\sigma(l)}}$ (for some $\sigma \in \mathfrak{S}_l$) is nonzero.*

Proof. If $x \rightarrow y$, then, for some a , there is a term $x_{(1)} \otimes \dots \otimes x_{(a)}$ in $\Delta^{[a]}(x)$ with $x_{(1)} \dots x_{(a)} = y$. So each $x_{(i)}$ must have factorisation $x_{(i)} = \prod_{j \in B_i} c_j$ for some set partition $B_1 | \dots | B_a$ of $\{1, 2, \dots, l\}$. In other words, there is some permutation $\sigma \in \mathfrak{S}_l$ and some $l_1, \dots, l_a \in \mathbb{N}$ such that $x_{(1)} = c_{\sigma(1)} \dots c_{\sigma(l_1)}$, $x_{(2)} = c_{\sigma(l_1+1)} \dots c_{\sigma(l_1+l_2)}$, \dots , $x_{(a)} = c_{\sigma(l_1+\dots+l_{a-1}+1)} \dots c_{\sigma(l)}$. Now $\Delta^{[l_1]}(x_{(1)})$ contains the term $c_{\sigma(1)} \otimes \dots \otimes c_{\sigma(l_1)}$, and similarly for $\Delta^{[l_2]}(x_{(2)}), \dots, \Delta^{[l_a]}(x_{(a)})$. So $\Delta^{[l]}(x) = (\Delta^{[l_1]} \otimes \dots \otimes \Delta^{[l_a]})\Delta^{[a]}(x)$ contains the term $c_{\sigma(1)} \otimes \dots \otimes c_{\sigma(l)}$. (This cannot cancel out with another term in $\Delta^{[l]}(x)$ because the coproduct structure constants are non-negative.) Hence $\eta_x^{c_{\sigma(1)}, \dots, c_{\sigma(l)}}$ is nonzero. \square

5.1.3 Probability Estimates from Eigenfunctions

The focus of this section is the right eigenfunctions, since they aid in measuring how far the chain is from being absorbed. But first, one observation about left eigenfunctions deserves a mention.

Recall that the eigenbasis $\{e(c_1) \dots e(c_k) | k \in \mathbb{N}, \{c_1, \dots, c_k\} \text{ a multiset in } \mathcal{C}\}$ from Theorem 2.5.1.A is “length-triangular” in the sense that $e(c_1) \dots e(c_k) = c_1 \dots c_k +$ terms of higher length; indeed, this allowed the conclusion that such vectors form a basis. Since the partial-order by \rightarrow refines the partial-ordering by length, it’s natural to wonder if this

basis is moreover “triangular” with respect to the \rightarrow partial-order. Proposition 5.1.8 below shows this is true: if

$$\mathbf{g}_{c_1 \dots c_k}(y) = \text{coefficient of } y \text{ in } \eta(y)e(c_1) \dots e(c_k),$$

then $\mathbf{g}_{c_1 \dots c_k}(c_1 \dots c_k) = \eta(c_1 \dots c_k)$, and $\mathbf{g}_{c_1 \dots c_k}(y) = 0$ if y is not accessible from $c_1 \dots c_k$.

Proposition 5.1.8. *Let \mathcal{B} be a free-commutative basis of a graded connected Hopf algebra over \mathbb{R} . If $x \in \mathcal{B}$ has factorisation into generators $x = c_1 \dots c_k$, then*

$$e(c_1) \dots e(c_k) = x + \sum_{\substack{x \rightarrow y \\ y \neq x}} \alpha_{xy} y$$

for some constants α_{xy} .

Proof. The proof of Theorem 2.5.1.A already shows that the coefficient of x in $e(c_1) \dots e(c_k)$ is 1, so it suffices to show that all y that appear in $e(c_1) \dots e(c_k)$ must satisfy $x \rightarrow y$.

First consider the case where $k = 1$. By definition of the Eulerian idempotent, each term y of $e(c_1)$ appears in $\frac{(-1)^{a-1}}{a} m^{[a]} \bar{\Delta}^{[a]}(c_1)$ for some a , and hence in $\Psi^a(c_1)$, so $c_1 \rightarrow y$ as required. Now for $k > 1$,

$$e(c_1) \dots e(c_k) = \left(\sum_{c_1 \rightarrow c'_1} \alpha_{c_1 c'_1} c'_1 \right) \dots \left(\sum_{c_k \rightarrow c'_k} \alpha_{c_k c'_k} c'_k \right),$$

and Lemma 5.1.6 precisely concludes that $x = c_1 \dots c_k \rightarrow c'_1 \dots c'_k$. \square

And now onto right eigenfunctions. By Proposition 3.2.1.R, these come from eigenvectors of the Hopf-power on the dual algebra \mathcal{H}^* . As \mathcal{H} is commutative, its dual \mathcal{H}^* is cocommutative, so Theorem 2.5.1.B generates an eigenbasis of Ψ^a on \mathcal{H}^* from a basis of primitives of \mathcal{H}^* , namely by taking symmetrised products. When the state space basis is free-commutative, a convenient choice of such a basis of primitives is the duals of the free generating set. Then the eigenfunctions are simply sums of coproduct structure constants, and this has the advantage that their calculation do not explicitly involve \mathcal{H}^* . Their values

have a combinatorial interpretation as the numbers of ways to break x into the constituent “components” of y , divided by the number of ways to break x into singletons.

Theorem 5.1.9. *Let \mathcal{H} be a Hopf algebra over \mathbb{R} with free-commutative state space basis \mathcal{B} . For each $y \in \mathcal{B}_n$, define $\mathbf{f}_y : \mathcal{B}_n \rightarrow \mathcal{B}_n$ by:*

$$\mathbf{f}_y(x) := \frac{1}{l!Z(y)\eta(x)} \sum_{\sigma \in \mathfrak{S}_l} \eta_x^{c_{\sigma(1)}, \dots, c_{\sigma(l)}} = \frac{1}{l!\eta(x)} \sum_{\sigma \in \mathfrak{S}_y} \eta_x^{c_{\sigma(1)}, \dots, c_{\sigma(l)}}.$$

Here, $y = c_1 \dots c_l$ is the factorisation of y into generators; $\eta_x^{c_{\sigma(1)}, \dots, c_{\sigma(l)}}$ is the coproduct structure constant, equal to the coefficient of $c_{\sigma(1)} \otimes \dots \otimes c_{\sigma(l)}$ in $\Delta^{[l]}(x)$; $\eta(x)$ is the rescaling function in Definition 4.3.1; $Z(y)$ is the size of the stabiliser of the symmetric group \mathfrak{S}_l permuting (c_1, \dots, c_l) ; and \mathfrak{S}_y is a set of coset representatives of this stabiliser in \mathfrak{S}_l . Then \mathbf{f}_y is a right eigenfunction for the l th Hopf-power Markov chain on \mathcal{B}_n , with eigenvalue $a^{l(y)-n}$. This right eigenfunction has a triangular property

$$\begin{aligned} \mathbf{f}_y(x) &= 0 && \text{if } x \not\rightarrow y; \\ \mathbf{f}_y(x) &> 0 && \text{if } x \rightarrow y; \\ \mathbf{f}_y(y) &= \frac{1}{\eta(y)}. \end{aligned}$$

Furthermore, $\{\mathbf{f}_y | y \in \mathcal{B}_n\}$ is a basis of right eigenfunctions dual to the basis of left eigenfunctions coming from Theorem 2.5.1.A. In other words, if $\mathbf{g}_{c'_1 \dots c'_k}(x)$ is the coefficient of x in $\eta(x)e(c'_1) \dots e(c'_k)$, then

$$\sum_{x \in \mathcal{B}_n} \mathbf{g}_{c'_1 \dots c'_k}(x) \mathbf{f}_y(x) = \begin{cases} 1 & \text{if } y = c'_1 \dots c'_k \text{ (i.e. } \{c_1, \dots, c_l\} = \{c'_1, \dots, c'_k\} \text{ as multisets);} \\ 0 & \text{otherwise.} \end{cases}$$

The proof is delayed until the end of this section. See Equation 5.1 below for some special cases of this formula. Note that, if \mathcal{H} is in addition cocommutative, then it is unnecessary to symmetrise - just set $\mathbf{f}_y(x) := \frac{1}{Z(y)\eta(x)} \eta_x^{c_1, \dots, c_l}$.



Figure 5.1: The “two triangles with one common vertex” graph

Example 5.1.10. Recall from the opening of this chapter the Hopf algebra \mathcal{G} of isomorphism classes of graphs, whose associated Markov chain models edge-removal. \mathcal{G} is co-commutative, so the simpler formula $\mathbf{f}_y(x) := \frac{1}{Z(y)\eta(x)} \eta_x^{c_1, \dots, c_l}$ applies. As remarked in the opening of this chapter, the rescaling function $\eta(x) = (\deg x)!$ for all x , so $\mathbf{f}_y(x) = \frac{1}{Z(y)(\deg x)!} \eta_x^{c_1, \dots, c_l}$. This example will calculate $\mathbf{f}_y(x)$ in the case where x is “two triangles with one common vertex” as depicted in Figure 5.1, and y is the disjoint union of a path of length 3 and an edge. So $c_1 = P_3$, the path of length 3, and $c_2 = P_2$, a single edge (or vice versa, the order does not matter). Since these are distinct, $Z(y) = 1$. There are four ways to partition the vertex set of x into a triple and a pair so that the respective induced subgraphs are P_3 and P_2 . (The triples for these four ways are, respectively: the top three vertices; the top left, top middle and bottom right; the top right, top middle and bottom left; and the bottom two vertices and the top middle.) Thus $\mathbf{f}_y(x) = \frac{1}{5!} 4$.

The triangular property of the right eigenfunctions \mathbf{f}_y makes them ideal to use in Proposition 1.1.3.iii to bound the probability that the chain can still reach y . The result is recorded in Proposition 5.1.11 below, along with a few variants. (Bounds analogous to those in Part i hold for any unidirectional Markov chain, since these right eigenfunctions come from applying Perron-Frobenius to the *minors* of the transition matrix - that is, the submatrix with the rows and columns corresponding to states which can reach y . However, Part ii requires $x \not\rightarrow y$ for every pair of distinct states x, y whose eigenfunctions $\mathbf{f}_x, \mathbf{f}_y$ have the same eigenvalue.) Remark 1 after [DPR14, Cor 4.10] shows that, for the rock-breaking chain of the present Section 5.2, the bound for $y = (2, 1, \dots, 1)$ is an asymptotic equality.

Proposition 5.1.11. *Let $\{X_m\}$ be the a th Hopf-power Markov chain on a free-commutative state space basis \mathcal{B}_n . Fix a state $y \in \mathcal{B}_n$ and let \mathbf{f}_y be its corresponding right eigenfunction as defined in Theorem 5.1.9. Then the probability that the chain can still reach y after m steps has the following upper bounds:*

(i)

$$P\{X_m \rightarrow y | X_0 = x_0\} \leq \frac{a^{(l(y)-n)m} \mathbf{f}_y(x_0)}{\min_{x \in \mathcal{B}_n, x \rightarrow y} \mathbf{f}_y(x)} = \frac{a^{(l(y)-n)m} \frac{1}{\eta(x_0)} \sum_{\sigma \in \mathcal{G}_l} \eta_{x_0}^{c_{\sigma(1)}, \dots, c_{\sigma(l)}}}{\min_{x \in \mathcal{B}_n, x \rightarrow y} \frac{1}{\eta(x)} \sum_{\sigma \in \mathcal{G}_l} \eta_x^{c_{\sigma(1)}, \dots, c_{\sigma(l)}}}.$$

(ii)

$$P\{X_m \rightarrow y | X_0 = x_0\} = \eta(y) a^{(l(y)-n)m} \mathbf{f}_y(x_0) (1 + o(1)) \text{ as } m \rightarrow \infty.$$

(iii) For any starting distribution,

$$P\{X_m \rightarrow y\} \leq \frac{a^{(l(y)-n)m}}{\min_{x \in \mathcal{B}_n, x \rightarrow y} \mathbf{f}_y(x) Z(y)} \frac{1}{\eta(y)} \binom{n}{\deg c_1 \dots \deg c_l}.$$

In each case, $y = c_1 \dots c_l$ is the factorisation of y into generators.

Be careful that Parts ii and iii depend upon the scaling of \mathbf{f}_y - they need adjustment if used with a right eigenfunction which is a scalar multiple of \mathbf{f}_y . See Proposition 5.1.14. (This problem does not occur for the first bound as that involves only a ratio of eigenfunction values.)

Proof. Part i is a straightforward application of Proposition 1.1.3.iii, a fact of general Markov chains.

To see Part ii, first note that, from the triangularity properties of \mathbf{f}_y , the difference of functions $\mathbf{1}_{\{\rightarrow y\}} - \eta(y) \mathbf{f}_y$ is non-zero only on $S'_y := \{y' \in \mathcal{B}_n | y' \rightarrow y, y' \neq y\}$. (Here, $\mathbf{1}_{\{\rightarrow y\}}$ is the indicator function of being able to reach y .) Also by triangularity, such functions that are non-zero only on S'_y are spanned by the eigenfunctions $\{\mathbf{f}_{y'} | y' \in S'_y\}$. Hence the expansion of $\mathbf{1}_{\{\rightarrow y\}}$ into right eigenfunctions has the form

$$\mathbf{1}_{\{\rightarrow y\}} = \eta(y) \mathbf{f}_y + \sum_{y' \in S'_y} \alpha_{y'} \mathbf{f}_{y'}$$

for some constants $\alpha_{y'}$. By linearity of expectations, as in Proposition 1.1.3, this implies

$$\begin{aligned} P\{X_m \rightarrow y | X_0 = x_0\} &= \eta(y) a^{(l(y)-n)m} \mathbf{f}_y(x_0) + \sum_{y' \in \mathcal{S}'_y} a^{(l(y')-n)m} \alpha_{y'} \mathbf{f}_{y'}(x_0) \\ &= \eta(y) a^{(l(y)-n)m} \mathbf{f}_y(x_0) \left(1 + \sum_{y' \in \mathcal{S}'_y} a^{(l(y')-l(y))m} \frac{\alpha_{y'} \mathbf{f}_{y'}(x_0)}{\eta(y) \mathbf{f}_y(x_0)} \right). \end{aligned}$$

Now use Proposition 5.1.4: all $y' \in \mathcal{S}'_y$ satisfies $y' \rightarrow y$ and $y' \neq y$, which forces $l(y') \leq l(y)$. So the ratios of eigenvalues $a^{(l(y')-l(y))m}$ is less than 1, and hence the sum tends to zero as $m \rightarrow \infty$.

Now turn to Part iii, the bound independent of the starting state. It suffices to show that

$$Z(y) \mathbf{f}_y(x_0) = \frac{1}{l! \eta(x_0)} \sum_{\sigma \in \mathfrak{S}_l} \eta_{x_0}^{c_{\sigma(1)}, \dots, c_{\sigma(l)}} \leq \frac{1}{\eta(y)} \binom{n}{\deg c_1 \dots \deg c_l}$$

for all states $x_0 \in \mathcal{B}_n$. For any composition $d_1 + \dots + d_l = n$, coassociativity says that $\Delta^{[n]} = (\Delta^{[d_1]} \otimes \dots \otimes \Delta^{[d_l]}) \Delta^{[l]}$, so

$$\eta(x_0) = \sum_{\deg(c'_i)=d_i} \eta(c'_1) \dots \eta(c'_l) \eta_{x_0}^{c'_1, \dots, c'_l}.$$

All summands on the right hand side are non-negative, so choosing $d_i = \deg(c_{\sigma(i)})$ shows that, for each $\sigma \in \mathfrak{S}_l$:

$$\eta(x_0) \geq \eta_{x_0}^{c_{\sigma(1)}, \dots, c_{\sigma(l)}} \eta(c_1) \dots \eta(c_l) = \eta_{x_0}^{c_{\sigma(1)}, \dots, c_{\sigma(l)}} \frac{\eta(y)}{\binom{n}{\deg c_1 \dots \deg c_l}},$$

using Lemma 5.1.2 for the last equality. □

In many common situations, including all examples in this thesis, all coproduct structure constants are integral. Then

$$\mathbf{f}_y(x) = \frac{1}{l! \eta(x)} \sum_{\sigma \in \mathfrak{S}_y} \eta_x^{c_{\sigma(1)}, \dots, c_{\sigma(l)}} \geq \frac{1}{l(y)! \eta(x)},$$

so replacing $(\min_{x \in \mathcal{B}_n, x \rightarrow y} \mathbf{f}_y(x))^{-1}$ with $l(y)! \max_{x \in \mathcal{B}_n, x \rightarrow y} \eta(x)$ in either inequality of Proposition 5.1.11 gives a looser but computationally easier bound.

Example 5.1.12. Continue from Example 5.1.10. Let $y = P_3 P_2$, the disjoint union of a path of length 3 and an edge, and x be “two triangles with one common vertex” as in Figure 5.1. Example 5.1.10 calculated $\mathbf{f}_y(x)$ to be $\frac{4}{5!}$. Then, using the looser bound in the last paragraph, the probability that, after m steps of the Hopf-square Markov chain starting at x , the graph still contains three vertices on which the induced subgraph is a path, and the other two vertices are still connected, is at most $2^{(2-5)m} \frac{4}{5!} 2! 5! = 2^{1-3m} 4$.

The previous example of bounding the probability of having three vertices on which the induced subgraph is a path, and the other two vertices connected feels a little contrived. It is more natural to ask for the probability that at least three vertices are still in the same connected component. This equates to being at a state which can reach either $P_3 \bullet^2$ or $K_3 \bullet^2$, since the only connected graphs on three vertices are P_3 , the path of length 3, and K_3 , the complete graph on 3 vertices. Similarly, being at a state which can reach $P_2 \bullet^3$, the graph with one edge and three isolated vertices, is synonymous with not yet being absorbed. So the most important probabilities of the form “in a state which can still reach y ” are when y has factorisation $y = c \bullet \cdots \bullet$ for some generator $c \neq \bullet$. In this case, it will be convenient to scale this eigenvector by $\frac{\deg y!}{\deg c!}$. So abuse notation and write \mathbf{f}_c for the eigenvector $\frac{\deg y!}{\deg c!} \mathbf{f}_y$ (note that the two notations agree when $y = c$), and extend it to degrees lower than $\deg(c)$ by declaring it to be the zero function there. In other words, for all $x \in \mathcal{B}$:

$$\begin{aligned}
 \mathbf{f}_c(x) &:= \frac{\deg x!}{\deg c!} \mathbf{f}_{c \bullet^{\deg(x) - \deg(c)}}(x) & (5.1) \\
 &= \frac{\binom{\deg x}{\deg c}}{\eta(x)(\deg x - \deg c + 1)} (\eta_x^{c, \bullet, \dots, \bullet} + \cdots + \eta_x^{\bullet, \dots, \bullet, c}) \\
 &= \frac{\binom{\deg x}{\deg c}}{\eta(x)(\deg x - \deg c + 1)} \sum_{w, z} \eta(w) \eta(z) \eta_x^{wcz} & \text{if } \deg(x) \geq \deg(c); \\
 \mathbf{f}_c(x) &:= 0. & \text{if } \deg(x) < \deg(c).
 \end{aligned}$$

Here, the second equality is by definition of \mathbf{f}_y , and the third equality is a consequence of the following coassociativity equation (which holds for any choice of d_i summing to

$\deg(x)$)

$$\sum_{w_i \in \mathcal{B}_{d_i}} \eta_{w_1}^{z_1, \dots, z_i} \eta_{w_2}^{z_{i+1}, \dots, z_j} \eta_{w_3}^{z_{j+1}, \dots, z_a} \eta_x^{w_1, w_2, w_3} = \eta_x^{z_1, \dots, z_a},$$

in the cases where all but one z_i are \bullet . The eigenvalue of \mathbf{f}_c is $a^{-\deg c + 1}$. These are usually the easiest right eigenfunctions to calculate, as they behave well with “recursive structures” such as the trees of Section 5.3. The following Proposition is one general instance of this principle; it reduces the calculation of \mathbf{f}_c to its value on generators.

Proposition 5.1.13. *The right eigenfunction \mathbf{f}_c is additive in the sense that*

$$\mathbf{f}_c(xx') = \mathbf{f}_c(x) + \mathbf{f}_c(x').$$

Proof. This argument is much like that of Lemma 5.1.2 regarding $\eta(xx')$. Since $\Delta(xx') = \Delta(x)\Delta(x')$, a term in $\eta_{xx'}^{c, \bullet, \dots, \bullet} + \dots + \eta_{xx'}^{\bullet, \dots, \bullet, c}$ arises in one of two ways: from a term in $\eta_x^{c, \bullet, \dots, \bullet} + \dots + \eta_x^{\bullet, \dots, \bullet, c}$ and a term in $\eta(x')$, or from a term in $\eta_{x'}^{c, \bullet, \dots, \bullet} + \dots + \eta_{x'}^{\bullet, \dots, \bullet, c}$ and a term in $\eta(x)$. The first way involves a choice of $\deg x'$ tensor-factors amongst $\deg xx' - \deg c + 1$ in which to place the term from $\eta(x')$, and similarly a choice of $\deg x$ positions for the second way. Hence

$$\begin{aligned} & (\eta_x^{c, \bullet, \dots, \bullet} + \dots + \eta_x^{\bullet, \dots, \bullet, c}) \\ &= \binom{\deg xx' - \deg c + 1}{\deg x'} (\eta_x^{c, \bullet, \dots, \bullet} + \dots + \eta_x^{\bullet, \dots, \bullet, c}) \eta(x') \\ & \quad + \binom{\deg xx' - \deg c + 1}{\deg x} (\eta_{x'}^{c, \bullet, \dots, \bullet} + \dots + \eta_{x'}^{\bullet, \dots, \bullet, c}) \eta(x) \\ &= \eta(x) \eta(x') (\deg xx' - \deg c + 1)! \deg x'! \deg x! \deg c! (\mathbf{f}_c(x) + \mathbf{f}_c(x')). \end{aligned}$$

Combining this with the formula for $\eta(xx')$ in Lemma 5.1.2 gives

$$\begin{aligned} \mathbf{f}_c(xx') &= \frac{\binom{\deg xx'}{\deg c}}{\eta(xx') (\deg xx' - \deg c + 1)} (\eta_x^{c, \bullet, \dots, \bullet} + \dots + \eta_x^{\bullet, \dots, \bullet, c}) \\ &= \left(\binom{\deg xx'}{\deg x} \eta(x) \eta(x') \right)^{-1} \frac{\binom{\deg xx'}{\deg c}}{(\deg xx' - \deg c + 1)} (\eta_x^{c, \bullet, \dots, \bullet} + \dots + \eta_x^{\bullet, \dots, \bullet, c}) \\ &= \mathbf{f}_c(x) + \mathbf{f}_c(x'). \end{aligned}$$

□

Because the \mathbf{f}_c are non-trivial multiples of the \mathbf{f}_y when $y \neq c$, the bound in Proposition 5.1.13.iii, which is independent of the starting state, does not apply verbatim. Here is the modified statement (which uses the fact that $Z(c \bullet^{n-\deg c}) = (n - \deg c)!$, and $\eta(y) = \binom{n}{\deg c} \eta(c)$ as per Lemma 5.1.2).

Proposition 5.1.14. *Let $\{X_m\}$ be the a th Hopf-power Markov chain on a free-commutative state space basis \mathcal{B}_n . Let c be a generator of the underlying Hopf algebra \mathcal{H} , and let \mathbf{f}_c be its corresponding right eigenfunction as defined in Equation 5.1. Then, for any starting distribution, the probability that the chain can still reach $c \bullet^{n-\deg c}$ after m steps has the following upper bound:*

$$P\{X_m \rightarrow c \bullet^{n-\deg c}\} \leq \frac{a^{(l(y)-n)m}}{\min_{x \in \mathcal{B}_n, x \rightarrow y} \mathbf{f}_c(x)(n - \deg c + 1)} \frac{1}{\eta(c)} \binom{n}{\deg c}.$$

□

In the case of (isomorphism classes of) graphs, $\eta_x^{c, \bullet, \dots, \bullet} = \dots = \eta_x^{\bullet, \dots, \bullet, c}$ is the number of induced subgraphs of x isomorphic to c , multiplied by the number of orders in which to choose the singletons, which is $(\deg x - \deg c)!$. Recall that $\eta(x) = (\deg x)!$. So

$$\mathbf{f}_c(x) = \frac{1}{\deg c!} |\{\text{induced subgraphs of } x \text{ isomorphic to } c\}|.$$

The analogous statement holds for other species-with-restrictions. Note that summing these over all connected graphs c on j vertices gives another right eigenfunction, with eigenvalue a^{-j+1} :

$$\mathbf{f}_j(x) := \frac{1}{j!} |\{\text{connected induced subgraphs of } G \text{ with } j \text{ vertices}\}|.$$

Minor variations on Propositions 1.1.3 and 5.1.11.i with the \mathbf{f}_c s and \mathbf{f}_j s then imply the following facts. They have an alternative, elementary derivation: the chance that any one particular connected subgraph c survives one step of the edge-removal chain is $a^{-\deg c+1}$, since all vertices of c must receive the same colour. Since expectation is linear, summing these over all subgraphs of interest gives the expected number of these subgraphs that survive.

Proposition 5.1.15. *Let $\{X_m\}$ be the a th Hopf-power Markov chain on graphs describing edge-removal. Let c be any connected graph. Then*

$$\begin{aligned} & E\{|\{\text{induced subgraphs of } X_m \text{ isomorphic to } c\}||X_0 = G\} \\ & = a^{(-\deg c + 1)m} |\{\text{induced subgraphs of } G \text{ isomorphic to } c\}|; \end{aligned}$$

$$\begin{aligned} & P\{X_m \text{ has a connected component with } \geq j \text{ vertices} | X_0 = G\} \\ & \leq E\{|\{\text{connected components of } X_m \text{ with } \geq j \text{ vertices}\}||X_0 = G\} \\ & \leq E\{|\{\text{connected induced subgraphs of } X_m \text{ with } j \text{ vertices}\}||X_0 = G\} \\ & = a^{(-j+1)m} |\{\text{connected induced subgraphs of } G \text{ with } j \text{ vertices}\}|. \end{aligned}$$

In particular, the case $j = 2$ gives

$$\begin{aligned} & P\{X_m \text{ is not absorbed} | X_0 = G\} \\ & \leq E\{|\{\text{edges in } X_m\}||X_0 = G\} = a^{-m} |\{\text{edges in } G\}|. \end{aligned}$$

□

Example 5.1.16. Take x_0 to be the “two triangles with one common vertex” graph of Figure 5.1 above. It has four induced subgraphs that are paths of length 3 (Example 5.1.10 identified these), and the two obvious induced subgraphs that are triangles. So the probability of having a connected component of size at least 3 after m steps of the Hopf-square Markov chain is less than $2^{-2m}6$, which is also the expected number of triples that remain connected.

Proof of Theorem 5.1.9, right eigenfunctions in terms of coproduct structure constants .

By definition of the coproduct structure constant, and of the product structure on the dual

Hopf algebra,

$$\begin{aligned}
\mathbf{f}_y(x) &:= \frac{1}{l!Z(y)\eta(x)} \sum_{\sigma \in \mathfrak{S}_l} \eta_x^{c_{\sigma(1)}, \dots, c_{\sigma(l)}} \\
&= \frac{1}{l!Z(y)\eta(x)} \sum_{\sigma \in \mathfrak{S}_l} c_{\sigma(1)}^* \otimes \dots \otimes c_{\sigma(l)}^* (\Delta^{[l]}x) \\
&= \frac{1}{l!Z(y)\eta(x)} \sum_{\sigma \in \mathfrak{S}_l} c_{\sigma(1)}^* \dots c_{\sigma(l)}^*(x).
\end{aligned}$$

So, thanks to Proposition 3.2.1.R, \mathbf{f}_y being a right eigenfunction of the Hopf-power Markov chain with eigenvalue a^{l-n} equates to $f_y := \frac{1}{l!Z(y)} \sum_{\sigma \in \mathfrak{S}_l} c_{\sigma(1)}^* \dots c_{\sigma(l)}^*$ being an eigenvector of Ψ^a on \mathcal{H}^* with eigenvalue a^l . This will follow from the Symmetrisation Lemma (Theorem 2.3.2) once it is clear that the c_i^* are primitive.

To establish that each c^* is primitive, proceed by contradiction. Take a term $w^* \otimes z^*$ in $\bar{\Delta}(c^*) = \Delta(c^*) - 1 \otimes c^* - c^* \otimes 1$, with $w, z \in \mathcal{B}$. Then $\Delta(c^*)(w \otimes z)$ is non-zero. Since comultiplication in \mathcal{H}^* is dual to multiplication in \mathcal{H} , $\Delta(c^*)(w \otimes z) = c^*(wz)$. Now \mathcal{B} is free-commutative so $wz \in \mathcal{B}$, thus $c^*(wz)$ is only non-zero if $wz = c$. But, by the counit axiom for graded connected Hopf algebras, $\bar{\Delta}(c^*) \in \bigoplus_{j=1}^{\deg c-1} \mathcal{H}_j^* \otimes \mathcal{H}_{\deg c-j}^*$, so both w and z have strictly positive degree. So $c = wz$ contradicts the assumption that c is a generator, and hence no term $w^* \otimes z^*$ can exist in $\bar{\Delta}(c^*)$, i.e. $\bar{\Delta}(c^*) = 0$.

To see the triangularity properties, note that $\mathbf{f}_y(x)$ is non-zero only if $\eta_x^{c_{\sigma(1)}, \dots, c_{\sigma(l)}}$ is non-zero for some $\sigma \in \mathfrak{S}_l$, which forces $x \rightarrow c_{\sigma(1)} \dots c_{\sigma(l)} = y$. Conversely, if $x \rightarrow y$, then by Lemma 5.1.7 $\eta_x^{c_{\sigma(1)}, \dots, c_{\sigma(l)}}$ is non-zero for some σ , and all other coproduct structure constants are non-negative, so $\mathbf{f}_y(x) > 0$. To show that $\mathbf{f}_y(y) = \frac{1}{\eta(y)}$, it suffices to show that $\sum_{\sigma \in \mathfrak{S}_l} \eta_y^{c_{\sigma(1)}, \dots, c_{\sigma(l)}} = Z(y)$ for each $\sigma \in \mathfrak{S}_l$. Rewrite the left hand side using the dual Hopf

algebra:

$$\begin{aligned}
\sum_{\sigma \in \mathfrak{S}_l} \eta_y^{c_{\sigma(1)}, \dots, c_{\sigma(l)}} &= \left(c_{\sigma(1)}^* \otimes \dots \otimes c_{\sigma(l)}^* \right) \Delta^{[l]}(y) \\
&= \left(c_{\sigma(1)}^* \otimes \dots \otimes c_{\sigma(l)}^* \right) \Delta^{[l]}(c_1 \dots c_l) \\
&= \left(c_{\sigma(1)}^* \dots c_{\sigma(l)}^* \right) (c_1 \dots c_l) \\
&= \left(\Delta^{[l]} \left(c_{\sigma(1)}^* \dots c_{\sigma(l)}^* \right) \right) (c_1 \otimes \dots \otimes c_l) \\
&= \left(\Delta^{[l]}(c_{\sigma(1)}^*) \dots \Delta^{[l]}(c_{\sigma(l)}^*) \right) (c_1 \otimes \dots \otimes c_l).
\end{aligned}$$

As each $c_{\sigma(i)}^*$ is primitive,

$$\left(\Delta^{[l]}(c_{\sigma(1)}^*) \dots \Delta^{[l]}(c_{\sigma(l)}^*) \right) = \sum_{A_1 \amalg \dots \amalg A_l = \{1, 2, \dots, l\}} \sum_{\sigma \in S_l} \left(\prod_{i \in A_1} c_{\sigma(i)}^* \right) \otimes \dots \otimes \left(\prod_{i \in A_l} c_{\sigma(i)}^* \right).$$

Hence its evaluation on $c_1 \otimes \dots \otimes c_l$ is

$$\begin{aligned}
&\left| \left\{ (A_1, \dots, A_l) \mid A_1 \amalg \dots \amalg A_l = \{1, 2, \dots, l\}, \prod_{i \in A_1} c_{\sigma(i)}^* = c_1^*, \dots, \prod_{i \in A_l} c_{\sigma(i)}^* = c_l^* \right\} \right| \\
&= |\{\tau \in \mathfrak{S}_l \mid c_{\tau\sigma(i)} = c_i\}| \\
&= |\{\tau \in \mathfrak{S}_l \mid c_{\tau(i)} = c_i\}| = Z(y).
\end{aligned}$$

The last claim of Theorem 5.1.9 is that $\sum_{x \in \mathcal{B}_n} \mathbf{g}_{c'_1 \dots c'_k}(x) \mathbf{f}_y(x) = 1$ when $y = c'_1 \dots c'_k$ and is 0 otherwise; it follows from this duality statement that \mathbf{f}_y is a basis. First take the case where $l(y) \neq k$; then \mathbf{f}_y and $\mathbf{g}_{c'_1 \dots c'_k}$ are eigenvectors of dual maps with different eigenvalues, so the required sum must be zero, by the following simple linear algebra argument (recall that \check{K} is the transition matrix):

$$a^k \sum_{x \in \mathcal{B}_n} \mathbf{g}_{c'_1 \dots c'_k}(x) \mathbf{f}_y(x) = \sum_{x, z \in \mathcal{B}_n} \mathbf{g}_{c'_1 \dots c'_k}(z) \check{K}(z, x) \mathbf{f}_y(x) = a^l \sum_{x \in \mathcal{B}_n} \mathbf{g}_{c'_1 \dots c'_k}(x) \mathbf{f}_y(x).$$

So take $l(y) = k$. Recall from Proposition 5.1.8 that $\mathbf{g}_{c'_1 \dots c'_k}(x)$ is non-zero only if $c'_1 \dots c'_k \rightarrow x$, and earlier in this proof showed that $\mathbf{f}_y(x)$ is non-zero only if $x \rightarrow y$. So

the only terms x which contribute to $\sum_{x \in \mathcal{B}_n} \mathbf{g}_{c'_1 \dots c'_k}(x) \mathbf{f}_y(x)$ must satisfy $c'_1 \dots c'_k \rightarrow x \rightarrow y$. By Proposition 5.1.4, this implies $k = l(c'_1 \dots c'_k) \geq l(x) \geq l(y)$ with equality if and only if $c'_1 \dots c'_k = x = y$. As the current assumption is that $k = l(y)$, no x 's contribute to the sum unless $c'_1 \dots c'_k = y$. In this case, the sum is $\mathbf{g}_y(y) \mathbf{f}_y(y) = \eta(y) \frac{1}{\eta(y)} = 1$. \square

5.1.4 Probability Estimates from Quasisymmetric Functions

The previous section provided upper bounds for the probabilities that a Hopf-power Markov chain is “far from absorbed”. This section connects the complementary probabilities, of being “close to absorbed”, to the following result of Aguiar, Bergeron and Sottile, that the algebra of quasisymmetric functions (Example 4.1.6) is terminal in the category of combinatorial Hopf algebras with a character. (For this section, elements of $QSym$ will be in the variables t_1, t_2, \dots to distinguish from the states x of the Markov chain.)

Theorem 5.1.17. [ABS06, Th. 4.1] *Let \mathcal{H} be a graded, connected Hopf algebra over \mathbb{R} , and let $\zeta : \mathcal{H} \rightarrow \mathbb{R}$ be a multiplicative linear functional (i.e. $\zeta(wz) = \zeta(w)\zeta(z)$). Then there is a unique Hopf-morphism $\chi^\zeta : \mathcal{H} \rightarrow QSym$ such that, for each $x \in \mathcal{H}$, the quasisymmetric function $\chi^\zeta(x)$ evaluates to $\zeta(x)$ when $t_1 = 1$ and $t_2 = t_3 = \dots = 0$. To explicitly construct χ^ζ , set the coefficient of the monomial quasisymmetric function M_I in $\chi^\zeta(x)$ to be the image of x under the composite*

$$\mathcal{H} \xrightarrow{\Delta^{[l(I)]}} \mathcal{H}^{\otimes l(I)} \xrightarrow{\pi_{i_1} \otimes \dots \otimes \pi_{i_{l(I)}}} \mathcal{H}_{i_1} \otimes \dots \otimes \mathcal{H}_{i_{l(I)}} \xrightarrow{\zeta^{\otimes l(I)}} \mathbb{R}.$$

where, in the middle map, π_{i_j} denotes the projection to the subspace of degree i_j .

One motivating example from the authors [ABS06, Ex. 4.5] is $\mathcal{H} = \bar{\mathcal{G}}$, the algebra of isomorphism classes of graphs. For a graph G , set $\zeta(G)$ to be 1 if G has no edges, and 0 otherwise. Then $\chi^\zeta(G)$ is Stanley’s *chromatic symmetric function* [Sta95]: the coefficient of $x_1^{r_1} \dots x_n^{r_n}$ in $\chi^\zeta(G)$ counts the proper colourings of G where r_i vertices receive colour i . (A *proper colouring* of G is an assignment of colours to the vertices of G so that no two vertices on an edge have the same colour.) Note that $\chi^\zeta(G)$ evaluated at $t_1 = \dots = t_a = 1$, $t_{a+1} = t_{a+2} = \dots = 0$ is then precisely the number of proper colourings of G in a colours (not necessarily using all of them). Equivalently, $\chi^\zeta(G)$ evaluated at $t_1 = \dots = t_a =$

$\frac{1}{a}$, $t_{a+1} = t_{a+2} = \dots = 0$ is the probability that uniformly and independently choosing one of a colours for each vertex of G produces a proper colouring. According to the description of the Hopf-power Markov chain on graphs (Example 4.1.3), this is precisely the probability of absorption after a single step. The same is true of other Hopf-power Markov chains on free-commutative bases. Note that it is enough to consider absorption in one step because, by the power rule, m steps of the a th Hopf-power Markov chain on a commutative Hopf algebra is equivalent to one step of the a^m th Hopf-power Markov chain.

In the results below, $[f]_{1/a}$ denotes evaluating the quasisymmetric function f at $t_1 = \dots = t_a = \frac{1}{a}$, $t_{a+1} = t_{a+2} = \dots = 0$.

Proposition 5.1.18 (Probability of absorption). *Let \mathcal{B} be a free-commutative state space basis of \mathcal{H} , and $\zeta : \mathcal{H} \rightarrow \mathbb{R}$ be the indicator function of absorption, extended linearly. (In other words, $\zeta(x) = 1$ if x is an absorbing state, and 0 for other states x .) Then the probability that the a th Hopf-power Markov chain on \mathcal{B}_n is absorbed in a single step starting from x_0 is*

$$\sum_{y:l(y)=n} \check{K}_{a,n}(x_0, y) = \left[\frac{n!}{\eta(x_0)} \chi^\zeta(x_0) \right]_{1/a}.$$

It is natural to ask whether χ^ζ will analogously give the probability of landing in some subset \mathcal{Y} of states if ζ is the indicator function on \mathcal{Y} . The obstacle is that such a ζ might not be multiplicative. The first theorem below gives one class of \mathcal{Y} s for which ζ is clearly multiplicative, and the second indicates the best one can hope for in a completely general setting, when \mathcal{B} might not even be free-commutative.

Theorem 5.1.19. *Let \mathcal{B} be a free-commutative state space basis of \mathcal{H} , and \mathcal{C}' a subset of the free generators. Let $\zeta : \mathcal{H} \rightarrow \mathbb{R}$ be the multiplicative linear functional with $\zeta(c) = \frac{\eta(c)}{(\deg c)!}$ if $c \in \mathcal{C}'$, and $\zeta(c) = 0$ for other free generators c . Then, for the a th Hopf-power Markov chain $\{X_m\}$ on \mathcal{B}_n ,*

$$P\{\text{all factors of } X_1 \text{ are in } \mathcal{C}' \mid X_0 = x_0\} = \left[\frac{(\deg x_0)!}{\eta(x_0)} \chi^\zeta(x_0) \right]_{1/a}.$$

Theorem 5.1.20. *Let \mathcal{B} be any state space basis of \mathcal{H} , and $\{X_m\}$ the a th Hopf-power Markov chain on \mathcal{B}_n . Let $\mathcal{Y} \subseteq \mathcal{B}_n$, and $\zeta : \mathcal{H} \rightarrow \mathbb{R}$ be a multiplicative linear functional*

satisfying $\zeta(y) > 0$ for $y \in \mathcal{B}_n \cap \mathcal{Y}$, $\zeta(y) = 0$ for $y \in \mathcal{B}_n \setminus \mathcal{Y}$. Then

$$\left(\min_{y \in \mathcal{Y}} \frac{\eta(y)}{\zeta(y)} \right) \left[\frac{1}{\eta(x_0)} \chi^\zeta(x_0) \right]_{1/a} \leq P\{X_1 \in \mathcal{Y} | X_0 = x_0\} \leq \left(\max_{y \in \mathcal{Y}} \frac{\eta(y)}{\zeta(y)} \right) \left[\frac{1}{\eta(x_0)} \chi^\zeta(x_0) \right]_{1/a}.$$

Example 5.1.21. Work in $\bar{\mathcal{G}}$, the algebra of isomorphism classes of graphs, where $\eta(x) = \deg(x)!$. Let $\mathcal{C}' = \mathcal{B}_1 \amalg \cdots \amalg \mathcal{B}_{j-1}$. Then the function ζ of Theorem 5.1.19 takes value 1 on graphs each of whose connected components have fewer than j vertices, and value 0 on graphs with a connected component of at least j vertices. Then $\left[\chi^\zeta(G) \right]_{1/a}$ yields the probability that, after one step of the edge-removal chain started at G , all connected components have size at most $j - 1$.

Proofs of Proposition 5.1.18, Theorems 5.1.19 and 5.1.20. First rewrite the definition of χ^ζ in terms of coproduct structure constants:

$$\begin{aligned} \chi^\zeta(x_0) &= \sum_{l=1}^n \sum_{\deg(z_i) > 0} \eta_{x_0}^{z_1, \dots, z_l} \zeta(z_1) \cdots \zeta(z_l) M_{(\deg z_1, \dots, \deg z_l)} \\ &= \sum_{l=1}^{\infty} \sum_{\substack{z_1, \dots, z_l \\ \deg(z_i) > 0}} \eta_{x_0}^{z_1, \dots, z_l} \zeta(z_1 \cdots z_l) t_1^{\deg z_1} \cdots t_l^{\deg z_l}. \end{aligned}$$

So, when $t_1 = \cdots = t_a = \frac{1}{a}$, $t_{a+1} = t_{a+2} = \cdots = 0$, the quasisymmetric function $\chi^\zeta(x_0)$ evaluates to

$$\sum_{z_1, \dots, z_a} \eta_{x_0}^{z_1, \dots, z_a} \zeta(z_1 \cdots z_a) a^{-n} = \eta(x_0) \sum_{y \in \mathcal{B}_n} \check{K}_{a,n}(x_0, y) \frac{\zeta(y)}{\eta(y)}.$$

Now, in the setup of Theorem 5.1.20,

$$\begin{aligned} P\{X_1 \in \mathcal{Y} | X_0 = x_0\} &= \sum_{y \in \mathcal{Y}} \check{K}_{a,n}(x_0, y) \\ &\leq \left(\max_{y \in \mathcal{Y}} \frac{\eta(y)}{\zeta(y)} \right) \frac{1}{\eta(x_0)} \left(\eta(x_0) \sum_{y \in \mathcal{Y}} \check{K}_{a,n}(x_0, y) \frac{\zeta(y)}{\eta(y)} \right) \\ &= \left(\max_{y \in \mathcal{Y}} \frac{\eta(y)}{\zeta(y)} \right) \left[\frac{1}{\eta(x_0)} \chi^\zeta(x_0) \right]_{1/a}, \end{aligned}$$

and an analogous argument gives the lower bound.

In the specialisation of Theorem 5.1.19, the character ζ has value $\frac{\eta(c)}{(\deg c)!}$ for $c \in \mathcal{C}'$ and is zero on other generators. By Lemma 5.1.2 on η of products, $\frac{\zeta(y)}{\eta(y)} = \frac{1}{(\deg y)!}$ if all factors of y are in \mathcal{C}' , and is 0 otherwise. Hence $\left[\chi^\zeta(x_0) \right]_{1/a}$ is precisely

$$\frac{\eta(x_0)}{(\deg y)!} \sum_y \check{K}_{a,n}(x_0, y),$$

summing over all y whose factors are in \mathcal{C}' . Proposition 5.1.18 is then immediate on taking $\mathcal{C}' = \mathcal{B}_1$. \square

5.2 Rock-Breaking

This section investigates a model of rock-breaking, one of two initial examples of a Hopf-power Markov chain in [DPR14, Sec. 4], which gives references to Kolmogorov's study of similar breaking models. The states of this Markov chain are *partitions* $\lambda = (\lambda_1, \dots, \lambda_l)$, a multiset of positive integers recording the sizes of a collection of rocks. (It is best here to think of the *parts* λ_i as unordered, although the standard notation is to write $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{l(\lambda)}$.) In what follows, $|\lambda| := \lambda_1 + \dots + \lambda_{l(\lambda)}$ is the total size of the rocks in the collection λ , and the number of rocks in the collection is $l(\lambda)$, the *length* of the partition. $Z(\lambda)$ is the size of the stabiliser of $\mathfrak{S}_{l(\lambda)}$ permuting the parts of λ . If $a_i(\lambda)$ is the number of parts of size i in λ , then $Z(\lambda) = \prod_i a_i(\lambda)!$. For example, if $\mu = (2, 1, 1, 1)$, then $|\mu| = 5$, $l(\mu) = 4$ and $Z(\mu) = 6$.

At each step of the Markov chain, each rock breaks independently into a pieces whose sizes follow a symmetric multinomial distribution. (This may result in some pieces of zero size.) Section 5.2.1 phrases this process as the Hopf-power Markov chain on the homogeneous symmetric functions $\{h_\lambda\}$. Sections 5.2.2 and 5.2.3 then leverage the machinery of Section 5.1.3 and Chapter 2 to deduce a full right and left eigenbasis respectively. These eigenbases correspond (up to scaling) to the power sum symmetric functions $\{p_\lambda\}$, so the

explicit expressions for the eigenfunctions recover well-known formulae for the change-of-basis between $\{h_\lambda\}$ and $\{p_\lambda\}$. Section 5.2.4 gives a numerical example of the transition matrix and full eigenbases, for the case $n = 4$.

5.2.1 Constructing the Chain

The goal of this section is to interpret the Hopf-power Markov chain on the homogeneous symmetric functions $\{h_\lambda\}$ as independent multinomial breaking. ([DPR14] took instead the elementary symmetric functions $\{e_\lambda\}$ as their state space basis, which is equivalent as there is a Hopf-involution on Λ exchanging $\{h_\lambda\}$ and $\{e_\lambda\}$ [Sta99, Sec. 7.6]. This thesis chooses to use $\{h_\lambda\}$ because its dual basis is $\{m_\lambda\}$, the monomial symmetric functions, while the dual of $\{e_\lambda\}$ is less studied.)

Recall from Example 4.1.5 that, as an algebra, Λ is the subalgebra of the power series algebra $\mathbb{R}[[x_1, x_2, \dots]]$ generated by

$$h_{(n)} := \sum_{i_1 \leq \dots \leq i_n} x_{i_1} \dots x_{i_n},$$

which has degree n . There is a large swathe of literature on Λ - the standard references are [Sta97, Chap. 7; Mac95, Chap. 1]. Only two facts are essential for building the present chain: first,

$$h_\lambda := h_{(\lambda_1)} \dots h_{(\lambda_{l(\lambda)})}$$

is a basis for Λ ; second, the coproduct satisfies $\Delta(h_{(n)}) = \sum_{i=0}^n h_{(i)} \otimes h_{(n-i)}$, with the convention $h_{(0)} = 1$. It follows from the compatibility axiom of Hopf algebras that

$$\Delta(h_\lambda) = \sum_{i_1, \dots, i_l=0}^{i_j=\lambda_j} h_{(i_1, \dots, i_l)} \otimes h_{(\lambda_1-i_1, \dots, \lambda_l-i_l)}.$$

(Here, it is not necessarily true that $i_1 \geq i_2 \geq \dots \geq i_l$. This is one instance where it is useful to think of the parts as unordered.) Then it is obvious that $\{h_\lambda\}$ is a state space basis - the product and coproduct structure constants of Λ with respect to $\{h_\lambda\}$ are non-negative, and $\bar{\Delta}(h_\lambda) \neq 0$ if $\deg(\lambda) > 1$. In the sequel, it will be convenient to write λ for h_λ . For

example, the above equation translates in this notation to

$$\Delta(\lambda) = \sum_{i_1, \dots, i_l=0}^{i_j=\lambda_j} (i_1, \dots, i_l) \otimes (\lambda_1 - i_1, \dots, \lambda_l - i_l).$$

Recall that Theorem 4.4.1 gives a three-step interpretation of a Hopf-power Markov chain. To apply this to the chain on $\{h_\lambda\}$, it is first necessary to compute the rescaling function η . A simple induction shows that

$$\Delta^{[r]}((n)) = \sum_{i_1 + \dots + i_r = n} (i_1) \otimes \dots \otimes (i_r), \quad (5.2)$$

so

$$\bar{\Delta}^{[n]}((n)) = (1) \otimes \dots \otimes (1),$$

and $\eta((n)) = 1$. Lemma 5.1.2 then shows that

$$\eta(\lambda) = \binom{|\lambda|}{\lambda_1 \dots \lambda_l} \eta((\lambda_1)) \dots \eta((\lambda_l)) = \binom{|\lambda|}{\lambda_1 \dots \lambda_l}.$$

Note that $\{h_\lambda\}$ is a free-commutative basis, so, by Theorem 5.1.1, each rock in the collection breaks independently. Thus it suffices to understand the chain starting at (n) . By Equation 5.2, the coproduct structure constant $\eta_{(n)}^{\mu^1, \dots, \mu^a} = 1$ if μ^1, \dots, μ^a are all partitions of single parts with $|\mu^1| + \dots + |\mu^a| = n$, and is 0 for all other a -tuples of partitions. As a result, the three-step description of Theorem 4.4.1 simplifies to:

1. Choose i_1, \dots, i_a according to a symmetric multinomial distribution.
2. Choose the a -tuple of one part partitions $(i_1), \dots, (i_a)$, some of which may be the zero partition.
3. Move to (i_1, \dots, i_a) .

Thus each rock breaks multinomially. Section 5.2.4 below displays the transition matrix for the case $a = 2$ and $n = 4$, describing binomial breaking of rocks of total size four.

5.2.2 Right Eigenfunctions

Begin with the simpler eigenfunctions $\mathbf{f}_{(j)}$ for $j > 1$, defined in Equation 5.1 to be

$$\mathbf{f}_{(j)}(\lambda) := \binom{|\lambda|}{j} \frac{\eta_{\lambda}^{(j),(1),\dots,(1)}}{\eta(\lambda)}.$$

By Proposition 5.1.13, these eigenfunctions satisfy $\mathbf{f}_{(j)}(\lambda) = \mathbf{f}_{(j)}(\lambda_1) + \dots + \mathbf{f}_{(j)}(\lambda_l)$; since $\eta_{(n)}^{(j),(1),\dots,(1)} = 1$ and $\eta((n)) = 1$, it follows that

$$\mathbf{f}_{(j)}(\lambda) = \sum_{i=1}^l \binom{\lambda_i}{j}.$$

The corresponding eigenvalue is a^{-j+1} .

Recall from Section 5.1.3 that the main use of these eigenfunctions is to measure how far the chain is from being absorbed. For the rock-breaking chain, this measure takes the form of “expected number of large rocks”. Note that each part of λ of size j or greater contributes at least 1 to $\mathbf{f}_{(j)}(\lambda)$; a simple application of Proposition 1.1.3 then gives the Proposition below. The analogous result for the more general Markov chain of removing edges from graphs is Proposition 5.1.15, from which this also follows easily.

Proposition 5.2.1. *Let $\{X_m\}$ denote the rock-breaking chain. Then, for any $j > 1$,*

$$\begin{aligned} & P\{X_m \text{ contains a rock of size } \geq j | X_0 = \lambda\} \\ & \leq E\{\{ \text{rocks of size } \geq j \text{ in } X_m \} | X_0 = \lambda\} \\ & \leq a^{(-j+1)m} \sum_{i=1}^l \binom{\lambda_i}{j}. \end{aligned}$$

In particular, the case $j = 2$ shows

$$P\{X_m \neq (1, 1, \dots, 1)\} \leq a^{-m} \sum_{i=1}^l \binom{\lambda_i}{2}.$$

□

Theorem 5.1.9 gives this formula for the full right eigenbasis:

Theorem 5.2.2. A basis $\{\mathbf{f}_\mu\}$ of right eigenfunctions of the rock-breaking chain is

$$\mathbf{f}_\mu(\lambda) := \frac{1}{\binom{|\lambda|}{\lambda_1 \dots \lambda_{l(\lambda)}}} \sum \frac{1}{Z(\mu^1) \dots Z(\mu^{l(\lambda)})}$$

where the sum is over all $l(\lambda)$ -tuples of partitions $\{\mu^j\}$ such that μ^j is a partition of λ_j and the disjoint union $\amalg_j \mu^j = \mu$, and $Z(\mu^j)$ is the size of the stabiliser of $\mathfrak{S}_{l(\mu^j)}$ permuting the parts of μ^j . In particular, $\mathbf{f}_\mu(\mu) = \left(\binom{|\mu|}{\mu_1 \dots \mu_{l(\mu)}} \right)^{-1}$, and $\mathbf{f}_\mu(\lambda)$ is positive if μ is a refinement of λ , and is otherwise 0. The corresponding eigenvalue is $a^{l(\mu)-n}$.

From this right eigenfunction formula, one can recover the expansion of the power sums in terms of monomial symmetric functions [Sta99, Prop. 7.7.1]:

$$p_\mu = Z(\mu) \sum_\lambda \binom{|\lambda|}{\lambda_1 \dots \lambda_{l(\lambda)}} \mathbf{f}_\mu(\lambda) m_\lambda = \sum_{\amalg \mu^i = \mu} \frac{Z(\mu)}{Z(\mu^1) \dots Z(\mu^l)} m_{(|\mu^1|, \dots, |\mu^l|)}.$$

Here is an illustration of how to compute with this formula; the proof will follow.

Example 5.2.3. Take $\mu = (2, 1, 1, 1)$, $\lambda = (3, 2)$. Then the possible $\{\mu^j\}$ are

$$\begin{aligned} \mu^1 &= (2, 1), & \mu^2 &= (1, 1); \\ \mu^1 &= (1, 1, 1), & \mu^2 &= (2). \end{aligned}$$

Hence

$$\mathbf{f}_\mu(\lambda) = \frac{1}{\binom{5}{3}} \left(\frac{1}{2} + \frac{1}{3} \right) = \frac{1}{12}.$$

The full basis of right eigenfunctions for the case $n = 4$ is in Section 5.2.4.

Proof. For concreteness, take $l(\lambda) = 2$ and $l(\mu) = 3$. Then the simplification of Theorem 5.1.9 for cocommutative Hopf algebras gives

$$\mathbf{f}_\mu(\lambda) = \frac{1}{Z(\mu)\eta(\lambda)} \eta_\lambda^{(\mu_1), (\mu_2), (\mu_3)} = \frac{1}{Z(\mu) \binom{|\lambda|}{\lambda_1 \lambda_2}} \eta_\lambda^{(\mu_1), (\mu_2), (\mu_3)},$$

where the parts of μ are ordered so $\mu_1 \geq \mu_2 \geq \mu_3$. To calculate the coproduct structure constant $\eta_\lambda^{(\mu_1),(\mu_2),(\mu_3)}$, recall that

$$\Delta^{[3]}(\lambda) = \Delta^{[3]}(\lambda_1)\Delta^{[3]}(\lambda_2) = \sum_{\substack{i_1+j_1+k_1=\lambda_1 \\ i_2+j_2+k_2=\lambda_2}} (i_1, i_2) \otimes (j_1, j_2) \otimes (k_1, k_2).$$

So $\eta_\lambda^{(\mu_1),(\mu_2),(\mu_3)}$ enumerates the sextuples $(i_1, j_1, k_1, i_2, j_2, k_2)$ such that $i_1 + j_1 + k_1 = \lambda_1$, $i_2 + j_2 + k_2 = \lambda_2$, and i_1 and i_2 are μ_1 and 0 in either order, and similarly for j_1, j_2 and k_1, k_2 . Set $\mu^1 := (i_1, j_1, k_1)$, $\mu^2 = (i_2, j_2, k_2)$; then these sextuples are precisely the case where $|\mu^1| = \lambda_1$, $|\mu^2| = \lambda_2$, and the disjoint union $\mu^1 \amalg \mu^2 = \mu$. If the parts of μ are distinct (i.e. $\mu_1 > \mu_2 > \mu_3$), then one can reconstruct a unique sextuple from such a pair of partitions: if μ^1 has a part of size μ_1 , then $i_1 = \mu_1$ and $i_2 = 0$; else μ^2 has a part of size μ_1 , and $i_2 = \mu_1$, $i_1 = 0$; and similarly for j_1, j_2, k_1, k_2 . If, however, $\mu_1 = \mu_2 > \mu_3$, and μ^1, μ^2 both have one part of the common size $\mu_1 = \mu_2$, then there are two sextuples corresponding to (μ^1, μ^2) : both $i_1 = j_2 = \mu_1$, $i_2 = j_1 = 0$ and $i_2 = j_1 = \mu_1$, $i_1 = j_2 = 0$ are possible. In general, this multiplicity is the product of multinomial coefficients

$$\prod_i \binom{a_i(\mu)}{a_i(\mu^1) \dots a_i(\mu^l(\lambda))},$$

where $a_i(\mu)$ is the number of parts of μ of size i . Since $\prod_i a_i(\mu)! = Z(\mu)$, the expression in the theorem follows.

Now show that $p_\mu = Z(\mu) \sum_{\lambda} \binom{|\lambda|}{\lambda_1 \dots \lambda_{l(\lambda)}} \mathbf{f}_\mu(\lambda) m_\lambda$. Theorem 5.1.9 and Proposition 3.2.1.R constructs $\mathbf{f}_\mu(\lambda)$ as $\frac{1}{Z(\mu)\eta(\lambda)} [(\mu_1)^* \dots (\mu_l)^*](\lambda)$, or the coefficient of λ^* in $\frac{1}{Z(\mu)\eta(\lambda)} (\mu_1)^* \dots (\mu_l)^*$. Viewing the algebra of symmetric functions as its own dual via the Hall inner product, λ^* is the monomial symmetric function m_λ . So $\mathbf{f}_\mu(\lambda)$ is the coefficient of m_λ in

$$\frac{1}{Z(\mu)\eta(\lambda)} m_{(\mu_1)} \dots m_{(\mu_l)} = \frac{1}{Z(\mu) \binom{|\lambda|}{\lambda_1 \dots \lambda_{l(\lambda)}}} p_\mu.$$

□

5.2.3 Left Eigenfunctions

Applying Theorem 2.5.1.A to the rock-breaking chain, taking the single-part partitions as the free generating set, gives the following basis of left eigenfunctions.

Theorem 5.2.4. *A basis $\{\mathbf{g}_\mu\}$ of left eigenfunctions of the rock-breaking chain is*

$$\mathbf{g}_\mu(\lambda) = (-1)^{l(\mu)-l(\lambda)} \binom{|\lambda|}{\lambda_1 \dots \lambda_{l(\mu)}} \sum \frac{(l(\lambda^1) - 1)! \dots (l(\lambda^{l(\mu)}) - 1)!}{Z(\lambda^1) \dots Z(\lambda^{l(\mu)})}$$

where the sum is over all $l(\mu)$ -tuples of partitions $\{\lambda^j\}$ such that λ^j is a partition of μ_j and the disjoint union $\amalg_j \lambda^j = \lambda$, and $Z(\lambda^j)$ is the size of the stabiliser of $\mathfrak{S}_{l(\lambda^j)}$ permuting the parts of λ^j . In particular, $\mathbf{g}_\mu(\mu) = \binom{|\mu|}{\mu_1 \dots \mu_{l(\mu)}}$, and $\mathbf{g}_\mu(\lambda)$ is non-zero only if λ is a refinement of μ . The corresponding eigenvalue is $a^{l(\mu)-n}$.

From this left eigenfunction formula, one can recover the expansion of the power sums in terms of complete symmetric functions:

$$\begin{aligned} p_\mu &= \mu_1 \dots \mu_l \sum_\lambda \frac{1}{\binom{|\lambda|}{\lambda_1 \dots \lambda_{l(\lambda)}}} \mathbf{g}_\mu(\lambda) h_\lambda \\ &= \sum_r (-1)^{l(\mu)-r} \mu_1 \dots \mu_l \sum_{|\lambda^j|=\mu_j} \frac{(l(\lambda^1) - 1)! \dots (l(\lambda^r) - 1)!}{Z(\lambda^1) \dots Z(\lambda^r)} h_{\amalg \lambda^j}. \end{aligned}$$

As previously, here is a calculational example.

Example 5.2.5. Take $\lambda = (2, 1, 1, 1)$, $\mu = (3, 2)$. Then the possible $\{\lambda^j\}$ are

$$\begin{aligned} \lambda^1 &= (2, 1), & \lambda^2 &= (1, 1); \\ \lambda^1 &= (1, 1, 1), & \lambda^2 &= (2). \end{aligned}$$

Hence

$$\mathbf{g}_\mu(\lambda) = (-1)^2 \binom{5}{2} \left(\frac{1!1!}{2!} + \frac{2!0!}{3!} \right) = \frac{25}{3}.$$

The full basis of left eigenfunctions for the case $n = 4$ is in Section 5.2.4.

Proof. By Theorem 2.5.1.A and Proposition 3.2.1.L,

$$\begin{aligned} \mathbf{g}_\mu(\lambda) &= \text{coefficient of } \lambda \text{ in } \eta(\lambda)e((\mu_1)) \dots e((\mu_{l(\mu)})) \\ &= \text{coefficient of } \lambda \text{ in } \binom{|\lambda|}{\lambda_1 \dots \lambda_{l(\lambda)}} e((\mu_1)) \dots e((\mu_{l(\mu)})). \end{aligned}$$

Every occurrence of λ in $e((\mu_1)) \dots e((\mu_{l(\mu)}))$ is a product of a λ^1 term in $e((\mu_1))$, a λ^2 term in $e((\mu_2))$, etc., for some choice of partitions λ^j with $|\lambda^j| = \mu_j$ for each j , and $\prod_j \lambda^j = \lambda$. Hence it suffices to show that the coefficient of a fixed λ^j in $e((\mu_j))$ is

$$\frac{(-1)^{l(\lambda^j)-1} (l(\lambda^j) - 1)!}{Z(\lambda^j)}.$$

Recall that

$$\begin{aligned} e((\mu_j)) &= \sum_{r \geq 1} \frac{(-1)^{r-1}}{r} m^{[r]} \bar{\Delta}^{[r]}((\mu_j)) \\ &= \sum_{r \geq 1} \frac{(-1)^{r-1}}{r} \sum_{\substack{i_1 + \dots + i_r = \mu_j \\ i_1, \dots, i_r > 0}} (i_1, \dots, i_r), \end{aligned}$$

so λ^j only appears in the summand with $r = l(\lambda^j)$. Hence the required coefficient is $\frac{(-1)^{l(\lambda^j)-1}}{l(\lambda^j)}$ multiplied by the number of distinct orderings of the parts of λ^j , which is $\frac{l(\lambda^j)!}{Z(\lambda^j)}$.

To deduce the h_λ -expansion of p_μ , recall from above that $\frac{\mathbf{g}_\mu(\lambda)}{\eta(\lambda)}$ is the coefficient of h_λ in the symmetric function $e(h_{(\mu_1)}) \dots e(h_{(\mu_l)})$. Since the algebra of symmetric functions is cocommutative, the Eulerian idempotent map e is a projection onto the subspace of primitives. So $e(h_{(n)})$ is a primitive symmetric function of degree n . But, up to scaling, the power sum $p_{(n)}$ is the only such symmetric function, so $e(h_{(n)})$ is necessarily $\alpha_n p_{(n)}$ for some number α_n . Thus $\frac{\mathbf{g}_\mu(\lambda)}{\eta(\lambda)}$ is the coefficient of h_λ in $\alpha_{\mu_1} \dots \alpha_{\mu_l} p_\mu$, and it suffices to show that $\alpha_n = \frac{1}{n}$.

As usual, let f_μ, g_μ be the symmetric functions inducing the eigenfunctions $\mathbf{f}_\mu, \mathbf{g}_\mu$ respectively. Then $\langle f_\mu, g_\mu \rangle = \sum_\lambda \mathbf{f}_\mu(\lambda) \mathbf{g}_\mu(\lambda)$, where the left hand side is the Hall inner

product. By Theorem 5.1.9, the right hand side is 1 for all μ . Take $\mu = (n)$, then

$$n\alpha_n = \alpha_n \langle p_{(n)}, p_{(n)} \rangle = \langle f_{(n)}, g_{(n)} \rangle = 1,$$

so $\alpha_n = \frac{1}{n}$ as desired. □

Remark. This calculation is greatly simplified for the algebra of symmetric functions, compared to other Hopf algebras. The reason is that, for a generator c , it is in general false that all terms of $m^{[a]}\bar{\Delta}^{[a]}(c)$ have length a , or equivalently that all tensor-factors of a term of $\bar{\Delta}^{[a]}(c)$ are generators. See the fourth summand of the coproduct calculation in Figure 4.1 for one instance of this, in the Hopf algebra of graphs. Then terms of length say, three, in $e(c)$ may show up in both $m^{[2]}\bar{\Delta}^{[2]}(c)$ and $m^{[3]}\bar{\Delta}^{[3]}(c)$, so determining the coefficient of this length three term in $e(c)$ is much harder, due to these potential cancellations in $e(c)$. Hence much effort [Fis10; AS06; AS05b] has gone into developing cancellation-free expressions for primitives, as alternatives to $e(c)$.

5.2.4 Transition Matrix and Eigenfunctions when $n = 4$

The Hopf-square Markov chain on partitions of four describes independent binomial breaking of a collection of rocks with total size four. Its transition matrix $K_{2,4}$ is the following matrix:

	(4)	(3, 1)	(2, 2)	(2, 1, 1)	(1, 1, 1, 1)
(4)	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{3}{8}$	0	0
(3, 1)	0	$\frac{1}{4}$	0	$\frac{3}{4}$	0
(2, 2)	0	0	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$
(2, 1, 1)	0	0	0	$\frac{1}{2}$	$\frac{1}{2}$
(1, 1, 1, 1)	0	0	0	0	1

Its basis of right eigenfunctions, as determined by Theorem 5.2.2, are the columns of the following matrix:

	$\mathbf{f}_{(4)}$	$\mathbf{f}_{(3,1)}$	$\mathbf{f}_{(2,2)}$	$\mathbf{f}_{(2,1,1)}$	$\mathbf{f}_{(1,1,1,1)}$
(4)	1	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{24}$
(3,1)	0	$\frac{1}{4}$	0	$\frac{1}{4}$	$\frac{1}{24}$
(2,2)	0	0	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{24}$
(2,1,1)	0	0	0	$\frac{1}{12}$	$\frac{1}{24}$
(1,1,1,1)	0	0	0	0	$\frac{1}{24}$

Its basis of left eigenfunctions, as determined by Theorem 5.2.4, are the rows of the following matrix:

	(4)	(3,1)	(2,2)	(2,1,1)	(1,1,1,1)
$\mathbf{g}_{(4)}$	1	-4	-3	12	-6
$\mathbf{g}_{(3,1)}$	0	4	0	-12	8
$\mathbf{g}_{(2,2)}$	0	0	6	-12	6
$\mathbf{g}_{(2,1,1)}$	0	0	0	12	-12
$\mathbf{g}_{(1,1,1,1)}$	0	0	0	0	24

5.3 Tree-Pruning

This section examines the Hopf-power Markov chain whose underlying Hopf algebra is the Connes-Kreimer algebra of rooted trees. This is one of many Hopf algebras arising from quantum field theory during the surge in the relationship between the two fields in the late 1990s. Its definition as a Hopf algebra first appeared in [Kre98; CK98], though they note that it is essentially the same data as the Butcher group for Runge-Kutta methods of solving ordinary differential equations [But72]. A textbook exposition of the use of trees in Runge-Kutta methods is in [But08, Chap. 3].

In his thesis, Foissy [Foi02a; Foi02b; Foi13] constructs a noncommutative version of the Connes-Kreimer Hopf algebra, which turns out to be isomorphic to **PBT**, the Loday-Ronco Hopf algebra of planar binary trees [LR98]. [AS05a; FNT11] then relate it (and its

dual $YSym$) to other Hopf algebras of trees, and well-known Hopf algebras coming from polynomial realisations.

The main purpose of this example is to illustrate how to interpret the chain and to calculate simple right eigenfunctions and probability bounds using the “recursive structure” of trees. The exposition below should serve as a prototype for studying Hopf-power Markov chains on other Hopf algebras of trees.

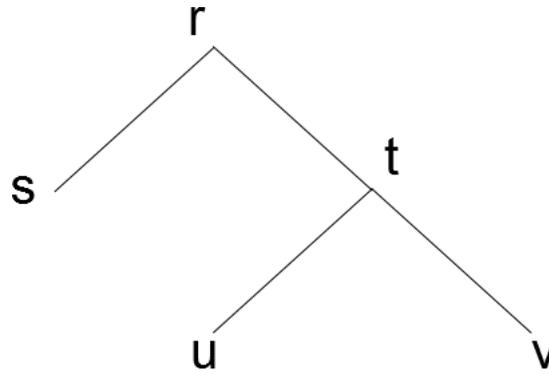
5.3.1 The Connes-Kreimer Hopf algebra

A *tree* is a connected graph (unlabelled) without cycles; a tree T is *rooted* if it has a distinguished vertex $\text{root}(T)$. (The embedding of a tree in the plane - e.g. whether an edge runs to the left or the right - is immaterial). A *rooted forest* is a disjoint union of rooted trees - so each of its components has a root. All trees and forests in this section are rooted unless specified otherwise. Following [CK98], all diagrams below will show $\text{root}(T)$ as the uppermost vertex, and edges will flow downwards from a *parent* to a *child*. A *leaf* is a vertex with no children. More rigorous definitions of these and related terms are in [But08, Sec. 300]; the trees here he calls “abstract trees” as their vertices are not labelled.

Some non-standard notation (see Example 5.3.1 below): $\text{deg}(T)$ is the number of vertices in the tree T . A tree T' is a subtree of T if the subgraph which T induces on the vertex set of T' is connected. Denote this by $T' \subseteq T$. Subtrees containing $\text{root}(T)$ are *trunks*; otherwise, the root of T' is the vertex which was closest to $\text{root}(T)$. If v is a vertex of T' , written $v \in T'$, then $\text{desc}_{T'}(v)$ is the number of *descendants* of v in T' , including v itself, and $\text{anc}_{T'}(v)$ is the number of *ancestors* of v in T' , including v itself.

Two families of graphs are of special interest here: let P_n be the *path* of degree n , where all but one vertex has precisely one child, and Q_n be the *star* of degree n , where the root has $n - 1$ children, and all non-root vertices have no children. (Again this notation is non-standard.) In line with the Hopf algebra notation in previous chapters, \bullet indicates the unique tree with one vertex.

Example 5.3.1. Let T be the tree in Figure 5.2. (The vertex labels are not part of the tree data, they are merely for easy reference.) Then $\text{deg}(T) = 5$. Vertex t has two children, namely u and v ; these are both leaves. The star Q_3 is a subtree of T in two ways: from

Figure 5.2: The tree $[\bullet Q_3]$

the vertices $\{r, s, t\}$, for which r is the root, and from the vertices $\{t, u, v\}$, for which t is the root. Only the first of these is a trunk. If T' is this first copy of Q_3 , then $\text{desc}_{T'}(r) = 3, \text{desc}_{T'}(t) = 1$. The ancestors of u are t and r , so $\text{anc}_T(v) = 2$.

Most results concerning the tree-pruning Markov chain will have an inductive proof, and the key to such arguments is this: given a tree $T \neq \bullet$, let T_1, \dots, T_f be the connected components of T after removing the root. (The ordering of the T_i are immaterial.) Following [But08], write $T := [T_1 \dots T_f]$; his Table 300(I) demonstrates how to write every tree in terms of \bullet (which he calls τ) and repeated applications of this operator. For example, $Q_3 = [\bullet\bullet]$, $P_3 = [\bullet[\bullet]]$, and $P_n = [P_{n-1}]$. The degree 5 tree in Figure 5.2 is $[\bullet Q_3] = [\bullet[\bullet\bullet]]$.

[Kre99, Sec. 5] then defines the *tree factorial* recursively:

$$\bullet! = 1, \quad T! = \deg(T)T_1! \dots T_f!$$

[But08] calls this the “density” $\gamma(T)$ and gives the following equivalent non-recursive expression:

Proposition 5.3.2. [But08, Thm. 301A.c]

$$T! = \prod_{v \in T} \text{desc}_T(v)$$

Proof. When $T = \bullet$, this is immediate. For $T \neq \bullet$, each non-root vertex $v \in T$ is a vertex of precisely one T_i , and $\text{desc}_T(v) = \text{desc}_{T_i}(v)$, so, by inductive hypothesis,

$$\begin{aligned} T! &= \text{deg } T \prod_{v_1 \in T_1} \text{desc}_{T_1}(v_1) \dots \prod_{v_f \in T_f} \text{desc}_{T_f}(v_f) \\ &= \prod_{v \in T} \text{desc}_T(v) \end{aligned}$$

as the root of T has $\text{deg } T$ descendants. □

It is clear from this alternative expression that $P_n! = n!$ (which inspired this notation) and $Q_n! = n$. Note that these are respectively the largest and smallest possible values for $T!$.

Example 5.3.3. Take $T = [\bullet Q_3]$ as pictured in Figure 5.2. Then $T! = 5 \bullet! Q_3! = 5 \cdot 1 \cdot 3 = 15$. Note that this is also $\text{desc}_T(r) \text{desc}_T(s) \text{desc}_T(t) \text{desc}_T(u) \text{desc}_T(v) = 5 \cdot 1 \cdot 3 \cdot 1 \cdot 1$.

Finally we are ready to define the Hopf structure on these trees. The basis \mathcal{B}_n for the subspace of degree n is the set of forests with n vertices. The product of two forests is their disjoint union, thus \mathcal{B} is a free-commutative basis, and the corresponding free generating set is the rooted trees. The coproduct of a tree T is given by

$$\Delta(T) = \sum T \setminus S \otimes S,$$

where the sum runs over all trunks S of T , including the empty tree and T itself, and $T \setminus S$ is the forest produced by removing from T all edges incident with S (each component is a *cut branch*). The root of each cut branch is the vertex which was closest to the root of T . Extend this definition multiplicatively to define the coproduct on forests: $\Delta(T_1 \amalg \dots \amalg T_l) = \Delta(T_1) \dots \Delta(T_l)$. Note that the trunk is always connected, but there may be several cut branches. Hence \mathcal{H} is noncocommutative.

It is not hard to derive a recursive formula for the coproduct of a tree. As above, write $T = [T_1 \dots T_f]$, where T_1, \dots, T_f are the connected components of T after removing the root. Then each non-empty trunk S of T has the form $S = [S_1 \dots S_f]$ for (possibly empty) trunks S_i of each T_i . The cut branches $T \setminus S$ are then the disjoint union $T_1 \setminus S_1 \amalg \dots \amalg T_f \setminus S_f$. So, in

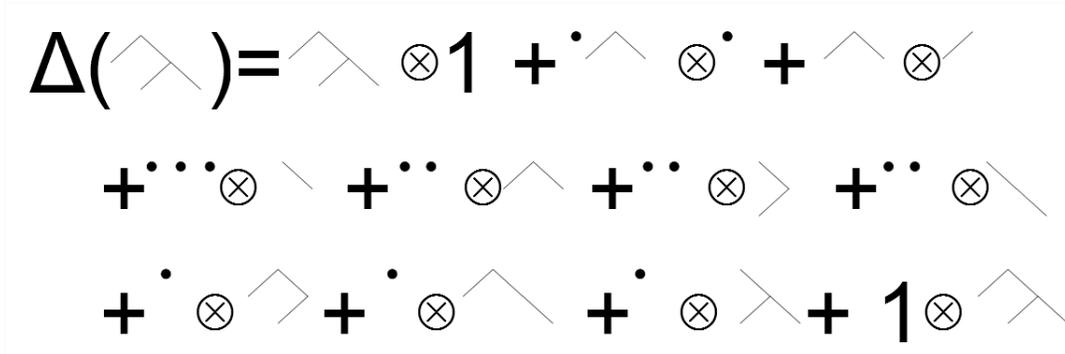


Figure 5.3: Coproduct of $[\bullet Q_3]$

Sweedler notation (so $\Delta(T_i) = \sum_{(T_i)} (T_i)_{(1)} \otimes (T_i)_{(2)}$), the following holds [CK98, Eq. 50, 51]:

$$\Delta([T_1 \dots T_f]) = T \otimes 1 + \sum_{(T_1), \dots, (T_f)} (T_1)_{(1)} \dots (T_f)_{(1)} \otimes [(T_1)_{(2)} \dots (T_f)_{(2)}]. \quad (5.3)$$

Example 5.3.4. Figure 5.3 calculates the coproduct for the tree $[\bullet Q_3]$ from Figure 5.2 above. Check this using Equation 5.3. By definition, $Q_3 = [\bullet\bullet]$ so

$$\Delta(Q_3) = Q_3 \otimes 1 + \bullet^2 \otimes \bullet + 2 \bullet \otimes P_2 + 1 \otimes Q_3.$$

(This made use of $[\bullet] = P_2$.) Then (recall $P_3 = [P_2]$),

$$\begin{aligned} \Delta([\bullet Q_3]) &= [\bullet Q_3] \otimes 1 + \bullet Q_3 \otimes \bullet + \bullet^3 \otimes P_2 + 2 \bullet^2 \otimes P_3 + \bullet \otimes [Q_3] \\ &\quad + Q_3 \otimes P_2 + \bullet^2 \otimes Q_3 + 2 \bullet \otimes [P_2 \bullet] + 1 \otimes [\bullet Q_3]. \end{aligned}$$

Example 5.3.5. Consider P_n , the path with n vertices. Its trunks are P_i , $0 \leq i \leq n$, and the sole cut branch corresponding to P_i is P_{n-i} . Hence $\Delta(P_n) = \sum_{i=0}^n P_{n-i} \otimes P_i$, which recovers the independent multinomial rock-breaking process of Section 5.2. Equivalently, $h_{(n)} \rightarrow P_n$ defines an embedding of the algebra of symmetric functions into \mathcal{H} .

5.3.2 Constructing the Chain

To describe the Hopf-power Markov chain on \mathcal{H} , it is necessary to first calculate the rescaling function η .

Theorem 5.3.6. *For a tree T , the rescaling function has the following “hook-length” formula*

$$\eta(T) = \frac{(\deg T)!}{T!}.$$

Proof. Proceed by induction on the number of vertices of T . The base case: $\eta(\bullet) = 1 = \frac{1!}{1!}$.

Now take $T \neq \bullet$. As previously, write $T = [T_1 \dots T_f]$, where T_1, \dots, T_f are the connected components of T after removing the root. View $\Delta^{[n]}$ as $(\iota \otimes \dots \otimes \iota \otimes \Delta)\Delta^{[n-1]}$; then the rescaling function η counts the ways to break T into singletons by pruning the vertices off one-by-one. Each such sequence of prunings is completely determined by the sequence of prunings (also one vertex off at a time) induced on each T_i , and a record of which T_i each of the first $\deg T - 1$ vertices came from (as the last vertex removed is the root). Hence

$$\begin{aligned} \eta(T) &= \binom{\deg T - 1}{\deg T_1 \dots \deg T_f} \eta(T_1) \dots \eta(T_f) \\ &= (\deg T - 1)! \frac{1}{T_1!} \dots \frac{1}{T_f!} = \frac{(\deg T)!}{T!}. \end{aligned}$$

□

As each tree in a forest breaks independently (Theorem 5.1.1), it suffices to understand the Markov chain starting from a tree. The below will give two descriptions of this: the second (Theorem 5.3.8) is generally more natural, but the first view may be useful for some special starting states; see Example 5.3.7 for the case where the starting states are stars Q_n . Depending on the starting state, one or the other interpretation may be easier to implement computationally.

The first interpretation is a straightforward application of the three-step description (Theorem 4.4.1). First take $a = 2$. Then, starting at a tree T of degree n , one step of the Hopf-square Markov chain is:

1. Choose i ($0 \leq i \leq n$) according to a symmetric binomial distribution.

2. Pick a trunk S of T of degree i with probability

$$\frac{\eta(S)\eta(T\setminus S)}{\eta(T)} = \frac{1}{\binom{n}{i}} \frac{T!}{S!(T\setminus S)!} = \frac{1}{\binom{n}{i}} \prod_{v \in S} \frac{\text{desc}_T(v)}{\text{desc}_S(v)}.$$

(The second equality holds because, for $v \notin S$, $\text{desc}_{T\setminus S}(v) = \text{desc}_T(v)$.)

3. Move to $T\setminus S \amalg S$.

Though it may be more mathematically succinct to combine the first two steps and simply choose a trunk S (of any degree) with probability $2^{-n} \prod_{v \in S} \frac{\text{desc}_T(v)}{\text{desc}_S(v)}$, the advantage of first fixing the trunk size i is that then one only needs to compute $\text{desc}_S(v)$ for trunks S of size i , not for all trunks.

Example 5.3.7. The star Q_n has $\binom{n-1}{i-1}$ trunks isomorphic to Q_i ($2 \leq i \leq n$), whose cut branches are respectively \bullet^{n-i} . The empty tree and \bullet are also legal trunks. Since the non-isomorphic trunks all have different degree, the second step above is trivial: the Hopf-square Markov chain sees Q_n move to $Q_i \bullet^{n-i}$ binomially. This corresponds to marking a corner of a rock and tracking the size of the marked piece under the rock-breaking process of Section 5.2. Note that this is not the same as removing the leaves of Q_n independently, as Q_n has $n-1$ leaves, not n .

To generalise this interpretation of the a th Hopf-power Markov chain to higher a , make use of coassociativity: $\Delta^{[a]} = (\mathbf{1} \otimes \cdots \otimes \mathbf{1} \otimes \Delta) \Delta^{[a-1]}$.

1. Choose the trunk sizes i_1, \dots, i_a (with $i_1 + \cdots + i_a = n$) according to a symmetric multinomial distribution.
2. Choose a trunk $S'_{(2)}$ of T of degree $i_2 + \cdots + i_a$, with probability $\frac{1}{\binom{n}{i_1}} \prod_{v \in S'_{(2)}} \frac{\text{desc}_T(v)}{\text{desc}_{S'_{(2)}}(v)}$.
3. Choose a trunk $S'_{(3)}$ of $S'_{(2)}$ of degree $i_3 + \cdots + i_a$, with probability $\frac{1}{\binom{n-i_1}{i_2}} \prod_{v \in S'_{(3)}} \frac{\text{desc}_{S'_{(2)}}(v)}{\text{desc}_{S'_{(3)}}(v)}$.
4. Continue choosing trunks $S'_{(4)}, S'_{(5)}, \dots, S'_{(a)}$ in the same way, and move to $T \setminus S'_{(2)} \amalg S'_{(2)} \setminus S'_{(3)} \amalg \cdots \amalg S'_{(a-1)} \setminus S'_{(a)} \amalg S'_{(a)}$.

Here is the second, more natural description of the tree-pruning chain, with a Jeu-de-Taquin flavour. Its inductive proof is at the end of this section.

Theorem 5.3.8. *One step of the a th Hopf-power Markov chain on rooted forests, starting at a tree T of degree n , is the following process:*

1. *Uniformly and independently assign one of a colours to each vertex of T .*
2. *If the root did not receive colour a , but there are some vertices in colour a , then uniformly select one of these to exchange colours with the root.*
3. *Look at the vertices v with $\text{anc}_T(v) = 2$ (i.e. the children of the root). Are there any of these which did not receive colour a , but has descendants in colour a ? Independently for each such v , uniformly choose a vertex u amongst its descendants in colour a , and switch the colours of u and v .*
4. *Repeat step 3 with vertices v where $\text{anc}_T(v) = 3, 4, \dots$ until the vertices of colour a form a trunk $S_{(a)}$ (i.e. no vertex of colour a is a descendant of a vertex of a different colour).*
5. *Repeat steps 2,3,4 with colours $a-1, a-2, \dots, 1$ on the cut branches $T \setminus S_{(a)}$ to obtain $S_{(a-1)}$. ($S_{(a-1)}$ is equivalent to $S'_{(a-1)} \setminus S'_{(a)}$ in the alternative “artificial” description above.)*
6. *Repeat step 5 to obtain $S_{(a-1)}, S_{(a-2)}, \dots, S_{(1)}$. Then move to $S_{(1)} \amalg S_{(2)} \amalg \dots \amalg S_{(a)}$.*

This colour exchange process is very natural if T describes the structure of an organisation, and if $a = 2$, where colour 1 indicates the members who leave, and colour 2 the members that stay. Then the recolourings are simply the promotion of members to fill deserted positions, with the assumption that the highest positions are replaced first, and that all members working under the departing member are equally qualified to be his or her replacement. [Pro09, Sec. 1] describes a related algorithm in a similar way.

Example 5.3.9. Take $T = [\bullet Q_3]$, as labelled in Figure 5.2: the root r has two children s and t , and t has two children u and v . Set $a = 2$, and let usual typeface denote colour 1, and **boldface** denote colour 2. Suppose step 1 above resulted in $r\mathbf{s}t\mathbf{u}v$. The root did not receive

colour 2, so, by step 2, either s or u must exchange colours with r . With probability $\frac{1}{2}$, u is chosen, and the resulting recolouring is \mathbf{rstuv} . As $\{r, s\}$ is a trunk of T , no more colour switching is necessary, and the chain moves to Q_3P_2 . If instead s had exchanged colours with r , then the recolouring would be \mathbf{rstuv} . Now step 3 is non-trivial, as $\text{anc}_T(t) = 1$, and t is not in colour 2, whilst its descendant u is. Since u is the only descendant of t in colour 2, t must switch colours with u , resulting in \mathbf{rstuv} . In this case, the chain moves to $(\bullet^3)P_2$.

Proof of Theorem 5.3.8, more natural description of the chain. Let $S'_{(a)} \subseteq S'_{(a-1)} \subseteq \cdots \subseteq S'_{(1)} = T$ be nested trunks, and write $S_{(j)}$ for the cut branches $S'_{(j)} \setminus S'_{(j+1)}$. The goal is to show that, after all colour exchanges,

$$P\{S_{(j)} \text{ ends up with colour } j \text{ for all } j\} = a^{-n} \prod_{j=1}^a \prod_{v \in S'_{(j)}} \frac{\text{desc}_{S'_{(j)}}(v)}{\text{desc}_{S'_{(j+1)}}(v)},$$

as this is the probability given by the previous, more artificial, description. Let a' be maximal so that $S_{(a')} \neq \emptyset$, so a' is the last colour which appears.

The key is to condition on the colouring of T after the root acquires colour a' (in the generic case where $a' = a$, this will be after step 2). Call this colouring χ , and notice that it can be any colouring where the root has colour a' , and $\deg S_{(j)}$ vertices have colour j . To reach this colouring after step 2, one of two things must have happened: either the starting colouring was already χ , or some vertex v that has colour $k \neq a'$ in χ originally had colour a' , and the root had colour k , and these colours were switched in step 2. For the second scenario, there are $\deg T - \deg S_{(a')}$ possible choices of v , and the chance that the root switched colours with v is $\frac{1}{\deg S_{(a')}}$. So

$$P\{\text{colouring after step 2 is } \chi\} = a^{-n} \left(1 + \frac{\deg T - \deg S_{(a')}}{\deg S_{(a')}} \right) = a^{-n} \frac{\deg T}{\deg S_{(a')}},$$

which depends only on $\deg S_{(a')}$, the number of vertices with the “last used colour” in χ , and not on which colour χ assigns each specific vertex. Consequently,

$$\begin{aligned}
& P\{S_{(j)} \text{ ends up with colour } j \text{ for all } j\} \\
&= \sum_{\chi} P\{S_{(j)} \text{ ends up with colour } j \text{ for all } j \mid \text{colouring after step 2 is } \chi\} \\
&\quad \times P\{\text{colouring after step 2 is } \chi\} \\
&= \sum_{\chi} P\{S_{(j)} \text{ ends up with colour } j \text{ for all } j \mid \text{colouring after step 2 is } \chi\} \left(a^{-n} \frac{\deg T}{\deg S_{(a')}} \right). \tag{5.4}
\end{aligned}$$

To calculate the sum on the right hand side, proceed by induction on $\deg T$. Write $T = [T_1 \dots T_f]$ as usual, and let χ_i be the induced colourings on the T_i . Then, because all colour exchanges after step 2 are between a non-root vertex and its descendant,

$$\begin{aligned}
& \sum_{\chi} P\{S_{(j)} \text{ ends up with colour } j \text{ for all } j \mid \text{colouring after step 2 is } \chi\} \\
&= \prod_{i=1}^f \sum_{\chi_i} P\{S_{(j)} \cap T_i \text{ ends up with colour } j \text{ for all } j \mid \text{starting colouring is } \chi_i\}.
\end{aligned}$$

Now note that each starting colouring of T_i has probability $a^{-\deg T_i}$, so, for each i ,

$$\begin{aligned}
& P\{S_{(j)} \cap T_i \text{ ends up with colour } j \text{ for all } j\} \\
&= \sum_{\chi} P\{S_{(j)} \cap T_i \text{ ends up with colour } j \text{ for all } j \mid \text{starting colouring is } \chi_i\} \\
&\quad \times P\{\text{starting colouring is } \chi_i\} \\
&= a^{-\deg T_i} \sum_{\chi} P\{S_{(j)} \cap T_i \text{ ends up with colour } j \text{ for all } j \mid \text{starting colouring is } \chi_i\}.
\end{aligned}$$

By inductive hypothesis, the left hand side is

$$a^{-\deg T_i} \prod_{j=1}^a \prod_{v \in S'_{(j)} \cap T_i} \frac{\text{desc}_{S'_{(j)} \cap T_i}(v)}{\text{desc}_{S'_{(j+1)} \cap T_i}(v)} = a^{-\deg T_i} \prod_{j=1}^a \prod_{v \in S'_{(j)} \cap T_i} \frac{\text{desc}_{S'_{(j)}}(v)}{\text{desc}_{S'_{(j+1)}}(v)}.$$

So, returning to (5.4),

$$\begin{aligned} & P\{S_{(j)} \text{ ends up with colour } j \text{ for all } j\} \\ &= \sum_{\chi} P\{S_{(j)} \text{ ends up with colour } j \text{ for all } j \mid \text{colouring after step 2 is } \chi\} \left(a^{-n} \frac{\deg T}{\deg S_{(a')}} \right) \\ &= \prod_{i=1}^f \sum_{\chi_i} P\{S_{(j)} \cap T_i \text{ ends up with colour } j \text{ for all } j \mid \text{starting colouring is } \chi_i\} \left(a^{-n} \frac{\deg T}{\deg S_{(a')}} \right) \\ &= \left(\prod_{i=1}^f \prod_{j=1}^a \prod_{v \in S'_{(j)} \cap T_i} \frac{\text{desc}_{S'_{(j)}}(v)}{\text{desc}_{S'_{(j+1)}}(v)} \right) \left(a^{-n} \frac{\deg T}{\deg S_{(a')}} \right) \\ &= a^{-n} \left(\prod_{j=1}^a \prod_{v \in S'_{(j)} \cap (\cup T_i)} \frac{\text{desc}_{S'_{(j)}}(v)}{\text{desc}_{S'_{(j+1)}}(v)} \right) \frac{\deg T}{\deg S_{(a')}} \\ &= a^{-n} \prod_{j=1}^a \prod_{v \in S'_{(j)}} \frac{\text{desc}_{S'_{(j)}}(v)}{\text{desc}_{S'_{(j+1)}}(v)}, \end{aligned}$$

since the root is the only vertex not in any T_i , and it is necessarily in $S'_{(a')}$. \square

5.3.3 Right Eigenfunctions

The aim of this section is to apply Proposition 5.1.14 to the special right eigenfunctions \mathbf{f}_C (C a tree) to bound the probability that the tree-pruning Markov chain can still reach $C \bullet \dots \bullet$ after a large number of steps. (C is in capital here in contrast to Section 5.1.3 as lowercase letters typically indicate vertices of trees.) Observe that being able to reach $C \bullet \dots \bullet$ is equivalent to containing C as a subtree.

More non-standard notation: for a vertex v in a forest T , let $\text{Anc}_T(v)$ denote the set of ancestors of v in T , including v itself. So $\text{Anc}_T(v)$ comprises the vertices on the path from v to the root of the connected component of T containing v , including both endpoints.

Theorem 5.3.10. *Let $C \neq \bullet$ be a tree, and T a forest. Then the right eigenfunction \mathbf{f}_C , of eigenvalue $a^{-\deg C+1}$, is*

$$\mathbf{f}_C(T) = \frac{1}{\deg C!} \sum_{C \subseteq T} \left(\left(\prod_{v \in \text{Anc}_T(\text{root}(C))} \frac{\text{desc}_T(v)}{\text{desc}_T(v) - \deg C + 1} \right) \left(\prod_{v \in C, v \neq \text{root}(C)} \text{desc}_T(v) \right) \right),$$

where the sum is over all subtrees of T isomorphic to C , though not necessarily with the same root. Moreover,

$$\frac{C! |\{C \subseteq T\}|}{\deg C! \deg C} \leq \mathbf{f}_C(T) \leq \binom{n'}{\deg C} \frac{|\{C \subseteq T\}|}{(n' - \deg C + 1)}$$

where $|\{C \subseteq T\}|$ is the number of subtrees of T isomorphic to C (not necessarily with the same root), and n' is the degree of the largest component of T .

The proof is fairly technical, so it is at the end of this section.

Remarks.

1. The second product in the expression for $\mathbf{f}_C(T)$ is not $C!$, since the product is over vertices of C , but the count is of the descendants in T .
2. The denominators $\text{desc}_T(v) - \deg C + 1$ are positive, since, if $v \in \text{Anc}_T(\text{root}(C))$, then all vertices of C are descendants of v .
3. The lower bound above is sharp: let $C = [Q_3]$, $T = [Q_3 P_{n-4}]$. Then $\mathbf{f}_C(T) = \frac{1}{4!} \frac{n}{n-3} \cdot 1 \cdot 1 \cdot 2 = \frac{1}{12} \frac{n}{n-3}$, which has limit $\frac{1}{12} = \frac{8 \cdot 1}{4!4} = \frac{C! |\{C \subseteq T\}|}{\deg C! \deg C}$, equal to the above lower bound, as $n \rightarrow \infty$.
4. The upper bound above is attained whenever C and T are both paths. In this case, the contribution to $\mathbf{f}_C(T)$ from the copy of C whose root is distance $n - i$ from $\text{root}(T)$ ($0 \leq i \leq n - \deg C$) is

$$\begin{aligned} & \frac{1}{\deg C!} \frac{n}{n - \deg C + 1} \frac{n-1}{n - \deg C} \cdots \frac{i}{i - \deg C + 1} (i-1) \cdots (i - \deg C + 1) \\ &= \frac{1}{\deg C!} n(n-1) \cdots (n - \deg C + 2) = \binom{n}{\deg C} \frac{1}{(n - \deg C + 1)}. \end{aligned}$$

Combining these bounds on $\mathbf{f}_C(T)$ with Proposition 1.1.3.ii gives the first of the two probability bounds below. The second result uses the universal bound of Proposition 5.1.14.

Corollary 5.3.11. *Let $\{X_m\}$ be the a th Hopf-power tree-pruning chain, started at a forest T . Write n' for the degree of the largest component of T . Then*

$$\begin{aligned} & P\{X_m \supseteq C | X_0 = T\} \\ & \leq E\{|\{\text{subtrees of } X_m \text{ isomorphic to } C\}| | X_0 = T\} \\ & \leq \frac{a^{(-\deg C + 1)m} \deg C! \deg C}{C!} \mathbf{f}_C(T) \\ & \leq \frac{a^{(-\deg C + 1)m} \deg C! \deg C}{C!} \binom{n}{\deg C} \frac{|\{C \subseteq T\}|}{(n - \deg C + 1)}. \end{aligned}$$

Besides, for any starting distribution on forests of n vertices,

$$P\{X_m \supseteq C\} \leq a^{-\deg C + 1} \deg C \binom{n}{\deg C}.$$

□

Example 5.3.12. Here is a demonstration of how to calculate with the formulae. Take $T = [\bullet Q_3]$ as in Figure 5.2, and calculate $\mathbf{f}_{Q_3}(T)$. As noted in Example 5.3.1, T has two subgraphs isomorphic to Q_3 , namely that spanned by $\{r, s, t\}$ and by $\{t, u, v\}$. The set of ancestors $\text{Anc}_T[\text{root}(Q_3)]$ is solely r for the first copy of Q_3 , and for the second copy of Q_3 , it is $\{r, t\}$. Hence

$$\begin{aligned} \mathbf{f}_{Q_3}(T) &= \frac{1}{\deg Q_3!} \left(\frac{\text{desc}_T(r)}{\text{desc}_T(r) - \deg Q_3 + 1} \text{desc}_T(s) \text{desc}_T(t) \right. \\ & \quad \left. + \frac{\text{desc}_T(r)}{\text{desc}_T(r) - \deg Q_3 + 1} \frac{\text{desc}_T(t)}{\text{desc}_T(z) - \deg Q_2 + 1} \text{desc}_T(u) \text{desc}_T(v) \right) \\ &= \frac{1}{6} \left(\frac{5}{3} \cdot 1 \cdot 3 + \frac{5}{3} \frac{3}{1} \cdot 1 \cdot 1 \right) = \frac{5}{3}. \end{aligned}$$

So, after m steps of the Hopf-square pruning chain started at T , the probability that there is still a vertex with at least two children is at most $\frac{2^{-2m} 3! 2}{3} \frac{5}{3} = 2^{-2m} \frac{20}{3}$.

Example 5.3.13. Specialise Corollary 5.3.11 to $C = P_j$, a path with j vertices. The copies of P_j in a tree T are in bijection with the vertices of T with at least j ancestors, by sending a path to its “bottommost” vertex (the one furthest from the root). There can be at most $\deg T - j + 1$ vertices with j or more ancestors, so by Corollary 5.3.11,

$$\begin{aligned} & P\{X_m \text{ has a vertex with } \geq j \text{ ancestors} | X_0 = T\} \\ & \leq E\{|\{\text{vertices of } X_m \text{ with } \geq j \text{ ancestors}\}| | X_0 = T\} \\ & \leq \frac{a^{(-j+1)m} j}{(\deg T - j + 1)} \binom{\deg T}{j} |\{\text{vertices of } T \text{ with } \geq j \text{ ancestors}\}| \\ & \leq a^{(-j+1)m} j \binom{\deg T}{j}. \end{aligned}$$

This result holds for any starting state T . In the particular case where T is the path P_n , this shows that, for the multinomial rock-breaking process of Section 5.2 started at a single rock of size n ,

$$P\{X_m \text{ contains a piece of size } \geq j | X_0 = (n)\} \leq a^{(-j+1)m} j \binom{n}{j},$$

which is looser than the bound in Proposition 5.2.1 by a factor of j .

Example 5.3.14. Take $C = Q_j$, the star with j vertices. Then $X_m \supseteq Q_j$ if and only if X_m has a vertex with at least $j - 1$ children. Each vertex with d children is responsible for $\binom{d}{j-1}$ copies of Q_j , so the two bounds in Corollary 5.3.11 are

$$\begin{aligned} & P\{X_m \text{ has a vertex with } \geq j - 1 \text{ children} | X_0 = T\} \\ & \leq E\{|\{\text{vertices of } X_m \text{ with } \geq j - 1 \text{ children}\}| | X_0 = T\} \\ & \leq E\{|\{\text{subtrees of } X_m \text{ isomorphic to } Q_j\}| | X_0 = T\} \\ & \leq \frac{a^{(-j+1)m} j!}{\deg T - j + 1} \binom{\deg T}{j} \sum_{v \in T} \binom{|\{\text{children of } v\}|}{j-1}, \\ & P\{X_m \text{ has a vertex with } \geq j - 1 \text{ children}\} \leq a^{(-j+1)m} j \binom{\deg T}{j}. \end{aligned}$$

The first bound is tighter if T has high degree compared to j , and has few vertices with at least j children.

Proof of Theorem 5.3.10. The following inductive argument proves both the expression for $\mathbf{f}_C(T)$ and the upper bound. To then obtain the lower bound, note that, for any vertex v , $\frac{\text{desc}_T(v)}{\text{desc}_T(v) - \text{deg } C + 1} \geq 1$, and for a subtree $C \subseteq T$,

$$\prod_{v \in C, v \neq \text{root}(T)} \text{desc}_T(v) \geq \prod_{v \in C, v \neq \text{root}(T)} \text{desc}_C(v) = \frac{C!}{\text{deg } C}.$$

To simplify notation, write $C_T!$ for $\prod_{v \in C, v \neq \text{root}(T)} \text{desc}_T(v)$, since $C_C! = \frac{1}{\text{deg } C} C!$. First, reduce both the expression for $\mathbf{f}_C(T)$ and the upper bound to the case when T is a tree: the claimed expression for $\mathbf{f}_C(T)$ is additive in the sense of Proposition 5.1.13, and $\binom{n}{\text{deg } C} \frac{1}{n - \text{deg } C + 1} = n(n-1) \dots (n - \text{deg } C + 2)$ is increasing in n . By definition of \mathbf{f}_C in Equation 5.1 and the calculation of $\eta(T)$ in Theorem 5.3.6, the goal is to prove

$$\begin{aligned} & \eta_T^{C, \bullet, \dots, \bullet} + \eta_T^{\bullet, C, \bullet, \dots, \bullet} + \dots + \eta_T^{\bullet, \dots, \bullet, C} \\ &= \frac{(\text{deg } T - \text{deg } C + 1)!}{T!} \sum_{C \subseteq T} \left(\left(\prod_{v \in \text{Anc}_T(\text{root}(C))} \frac{\text{desc}_T(v)}{\text{desc}_T(v) - \text{deg } C + 1} \right) C_T! \right) \\ &\leq |\{C \subseteq T\}| \frac{\text{deg } T!}{T!}. \end{aligned} \tag{5.5}$$

The key is again to write $T = [T_1 \dots T_f]$ and induct on degree. (The base case: when $T = \bullet$, both sides are zero, as there are no copies of C in T since $C \neq \bullet$.) The left hand side of (5.5) counts the ways to prune T successively so that it results in one copy of C and singletons. Divide this into two cases: $\eta_T^{\bullet, \dots, \bullet, C}$ counts the successive pruning processes where $C \ni \text{root}(T)$; the sum of the other coproduct structure constants in (5.5) counts the successive pruning processes where $C \not\ni \text{root}(T)$, so $C \subseteq T_i$ for some i . The inductive proof

below handles these cases separately, to show that

$$\begin{aligned}
 & \eta_T^{\bullet, \dots, \bullet, C} \\
 &= \frac{(\deg T - \deg C + 1)!}{T!} \sum_{\substack{C \subseteq T \\ C \ni \text{root}(T)}} \left(\left(\prod_{v \in \text{Anc}_T(\text{root}(C))} \frac{\text{desc}_T(v)}{\text{desc}_T(v) - \deg C + 1} \right) C_T! \right) \\
 &\leq |\{C \subseteq T \mid C \ni \text{root}(T)\}| \frac{\deg T!}{T!};
 \end{aligned} \tag{5.6}$$

$$\begin{aligned}
 & \eta_T^{C, \bullet, \dots, \bullet} + \dots + \eta_T^{\bullet, \dots, C, \bullet} \\
 &= \frac{(\deg T - \deg C + 1)!}{T!} \sum_{\substack{C \subseteq T \\ C \not\ni \text{root}(T)}} \left(\left(\prod_{v \in \text{Anc}_T(\text{root}(C))} \frac{\text{desc}_T(v)}{\text{desc}_T(v) - \deg C + 1} \right) C_T! \right) \\
 &\leq |\{C \subseteq T \mid C \not\ni \text{root}(T)\}| \frac{\deg T!}{T!}.
 \end{aligned} \tag{5.7}$$

Adding these together then gives (5.5).

The argument for (5.7) is simpler (though it relies on (5.5) holding for T_1, \dots, T_f). The ways to successively prune T into singletons and one copy of C not containing $\text{root}(T)$ correspond bijectively to the ways to prune some T_i into singletons and one copy of C (which may contain $\text{root}(T_i)$) and all other T_j into singletons, keeping track of which T_j was pruned at each step. Hence, writing d_i for $\deg T_i$,

$$\begin{aligned}
 & \eta_T^{C, \bullet, \dots, \bullet} + \dots + \eta_T^{\bullet, \dots, C, \bullet} \\
 &= \sum_i \binom{\deg T - \deg C}{d_1 \dots d_{i-1} \quad d_i - \deg C + 1 \quad d_{i+1} \dots d_f} \left(\eta_{T_i}^{C, \bullet, \dots, \bullet} + \dots + \eta_{T_i}^{\bullet, \dots, C, \bullet} \right) \prod_{j \neq i} \eta(T_j).
 \end{aligned}$$

Use Theorem 5.3.6 and the inductive hypothesis of (5.5) to substitute for $\eta(T_j)$ and $\eta_{T_i}^{C, \bullet, \dots, \bullet} + \dots + \eta_{T_i}^{\bullet, \dots, \bullet, C}$ respectively:

$$\begin{aligned} & \eta_T^{C, \bullet, \dots, \bullet} + \dots + \eta_T^{\bullet, \dots, \bullet, C} \\ &= \frac{(\deg T - \deg C)!}{T_1! \dots T_f!} \sum_i \sum_{\substack{C \subseteq T_i \\ C \not\subseteq \text{root}(T)}} \left(\left(\prod_{v \in \text{Anc}_{T_i}(\text{root}(C))} \frac{\text{desc}_{T_i}(v)}{\text{desc}_{T_i}(v) - \deg C + 1} \right) C_{T_i}! \right), \end{aligned} \quad (5.8)$$

and

$$\begin{aligned} & \eta_T^{C, \bullet, \dots, \bullet} + \dots + \eta_T^{\bullet, \dots, \bullet, C} \\ & \leq \frac{(\deg T - \deg C)!}{T_1! \dots T_f!} \sum_i \frac{d_i!}{(d_i - \deg C + 1)!} |\{C \subseteq T_i\}|. \end{aligned}$$

To deduce the equality in (5.7), first rewrite the fraction outside the sum in (5.8) as

$$\frac{\deg T}{\deg T - \deg C + 1} \frac{(\deg T - \deg C + 1)!}{T!}.$$

Then it suffices to show that

$$\begin{aligned} & \frac{\deg T}{\deg T - \deg C + 1} \sum_i \sum_{\substack{C \subseteq T_i \\ C \not\subseteq \text{root}(T)}} \left(\left(\prod_{v \in \text{Anc}_{T_i}(\text{root}(C))} \frac{\text{desc}_{T_i}(v)}{\text{desc}_{T_i}(v) - \deg C + 1} \right) C_{T_i}! \right) \\ &= \sum_{\substack{C \subseteq T \\ C \not\subseteq \text{root}(T)}} \left(\left(\prod_{v \in \text{Anc}_T(\text{root}(C))} \frac{\text{desc}_T(v)}{\text{desc}_T(v) - \deg C + 1} \right) C_T! \right). \end{aligned}$$

Now note that, for $C \subseteq T_i$, $\text{Anc}_T(\text{root}(C)) = \text{Anc}_{T_i}(\text{root}(C)) \cup \text{root}(T)$. For each $v \in \text{Anc}_{T_i}(\text{root}(C))$, $\text{desc}_{T_i}(v) = \text{desc}_T(v)$, and for $v = \text{root}(T)$, $\frac{\text{desc}_{T_i}(v)}{\text{desc}_{T_i}(v) - \deg C + 1} = \frac{\deg T}{\deg T - \deg C + 1}$. As for the inequality: for each i , $d_i = \deg T_i \leq \deg T - 1$, so

$$\frac{(\deg T - \deg C)! d_i!}{(d_i - \deg C + 1)!} = d_i(d_i - 1) \dots (d_i - \deg C + 2)(\deg T - \deg C)! \leq (\deg T - 1)!.$$

Now turn to the case where $C \ni \text{root}(T)$. Then $\text{Anc}_T(\text{root}(C)) = \text{root}(T)$, so (5.6) simplifies to

$$\eta_T^{\bullet, \dots, \bullet, C} = \frac{(\deg T - \deg C)!}{T!} \sum_{\substack{C \subseteq T \\ C \ni \text{root}(T)}} \prod_{v \in C} \text{desc}_T(v) \leq |\{C \subseteq T \mid C \ni \text{root}(T)\}| \frac{\deg T!}{T!}. \quad (5.9)$$

Here $C \ni \text{root}(T)$ means that $C \not\subseteq T_i$ for any i , hence a proof based on $T = [T_1 \dots T_f]$ will need to consider several C 's (in contrast to the previous paragraph when C did not contain $\text{root}(T)$). Note first that (5.9) does hold for $C = \emptyset$ (both sides are zero) and $C = \bullet$ (the formula for $\eta(T)$ as in Theorem 5.3.6) - these cases are not part of the theorem, but are useful for the proof. For $C \neq \emptyset, \bullet$, write $C = [C_1 \dots C_{f'}]$; necessarily $f' \leq f$ or there would be no copy of C in T with $C \ni \text{root}(T)$ (then both sides of (5.9) are zero). For ease of notation, let $C_{f'+1} = \dots = C_f = \emptyset$. Recall that $\eta_T^{\bullet, \dots, \bullet, C}$ counts the number of ways to successively prune vertices from T to leave C . This is equivalent to successively pruning vertices from each T_i to leave C_1 in some $T_{\sigma(1)}$, C_2 in some $T_{\sigma(2)}$, etc, and keeping track of which T_i was pruned at each step. Thus

$$\eta_T^{\bullet, \dots, \bullet, C} = \sum_{\sigma} \binom{\deg T - \deg C}{\deg T_{\sigma(1)} - \deg C_1 \dots \deg T_{\sigma(f')} - \deg C_{f'}} \eta_{T_{\sigma(1)}}^{\bullet, \dots, \bullet, C_1} \dots \eta_{T_{\sigma(f')}}^{\bullet, \dots, \bullet, C_{f'}},$$

where the sum is over one choice of $\sigma \in S_f$ for each distinct multiset of pairs $\{(C_1, T_{\sigma(1)}), \dots, (C_{f'}, T_{\sigma(f')})\}$. The inductive hypothesis of (5.7) for $(C_i, T_{\sigma(i)})$ then yields

$$\begin{aligned} \eta_T^{\bullet, \dots, \bullet, C} &= \sum_{\sigma} (\deg T - \deg C)! \prod_i \sum_{\substack{C_i \subseteq T_{\sigma(i)} \\ C_i \ni \text{root}(T_{\sigma(i)})}} \frac{1}{T_{\sigma(i)}!} \prod_{v \in C_i} \text{desc}_{T_{\sigma(i)}}(v) \\ &= \frac{\deg T}{T!} \sum_{\substack{C \subseteq T \\ C \ni \text{root}(T)}} \frac{(\deg T - \deg C)!}{\deg T!} \prod_{v \in C, v \neq \text{root}(T)} \text{desc}_T(v), \end{aligned}$$

since $\text{root}(T) = C \setminus \coprod C_i$, and for each $v \in C_i$, $\text{desc}_{T_{\sigma(i)}}(v) = \text{desc}_T(v)$. To conclude the equality in (5.9), simply absorb the factor of $\deg T$ at the front into the product as $\text{desc}_T(\text{root}(T))$.

Also by the inductive hypothesis,

$$\eta_T^{\bullet, \dots, \bullet, C} \leq \sum_{\sigma} (\deg T - \deg C)! \prod_i \frac{\deg T_{\sigma(i)}!}{(\deg T_{\sigma(i)} - \deg C_i)! T_{\sigma(i)}!} |\{C_i \subseteq T_{\sigma(i)} | C_i \ni \text{root}(T_{\sigma(i)})\}|$$

Now $\frac{\deg T_{\sigma(i)}!}{(\deg T_{\sigma(i)} - \deg C_i)!}$ enumerates the ways to choose $\deg C_i$ ordered objects amongst $\deg T_{\sigma(i)}$; choosing such objects for each i is a subset of the ways to choose $\deg C - 1$ objects from $\deg T - 1$. Hence

$$\begin{aligned} \eta_T^{\bullet, \dots, \bullet, C} &\leq \frac{(\deg T - 1)!}{(\deg T - \deg C)!} \sum_{\sigma} (\deg T - \deg C)! \prod_i \frac{1}{T_{\sigma(i)}!} |\{C_i \subseteq T_{\sigma(i)} | C_i \ni \text{root}(T_{\sigma(i)})\}| \\ &= \frac{\deg T!}{T!} |\{C \subseteq T | C \ni \text{root}(T)\}| \end{aligned}$$

as claimed. □

Chapter 6

Hopf-power Markov Chains on Cofree Commutative Algebras

Sections 6.1 and 6.2 study in detail respectively the chains of riffle-shuffling and of the descent set under riffle-shuffling. These arise from the shuffle algebra and the algebra of quasisymmetric functions, which are both cofree and commutative.

6.1 Riffle-Shuffling

Recall from Chapter 1 the Gilbert-Shannon-Reeds model of riffle-shuffling of a deck of cards: cut the deck binomially into two piles, then choose uniformly an interleaving of the two piles. The first extensive studies of this model are [AD86, Sec. 4] and [BD92]. They give explicit formulae for all the transition probabilities and find that $\frac{3}{2} \log n$ shuffles are required to mix a deck of n distinct cards. More recently, [ADS11] derives the convergence rate for decks of repeated cards, which astonishingly depends almost entirely on the total number of cards and the number of distinct values that they take. The number of cards of each value hardly influences the convergence rate.

One key notion introduced in [BD92] is the generalisation of the GSR model to a -handed shuffles, which cuts the deck into a piles multinomially before uniformly interleaving. As Example 4.4.2 showed, a -handed shuffling is exactly the a th Hopf-power Markov chain on the shuffle algebra \mathcal{S} , with respect to its basis of words. In \mathcal{S} , the product of two

words is the sum of their interleavings (with multiplicity), and the coproduct of a word is the sum of its deconcatenations - see Example 4.1.1. As mentioned in Section 4.1, \mathcal{S} has a multigrading: for a sequence of non-negative integers \mathbf{v} , the subspace $\mathcal{S}_{\mathbf{v}}$ is spanned by words where 1 appears v_1 times, 2 appears v_2 times, etc. The Hopf-power Markov chain on $\mathcal{S}_{\mathbf{v}}$ describes shuffling a deck of *composition* \mathbf{v} , where there are v_i cards with face value i . For example, $\mathbf{v} = (1, 1, \dots, 1)$ corresponds to a deck where all cards are distinct, and $\mathbf{v} = (n-1, 1)$ describes a deck with one distinguished card, as studied in [ADS12, Sec. 2]. Work with the following partial order on deck compositions: $\mathbf{v} \geq \mathbf{v}'$ if $v_i \geq v'_i$ for all i . Write $|\mathbf{v}|$ for the sum of the entries of \mathbf{v} - this is the number of cards in a deck with composition \mathbf{v} . For a word w , let $\deg(w)$ denote its corresponding deck composition (this is also known as the *evaluation* $ev(w)$), and $|w| = |\deg w|$ the total number of cards in the deck. For example, $\deg((1233212)) = (2, 3, 2)$, and $|1233212| = 7$. Since the cards behave equally independent of their values, there is no harm in assuming $v_1 \geq v_2 \geq \dots$. In other words, it suffices to work with $\mathcal{H}_{\mathbf{v}}$ for partitions \mathbf{v} , though what follows will not make use of this reduction.

A straightforward application of Theorem 4.5.1 shows that the stationary distribution of riffle-shuffling is the uniform distribution, for all powers a and all deck compositions \mathbf{v} .

Sections 6.1.1 and 6.1.2 construct some simple right and left eigenfunctions using Parts B' and A' of Theorem 2.5.1 respectively, and Section 6.1.3 gives a partial duality result. All this relies on the Lyndon word terminology of Section 2.4. Much of the right eigenfunction analysis is identical to [DPR14, Sec. 5], which studies inverse shuffling as the Hopf-power Markov chain on the free associative algebra; the left eigenfunction derivations here are new. In the case of distinct cards, these right and left eigenfunctions have previously appeared in [Sal12, Sec. 4; Den12, Th. 3.6] and [Pik13] respectively. All these examine the time-reversal of riffle-shuffling in the context of walks on hyperplane arrangements and their generalisation to left regular bands.

6.1.1 Right Eigenfunctions

Recall from Proposition 3.2.1.R that the right eigenfunctions of a Hopf-power Markov chain come from diagonalising Ψ^a on the dual of the underlying Hopf algebra. For the case

of riffle-shuffling, this dual is the free associative algebra (Example 2.1.2), with concatenation product and deshuffling coproduct. The word basis of the free associative algebra fits the hypothesis of Theorem 2.5.1.B'. All single letters have degree 1, so there is no need to apply the Eulerian idempotent, which simplifies the algorithm a little. In order to achieve $\sum_w \mathbf{f}_w(v) \mathbf{g}_w(v) = 1$ for some w , with the left eigenbasis \mathbf{g}_w in Section 6.1.2 below, it will be necessary to divide the output of Theorem 2.5.1.B' by an extra factor $Z(w)$, the size of the stabiliser of $\mathfrak{S}_{k(w)}$ permuting the Lyndon factors of w . For example, (31212) has Lyndon factorisation $(3 \cdot 12 \cdot 12)$, and the stabiliser of \mathfrak{S}_3 permuting these factors comprises the identity and the transposition of the last two elements, so $Z((31212)) = 2$.

Coupling this rescaled version of Theorem 2.5.1.B' with Proposition 3.2.1.R, $\mathbf{f}_{w'}(w)$ is the coefficient of w in:

$$\begin{aligned} f_{w'} &:= w' && \text{if } w' \text{ is a single letter;} \\ f_{w'} &:= f_{u_1} f_{u_2} - f_{u_2} f_{u_1} && \text{if } w' \text{ is Lyndon with standard factorisation } w' = u_1 \cdot u_2; \\ f_{w'} &:= \frac{1}{Z(w')k!} \sum_{\sigma \in \mathfrak{S}_k} f_{u_{\sigma(1)}} \cdots f_{u_{\sigma(k)}} && \text{if } w' \text{ has Lyndon factorisation } w' = u_1 \cdots u_k. \end{aligned}$$

(The second line is a recursive definition for the *standard bracketing*.) A visual description of $\mathbf{f}_{w'}(w)$ is two paragraphs below.

$\mathbf{f}_{w'}$ is a right eigenfunction of eigenvalue $a^{-|w'|+k(w')}$, where $k(w')$ is the number of Lyndon factors of w' . Since the $\mathbf{f}_{w'}$ form an eigenbasis, the multiplicity of the eigenvalue $a^{-|\nu|+k}$ when shuffling a deck of composition ν is the number of words of degree ν with k Lyndon factors. This has two consequences of note. Firstly, when $\nu = (1, 1, \dots, 1)$, this multiplicity is $c(|\nu|, k)$, the signless Stirling number of the first kind. Its usual definition is the number of permutations of $|\nu|$ objects with k cycles, which is easily equivalent [Sta97, Prop. 1.3.1] to the number of words of $\deg(\nu)$ with k record minima. (The letter i is a *record minima* of w if all letters appearing before i in w are greater than i .) This is the eigenvalue multiplicity because a word with distinct letters is Lyndon if and only if its first letter is minimal, so the Lyndon factors of a word with distinct letters start precisely at the record minima.

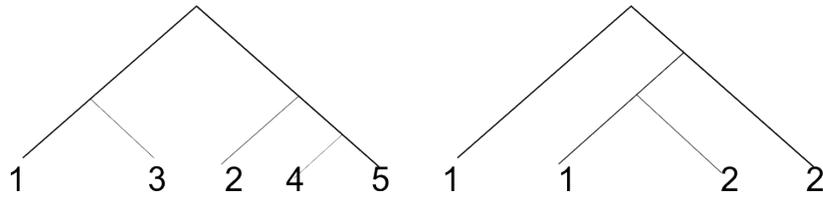


Figure 6.1: The trees $T_{(13245)}$ and $T_{(1122)}$

Secondly, for general ν , the eigenvalues $1, a^{-1}, \dots, a^{-|\nu|+1}$ all occur. Each eigenfunction \mathbf{f}_w of eigenvalue $a^{-|\nu|+k}$ corresponds to a word of degree ν with k Lyndon factors, or equivalently, k Lyndon words whose degrees sum to ν . One way to find such a k -tuple is to choose a Lyndon word of length $|\nu| - k + 1$ in which letter i occurs at most ν_i times, and take the remaining $k - 1$ letters of ν as singleton Lyndon factors. How to construct the non-singleton Lyndon factor depends on ν and k : if $|\nu| - k > \nu_1$, one possibility is the smallest $|\nu| - k + 1$ values in increasing order. For $|\nu| - k \leq \nu_1$, take the word with $|\nu| - k$ 1s followed by a 2.

As for the eigenvectors, [GR89, Sec. 2] and [BB90] provide a way to calculate them graphically, namely via *decreasing Lyndon hedgerows*. For a Lyndon word u with standard factorisation $u = u_1 \cdot u_2$, inductively draw a rooted binary tree T_u by taking T_{u_1} as the left branch and T_{u_2} as the right branch. Figure 6.1 shows $T_{(13245)}$ and $T_{(1122)}$.

For a Lyndon word u , it follows from the recursive definition of f_u above that $\mathbf{f}_u(w)$ is the signed number of ways to exchange the left and right branches at some vertices of T_u so that the leaves of T_u , reading from left to right, spell out w (the sign is the parity of the number of exchanges required). For example,

- $\mathbf{f}_{(13245)}((25413)) = 1$ since the unique way to rearrange $T_{(13245)}$ so the leaves spell (25413) is to exchange the branches at the root and the lowest interior vertex;
- $\mathbf{f}_{(13245)}((21345)) = 0$ since in all legal rearrangements of $T_{(13245)}$, 2 appears adjacent to either 4 or 5, which does not hold for (21345);
- $\mathbf{f}_{(1122)}((1221)) = 0$ as there are two ways to make the leaves of $T_{(1122)}$ spell (1221): either exchange branches at the root, or exchange branches at both of the other interior vertices. These two rearrangements have opposite signs, so the signed count of rearrangements is 0.

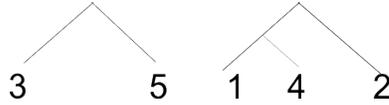


Figure 6.2: The Lyndon hedgerow $T_{(35142)}$

Now for general w' with Lyndon factorisation $w = u_1 \cdots u_k$, set $T_{w'}$ to be simply $T_{u_1}, T_{u_2}, \dots, T_{u_k}$ placed in a row. So $T_{(35142)}$ is the hedgerow in Figure 6.2.

Again $\mathbf{f}_{w'}(w)$ is the signed number of ways to rearrange $T_{w'}$ so the leaves spell w , divided by $k!$. Now there are two types of allowed moves: exchanging the left and right branches at a vertex (as before), and permuting the trees of the hedgerow. The latter move does not come with a sign. Thus $\mathbf{f}_{(35142)}((14253)) = \frac{1}{2!}(-1)$, as the unique rearrangement of $T_{(35142)}$ which spells (14253) requires transposing the trees and permuting the branches of 3 and 5. The division by $Z(w')$ in the definition of $\mathbf{f}_{w'}$ means that, if w' has a repeat Lyndon factor, then the multiple trees corresponding to this repeated factor are not distinguished, and transposing them does not count as a valid rearrangement. So if $w' = (31212) = (3 \cdot 12 \cdot 12)$, then $\mathbf{f}_{w'}((12312)) = \frac{1}{3!}$.

Writing \overleftarrow{w} for the reverse of w , this graphical calculation method shows that $\mathbf{f}_{w'}(\overleftarrow{w})$ and $\mathbf{f}_{w'}(w)$ differ only in possibly a sign, since switching branches at every interior vertex and arranging the trees in the opposite order reverses the word spelt by the leaves. The number of interior vertices of a tree is one fewer than the number of leaves, hence the sign change is $(-1)^{|w'| - k(w')}$, which depends only on the corresponding eigenvalue. In conclusion,

Proposition 6.1.1. *Let \overleftarrow{w} denote the reverse of w . Then, if \mathbf{f} is any right eigenfunction of a -handed shuffling with eigenvalue a^j then $\mathbf{f}(w) = (-1)^j \mathbf{f}(\overleftarrow{w})$. \square*

Let u be a Lyndon word. In similar notation abuse as in Section 5.1.3, write $\mathbf{f}_u(w)$ for the sum of \mathbf{f}_u evaluated on all consecutive subwords of w whose degree is $\deg u$ (i.e. on all consecutive subwords of w whose constituent letters are those of u). For example, in calculating $\mathbf{f}_{(12)}((1233212))$, the relevant subwords are 1233212, 1233212 and 1233212, so $\mathbf{f}_{(12)}((1233212)) = \mathbf{f}_{(12)}((12)) + \mathbf{f}_{(12)}((21)) + \mathbf{f}_{(12)}((12)) = 1 - 1 + 1 = 1$. It is clear from the graphical calculation of eigenfunctions that, on any subspace \mathcal{S}_v with $v \geq \deg(u)$, this new function \mathbf{f}_u is $(|v| - |u| + 1)! \mathbf{f}_{w'}$ where w' has degree v and u is its only non-singleton Lyndon factor. The corresponding eigenvalue is $a^{-|u|+1}$. For the example above,

$\mathbf{f}_{(12)} = (7 - 2 + 1)! \mathbf{f}_{(3322121)}$, since (3322121) has Lyndon factorisation $(3 \cdot 3 \cdot 2 \cdot 2 \cdot 12 \cdot 1)$. The pointwise products of certain \mathbf{f}_u s are also right eigenfunctions, see Proposition 6.1.6 at the end of this section.

Example 6.1.2. Take the simplest case of $u = (ij)$, where $i < j$. Then $\mathbf{f}_{(ij)}(w)$ is the number of consecutive subwords (ij) occurring in w , subtract the occurrences of (ji) as a consecutive subword. In particular, if w has distinct letters, then

$$\mathbf{f}_{(ij)}(w) = \begin{cases} 1, & \text{if } (ij) \text{ occurs as a consecutive subword of } w; \\ -1, & \text{if } (ji) \text{ occurs as a consecutive subword of } w; \\ 0, & \text{otherwise.} \end{cases}$$

The corresponding eigenvalue is $\frac{1}{a}$.

Summing the $\mathbf{f}_{(ij)}$ over all pairs $i < j$ gives another right eigenfunction \mathbf{f}_\setminus , also with eigenvalue $\frac{1}{a}$: $\mathbf{f}_\setminus(w)$ counts the increasing 2-letter consecutive subwords of w , then subtracts the number of decreasing 2-letter consecutive subwords. These subwords are respectively the ascents and descents of w , so denote their number by $\text{asc}(w)$ and $\text{des}(w)$ respectively. Note that reversing w turns an ascent into a descent, so $\mathbf{f}_\setminus(w) = -\mathbf{f}_\setminus(\overleftarrow{w})$, as predicted by Proposition 6.1.1. If w has all letters distinct then the non-ascents are precisely the descents; this allows Proposition 6.1.3 below to express \mathbf{f}_\setminus solely in terms of $\text{des}(w)$. (This explains the notation \mathbf{f}_\setminus . Descents typically receive more attention in the literature than ascents.) The claims regarding expected values follow from Proposition 1.1.3.i.

Proposition 6.1.3. *The function $\mathbf{f}_\setminus : \mathcal{B}_V \rightarrow \mathbb{R}$ with formula*

$$\mathbf{f}_\setminus(w) := \text{asc}(w) - \text{des}(w)$$

is a right eigenfunction of a -handed shuffling of eigenvalue $\frac{1}{a}$. Hence, if X_m denotes the deck order after m shuffles,

$$E\{\text{asc}(X_m) - \text{des}(X_m) | X_0 = w_0\} = a^{-m}(\text{asc}(w_0) - \text{des}(w_0)).$$

If $\mathbf{v} = (1, 1, \dots, 1)$, then \mathbf{f}_\backslash is a multiple of the “normalised number of descents”:

$$\mathbf{f}_\backslash(w) := -2 \left(\text{des}(w) - \frac{n-1}{2} \right).$$

So, if a deck of distinct cards started in ascending order (i.e. $\text{des}(w_0) = 0$), then

$$E\{\text{des}(X_m) | X_0 = w_0\} = (1 - a^{-m}) \left(\frac{n-1}{2} \right).$$

□

Similar analysis applies to Lyndon words with three letters:

Example 6.1.4. Fix three letters $i < j < k$. There are two Lyndon words with three distinct letters: (ijk) and (ikj) . Their standard factorisations are $(i \cdot jk)$ and $(ik \cdot j)$, so

- $\mathbf{f}_{(ijk)}$ counts the consecutive subwords (ijk) and (kji) with weight 1, and (ikj) and (jki) with weight -1;
- $\mathbf{f}_{(ikj)}$ counts the consecutive subwords (ikj) and (jki) with weight 1, and (kij) and (jik) with weight -1.

By inspection, $\mathbf{f}_{(ijk)} = \mathbf{f}_{(ij)}\mathbf{f}_{(jk)} + \mathbf{f}_{(ik)}\mathbf{f}_{(jk)}$, $\mathbf{f}_{(ikj)} = -\mathbf{f}_{(ik)}\mathbf{f}_{(jk)} + \mathbf{f}_{(ik)}\mathbf{f}_{(ij)}$. (This is unrelated to Proposition 6.1.6.) These have eigenvalue a^{-2} .

When all cards in the deck are distinct, certain linear combinations of these again have a neat interpretation in terms of well-studied statistics on words. The table below lists the definition of the four relevant statistics in terms of 3-letter consecutive subwords, and the (non-standard) notation for their number of occurrences in a given word w .

peak	$\text{peak}(w)$	middle letter is greatest
valley	$\text{vall}(w)$	middle letter is smallest
double ascent	$\text{aasc}(w)$	letters are in increasing order
double descent	$\text{ddes}(w)$	letters are in decreasing order

For example, if $w = (1233212)$, then $\text{vall}(w) = \text{aasc}(w) = \text{ddes}(w) = 1$ and $\text{peak}(w) = 0$.

Proposition 6.1.5. *The function $\mathbf{f}_{\wedge\vee} : \mathcal{B}_v \rightarrow \mathbb{R}$ with formula*

$$\mathbf{f}_{\wedge\vee}(w) := \text{peak}(w) - \text{vall}(w)$$

is a right eigenfunction of a -handed shuffling of eigenvalue a^{-2} . Hence, if X_m denotes the deck order after m shuffles,

$$E\{\text{peak}(X_m) - \text{vall}(X_m) | X_0 = w_0\} = a^{-2m}(\text{peak}(w_0) - \text{vall}(w_0)).$$

If $v = (1, 1, \dots, 1)$, then the following are also right eigenfunctions of a -handed shuffling of eigenvalue a^{-2} :

$$\begin{aligned} \mathbf{f}_{\wedge}(w) &:= \text{peak}(w) - \frac{n-2}{3}; \\ \mathbf{f}_{\vee}(w) &:= \text{vall}(w) - \frac{n-2}{3}; \\ \mathbf{f}_{-}(w) &:= \text{aasc}(w) + \text{ddes}(w) - \frac{n-2}{3}. \end{aligned}$$

So, if a deck of distinct cards started in ascending order (i.e. $\text{peak}(w_0) = \text{vall}(w_0) = 0$), then

$$\begin{aligned} E\{\text{peak}(X_m) | X_0 = w_0\} &= E\{\text{vall}(X_m) | X_0 = w_0\} = (1 - a^{-2m})\frac{n-2}{3}; \\ E\{\text{aasc}(X_m) + \text{ddes}(X_m) | X_0 = w_0\} &= (1 + 2a^{-2m})\frac{n-2}{3}. \end{aligned}$$

Proof. From Example 6.1.4 above,

$$\begin{aligned} \sum_{i < j < k} \mathbf{f}_{(ijk)}(w) &= \text{aasc}(w) + \text{ddes}(w) - \text{peak}(w), \\ \sum_{i < j < k} \mathbf{f}_{(ikj)}(w) &= \text{peak}(w) - \text{vall}(w) = \mathbf{f}_{\wedge\vee}(w) \end{aligned}$$

are right eigenfunctions of eigenvalue a^{-2} . If all cards in the deck are distinct, then

$$\text{peak}(w) + \text{vall}(w) + \text{aasc}(w) + \text{ddes}(w) = n - 2,$$

so $\mathbf{f}_\wedge = \frac{1}{3}(\mathbf{f}_{\wedge\vee} - \sum_{i<j<k} \mathbf{f}_{(ijk)})$, $\mathbf{f}_\vee = \frac{-1}{3}(\sum_{i<j<k} \mathbf{f}_{(ijk)} + 2\mathbf{f}_{\wedge\vee})$, $\mathbf{f}_- = \frac{1}{3}(2\sum_{i<j<k} \mathbf{f}_{(ijk)} + \mathbf{f}_{\wedge\vee})$ are also right eigenfunctions. The statements on expectations follow from Proposition 1.1.3. \square

Linear combinations of \mathbf{f}_u for Lyndon u with $|u| = 4$ provides right eigenfunctions of eigenvalue a^{-3} which are weighted counts of consecutive 4-letter subwords of each “pattern”, but these are more complicated.

Here is one final fact about right eigenfunctions, deducible from the graphical calculation:

Proposition 6.1.6. *Let u_1, \dots, u_j be Lyndon words each with distinct letters, such that no letter appears in more than one u_i . Then, for any $v \geq \deg(u_1) + \dots + \deg(u_j)$, the pointwise product $\mathbf{f}(w) := \mathbf{f}_{u_1}(w) \dots \mathbf{f}_{u_j}(w)$ is a right eigenfunction on \mathcal{S}_v of eigenvalue $a^{-|u_1| - \dots - |u_j| + j}$; in fact, $\mathbf{f} = (|v| - |u_1| - \dots - |u_j| + j)! \mathbf{f}_{w'}$, where the only non-singleton Lyndon factors of w' are precisely u_1, \dots, u_j , each occurring exactly once.* \square

Under these same conditions, the corresponding relationship for the left eigenfunctions of the following section is $\mathbf{g}_{u_1}(w) \dots \mathbf{g}_{u_j}(w) = \frac{1}{Z(w')} \mathbf{g}_{w'}$, where again w' is the word whose only non-singleton Lyndon factors are precisely u_1, \dots, u_j .

Example 6.1.7. Let $v = (2, 1, 1, 1, 1)$ and let $w' = (352141)$ which has Lyndon factorisation $(35 \cdot 2 \cdot 14 \cdot 1)$. The two non-singleton Lyndon factors (35) and (14) combined have distinct letters, so $\mathbf{f}_{w'} = \frac{1}{(6-2-2+2)!} \mathbf{f}_{(35)} \mathbf{f}_{(14)}$. For instance, $\mathbf{f}_{w'}((114253)) = \frac{1}{24} \mathbf{f}_{(35)}((114253)) \mathbf{f}_{(14)}((114253)) = \frac{1}{24} (-1) 1 = -\frac{1}{24}$.

6.1.2 Left Eigenfunctions

Now comes a parallel analysis of the left eigenfunctions, which arise from diagonalising Ψ^a on the shuffle algebra \mathcal{S} . Apply Theorem 2.5.1.A' to the word basis of \mathcal{S} and use Proposition 3.2.1.L to translate the result: if w' has Lyndon factorisation $u_1 \dots u_k$, then the left eigenfunction $\mathbf{g}_{w'}(w) = \text{coefficient of } w \text{ in } e(u_1) \dots e(u_k)$, where e is the Eulerian idempotent map:

$$e(x) = \sum_{r \geq 1} \frac{(-1)^{r-1}}{r} m^{[r]} \bar{\Delta}^{[r]}(x).$$

Again, concentrate on the case where only one of the factors is not a single letter. For a Lyndon word u , let $\mathbf{g}_u(w)$ be the sum of \mathbf{g}_u evaluated on all subwords (not necessarily consecutive) of w whose degree is $\deg u$ (i.e. on all subwords of w whose constituent letters are those of u). For example, the relevant subwords for calculating $\mathbf{g}_{(12)}((1233212))$ are $\underline{1233212}$, $\underline{12332}12$, $\underline{123321}\underline{2}$, $\underline{123321}2$, $\underline{123321}2$, and $\underline{123321}2$. Because $e(12) = (12) - \frac{1}{2}(1)(2) = \frac{1}{2}((12) - (21))$, it follows that $\mathbf{g}_{(12)}((1233212)) = 4\mathbf{g}_{(12)}((12)) + 2\mathbf{g}_{(12)}((21)) = \frac{1}{2}(4 - 2) = 1$. It follows from the definition of $\mathbf{g}_{w'}$ for general w' that, on any subspace \mathcal{S}_v with $v \geq \deg(u)$, this new function $\mathbf{g}_u = \frac{1}{Z(w')} \mathbf{g}_{w'}$ for w' with degree v and u its only non-singleton Lyndon factor, as was the case with right eigenfunctions. (Recall that $Z(w')$ is the size of the stabiliser in \mathfrak{S}_k permuting the Lyndon factors of w' .) The corresponding eigenvalue is $a^{-|u|+1}$. For the example above, $\mathbf{g}_{(12)} = \frac{1}{2!} \mathbf{g}_{(3322121)}$.

Example 6.1.8. Again, start with $u = (ij)$, with $i < j$. Because $e(ij) = (ij) - \frac{1}{2}(i)(j) = \frac{1}{2}((ij) - (ji))$, the left eigenfunction $\mathbf{g}_{(ij)}$ counts the pairs (i, j) with i appearing before j , subtracts the number of pairs (i, j) with i occurring after j , then divides by 2. In particular, if w has distinct letters, then

$$\mathbf{g}_{(ij)}(w) = \begin{cases} \frac{1}{2}, & \text{if } i \text{ occurs before } j \text{ in } w; \\ -\frac{1}{2}, & \text{if } i \text{ occurs after } j \text{ in } w. \end{cases}$$

The corresponding eigenvalue is $\frac{1}{a}$. In general, \mathbf{f}_u and \mathbf{g}_u do not count the same subwords.

As before, sum the $\mathbf{g}_{(ij)}$ over all pairs $i < j$ to obtain a more “symmetric” left eigenfunction \mathbf{g}_\setminus , also with eigenvalue $\frac{1}{a}$: $\mathbf{g}_\setminus(w)$ halves the number of pairs appearing in increasing order in w minus the number of *inversions* $\text{inv}(w)$, when a pair appears in decreasing order. These eigenfunctions also feature in [Pik13, Th. 3.2.1]. Out of the $\binom{|w|}{2}$ pairs of letters in w , there are $\sum_i \binom{(\deg w)_i}{2}$ pairs of the same letter, and all other pairs must either appear in increasing order or be an inversion. This explains:

Proposition 6.1.9. *The function $\mathbf{g}_\setminus : \mathcal{B}_v \rightarrow \mathbb{R}$ with formula*

$$\mathbf{g}_\setminus(w) := \frac{1}{2} \binom{|v|}{2} - \frac{1}{2} \sum_i \binom{v_i}{2} - \text{inv}(w)$$

is a left eigenfunction of a -handed shuffling of eigenvalue $\frac{1}{a}$. \square

There is no terminology for a “non-consecutive peak” in the same way that an inversion is a “non-consecutive descent”, so it is not too interesting to derive an analogue of Proposition 6.1.5 from $\mathbf{g}_{(ijk)}$ and $\mathbf{g}_{(ikj)}$.

6.1.3 Duality of Eigenfunctions

Recall from Proposition 1.1.4 that explicit diagonalisation of Markov chains is most useful when the right and left eigenbases obtained are dual bases. This is almost true of $\{\mathbf{f}_w\}$ and $\{\mathbf{g}_w\}$: $\sum_{v \in \mathcal{S}_v} \mathbf{f}_{w'}(v) \mathbf{g}_w(v) = 0$ for the large majority of pairs of distinct words w and w' , but, for $v \geq (1, 1, 1, 0, 0, \dots)$, there will always be $w \neq w' \in \mathcal{S}_v$ with $\sum_v \mathbf{f}_{w'}(v) \mathbf{g}_w(v) \neq 0$, essentially because of Example 6.1.13 below. For ease of notation, write the inner product $\langle \mathbf{f}_{w'}, \mathbf{g}_w \rangle$ for $\sum_v \mathbf{f}_{w'}(v) \mathbf{g}_w(v)$.

Theorem 6.1.10. *Let w, w' be words with Lyndon factorisations $w = u_1 \cdots u_k$, $w' = u'_1 \cdots u'_k$ respectively. Then*

$$\langle \mathbf{f}_{w'}, \mathbf{g}_w \rangle = \begin{cases} 0 & \text{if } k \neq k'; \\ \frac{1}{Z(w')} \sum_{\sigma \in \mathfrak{S}_k} \mathbf{f}_{u'_{\sigma(1)}}(u_1) \cdots \mathbf{f}_{u'_{\sigma(k)}}(u_k) = \frac{1}{Z(w')} \mathbf{f}_{w'}(u_1 \dots u_k) & \text{if } k = k'. \end{cases}$$

(Note that $u_1 \dots u_k$ is the shuffle product of the Lyndon factors, not the concatenation, and is therefore not equal to w .) In particular, $\langle \mathbf{f}_w, \mathbf{g}_w \rangle = 1$, and $\langle \mathbf{f}_{w'}, \mathbf{g}_w \rangle$ is non-zero only when there is a permutation $\sigma \in \mathfrak{S}_k$ such that $\deg(u'_{\sigma(i)}) = \deg(u_i)$ for each i , and each u_i is equal to or lexicographically larger than $u'_{\sigma(i)}$.

Example 6.1.11. $\langle \mathbf{f}_{(23113)}, \mathbf{g}_{(13123)} \rangle = 0$: the Lyndon factorisations are $(23 \cdot 113)$, which has degrees $(0, 1, 1)$ and $(2, 0, 1)$; and $(13 \cdot 123)$, which has degrees $(1, 0, 1)$ and $(1, 1, 1)$. These degrees do not agree, so the inner product is 0.

Example 6.1.12. $\langle \mathbf{f}_{(13213)}, \mathbf{g}_{(13123)} \rangle = 0$: the Lyndon factorisations are $(132 \cdot 13)$ and $(13 \cdot 123)$, so $\deg(u'_{\sigma(i)}) = \deg(u_i)$ is true for $i = 1, 2$ if σ is the transposition. But (123) is lexicographically smaller than (132) .

Example 6.1.13. Using the Lyndon factorisations in the previous example, $\langle \mathbf{f}_{(13123)}, \mathbf{g}_{(13213)} \rangle = \frac{1}{1} \mathbf{f}_{(13)}(13) \mathbf{f}_{(123)}(132) = 1 \cdot 1(-1) = -1$.

Proof of Theorem 6.1.10. As usual, write $f_{w'}, g_w$ for the eigenvectors in the free associative algebra and the shuffle algebra respectively corresponding to $\mathbf{f}_{w'}, \mathbf{g}_w$ under Proposition 3.2.1. So

$$f_{w'} = \frac{1}{k'!Z(w')} \sum_{\sigma \in \mathfrak{S}_k} f_{u'_{\sigma(1)}} \cdots f_{u'_{\sigma(k')}},$$

$$g_w = e(u_1) \cdots e(u_k).$$

If $k \neq k'$, so w and w' have different numbers of Lyndon factors, then $\mathbf{f}_{w'}$ and \mathbf{g}_w are eigenfunctions with different eigenvalues, so from pure linear algebra, $\langle \mathbf{f}_{w'}, \mathbf{g}_w \rangle = 0$. (A more detailed explanation is in the penultimate paragraph of the proof of Theorem 5.1.9, at the end of Section 5.1.3.)

Now assume $k = k'$. First, take $k = 1$, so w, w' are both Lyndon. Then

$$\begin{aligned} \langle \mathbf{f}_{w'}, \mathbf{g}_w \rangle &= f_{w'}(e(w)) \\ &= f_{w'} \left(w - \frac{1}{2} m \bar{\Delta} w + \frac{1}{3} m^{[3]} \bar{\Delta}^{[3]} w - \dots \right) \\ &= f_{w'}(w) - \frac{1}{2} (\Delta f_{w'}) (\bar{\Delta} w) + \frac{1}{3} (\Delta^{[3]} f_{w'}) (\bar{\Delta}^{[3]} w) - \dots \\ &= f_{w'}(w). \end{aligned}$$

The third equality uses that comultiplication in the free associative algebra is dual to multiplication in \mathcal{S} , and the last step is because $f_{w'}$ is primitive, by construction.

For the case $k > 1$, the argument is similar to the third paragraph of the proof of Theorem 5.1.9.

$$\begin{aligned} \langle \mathbf{f}_{w'}, \mathbf{g}_w \rangle &= \left(\frac{1}{k!Z(w')} \sum_{\sigma \in \mathfrak{S}_k} f_{u'_{\sigma(1)}} \cdots f_{u'_{\sigma(k)}} \right) (e(u_1) \cdots e(u_k)) \\ &= \left(\frac{1}{k!Z(w')} \sum_{\sigma \in \mathfrak{S}_k} \Delta^{[k]} f_{u'_{\sigma(1)}} \cdots \Delta^{[k]} f_{u'_{\sigma(k)}} \right) (e(u_1) \otimes \cdots \otimes e(u_k)), \end{aligned}$$

as comultiplication in the free associative algebra is dual to multiplication in \mathcal{S} . Each $f_{u'_{\sigma(r)}}$ is primitive, so the terms of $\Delta^{[k]} f_{u'_{\sigma(r)}}$ are all possible ways to have $f_{u'_{\sigma(r)}}$ in one tensor-factor and 1 in all other tensor-factors. Hence the right hand side above simplifies to

$$\begin{aligned} & \frac{1}{k!Z(w')} \sum_{\sigma \in \mathfrak{S}_k} \sum_{\tau \in \mathfrak{S}_k} f_{u'_{\tau\sigma(1)}}(e(u_1)) \dots f_{u'_{\tau\sigma(k)}}(e(u_k)) \\ &= \frac{1}{Z(w')} \sum_{\sigma \in \mathfrak{S}_k} f_{u'_{\sigma(1)}}(e(u_1)) \dots f_{u'_{\sigma(k)}}(e(u_k)) \\ &= \frac{1}{Z(w')} \sum_{\sigma \in \mathfrak{S}_k} f_{u'_{\sigma(1)}}(u_1) \dots f_{u'_{\sigma(k)}}(u_k), \end{aligned}$$

using the $k = 1$ case in the last step. Running this calculation with u_i instead of $e(u_i)$ reaches the same conclusion, so $\langle \mathbf{f}_{w'}, \mathbf{g}_w \rangle$ must also equal $f_{w'}(u_1 \dots u_k)$, which is $\mathbf{f}_{w'}(u_1 \dots u_k)$ by definition (because riffle-shuffling does not require any basis rescaling via η).

Clearly $\mathbf{f}_{u'}(u)$ is non-zero only if u and u' have the same constituent letters, i.e. $\deg(u') = \deg(u)$. Also, [Reu93, Th. 5.1] claims that, for Lyndon u' and any word u , the right eigenfunction value $\mathbf{f}_{u'}(u)$ is non-zero only if u is lexicographically larger than or equal to u' .

If $w = w'$, then $u'_i = u_i$ for each i , so the condition that each u_i is equal to or lexicographically larger than $u'_{\sigma(i)}$ can only hold when $u_i = u'_{\sigma(i)}$ for all i . The set of $\sigma \in \mathfrak{S}_k$ which achieves this is precisely the stabiliser in \mathfrak{S}_k permuting the u_i . So

$$\langle \mathbf{f}_w, \mathbf{g}_w \rangle = f_{u_1}(u_1) \dots f_{u_k}(u_k),$$

and [Reu93, Th. 5.1] states that $f_u(u) = 1$ for all Lyndon words u . □

6.2 Descent Sets under Riffle-Shuffling

This section applies the Hopf-power Markov chain machinery to the algebra $QSym$ of quasisymmetric functions (Example 4.1.6) to refine a result of Diaconis and Fulman on the Markov chain tracking the number of descents under riffle-shuffling of a distinct deck of cards. (Recall that a descent is a high value card directly on top of a low value card.) The

result in question is the following interpretations of the right and left eigenfunctions \mathbf{f}_i and \mathbf{g}_i ($0 \leq i \leq n - 1$):

- [DF12, Th. 2.1] $\mathbf{f}_i(j)$ is the coefficient of any permutation with j descents in the i th Eulerian idempotent;
- [DF12, Cor. 3.2] $\mathbf{g}_i(j)$ is the value of the j th Foulkes character of the symmetric group on any permutation with i cycles.

[NT12] recovers these connections using the algebra \mathbf{Sym} of noncommutative symmetric functions, which is dual to $QSym$.

The first result of the present refinement is the existence of an “intermediate” chain between riffle-shuffling and the number of descents, namely the position of descents. (This also follows from the descent set being a “shuffle-compatible statistic”, which [Ges10] attributes to Stanley.) Theorem 6.2.2 identifies this chain as the Hopf-power Markov chain on the basis of fundamental quasisymmetric functions $\{F_I\}$. For a deck of n cards, the states of this descent-set chain naturally correspond to subsets of $n - 1$, though it will be more convenient here to instead associate them to compositions of n , recording the lengths between each pair of descents. A more detailed explanation is in Section 6.2.1. The right and left eigenfunctions for this chain, coming from Theorem 2.5.1.B' and 2.5.1.A' respectively, are also labelled by compositions. The subset of eigenfunctions with interpretations akin to the Diaconis-Fulman result correspond to non-decreasing compositions I , which may be viewed as partitions:

- (Theorem 6.2.3) $\mathbf{f}_I(J)$ is the coefficient of any permutation with descent set J in the Garsia-Reutenauer idempotent (of the descent algebra) corresponding to I ;
- (Theorem 6.2.10) $\mathbf{g}_I(J)$ is the value of the ribbon character (of the symmetric group) corresponding to J on any permutation of cycle type I .

Instructions for calculating these eigenfunctions are in Sections 6.2.4 and 6.2.6 respectively; the computations are entirely combinatorial so they only require the notation in Section 6.2.1 below, and are independent of all other sections. The eigenfunctions for general compositions are considerably more unwieldy.

The calculation and interpretation of eigenfunctions are but a small piece in the Diaconis-Fulman collaboration concerning the number of descents under riffle-shuffling. The first of their series of papers on the topic proves [DF09b, Th. 3.3, 3.4] that $\frac{1}{2} \log n$ steps are necessary and sufficient to randomise the number of descents. As an aside, they show that $\log n$ steps are sufficient to randomise the positions of descents, hence the descent-set Markov chain has a mixing time between $\frac{1}{2} \log n$ and $\log n$. Their second paper [DF09a] gives a neat combinatorial explanation that this number-of-descents Markov chain is the same as the carries observed while adding a list of numbers, a chain previously studied by [Hol97]. [NS14] finds a carries process which equates to the number of descents under *generalised riffle-shuffles*. Here the cards can have one of p colours, and the colours change during shuffling depending on which pile the cards fall into when the deck is cut. The notion of descent is modified to take into account the colours of the cards. The left eigenfunctions of the Markov chain on the number of descents correspond to a generalisation of Foulkes characters in [Mil14]; these are characters of wreath products $\mathbb{Z}/p\mathbb{Z} \wr \mathfrak{S}_n$. An interesting question for the future is whether the descent set of generalised riffle-shuffles also forms a Markov chain, with some refinement of these generalised Foulkes characters describing some of the left eigenfunctions.

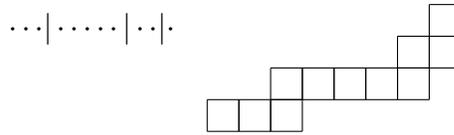
Returning to the present, the rest of the chapter is organised as follows: Section 6.2.1 establishes the necessary notation. Section 6.2.2 covers background on the algebra $QSym$ of quasisymmetric functions and its dual \mathbf{Sym} , the noncommutative symmetric functions, necessary for the proofs and for computing the “messy” eigenfunctions. Section 6.2.3 shows that the descent set is indeed a Markov statistic for riffle-shuffling, by creating a Hopf morphism $\mathcal{S} \rightarrow QSym$ and appealing to the projection theory of Hopf-power Markov chains (Section 4.7). Sections 6.2.4 and 6.2.6 detail the right and left eigenfunctions corresponding to partitions, while Sections 6.2.5 and 6.2.7 contain the full eigenbasis and the proofs of the relationships to ribbon characters and Garsia-Reutenauer idempotents. Section 6.2.8 addresses a partial duality between the two eigenbases, recovering a weak version of a result of Stanley on the probability that a deck in ascending order acquires a particular descent composition after m shuffles. Section 6.2.9 is an appendix containing the transition matrix and full eigenbases for the case $n = 4$. The main results of this section previously appeared in the extended abstract [Pan13].

6.2.1 Notation regarding compositions and descents

For easy reference, this section collects all notation relevant to the rest of this chapter.

A *composition* I is a list of positive integers $(i_1, i_2, \dots, i_{l(I)})$. Each i_j is a *part* of I . The sum $i_1 + \dots + i_{l(I)}$ is denoted $|I|$, and $l(I)$ is the number of parts in I . So $|(3, 5, 2, 1)| = 11$, $l((3, 5, 2, 1)) = 4$. Forgetting the ordering of the parts of I gives a multiset $\lambda(I) := \{i_1, \dots, i_{l(I)}\}$. Clearly $\lambda(I) = \lambda(I')$ if and only if I' has the same parts as I , but in a different order. I is a *partition* if its parts are non-increasing, that is, $i_1 \geq i_2 \geq \dots \geq i_{l(I)}$.

The following two pictorial descriptions of compositions will come in useful for calculating right and left eigenfunctions respectively. Firstly, the *diagram* of I is a string of $|I|$ dots with a division after the first i_1 dots, another division after the next i_2 dots, etc. Next, the *ribbon shape* of I is a skew-shape (in the sense of tableaux) with i_1 boxes in the bottommost row, i_2 boxes in the second-to-bottom row, etc, so that the rightmost square of each row is directly below the leftmost square of the row above. Hence this skew-shape contains no 2-by-2 square. The diagram and ribbon shape of $(3, 5, 2, 1)$ are shown below.



There is a natural partial ordering on the collection of compositions $\{I \mid |I| = n\}$ - define $J \geq I$ if J is a *refinement* of I . Then I is a *coarsening* of J .

Given compositions I, J with $|I| = |J|$, [Gel+95, Sec. 4.8] defines the *decomposition of J relative to I* as the $l(I)$ -tuple of compositions $(J_1^I, \dots, J_{l(I)}^I)$ such that $|J_r^I| = i_r$ and each $l(J_r^I)$ is minimal such that the concatenation $J_1^I \dots J_{l(I)}^I$ refines J . Pictorially, the diagrams of $J_1^I, \dots, J_{l(I)}^I$ are obtained by “splitting” the diagram of J at the points specified by the divisions in the diagram of I . For example, if $I = (4, 4, 3)$ and $J = (3, 5, 2, 1)$, then $J_1^I = (3, 1)$, $J_2^I = (4)$, $J_3^I = (2, 1)$.

It will be useful to identify the composition I with the word $i_1 \dots i_{l(I)}$; then it makes sense to talk of Lyndon compositions, factorisations into Lyndon compositions, and the other concepts from Section 2.4. Write $I = I_{(1)} \dots I_{(k)}$ for the Lyndon factorisation of I ; so, if $I = (3, 5, 2, 1)$, then $I_{(1)} = (3, 5)$, $I_{(2)} = (2)$, $I_{(3)} = (1)$. $k(I)$ will always denote the number of Lyndon factors in I . A composition I is a partition precisely when all its Lyndon

factors are singletons - this is what simplifies their corresponding eigenfunctions. $\lambda(I)$ is the multigrading of I as a word, and $l(I)$ is the integer grading, though neither agrees with the grading $|I|$ on $QSym$ so this view may be more confusing than helpful.

Finally, the *descent set* of a word $w = w_1 \dots w_n$ is defined to be $D(w) = \{j \in \{1, 2, \dots, |w| - 1\} \mid w_j > w_{j+1}\}$. As noted earlier, it is more convenient here to consider the associated composition of $D(w)$. Hence a word w has *descent composition* $\text{Des}(w) = I$ if i_j is the number of letters between the $j - 1$ th and j th descent, i.e. if $w_{i_1 + \dots + i_j} > w_{i_1 + \dots + i_{j+1}}$ for all j , and $w_r \leq w_{r+1}$ for all $r \neq i_1 + \dots + i_j$. For example, $D(4261) = \{1, 3\}$ and $\text{Des}(4261) = (1, 2, 1)$. Note that no information is lost in passing from $D(w)$ to $\text{Des}(w)$, as the divisions in the diagram of $\text{Des}(w)$ indicate the positions of descents in w .

6.2.2 Quasisymmetric Functions and Noncommutative Symmetric Functions

Recall from Example 4.1.6 the algebra $QSym$ of quasisymmetric functions: it is a subalgebra of the algebra of power series in infinitely-many commuting variables $\{x_1, x_2, \dots\}$ spanned by the *monomial quasisymmetric functions*

$$M_I = \sum_{j_1 < \dots < j_{l(I)}} x_{j_1}^{i_1} \dots x_{j_{l(I)}}^{i_{l(I)}}.$$

The basis runs over all compositions $I = (i_1, \dots, i_{l(I)})$. This, however, is not the state space basis of the Markov chain of interest; that basis is the *fundamental quasisymmetric functions*

$$F_I = \sum_{J \geq I} M_J$$

where the sum runs over all partitions J refining I . $QSym$ inherits a grading and a commutative algebra structure from the algebra of power series, so $\deg(M_I) = \deg(F_I) = |I|$. [MR95] extends this to a Hopf algebra structure using the ‘‘alphabet doubling’’ coproduct: take two sets of variables $X = \{x_1, x_2, \dots\}$, $Y = \{y_1, y_2, \dots\}$ that all commute, and totally-order $X \cup Y$ by setting $x_i < x_j$ if $i < j$, $y_i < y_j$ if $i < j$, and $x_i < y_j$ for all i, j . Then, if $F(x, y)$

denotes the quasisymmetric function F applied to $X \cup Y$, and $F(x, y) = \sum_i G_i(x)H_i(y)$, then $\Delta(F) = \sum_i G_i \otimes H_i$. For example, $\Delta(M_i) = M_i \otimes 1 + 1 \otimes M_i$, and

$$\Delta(M_I) = \sum_{j=0}^{l(I)} M_{(i_1, i_2, \dots, i_j)} \otimes M_{(i_{j+1}, \dots, i_{l(I)})}.$$

The graded dual Hopf algebra of $QSym$ is **Sym**, the algebra of noncommutative symmetric functions. (Some authors call this $NSym$. Beware that there are several noncommutative analogues of the symmetric functions, such as $NCSym$, and these are not all isomorphic.) A comprehensive reference on this algebra is [Gel+95] and its many sequels. The notation here follows this tome, except that all indices of basis elements will be superscripts, to distinguish from elements of $QSym$ which use subscripts. The duality of **Sym** and $QSym$ was first established in [MR95, Th. 2.1].

[NPT13, Sec. 2] frames **Sym** under the polynomial realisation viewpoint previously discussed in Section 4.1.3. The construction starts with the power series algebra in infinitely-many noncommuting variables. For simplicity, write the word $(i_1 \dots i_l)$ for the monomial $x_{i_1} \dots x_{i_l}$; so, for example, (12231) stands for $x_1 x_2^2 x_3 x_1$. As an algebra, **Sym** is a subalgebra of this power series algebra generated by

$$S^{(n)} := \sum_{w: \text{Des}(w)=(n)} w,$$

the sum over all words of length n with no descent. For example,

$$S^{(1)} = (1) + (2) + (3) + \dots = x_1 + x_2 + x_3 + \dots;$$

$$S^{(2)} = (11) + (12) + (13) + \dots + (22) + (23) + \dots$$

The algebra **Sym** inherits a concatenation product from the full power series algebra, and the alphabet doubling trick endows **Sym** with the coproduct

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

For any composition $I = (i_1, \dots, i_l)$, define the *complete noncommutative symmetric functions*

$$S^I := S^{(i_1)} \dots S^{(i_l)} = \sum_{w: \text{Des}(w) \leq I} w.$$

A moment's thought will convince that $\{S^I\}$ is linearly independent. So $\{S^I\}$ is a free basis in the sense of Theorem 2.5.1.B'; it is analogous to the $\{h_\lambda\}$ basis of the symmetric functions. Indeed, the abelianisation map from the noncommutative power series ring to $\mathbb{R}[[x_1, x_2, \dots]]$ (i.e. allowing the variables x_i to commute) sends each $S^{(n)}$ to $h_{(n)}$, and consequently sends S^I to $h_{\lambda(I)}$. The basis $\{S^I\}$ is dual to the monomial quasisymmetric functions $\{M_I\}$.

The dual basis to the fundamental quasisymmetric functions $\{F_I\}$ is the *ribbon noncommutative symmetric functions* $\{R^I\}$:

$$R^I := \sum_{w: \text{Des}(w) = I} w.$$

One more basis will be useful in the ensuing analysis. [Gel+95, Eq. 26] defines $\frac{\Phi^{(n)}}{n}$ to be the coefficient of t^n in the formal power series $\log(1 + \sum_{i>0} S^{(i)} t^i)$. Equivalently,

$$\Phi^{(n)} := ne(S^{(n)}) = n \sum_I \frac{(-1)^{l(I)}}{l(I)} \sum_{w: \text{Des}(w) \leq I} w,$$

where e is the Eulerian idempotent map. This is a noncommutative analogue of the relationship $e(h_{(n)}) = \frac{1}{n} p_{(n)}$, established in Section 5.2.3. Noncommutativity of the underlying variables means that there is sadly no formula for the $\Phi^{(n)}$ quite as convenient as $p_{(n)} = x_1^n + x_2^n + \dots$. Then the *power sum noncommutative symmetric functions of the second kind* are

$$\Phi^I := \Phi^{(i_1)} \dots \Phi^{(i_l)}.$$

[Gel+95] details explicitly the change-of-basis matrices of these and other bases in **Sym**; these will be extremely useful in Sections 6.2.4 and 6.2.5 for determining the right eigenfunctions of the associated Markov chain.

6.2.3 The Hopf-power Markov chain on $QSym$

Solely from the above definitions of the fundamental quasisymmetric functions, the product and the coproduct, it is unclear what process the Hopf-power Markov chain on $\{F_I\}$ might represent. The key to solving this mystery is the following Hopf morphism, which sends any word with distinct letters to the fundamental quasisymmetric function indexed by its descent set.

Theorem 6.2.1. *There is a morphism of Hopf algebras $\theta : \mathcal{S} \rightarrow QSym$ such that, if w is a word with distinct letters, then $\theta(w) = F_{\text{Des}(w)}$.*

The proof is at the end of this section. Applying the Projection Theorem for Hopf-power Markov Chains (Theorem 4.7.1) to the map θ shows that:

Theorem 6.2.2. *The Hopf-power Markov chain on the fundamental quasisymmetric functions $\{F_I\}$ tracks the descent set under riffle-shuffling of a distinct deck of cards. In particular, the descent set is a Markovian statistic of riffle-shuffling of a distinct deck of cards.*

In order to keep the algebra in the background, the subsequent sections will refer to this chain simply as the Hopf-power Markov chain on compositions, and the states of the chain will be labelled by compositions I instead of the corresponding quasisymmetric functions F_I . This is similar to the notation of Section 5.2.

Proof of Theorem 6.2.2. Follow the notation of the Projection Theorem and write \mathcal{B} for the word basis of the shuffle algebra, and $\bar{\mathcal{B}}$ for the fundamental quasisymmetric functions. Then, for any \mathbf{v} where each v_i is 0 or 1, $\mathcal{B}_{\mathbf{v}}$ consists of words with distinct letters, so the map θ from Theorem 6.2.1 satisfies $\theta(\mathcal{B}_{\mathbf{v}}) = \bar{\mathcal{B}}_{|\mathbf{v}|}$. Moreover, θ sends all single letters to $F_1 = \bar{\mathcal{B}}_1$. Hence the conditions of the Projection Theorem hold, and its application proves the result. \square

Proof of Theorem 6.2.1. By [ABS06, Th. 4.1], $QSym$ is the terminal object in the category of combinatorial Hopf algebras equipped with a multiplicative character. So, to define a Hopf morphism to $QSym$, it suffices to define the corresponding character ζ on the domain. By [Reu93, Th. 6.1.i], the shuffle algebra is freely generated by Lyndon words, so any

choice of the values of ζ on Lyndon words extends uniquely to a well-defined character on the shuffle algebra. For Lyndon u , set

$$\zeta(u) = \begin{cases} 1 & \text{if } u \text{ has all letters distinct and has no descents;} \\ 0 & \text{otherwise.} \end{cases} \quad (6.1)$$

I claim that, consequently, (6.1) holds for all words with distinct letters, even if they are not Lyndon. Assuming this for now, [ABS06, Th. 4.1] defines

$$\theta(w) = \sum_{I:|I|=|w|} \zeta(w_1 \cdots w_{i_1}) \zeta(w_{i_1+1} \cdots w_{i_1+i_2}) \cdots \zeta(w_{i_{l(I)-1}+1} \cdots w_n) M_I.$$

If w has distinct letters, then every consecutive subword $w_{i_1+\cdots+i_j+1} \cdots w_{i_1+\cdots+i_{j+1}}$ of w also has distinct letters, so

$$\zeta(w_1 \cdots w_{i_1}) \cdots \zeta(w_{i_{l(I)-1}+1} \cdots w_n) = \begin{cases} 1 & \text{if } \text{Des}(w) \leq I; \\ 0 & \text{otherwise.} \end{cases}$$

Hence $\theta(w) = \sum_{\text{Des}(w) \leq I} M_I = F_{\text{Des}(w)}$.

Now return to proving the claim that (6.1) holds whenever w has distinct letters. Proceed by induction on w , with respect to lexicographic order. [Reu93, Th. 6.1.ii], applied to a word w with distinct letters, states that: if w has Lyndon factorisation $w = u_1 \cdots u_k$, then the product of these factors in the shuffle algebra satisfies

$$u_1 \cdots u_k = w + \sum_{v < w} \alpha_v v$$

where α_v is 0 or 1. The character ζ is multiplicative, so

$$\zeta(u_1) \cdots \zeta(u_k) = \zeta(w) + \sum_{v < w} \alpha_v \zeta(v). \quad (6.2)$$

If w is Lyndon, then the claim is true by definition; this includes the base case for the induction. Otherwise, $k > 1$ and there are two possibilities:

- None of the u_i s have descents. Then the left hand side of (6.2) is 1. Since the u_i s together have all letters distinct, the only way to shuffle them together and obtain a word with no descents is to arrange the constituent letters in increasing order. This word is Lyndon, so it is not w , and, by inductive hypothesis, it is the only v in the sum with $\zeta(v) = 1$. So $\zeta(w)$ must be 0.
- Some Lyndon factor u_i has at least one descent. Then $\zeta(u_i) = 0$, so the left hand side of (6.2) is 0. Also, no shuffle of u_1, \dots, u_k has its letters in increasing order. Therefore, by inductive hypothesis, all v in the sum on the right hand side have $\zeta(v) = 0$. Hence $\zeta(w) = 0$ also.

□

Remarks.

1. From the proof, one sees that the conclusion $\theta(w) = F_{\text{Des}(w)}$ for w with distinct letters relies only on the value of ζ on Lyndon words with distinct letters. The proof took $\zeta(u) = 0$ for all Lyndon u with repeated letters, but any other value would also work. Alas, no definition of ζ will ensure $\theta(w) = F_{\text{Des}(w)}$ for all w :

$$\theta((11)) = \frac{1}{2}\theta((1)(1)) = \frac{1}{2}\theta(1)\theta(1) = \frac{1}{2}M_1^2 \neq F_2.$$

2. The map θ is inspired by, but ultimately mathematically unrelated to, the polynomial realisation of **Sym**. Dualising the algebra embedding $\mathbf{Sym} \subseteq \mathcal{S}^*$ gives a coalgebra map $\theta' : \mathcal{S} \rightarrow QSym$, with $\theta'(w) = F_{\text{Des}(w)}$ for all w , but this is not a Hopf algebra map. Mysteriously and miraculously, if all letters occurring in v and w together are distinct, then $\theta'(vw) = \theta'(v)\theta'(w)$, and doctoring the image of θ' on words with repeated letters can make this true for all v, w . I have yet to find another combinatorial Hopf algebra \mathcal{H} where the coalgebra map $\theta' : \mathcal{S} \rightarrow \mathcal{H}$ dual to a polynomial realisation $\mathcal{H}^* \subseteq \mathcal{S}^*$ satisfies $\theta'(vw) = \theta'(v)\theta'(w)$ for a large class of $v, w \in \mathcal{S}$.

6.2.4 Right Eigenfunctions Corresponding to Partitions

Throughout this subsection, let I be a partition. That is, $i_1 \geq i_2 \geq \dots \geq i_{l(I)}$. Set $n = |I|$.

All right eigenfunctions are essentially built from the function

$$\mathbf{f}(J) := \frac{1}{|J|} \frac{(-1)^{l(J)-1}}{\binom{|J|-1}{l(J)-1}}.$$

Note that $\mathbf{f}(J)$ depends only on $|J|$ and $l(J) - 1$, which are respectively the number of dots and the number of divisions in the diagram of J .

Theorem 6.2.3 below gives the formula for \mathbf{f}_I , the right eigenfunction corresponding to the partition I , in terms of \mathbf{f} . The proof is at the end of the following section, after establishing the full eigenbasis. The scaling of these eigenfunctions differs from that in Theorem 2.5.1.B' in order to connect them to the idempotents E_I defined by [GR89, Sec. 3], of the descent algebra. (The descent algebra is the subalgebra of the group algebra $\mathbb{Z}\mathfrak{S}_n$ spanned by sums of permutations with the same descent sets. Hence each E_I is a linear combination of permutations, where permutations with the same descent set have the same coefficient.)

Theorem 6.2.3. *With \mathbf{f} as defined above, the function*

$$\begin{aligned} \mathbf{f}_I(J) &:= \frac{1}{l(I)!} \sum_{I': \lambda(I') = \lambda(I)} \prod_{r=1}^{l(I')} \mathbf{f}(J_r) \\ &= \frac{1}{l(I)! i_1 \cdots i_{l(I)}} \sum_{I': \lambda(I') = \lambda(I)} \prod_{r=1}^{l(I')} \frac{(-1)^{l(J_r)-1}}{\binom{|J_r|-1}{l(J_r)-1}}, \end{aligned}$$

is a right eigenfunction of eigenvalue $a^{l(I)-n}$ of the a th Hopf-power Markov chain on compositions. (Here $(J_1, \dots, J_{l(I)})$ is the decomposition of J with respect to I' .) The numbers $\mathbf{f}_I(J)$ appear as coefficients in the Garsia-Reutenauer idempotent E_I :

$$E_I = \sum_{\sigma \in \mathfrak{S}_n} \mathbf{f}_I(\text{Des}(\sigma)) \sigma.$$

(Here, $\text{Des}(\sigma)$ is the descent composition of the word whose i th letter is $\sigma(i)$ - that is, the word given by σ in one-line notation.)

Remark. The sum of E_I across all I with i parts is the i th Eulerian idempotent, in which the coefficients of a permutation σ depend only on its number of descents. Hence $\sum_{l(I)=i} \mathbf{f}_I$ is a right eigenfunction of eigenvalue a^{i-n} whose value depends only on $l(J)$, the number of parts. The n such eigenfunctions descend to the right eigenbasis of [DF12, Th. 2.1] for the number of descents under riffle-shuffling.

Here is a more transparent description of how to calculate $\mathbf{f}_I(J)$:

1. Split the diagram of J into pieces whose numbers of dots are the parts of I , with multiplicity.
2. Calculate \mathbf{f} on each piece of J by counting the number of dots and divisions and multiply these \mathbf{f} values together.
3. Sum this number across all decompositions of J in step 1, then divide by $l(I)!$.

Note that \mathbf{f} itself, when restricted to compositions of a fixed size, is a right eigenfunction, that corresponding to the partition with a single part. Its eigenvalue is a^{1-n} , the smallest possible.

Example 6.2.4. Here’s how to apply the algorithm above to calculate $\mathbf{f}_{(4,4,3)}((3,5,2,1))$. There are three relevant decompositions of $(3,5,2,1)$:

$$\dots | \cdot \dots \cdot | \cdot \dots | \cdot \dots \cdot | \cdot \cdot | \cdot \dots \dots \cdot | \cdot \cdot | \cdot$$

so

$$\mathbf{f}_{(4,4,3)}((3,5,2,1)) = \frac{1}{3!} \left(\frac{-1}{4} \frac{1}{\binom{3}{1}} \frac{-1}{4} \frac{1}{\binom{2}{1}} + \frac{-1}{4} \frac{1}{\binom{3}{1}} \frac{1}{3} \frac{1}{4} \frac{1}{\binom{3}{2}} + \frac{1}{3} \frac{1}{4} \frac{1}{4} \frac{1}{\binom{3}{2}} \right) = \frac{7}{5184}.$$

Note that $\mathbf{f}((1)) = 1$, so pieces of size one do not contribute to step 2 of the algorithm above. This observation simplifies the calculation of $\mathbf{f}_{(i_1,1,1,\dots,1)}(J)$, in a similar way to the \mathbf{f}_u of Section 6.1.1: $\mathbf{f}_{(i_1,1,1,\dots,1)}(J)$ is the sum of \mathbf{f} evaluated on the “subcompositions” of J formed by i_1 consecutive dots. In other words, $\mathbf{f}_{(i_1,1,1,\dots,1)}$ is the weighted enumeration of “patterns” of length i_1 , where pattern J has weight $\frac{f(J)}{(n-i_1+1)!}$. In the similar notational abuse as Section 6.1.1, call this eigenfunction $\mathbf{f}_{(i)}$. (The parallels end here: products of $\mathbf{f}_{(i)}$ are not eigenfunctions, that fact is particular to riffle-shuffling.)

Each right eigenfunction \mathbf{f}_I has a lift to the riffle-shuffle chain: that is, the function $\tilde{\mathbf{f}}_I(w) := \mathbf{f}_I(\text{Des}(w))$ for words w with distinct letters is a right eigenfunction for riffle-shuffling. (This is a general fact about projections of Markov chains and is unrelated to Hopf algebras, see [LPW09, Lem. 12.8.ii]). As divisions correspond to descents, $\tilde{\mathbf{f}}_{(i)}$ is a weighted enumeration of “up-down-patterns” of length i .

Example 6.2.5. Take $i = 2$, then each subcomposition is either (2) or $(1, 1)$. Since $\mathbf{f}((2)) = \frac{1}{2}$ and $\mathbf{f}((1, 1)) = -\frac{1}{2}$, the right eigenfunction $\mathbf{f}_{(2)}$ counts a non-division with weight $\frac{1}{2(n-1)!}$ and a division with weight $\frac{-1}{2(n-1)!}$. Since the number of non-divisions and the number of divisions sum to $n - 1$,

$$\mathbf{f}_{(2)}(J) = \frac{1}{(n-1)!} \left(\frac{|J| - 1}{2} - (l(J) - 1) \right).$$

It will follow from the full eigenbasis description of Theorem 6.2.7 that this is the unique right eigenvector of eigenvalue $\frac{1}{a}$, up to scaling. Its lift $\tilde{\mathbf{f}}_{(2)}$ to the riffle-shuffling chain is (a multiple of) the “normalised number of descents” eigenvector of Proposition 6.1.3: $\tilde{\mathbf{f}}_{(2)} = \frac{1}{2(n-1)!} \mathbf{f}_{\setminus}$.

Example 6.2.6. Take $i = 3$. Then $\mathbf{f}((3)) = \mathbf{f}((1, 1, 1)) = \frac{1}{3}$, $\mathbf{f}((2, 1)) = \mathbf{f}((1, 2)) = -\frac{1}{6}$, so

$$\mathbf{f}_{(3)}(J) = \frac{1}{3(n-2)!} \left(\#(\text{two consecutive non-divisions}) + \#(\text{two consecutive divisions}) - \frac{1}{2} \#(\text{division followed by non-division}) - \frac{1}{2} \#(\text{non-division followed by division}) \right).$$

The associated eigenvalue is $\frac{1}{a^2}$. Its lift $\tilde{\mathbf{f}}_{(3)}$ to the riffle-shuffling chain is

$$\begin{aligned} \tilde{\mathbf{f}}_{(3)}(w) &= \frac{1}{3(n-2)!} \left(\text{aasc}(w) + \text{ddes}(w) - \frac{1}{2} \text{vall}(w) - \frac{1}{2} \text{peak}(w) \right) \\ &= \frac{1}{2(n-2)!} \mathbf{f}_{-} \end{aligned}$$

in the notation of Proposition 6.1.5.

6.2.5 A full Basis of Right Eigenfunctions

When I is not a partition, the calculation of $\mathbf{f}_I(J)$ is very similar to the previous three-step process, except that, in the last step, each summand is weighted by $\mathbf{f}_I^{\mathcal{S}}(I')$, the value on I' of the right eigenfunction $\mathbf{f}_I^{\mathcal{S}}$ of riffle-shuffling.

Theorem 6.2.7. *Given a composition $I = (i_1, \dots, i_l)$ with $k(I)$ Lyndon factors, define the function*

$$\mathbf{f}_I(J) := \frac{1}{i_1 \dots i_{l(I)}} \sum_{I': \lambda(I') = \lambda(I)} \mathbf{f}_I^{\mathcal{S}}(I') \prod_{r=1}^{l(I')} \frac{(-1)^{l(J_r)-1}}{\binom{|J_r|-1}{l(J_r)-1}},$$

where $(J_1, \dots, J_{l(I)})$ is the decomposition of J relative to I' , and $\mathbf{f}_I^{\mathcal{S}}$ is the right eigenfunction of riffle-shuffling corresponding to the word $i_1 \dots i_l$, as explained in Section 6.1.1. Then $\{\mathbf{f}_I \mid |I| = n, I \text{ has } k \text{ Lyndon factors}\}$ is a basis of right a^{k-n} -eigenfunctions for the a th Hopf-power Markov chain on compositions.

The proof is at the end of this section.

Example 6.2.8. Take $I = (1, 2, 1)$ and $J = (3, 1)$. Using the decreasing Lyndon hedgerows technique of Section 6.1.1, one finds that $\mathbf{f}_{(1,2,1)}^{\mathcal{S}}((1, 1, 2)) = \frac{1}{2}$, $\mathbf{f}_{(1,2,1)}^{\mathcal{S}}((2, 1, 1)) = -\frac{1}{2}$, and $\mathbf{f}_{(1,2,1)}^{\mathcal{S}}$ is zero on all other compositions. The decomposition of $(3, 1)$ relative to $(1, 1, 2)$ and $(2, 1, 1)$ are $((1), (1), (1, 1))$ and $((2), (1), (1))$ respectively. Putting all this information into the formula in Theorem 6.2.7 above yields

$$\mathbf{f}_{(1,2,1)}((3, 1)) = \frac{1}{1 \cdot 2 \cdot 1} \left(\frac{1}{2} \cdot 1(-1) - \frac{1}{2} \cdot 1 \cdot 1 \right) = -\frac{1}{2}.$$

The full right eigenbasis for the case $n = 4$, as specified by Theorem 6.2.7, is tabulated in Section 6.2.9.

The following property of the right eigenfunctions will be useful for proving Proposition 6.2.18. It essentially says that, if the starting state is the one-part partition, then only the right eigenfunctions corresponding to partitions are relevant. When interpreting this chain on compositions as the descent-set chain under riffle-shuffling, this scenario corresponds to starting the deck in ascending order.

Proposition 6.2.9. *If I is a partition, then $\mathbf{f}_I((n)) = \frac{1}{Z(I)i_1 \dots i_l}$, the proportion of permutations in \mathfrak{S}_n with cycle type I . For all other I , $\mathbf{f}_I((n)) = 0$.*

Proof. First note that the decomposition of (n) relative to any composition I is $((i_1), \dots, (i_{l(I)}))$, so

$$\mathbf{f}_I((n)) = \frac{1}{i_1 \dots i_{l(I)}} \sum_{I': \lambda(I') = \lambda(I)} \mathbf{f}_I^{\mathcal{S}}(I').$$

Recall from Section 6.1.1 that $k(I)! \mathbf{f}_I^{\mathcal{S}}(I')$ is the signed number of ways to rearrange the decreasing Lyndon hedgerow T_I so the leaves spell I' . So $k(I)! \sum_{I': \lambda(I') = \lambda(I)} \mathbf{f}_I^{\mathcal{S}}(I')$ is the total signed number of rearrangements of T_I . If I is a partition, then T_I consists only of singletons, so the rearrangements of T_I are exactly the orbit of $\mathfrak{S}_{k(I)}$ permuting the Lyndon factors of I , and these all have positive sign. Writing $Z(I)$ for the size of the stabiliser of this \mathfrak{S}_k action, it follows that

$$\mathbf{f}_I((n)) = \frac{1}{i_1 \dots i_{l(I)}} \frac{1}{k(I)!} \frac{k(I)!}{Z(I)} = \frac{1}{Z(I)i_1 \dots i_{l(I)}}.$$

By [Sta97, Prop. 1.3.2], $Z(I)i_1 \dots i_{l(I)}$ is the size of the centraliser in \mathfrak{S}_n of a permutation with cycle type I , so its reciprocal is the proportion of permutations with cycle type I .

If I is not a partition, then I has a Lyndon factor which is not a single part. So T_I has an internal vertex, allowing the following “signed involution” trick: exchanging the branches at this vertex gives a bijection between rearrangements of opposite signs. So $\sum_{I': \lambda(I') = \lambda(I)} \mathbf{f}_I^{\mathcal{S}}(I') = 0$. \square

Proof of Theorem 6.2.7, the full right eigenbasis. By Proposition 3.2.1.R, the right eigenfunctions of the Hopf-power Markov chain on compositions come from the eigenvectors of the Hopf-power map on the dual Hopf algebra \mathbf{Sym} . \mathbf{Sym} is cocommutative and has the complete noncommutative symmetric functions S^I as a word basis, so Theorem 2.5.1.B' applies. Specifically, use the alternate formulation of the eigenvectors in the ensuing Remark 3 involving the right eigenfunctions $\mathbf{f}^{\mathcal{S}}$ of riffle-shuffling, and input the result into Proposition 3.2.1.R. The resulting basis of right eigenfunctions for the descent-set chain is

$$\mathbf{f}_I(J) := \sum_{I': \lambda(I') = \lambda(I)} \mathbf{f}_I^{\mathcal{S}}(I') e(S^{(i_1)}) \dots e(S^{(i_{l(I)})}) \text{ evaluated at } F_J.$$

(Recall that e is the Eulerian idempotent map.) Since the basis of ribbon noncommutative symmetric functions $\{R^J\}$ is the dual basis to the fundamental quasisymmetric functions $\{F_J\}$, the above is equivalent to

$$\mathbf{f}_I(J) = \text{coefficient of } R^J \text{ in } \sum_{I': \lambda(I') = \lambda(I)} \mathbf{f}_I^{\mathcal{S}}(I') e(S^{(i'_1)}) \dots e(S^{(i'_{l(I)})})$$

Now Section 6.2.2 defines $\Phi^{(n)}$ to be $ne(S^{(n)})$, so

$$\begin{aligned} \mathbf{f}_I(J) &= \text{coefficient of } R^J \text{ in } \sum_{I': \lambda(I') = \lambda(I)} \mathbf{f}_I^{\mathcal{S}}(I') \frac{\Phi^{I'}}{i'_1 \dots i'_{l(I)}} \\ &= \text{coefficient of } R^J \text{ in } \frac{1}{i_1 \dots i_{l(I)}} \sum_{I': \lambda(I') = \lambda(I)} \mathbf{f}_I^{\mathcal{S}}(I') \Phi^{I'}, \end{aligned}$$

and [Gel+95, Cor. 4.28] gives the coefficient of R^J in $\Phi^{I'}$ as

$$\prod_{r=1}^{l(I')} \frac{(-1)^{l(J_r)-1}}{\binom{|J_r|-1}{l(J_r)-1}}.$$

□

Proof of Theorem 6.2.3, right eigenfunctions corresponding to partitions. Fix a partition I . The decreasing Lyndon hedgerow T_I consists only of singletons, so, for any I' with $\lambda(I') = \lambda(I)$, there is only one rearrangement of T_I spelling I' , and it has positive sign. So $\mathbf{f}_I^{\mathcal{S}}(I') = \frac{1}{k(I)!} = \frac{1}{l(I)!}$.

[KLT97, Sec. 3.3] then states that $\sum \mathbf{f}_I(J) R_J$ is the image of E_I under their map α from the descent algebra to \mathbf{Sym} sending $\sum_{\sigma: \text{Des}(\sigma)=I} \sigma$ to the ribbon noncommutative symmetric function R_I . As this map is injective, it must be that $E_I = \sum_{\sigma \in \mathfrak{S}_n} \mathbf{f}_I(\text{Des}(\sigma)) \sigma$. □

6.2.6 Left Eigenfunctions Corresponding to Partitions

Throughout this section, let I be a partition with $|I| = n$. The left eigenfunctions \mathbf{g}_I are most concisely defined using some representation theory of the symmetric group \mathfrak{S}_n , although their calculation is completely combinatorial. [CSST10, Sec. 3.5.2] describes a

representation of \mathfrak{S}_n for each skew-shape with n boxes; denote by χ^J the character of such a representation whose skew-shape is the ribbon shape of J .

Theorem 6.2.10. *Let I be a partition. Define $\mathbf{g}_I(J) := \chi^J(I)$, the character of \mathfrak{S}_n associated to the ribbon shape J evaluated at a permutation with cycle type I . Then \mathbf{g}_I is a left eigenfunction of the a th Hopf-power Markov chain on compositions with eigenvalue $a^{l(I)-n}$.*

Remark. Here's how to recover from this the left eigenfunctions of the chain tracking the number of descents. As observed in [Pik13, Th. 1.3.1.3], any left eigenfunction \mathbf{g} of a Markov chain induces a left eigenfunction $\bar{\mathbf{g}}$ on its projection, by summing over the values of \mathbf{g} on its preimage. Here, this construction gives

$$\bar{\mathbf{g}}_I(j) = \sum_{l(J)=j} \chi^J(I),$$

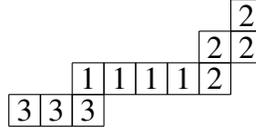
and $\sum_{l(J)=j} \chi^J$ is by definition the Foulkes character. Hence these induced eigenfunctions are precisely those calculated in [DF12, Cor. 3.2].

By Theorem 6.2.10, calculating the eigenfunctions \mathbf{g}_I for partitions I amounts to evaluating characters of the symmetric group, for which the standard method is the Murnaghan-Nakayama rule. This rule simplifies when the character in question corresponds to a ribbon shape; as noted in [CSST10, Rem. 3.5.18], finding $\chi^J(I)$ requires the following:

1. Find all possible ways of filling the ribbon shape of J with i_1 copies of 1, i_2 copies of 2, etc. such that all copies of each integer are in adjacent cells, and all rows and columns are weakly increasing.
2. Let l_r be the number of rows containing r . Sum over all the fillings found in step 1, weighted by $(-1)^{\sum(l_r-1)}$.

Example 6.2.11. Calculating $\mathbf{g}_{(4,4,3)}((3,5,2,1))$ requires filling the ribbon shape of $(3,5,2,1)$ with four copies of 1, four copies of 2 and three copies of 3, subject to the constraints in step 1 above. Observe that the top square cannot be 1, because then the top four squares must all contain 1, and the fifth square from the top must be equal to or smaller

than these. Similarly, the top square cannot be 3, because then the top three squares are all 3s, but the fourth must be equal or larger. Hence 2 must fill the top square, and the only legal way to complete this is



so

$$\mathbf{g}_{(4,4,3)}((3, 5, 2, 1)) = (-1)^{(0+2+0)} = 1.$$

Example 6.2.12. There is only one way to fill any given ribbon shape with n copies of 1, so

$$\mathbf{g}_{(n)}(J) = (-1)^{l(J)}.$$

Next, take $I = (1, 1, \dots, 1)$. Then $\mathbf{g}_{(1,1,\dots,1)}$ has eigenvalue $a^{n-n} = 1$, so $\mathbf{g}_{(1,1,\dots,1)}$ is a multiple of the stationary distribution. (The full left eigenbasis of Theorem 6.2.14 will show that the stationary distribution is unique). Following the algorithm for $\mathbf{g}_I(J)$ above, $\mathbf{g}_{(1,1,\dots,1)}$ is the signed enumeration of fillings of the ribbon shape of J by $1, 2, \dots, n$, each appearing exactly once. Reading the fillings from bottom left to top right gives a word of degree $(1, 1, \dots, 1)$ whose descent composition is exactly J . In conclusion:

Corollary 6.2.13. *The stationary distribution for the Hopf-power Markov chain on compositions is*

$$\pi(J) = \frac{1}{n!} |\{w \mid |w| = n, \deg(w) = (1, 1, \dots, 1), \text{Des}(w) = J\}|.$$

In other words, the stationary probability of J is the proportion of permutations with descent composition J . □

This also follows from the stationary distribution of riffle-shuffling being the uniform distribution.

6.2.7 A full Basis of Left Eigenfunctions

The definition of the full basis of left eigenfunctions involve an obscure basis of $QSym$, which [MR95, Cor. 2.2, Eq. 2.12] defines as the following sum of monomial quasisymmetric functions:

$$P_I := \sum_{J \leq I} (l(I_1)! \dots l(I_{l(J)}!)^{-1} M_J$$

Here the sum runs over all compositions J coarser than I , and $(I_1, \dots, I_{l(J)})$ is the decomposition of I relative to J . (This may be related to the basis of [Haz10].) Also recall that e is the Eulerian idempotent map

$$e(x) = \sum_{r=1}^{\deg x} \frac{(-1)^{r-1}}{r} m^{[r]} \bar{\Delta}^{[r]}(x).$$

Theorem 6.2.14. *Given a composition I with Lyndon factorisation $I = I_{(1)} \dots I_{(k)}$, define the function*

$$\mathbf{g}_I(J) := \text{coefficient of } F_J \text{ in } e(P_{I_{(1)}}) \dots e(P_{I_{(k)}}).$$

Then $\{\mathbf{g}_I \mid |I| = n, I \text{ has } k \text{ Lyndon factors}\}$ is a basis of left a^{k-n} -eigenfunctions for the a th Hopf-power Markov chain on compositions.

Example 6.2.15. Take $I = (1, 2, 1)$, $J = (3, 1)$. Then $I_{(1)} = (1, 2)$, $I_{(2)} = (1)$, so \mathbf{g}_I has eigenvalue a^{-2} , and is described by $e(P_{(1,2)})e(P_{(1)})$. Now

$$\begin{aligned} e(P_{(1,2)}) &= e\left(\frac{1}{1!1!}M_{(1,2)} + \frac{1}{2!}M_{(3)}\right) \\ &= \left(M_{(1,2)} - \frac{1}{2}M_{(1)}M_{(2)}\right) + \frac{1}{2}M_{(3)} \\ &= \frac{1}{2}(M_{(1,2)} - M_{(2,1)}), \end{aligned}$$

and

$$e(P_{(1)}) = e(M_{(1)}) = M_{(1)}.$$

So

$$\begin{aligned}
e(P_{1,2})e(P_1) &= \frac{1}{2}(M_{(1,2)} - M_{(2,1)})M_{(1)} \\
&= \frac{1}{2}(2M_{(1,1,2)} - 2M_{(2,1,1)} + M_{(1,3)} - M_{(3,1)}) \\
&= \frac{1}{2}(F_{(1,1,2)} - F_{(2,1,1)} + F_{(1,3)} - F_{(3,1)}).
\end{aligned}$$

Hence $\mathbf{g}_{(1,2,1)}((3,1)) = -\frac{1}{2}$. The full left eigenbasis for $n = 4$ is documented in Section 6.2.9.

Proof of Theorem 6.2.14, the full left eigenbasis. By Proposition 3.2.1.L and Theorem 2.5.1.A', it suffices to show that there is a (non-graded) algebra isomorphism $\mathcal{S} \rightarrow QSym$ sending the word $(i_1 \dots i_l)$ to $P_{(i_1, \dots, i_l)}$. This is the content of [MR95, Cor. 2.2]. The main idea of the proof goes as follows: the scaled power sum of the second kind $\{\frac{1}{i_1 \dots i_l} \Phi^I\}$ (which they call $\{P_I^*\}$) form a free basis for \mathbf{Sym} , and $\frac{1}{i} \Phi^{(i)}$ is primitive, so $\frac{1}{i_1 \dots i_l} \Phi^I \rightarrow (i_1 \dots i_l)$ is a Hopf-isomorphism from \mathbf{Sym} to the free associative algebra. Dualising this map gives a Hopf-isomorphism $\mathcal{S} \rightarrow QSym$. [MR95, Cor. 2.2] gives a generating function proof that the image of $(i_1 \dots i_l)$ under this map is indeed $P_{(i_1, \dots, i_l)}$ as defined in the theorem. \square

Proof of Theorem 6.2.10, left eigenfunctions corresponding to partitions. If I is a partition, then its Lyndon factors are all singletons, so

$$\mathbf{g}_I(J) = \text{coefficient of } F_J \text{ in } e(P_{(i_1)}) \dots e(P_{(i_l)}).$$

By definition, $P_{(i_r)} = M_{(i_r)}$ and this is primitive, so $\bar{\Delta}^{[a]}M_{(i_r)} = 0$ for all $a \geq 2$, and $e(M_{(i_r)}) = M_{(i_r)}$. So $\mathbf{g}_I(J)$ is the coefficient of F_J in $M_{(i_1)} \dots M_{(i_l)} = p_I$, the power sum symmetric function. As p_I is a symmetric function (as opposed to simply quasisymmetric), [Ges84, Th. 3] determines its coefficient of F_J to be the inner product $\langle p_I, s_J \rangle$, with s_J the skew-Schur function associated to the ribbon shape J . By the Murnaghan-Nakayama rule, $\langle p_I, s_J \rangle = \chi^J(I)$. \square

6.2.8 Duality of Eigenfunctions

The eigenfunctions $\{\mathbf{f}_I\}$ and $\{\mathbf{g}_I\}$ above are “almost dual” in the same sense as the riffle-shuffle eigenfunctions $\{\mathbf{f}_w^{\mathcal{S}}\}$, $\{\mathbf{g}_w^{\mathcal{S}}\}$ of Section 6.1, and this is enough to produce the neat Corollary 6.2.18. As before, write $\langle \mathbf{f}_{I'}, \mathbf{g}_I \rangle$ for $\sum_{J:|J|=n} \mathbf{f}_{I'}(J) \mathbf{g}_I(J)$.

Theorem 6.2.16. *Let I, I' be compositions of n . Then*

$$\langle \mathbf{f}_{I'}, \mathbf{g}_I \rangle = \langle \mathbf{f}_{I'}^{\mathcal{S}}, \mathbf{g}_I^{\mathcal{S}} \rangle.$$

In particular,

- (i) $\langle \mathbf{f}_I, \mathbf{g}_I \rangle = 1$;
- (ii) if I is a partition and I' is any composition different from I , then $\langle \mathbf{f}_{I'}, \mathbf{g}_I \rangle = \langle \mathbf{f}_I, \mathbf{g}_{I'} \rangle = 0$;
- (iii) in fact, $\langle \mathbf{f}_{I'}, \mathbf{g}_I \rangle = 0$ unless there is a permutation $\sigma \in \mathfrak{S}_{k(I)}$ such that $\lambda(I'_{(\sigma(r))}) = \lambda(I_{(r)})$ for each r , and each $I_{(r)}$ is equal to or lexicographically larger than $I'_{(\sigma(r))}$. (Here, $I = I_{(1)} \dots I_{(k)}$ is the Lyndon factorisation of I , and similarly for I' .)

Proof. Theorem 6.1.10, the partial duality of riffle-shuffle eigenfunctions, shows that

$$\langle \mathbf{f}_{w'}^{\mathcal{S}}, \mathbf{g}_w^{\mathcal{S}} \rangle = \frac{1}{Z(w')} \sum_{\sigma \in \mathfrak{S}_k} \mathbf{f}_{u'_{\sigma(1)}}^{\mathcal{S}}(u_1) \dots \mathbf{f}_{u'_{\sigma(k)}}^{\mathcal{S}}(u_k),$$

where $Z(w')$ is the size of the stabiliser of \mathfrak{S}_k acting on the Lyndon factors of w' , and $w = u_1 \dots u_k$ and $w' = u'_1 \dots u'_k$ are Lyndon factorisations. The same argument, with $P_{I_{(r)}}$ in place of u_r and $\mathbf{f}_{I'_{(r)}}$ in place of $\mathbf{f}_{u'_r}^{\mathcal{S}}$, proves

$$\langle \mathbf{f}_{I'}, \mathbf{g}_I \rangle = \frac{1}{Z(I')} \sum_{\sigma \in \mathfrak{S}_k} f_{I'_{(\sigma 1)}}(P_{I_{(1)}}) \dots f_{I'_{(\sigma k)}}(P_{I_{(k)}}).$$

So, for the main statement, it suffices to show $f_{I'}(P_I) = \mathbf{f}_{I'}^{\mathcal{S}}(I)$ for Lyndon compositions I, I' . Recall that

$$f_{I'} = \frac{1}{i'_1 \dots i'_{l(I)}} \sum_{J: \lambda(J) = \lambda(I')} \mathbf{f}_{I'}^{\mathcal{S}}(J) \Phi^J.$$

Now the basis $\{P_I\}$ was designed to be the dual basis to $\{\frac{1}{i_1 \dots i_l} \Phi^I\}$, so, when evaluating $f_{I'}$ at P_I , the only summand that contributes is $J' = I$. So indeed $f_{I'}(P_I) = \mathbf{f}_{I'}^{\mathcal{L}}(I)$.

Parts (i) and (iii) then follow from the analogous statements of Theorem 6.1.10. To deduce Part (ii), note that the Lyndon factors of a partition I are its parts, so the condition $\lambda(I'_{(\sigma(r))}) = \lambda(I_{(r)})$ reduces to $\lambda(I'_{(\sigma(r))}) = (i_{(r)})$. Hence $\langle \mathbf{f}_{I'}, \mathbf{g}_I \rangle$ or $\langle \mathbf{f}_I, \mathbf{g}_{I'} \rangle$ is nonzero only if all Lyndon factors of I' are singletons, which forces I' to also be a partition. Then the condition $i'_{(\sigma(r))} = i_{(r)}$ implies $I' = I$. \square

If I, I' are both partitions, then the interpretations of Theorems 6.2.3 and 6.2.10 translate Part ii of the previous Theorem to:

Corollary 6.2.17. *Let χ^J be the character corresponding to the ribbon shape J , and $E_\lambda(J)$ be the coefficient of any permutation with descent composition J in the Garsia-Reutenauer idempotent E_λ . Then*

$$\sum_J \chi^J(\sigma) E_\lambda(J) = \begin{cases} 1 & \text{if } \sigma \text{ has cycle type } \lambda; \\ 0 & \text{otherwise.} \end{cases}$$

\square

There is another consequence of Theorem 6.2.16.ii that is more relevant to the riffle-shuffle Markov chain:

Corollary 6.2.18. *Let $\{X_m\}$ be the Markov chain of a -handed riffle-shuffling for a deck of n distinct cards, starting in ascending order. Then*

$$P\{\text{Des}(X_m) = J\} = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} a^{m(-n+l(\sigma))} \chi^J(\sigma),$$

where $l(\sigma)$ is the number of cycles of σ .

This also follows from [Sta01, Th. 2.1]. In the present notation, his theorem reads

$$P\{Y_1 = w | Y_0 = (12 \dots n)\} = F_{\text{Des}(w^{-1})}(t_1, t_2, \dots)$$

where $\{Y_m\}$ is the biased riffle-shuffling chain: cut the deck multinomially with parameter (t_1, t_2, \dots) and interleave the piles uniformly as before. The usual a -handed shuffle is the case where $t_1 = t_2 = \dots = t_a = \frac{1}{a}$, $t_{a+1} = t_{a+2} = \dots = 0$. So, letting $[g]_{1/a}$ denote the evaluation of the quasisymmetric function g at $t_1 = \dots = t_a = \frac{1}{a}$, $t_{a+1} = t_{a+2} = \dots = 0$ as in Section 5.1.4,

$$P\{\text{Des}(X_1) = J\} = \left[\sum_{w \in \mathfrak{S}_n: \text{Des}(w) = J} F_{\text{Des}(w^{-1})} \right]_{1/a}.$$

According to [Sta99, Th. 7.19.7], $\sum_{w \in \mathfrak{S}_n: \text{Des}(w) = J} F_{\text{Des}(w^{-1})} = s_J$, the skew-Schur (symmetric) function of ribbon shape J . And checking on the power sums p_λ shows that the linear map of evaluating a symmetric function of degree n at $t_1 = \dots = t_a = \frac{1}{a}$, $t_{a+1} = t_{a+2} = \dots = 0$ is equivalent to taking its inner product with $\frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} a^{-n+l(\sigma)} p_{\lambda(\sigma)}$, where $\lambda(\sigma)$ is the cycle type of σ . So

$$P\{\text{Des}(X_1) = J\} = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} a^{-n+l(\sigma)} \langle p_{\lambda(\sigma)}, s_J \rangle = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} a^{-n+l(\sigma)} \chi^J(\sigma).$$

The case $m > 1$ follows from the power rule, as m iterations of a -handed shuffling is equivalent to one a^m -handed shuffle.

Below is the proof of Corollary 6.2.18 using the diagonalisation of the descent-set chain.

Proof. Write K for the transition matrix of the descent-set chain under riffle-shuffling. Then the left hand side is $K^m((n), J)$, which, by the change of coordinates in Proposition 1.1.4, is equal to

$$\sum_I a^{m(-n+l(I))} \mathbf{f}_I((n)) \check{\mathbf{g}}_I(J),$$

where $\{\check{\mathbf{g}}_I\}$ is the basis of left eigenfunctions dual to the right eigenbasis $\{\mathbf{f}_I\}$. By Proposition 6.2.9, $\mathbf{f}_I((n))$ is 0 unless I is a partition, in which case $\mathbf{f}_I((n))$ is the proportion of permutations in \mathfrak{S}_n with cycle type I . So

$$K^m((n), J) = \sum_{\sigma \in \mathfrak{S}_n} a^{m(-n+l(\sigma))} \frac{1}{n!} \mathbf{g}_{\lambda(\sigma)} \check{\mathbf{g}}_{\lambda(\sigma)}(J),$$

where $\lambda(\sigma)$ denotes the cycle type of σ . For a partition I , Theorem 6.2.16 asserts that $\langle \mathbf{f}_I, \mathbf{g}_I \rangle = 1$ and $\langle \mathbf{f}_{I'}, \mathbf{g}_I \rangle = 0$ for any composition I' different from I - this means $\check{\mathbf{g}}_I = \mathbf{g}_I$ when I is a partition. So

$$K^m((n), J) = \sum_{\sigma \in \mathfrak{S}_n} a^{m(-n+l(\sigma))} \frac{1}{n!} \mathbf{g}_{\lambda(\sigma)}(J),$$

and the conclusion follows from Theorem 6.2.10 relating the left eigenfunctions to the ribbon characters. \square

There is an intermediate statement, stronger than this Corollary and deducible from Stanley's theorem:

$$P\{\text{Des}(Y_1) = J\} = \sum_{w \in \mathfrak{S}_n: \text{Des}(w) = J} F_{\text{Des}(w^{-1})} = \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} \chi^J(\sigma) p_\sigma.$$

I conjecture that this can be proved independently of Stanley's result via an analogous diagonalisation of the descent-set Markov chain under biased riffle-shuffling. (It is not hard to define "biased Hopf-powers" to study deformations of the chains in this thesis, but I will not discuss it here as eigenbasis algorithms for these chains are still in development.)

6.2.9 Transition Matrix and Eigenfunctions when $n = 4$

The Hopf-square Markov chain on compositions of four describes the changes in descent set under the GSR riffle-shuffle of four distinct cards. By explicit calculation of $m\Delta$ for the fundamental quasisymmetric functions of degree four, the transition matrix $K_{2,4}$ is the

following matrix multiplied by $\frac{1}{16}$:

	(4)	(1,3)	(3,1)	(2,2)	(1,2,1)	(2,1,1)	(1,1,2)	(1,1,1,1)
(4)	5	3	3	4	1	0	0	0
(1,3)	1	5	2	3	2	1	2	0
(3,1)	1	2	5	3	2	2	1	0
(2,2)	1	2	2	6	3	1	1	0
(1,2,1)	0	1	1	3	6	2	2	1
(2,1,1)	0	1	2	2	3	5	2	1
(1,1,2)	0	2	1	2	3	2	5	1
(1,1,1,1)	0	0	0	1	4	3	3	5

Its basis of right eigenfunctions, as determined by Theorem 6.2.7, are the columns of the following matrix:

	$E_{(4)}$	$\mathbf{f}_{(1,3)}$	$E_{(3,1)}$	$E_{(2,2)}$	$\mathbf{f}_{(1,2,1)}$	$E_{(2,1,1)}$	$\mathbf{f}_{(1,1,2)}$	$E_{(1,1,1,1)}$
(4)	$\frac{1}{4}$	0	$\frac{1}{3}$	$\frac{1}{8}$	0	$\frac{1}{4}$	0	$\frac{1}{24}$
(1,3)	$-\frac{1}{12}$	$\frac{1}{2}$	$\frac{1}{12}$	$-\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{12}$	-1	$\frac{1}{24}$
(3,1)	$-\frac{1}{12}$	$-\frac{1}{2}$	$\frac{1}{12}$	$-\frac{1}{8}$	$-\frac{1}{2}$	$\frac{1}{12}$	-1	$\frac{1}{24}$
(2,2)	$-\frac{1}{12}$	0	$-\frac{1}{6}$	$\frac{1}{8}$	0	$\frac{1}{12}$	2	$\frac{1}{24}$
(1,2,1)	$\frac{1}{12}$	0	$-\frac{1}{6}$	$\frac{1}{8}$	0	$-\frac{1}{12}$	-2	$\frac{1}{24}$
(2,1,1)	$\frac{1}{12}$	$\frac{1}{2}$	$\frac{1}{12}$	$-\frac{1}{8}$	$-\frac{1}{2}$	$-\frac{1}{12}$	1	$\frac{1}{24}$
(1,1,2)	$\frac{1}{12}$	$-\frac{1}{2}$	$\frac{1}{12}$	$-\frac{1}{8}$	$\frac{1}{2}$	$-\frac{1}{12}$	1	$\frac{1}{24}$
(1,1,1,1)	$-\frac{1}{4}$	0	$\frac{1}{3}$	$\frac{1}{8}$	0	$-\frac{1}{4}$	0	$\frac{1}{24}$

Its basis of left eigenfunctions, as determined by Theorem 6.2.14, are the rows of the following matrix:

	(4)	(1,3)	(3,1)	(2,2)	(1,2,1)	(2,1,1)	(1,1,2)	(1,1,1,1)
$\chi(4)$	1	-1	-1	-1	1	1	1	-1
$\mathbf{g}_{(1,3)}$	0	$\frac{1}{2}$	$-\frac{1}{2}$	0	0	$\frac{1}{2}$	$-\frac{1}{2}$	0
$\chi(3,1)$	1	0	0	-1	-1	0	0	1
$\chi(2,2)$	1	-1	-1	1	1	-1	-1	1
$\mathbf{g}_{(1,2,1)}$	0	$\frac{1}{2}$	$-\frac{1}{2}$	0	0	$-\frac{1}{2}$	$\frac{1}{2}$	0
$\chi(2,1,1)$	1	1	1	1	-1	-1	-1	-1
$\mathbf{g}_{(1,1,2)}$	0	$-\frac{1}{12}$	$-\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{12}$	0
$\chi(1,1,1,1)$	1	3	3	5	5	3	3	1

Bibliography

- [ABS06] Marcelo Aguiar, Nantel Bergeron, and Frank Sottile. “Combinatorial Hopf algebras and generalized Dehn–Sommerville relations”. In: *Compos. Math.* 142.1 (2006), pp. 1–30. ISSN: 0010-437X. eprint: 10.1112/S0010437X0500165X.
- [ABT13] Marcelo Aguiar, Nantel Bergeron, and Nathaniel Thieme. “Hopf monoids from class functions on unitriangular matrices”. In: *Algebra Number Theory* 7.7 (2013), pp. 1743–1779. ISSN: 1937-0652. DOI: 10.2140/ant.2013.7.1743. URL: <http://dx.doi.org/10.2140/ant.2013.7.1743>.
- [AD10] Christos A. Athanasiadis and Persi Diaconis. “Functions of random walks on hyperplane arrangements”. In: *Adv. in Appl. Math.* 45.3 (2010), pp. 410–437. ISSN: 0196-8858. DOI: 10.1016/j.aam.2010.02.001. URL: <http://dx.doi.org/10.1016/j.aam.2010.02.001>.
- [AD86] David Aldous and Persi Diaconis. “Shuffling cards and stopping times”. In: *Amer. Math. Monthly* 93.5 (1986), pp. 333–348. ISSN: 0002-9890. DOI: 10.2307/2323590. URL: <http://dx.doi.org/10.2307/2323590>.
- [ADS11] Sami Assaf, Persi Diaconis, and K. Soundararajan. “A rule of thumb for riffle shuffling”. In: *Ann. Appl. Probab.* 21.3 (2011), pp. 843–875. ISSN: 1050-5164. DOI: 10.1214/10-AAP701. URL: <http://dx.doi.org/10.1214/10-AAP701>.
- [ADS12] Sami Assaf, Persi Diaconis, and Kannan Soundararajan. “Riffle shuffles with biased cuts”. In: *24th International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2012)*. Discrete Math. Theor. Comput.

- Sci. Proc., AR. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2012, pp. 445–456.
- [Agu+12] Marcelo Aguiar, Carlos André, Carolina Benedetti, Nantel Bergeron, Zhi Chen, Persi Diaconis, Anders Hendrickson, Samuel Hsiao, I. Martin Isaacs, Andrea Jedwab, Kenneth Johnson, Gizem Karaali, Aaron Lauve, Tung Le, Stephen Lewis, Huilan Li, Kay Magaard, Eric Marberg, Jean-Christophe Novelli, Amy Pang, Franco Saliola, Lenny Tevlin, Jean-Yves Thibon, Nathaniel Thiem, Vidya Venkateswaran, C. Ryan Vinroot, Ning Yan, and Mike Zabrocki. “Supercharacters, symmetric functions in noncommuting variables, and related Hopf algebras”. In: *Adv. Math.* 229.4 (2012), pp. 2310–2337. ISSN: 0001-8708. DOI: 10.1016/j.aim.2011.12.024. URL: <http://dx.doi.org/10.1016/j.aim.2011.12.024>.
- [AL13] Marcelo Aguiar and Aaron Lauve. “Antipode and Convolution Powers of the Identity in Graded Connected Hopf Algebras”. In: *25th International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2013)*. Discrete Math. Theor. Comput. Sci. Proc., AS. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2013, pp. 1083–1094.
- [AM10] Marcelo Aguiar and Swapneel Mahajan. *Monoidal functors, species and Hopf algebras*. Vol. 29. CRM Monograph Series. With forewords by Kenneth Brown and Stephen Chase and André Joyal. Providence, RI: American Mathematical Society, 2010, pp. lii+784. ISBN: 978-0-8218-4776-3.
- [And14] Scott Andrews. “The Hopf monoid on nonnesting supercharacters of pattern groups”. 2014. URL: <http://arxiv.org/abs/1405.5480>.
- [AS05a] Marcelo Aguiar and Frank Sottile. “Cocommutative Hopf algebras of permutations and trees”. In: *J. Algebraic Combin.* 22.4 (2005), pp. 451–470. ISSN: 0925-9899. DOI: 10.1007/s10801-005-4628-y. URL: <http://dx.doi.org/10.1007/s10801-005-4628-y>.
- [AS05b] Marcelo Aguiar and Frank Sottile. “Structure of the Malvenuto-Reutenauer Hopf algebra of permutations”. In: *Adv. Math.* 191.2 (2005), pp. 225–275.

- ISSN: 0001-8708. DOI: 10.1016/j.aim.2004.03.007. URL: <http://dx.doi.org/10.1016/j.aim.2004.03.007>.
- [AS06] Marcelo Aguiar and Frank Sottile. “Structure of the Loday-Ronco Hopf algebra of trees”. In: *J. Algebra* 295.2 (2006), pp. 473–511. ISSN: 0021-8693. DOI: 10.1016/j.jalgebra.2005.06.021. URL: <http://dx.doi.org/10.1016/j.jalgebra.2005.06.021>.
- [BB90] H el ene Barcelo and Nantel Bergeron. “The Orlik-Solomon algebra on the partition lattice and the free Lie algebra”. In: *J. Combin. Theory Ser. A* 55.1 (1990), pp. 80–92. ISSN: 0097-3165. DOI: 10.1016/0097-3165(90)90049-3. URL: [http://dx.doi.org/10.1016/0097-3165\(90\)90049-3](http://dx.doi.org/10.1016/0097-3165(90)90049-3).
- [BD92] Dave Bayer and Persi Diaconis. “Trailing the dovetail shuffle to its lair”. In: *Ann. Appl. Probab.* 2.2 (1992), pp. 294–313. ISSN: 1050-5164.
- [BHR99] Pat Bidigare, Phil Hanlon, and Dan Rockmore. “A combinatorial description of the spectrum for the Tsetlin library and its generalization to hyperplane arrangements”. In: *Duke Math. J.* 99.1 (1999), pp. 135–174. ISSN: 0012-7094. DOI: 10.1215/S0012-7094-99-09906-4. URL: <http://dx.doi.org/10.1215/S0012-7094-99-09906-4>.
- [BL09] Nantel Bergeron and Huilan Li. “Algebraic structures on Grothendieck groups of a tower of algebras”. In: *J. Algebra* 321.8 (2009), pp. 2068–2084. ISSN: 0021-8693. DOI: 10.1016/j.jalgebra.2008.12.005. URL: <http://dx.doi.org/10.1016/j.jalgebra.2008.12.005>.
- [BLL12] Nantel Bergeron, Thomas Lam, and Huilan Li. “Combinatorial Hopf algebras and towers of algebras—dimension, quantization and functorality”. In: *Algebr. Represent. Theory* 15.4 (2012), pp. 675–696. ISSN: 1386-923X. DOI: 10.1007/s10468-010-9258-y. URL: <http://dx.doi.org/10.1007/s10468-010-9258-y>.
- [BLL98] F. Bergeron, G. Labelle, and P. Leroux. *Combinatorial species and tree-like structures*. Vol. 67. Encyclopedia of Mathematics and its Applications.

- Translated from the 1994 French original by Margaret Readdy, With a foreword by Gian-Carlo Rota. Cambridge: Cambridge University Press, 1998, pp. xx+457. ISBN: 0-521-57323-8.
- [But08] J. C. Butcher. *Numerical methods for ordinary differential equations*. Second. Chichester: John Wiley & Sons Ltd., 2008, pp. xx+463. ISBN: 978-0-470-72335-7. DOI: 10.1002/9780470753767. URL: <http://dx.doi.org/10.1002/9780470753767>.
- [But72] J. C. Butcher. “An algebraic theory of integration methods”. In: *Math. Comp.* 26 (1972), pp. 79–106. ISSN: 0025-5718.
- [BV13] Carolina Benedetti Velasquez. *Combinatorial Hopf algebras of supercharacters of lie type D*. Thesis (Ph.D.)—York University (Canada). ProQuest LLC, Ann Arbor, MI, 2013, p. 116. ISBN: 978-0499-00126-9. URL: http://gateway.proquest.com/openurl?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&res_dat=xri:pqm&rft_dat=xri:pqdiss:NS00126.
- [Car07] Pierre Cartier. “A primer of Hopf algebras”. In: *Frontiers in Number Theory, Physics, and Geometry. II*. Berlin: Springer, 2007, pp. 537–615. eprint: 10.1007/978-3-540-30308-4_12.
- [CF69] P. Cartier and D. Foata. *Problèmes combinatoires de commutation et réarrangements*. Lecture Notes in Mathematics, No. 85. Springer-Verlag, Berlin-New York, 1969, pp. iv+88.
- [CK98] Alain Connes and Dirk Kreimer. “Hopf algebras, renormalization and non-commutative geometry”. In: *Comm. Math. Phys.* 199.1 (1998), pp. 203–242. ISSN: 0010-3616. DOI: 10.1007/s002200050499. URL: <http://dx.doi.org/10.1007/s002200050499>.
- [CSST10] Tullio Ceccherini-Silberstein, Fabio Scarabotti, and Filippo Tolli. *Representation theory of the symmetric groups*. Vol. 121. Cambridge Studies in

- Advanced Mathematics. The Okounkov-Vershik approach, character formulas, and partition algebras. Cambridge: Cambridge University Press, 2010, pp. xvi+412. ISBN: 978-0-521-11817-0.
- [Den12] Graham Denham. “Eigenvectors for a random walk on a hyperplane arrangement”. In: *Advances in Applied Mathematics* 48.2 (2012), pp. 312–324. ISSN: 0196-8858. DOI: 10.1016/j.aam.2010.09.009. URL: <http://www.sciencedirect.com/science/article/pii/S0196885811000856>.
- [DF09a] Persi Diaconis and Jason Fulman. “Carries, shuffling, and an amazing matrix”. In: *Amer. Math. Monthly* 116.9 (2009), pp. 788–803. ISSN: 0002-9890. DOI: 10.4169/000298909X474864. URL: <http://dx.doi.org/10.4169/000298909X474864>.
- [DF09b] Persi Diaconis and Jason Fulman. “Carries, shuffling, and symmetric functions”. In: *Adv. in Appl. Math.* 43.2 (2009), pp. 176–196. ISSN: 0196-8858. eprint: 10.1016/j.aam.2009.02.002.
- [DF12] Persi Diaconis and Jason Fulman. “Foulkes characters, Eulerian idempotents, and an amazing matrix”. In: *J. Algebraic Combin.* 36.3 (2012), pp. 425–440. ISSN: 0925-9899. DOI: 10.1007/s10801-012-0343-7. URL: <http://dx.doi.org/10.1007/s10801-012-0343-7>.
- [DHT02] Gérard Duchamp, Florent Hivert, and Jean-Yves Thibon. “Noncommutative symmetric functions. VI. Free quasi-symmetric functions and related algebras”. In: *Internat. J. Algebra Comput.* 12.5 (2002), pp. 671–717. ISSN: 0218-1967. DOI: 10.1142/S0218196702001139. URL: <http://dx.doi.org/10.1142/S0218196702001139>.
- [DI08] Persi Diaconis and I. M. Isaacs. “Supercharacters and superclasses for algebra groups”. In: *Trans. Amer. Math. Soc.* 360.5 (2008), pp. 2359–2392. ISSN: 0002-9947. DOI: 10.1090/S0002-9947-07-04365-6. URL: <http://dx.doi.org/10.1090/S0002-9947-07-04365-6>.

- [Dia88] Persi Diaconis. *Group representations in probability and statistics*. Institute of Mathematical Statistics Lecture Notes—Monograph Series, 11. Hayward, CA: Institute of Mathematical Statistics, 1988, pp. vi+198. ISBN: 0-940600-14-5.
- [DPR14] Persi Diaconis, C. Y. Amy Pang, and Arun Ram. “Hopf algebras and Markov chains: two examples and a theory”. In: *J. Algebraic Combin.* 39.3 (2014), pp. 527–585. ISSN: 0925-9899. DOI: 10.1007/s10801-013-0456-7. URL: <http://dx.doi.org/10.1007/s10801-013-0456-7>.
- [Fis10] Forest Fisher. “CoZinbiel Hopf algebras in combinatorics”. PhD thesis. The George Washington University, 2010, p. 127. ISBN: 978-1124-14567-9.
- [FNT11] Loïc Foissy, Jean-Christophe Novelli, and Jean-Yves Thibon. “Polynomial realizations of some combinatorial Hopf algebras”. 2011. URL: <http://arxiv.org/abs/1009.2067>.
- [Foi02a] L. Foissy. “Les algèbres de Hopf des arbres enracinés décorés. I”. In: *Bull. Sci. Math.* 126.3 (2002), pp. 193–239. ISSN: 0007-4497. DOI: 10.1016/S0007-4497(02)01108-9. URL: [http://dx.doi.org/10.1016/S0007-4497\(02\)01108-9](http://dx.doi.org/10.1016/S0007-4497(02)01108-9).
- [Foi02b] L. Foissy. “Les algèbres de Hopf des arbres enracinés décorés. II”. In: *Bull. Sci. Math.* 126.4 (2002), pp. 249–288. ISSN: 0007-4497. DOI: 10.1016/S0007-4497(02)01113-2. URL: [http://dx.doi.org/10.1016/S0007-4497\(02\)01113-2](http://dx.doi.org/10.1016/S0007-4497(02)01113-2).
- [Foi12] L. Foissy. “Free and cofree Hopf algebras”. In: *J. Pure Appl. Algebra* 216.2 (2012), pp. 480–494. ISSN: 0022-4049. DOI: 10.1016/j.jpaa.2011.07.010. URL: <http://dx.doi.org/10.1016/j.jpaa.2011.07.010>.
- [Foi13] Loïc Foissy. “An introduction to Hopf algebras of trees”. 2013. URL: <http://loic.foissy.free.fr/pageperso/preprint3.pdf>.

- [Ful05] Jason Fulman. “Stein’s method and Plancherel measure of the symmetric group”. In: *Trans. Amer. Math. Soc.* 357.2 (2005), pp. 555–570. ISSN: 0002-9947. DOI: 10.1090/S0002-9947-04-03499-3. URL: <http://dx.doi.org/10.1090/S0002-9947-04-03499-3>.
- [Gan59] F. R. Gantmacher. *The theory of matrices. Vols. 1, 2*. Translated by K. A. Hirsch. Chelsea Publishing Co., New York, 1959, Vol. 1, x+374 pp. Vol. 2, ix+276.
- [Gei77] Ladnor Geissinger. “Hopf algebras of symmetric functions and class functions”. In: *Combinatoire et représentation du groupe symétrique (Actes Table Ronde C.N.R.S., Univ. Louis-Pasteur Strasbourg, Strasbourg, 1976)*. Vol. 579. Lecture Notes in Math. Berlin: Springer, 1977, pp. 168–181.
- [Gel+95] Israel M. Gelfand, Daniel Krob, Alain Lascoux, Bernard Leclerc, Vladimir S. Retakh, and Jean-Yves Thibon. “Noncommutative symmetric functions”. In: *Adv. Math.* 112.2 (1995), pp. 218–348. ISSN: 0001-8708. DOI: 10.1006/aima.1995.1032. URL: <http://dx.doi.org/10.1006/aima.1995.1032>.
- [Ges10] Ira Gessel. “Descents, Peaks, and Shuffles of Permutations and Noncommutative Symmetric Functions”. Slides from a talk at Workshop on Quasisymmetric Functions, Banff International Research Station. 2010. URL: <http://www.birs.ca/workshops/2010/10w5031/files/gessel.pdf>.
- [Ges84] Ira M. Gessel. “Multipartite P -partitions and inner products of skew Schur functions”. In: *Combinatorics and algebra (Boulder, Colo., 1983)*. Vol. 34. Contemp. Math. Providence, RI: Amer. Math. Soc., 1984, pp. 289–317. DOI: 10.1090/conm/034/777705. URL: <http://dx.doi.org/10.1090/conm/034/777705>.
- [GR14] Darij Grinberg and Vic Reiner. “Hopf algebras in combinatorics”. Lecture notes from a course at University of Minnesota. 2014. URL: <http://www.math.umn.edu/~reiner/Classes/HopfComb.pdf>.

- [GR89] A. M. Garsia and C. Reutenauer. “A decomposition of Solomon’s descent algebra”. In: *Adv. Math.* 77.2 (1989), pp. 189–262. ISSN: 0001-8708. eprint: 10.1016/0001-8708(89)90020-0.
- [Haz10] Michiel Hazewinkel. “Explicit polynomial generators for the ring of quasymmetric functions over the integers”. In: *Acta Appl. Math.* 109.1 (2010), pp. 39–44. ISSN: 0167-8019. DOI: 10.1007/s10440-009-9439-z. URL: <http://dx.doi.org/10.1007/s10440-009-9439-z>.
- [Hiv07] Florent Hivert. “An introduction to combinatorial Hopf algebras—examples and realizations”. In: *Physics and theoretical computer science*. Vol. 7. NATO Secur. Sci. Ser. D Inf. Commun. Secur. Amsterdam: IOS, 2007, pp. 253–274.
- [Hol04] Ralf Holtkamp. “On Hopf algebra structures over operads”. 2004. URL: <http://arxiv.org/abs/math/0407074>.
- [Hol97] John M. Holte. “Carries, combinatorics, and an amazing matrix”. In: *Amer. Math. Monthly* 104.2 (1997), pp. 138–149. ISSN: 0002-9890. DOI: 10.2307/2974981. URL: <http://dx.doi.org/10.2307/2974981>.
- [Joy81] André Joyal. “Une théorie combinatoire des séries formelles”. In: *Adv. in Math.* 42.1 (1981), pp. 1–82. ISSN: 0001-8708. DOI: 10.1016/0001-8708(81)90052-9. URL: [http://dx.doi.org/10.1016/0001-8708\(81\)90052-9](http://dx.doi.org/10.1016/0001-8708(81)90052-9).
- [JR79] S. A. Joni and G.-C. Rota. “Coalgebras and bialgebras in combinatorics”. In: *Stud. Appl. Math.* 61.2 (1979), pp. 93–139. ISSN: 0022-2526.
- [Kas00] Yevgenia Kashina. “A generalized power map for Hopf algebras”. In: *Hopf algebras and quantum groups (Brussels, 1998)*. Vol. 209. Lecture Notes in Pure and Appl. Math. New York: Dekker, 2000, pp. 159–175.
- [Ker99] S. Kerov. “A differential model for the growth of Young diagrams”. In: *Proceedings of the St. Petersburg Mathematical Society, Vol. IV*. Vol. 188. Amer. Math. Soc. Transl. Ser. 2. Amer. Math. Soc., Providence, RI, 1999, pp. 111–130.

- [KLT97] D. Krob, B. Leclerc, and J.-Y. Thibon. “Noncommutative symmetric functions. II. Transformations of alphabets”. In: *Internat. J. Algebra Comput.* 7.2 (1997), pp. 181–264. ISSN: 0218-1967. DOI: 10 . 1142 / S0218196797000113. URL: <http://dx.doi.org/10.1142/S0218196797000113>.
- [KMLR82] K. H. Kim, L. G. Makar-Limanov, and F. W. Roush. “Graph monoids”. In: *Semigroup Forum* 25.1-2 (1982), pp. 1–7. ISSN: 0037-1912. DOI: 10 . 1007 / BF02573584. URL: <http://dx.doi.org/10.1007/BF02573584>.
- [Kna02] Anthony W. Knap. *Lie groups beyond an introduction*. Second. Vol. 140. Progress in Mathematics. Birkhäuser Boston, Inc., Boston, MA, 2002, pp. xviii+812. ISBN: 0-8176-4259-5.
- [Kre98] Dirk Kreimer. “On the Hopf algebra structure of perturbative quantum field theories”. In: *Adv. Theor. Math. Phys.* 2.2 (1998), pp. 303–334. ISSN: 1095-0761.
- [Kre99] D. Kreimer. “Chen’s iterated integral represents the operator product expansion”. In: *Adv. Theor. Math. Phys.* 3.3 (1999), pp. 627–670. ISSN: 1095-0761.
- [KS60] John G. Kemeny and J. Laurie Snell. *Finite Markov chains*. The University Series in Undergraduate Mathematics. D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto-London-New York, 1960, pp. viii+210.
- [KSK66] John G. Kemeny, J. Laurie Snell, and Anthony W. Knap. *Denumerable Markov chains*. D. Van Nostrand Co., Inc., Princeton, N.J.-Toronto, Ont.-London, 1966, pp. xi+439.
- [KT97] Daniel Krob and Jean-Yves Thibon. “Noncommutative symmetric functions. IV. Quantum linear groups and Hecke algebras at $q = 0$ ”. In: *J. Algebraic Combin.* 6.4 (1997), pp. 339–376. ISSN: 0925-9899. DOI: 10 . 1023 / A : 1008673127310. URL: <http://dx.doi.org/10.1023/A:1008673127310>.

- [LMS06] Rachel Landers, Susan Montgomery, and Peter Schauenburg. “Hopf powers and orders for some bismash products”. In: *J. Pure Appl. Algebra* 205.1 (2006), pp. 156–188. ISSN: 0022-4049. DOI: 10.1016/j.jpaa.2005.06.017. URL: <http://dx.doi.org/10.1016/j.jpaa.2005.06.017>.
- [Lot97] M. Lothaire. *Combinatorics on Words*. Cambridge Mathematical Library. With a foreword by Roger Lyndon and a preface by Dominique Perrin; corrected reprint of the 1983 original, with a new preface by Perrin. Cambridge: Cambridge University Press, 1997, pp. xviii+238. ISBN: 0-521-59924-5. eprint: 10.1017/CB09780511566097.
- [LPW09] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. With a chapter by James G. Propp and David B. Wilson. Providence, RI: American Mathematical Society, 2009, pp. xviii+371. ISBN: 978-0-8218-4739-8.
- [LR10] Jean-Louis Loday and María Ronco. “Combinatorial Hopf algebras”. In: *Quanta of maths*. Vol. 11. Clay Math. Proc. Providence, RI: Amer. Math. Soc., 2010, pp. 347–383.
- [LR98] Jean-Louis Loday and María O. Ronco. “Hopf algebra of the planar binary trees”. In: *Adv. Math.* 139.2 (1998), pp. 293–309. ISSN: 0001-8708. DOI: 10.1006/aima.1998.1759. URL: <http://dx.doi.org/10.1006/aima.1998.1759>.
- [Mac95] I. G. Macdonald. *Symmetric Functions and Hall Polynomials*. Second. Oxford Mathematical Monographs. With contributions by A. Zelevinsky, Oxford Science Publications. New York: The Clarendon Press Oxford University Press, 1995, pp. x+475. ISBN: 0-19-853489-2.
- [Mil14] Alexander Miller. “Foulkes characters for complex reflection groups”. To appear in Proc. AMS. 2014.
- [MR95] Claudi Malvenuto and Christophe Reutenauer. “Duality between quasi-symmetric functions and the Solomon descent algebra”. In: *J. Algebra* 177.3

- (1995), pp. 967–982. ISSN: 0021-8693. DOI: 10.1006/jabr.1995.1336. URL: <http://dx.doi.org/10.1006/jabr.1995.1336>.
- [NPT13] Jean-Christophe Novelli, Frédéric Patras, and Jean-Yves Thibon. “Natural endomorphisms of quasi-shuffle Hopf algebras”. In: *Bull. Soc. Math. France* 141.1 (2013), pp. 107–130. ISSN: 0037-9484.
- [NS14] Fumihiko Nakano and Taizo Sadahiro. “A generalization of carries process and a relation to riffle shuffles”. 2014. URL: <http://arxiv.org/abs/1403.5822>.
- [NT06] Jean-Christophe Novelli and Jean-Yves Thibon. “Polynomial realizations of some trialgebras”. Proc. FPSAC’06, San Diego. 2006. URL: <http://arxiv.org/abs/math.CO/0605061>.
- [NT12] Jean-Christophe Novelli and Jean-Yves Thibon. “Noncommutative symmetric functions and an amazing matrix”. In: *Adv. in Appl. Math.* 48.3 (2012), pp. 528–534. ISSN: 0196-8858. DOI: 10.1016/j.aam.2011.11.008. URL: <http://dx.doi.org/10.1016/j.aam.2011.11.008>.
- [Oko00] Andrei Okounkov. “Random matrices and random permutations”. In: *Internat. Math. Res. Notices* 20 (2000), pp. 1043–1095. ISSN: 1073-7928. DOI: 10.1155/S1073792800000532. URL: <http://dx.doi.org/10.1155/S1073792800000532>.
- [Pan13] C. Y. Amy Pang. “A Hopf-power Markov chain on compositions”. In: *25th International Conference on Formal Power Series and Algebraic Combinatorics (FPSAC 2013)*. Discrete Math. Theor. Comput. Sci. Proc., AS. Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 2013, pp. 499–510.
- [Pat93] Frédéric Patras. “La décomposition en poids des algèbres de Hopf”. In: *Ann. Inst. Fourier (Grenoble)* 43.4 (1993), pp. 1067–1087. ISSN: 0373-0956. eprint: http://www.numdam.org/item?id=AIF_1993__43_4_1067_0.
- [Pik13] John Pike. *Eigenfunctions for Random Walks on Hyperplane Arrangements*. Thesis (Ph.D.) ProQuest LLC, Ann Arbor, MI, 2013.

- [Pin14] Servando Pineda. *Hopf random walks on faces of permutahedra*. Thesis. 2014.
- [PR95] Stéphane Poirier and Christophe Reutenauer. “Algèbres de Hopf de tableaux”. In: *Ann. Sci. Math. Québec* 19.1 (1995), pp. 79–90. ISSN: 0707-9109.
- [Pri13] Jean-Baptiste Priez. “A lattice of combinatorial Hopf algebras, Application to binary trees with multiplicities”. In: *ArXiv e-prints* (Mar. 2013). arXiv: 1303.5538.
- [Pro09] Robert A. Proctor. “d-Complete Posets Generalize Young Diagrams for the Jeu de Taquin Property”. 2009. URL: <http://arxiv.org/abs/0905.3716>.
- [Ree58] Rimhak Ree. “Lie elements and an algebra associated with shuffles”. In: *Ann. of Math. (2)* 68 (1958), pp. 210–220. ISSN: 0003-486X.
- [Reu93] Christophe Reutenauer. *Free Lie Algebras*. Vol. 7. London Mathematical Society Monographs. New Series. Oxford Science Publications. New York: The Clarendon Press Oxford University Press, 1993, pp. xviii+269. ISBN: 0-19-853679-8.
- [Sal12] Franco Saliola. “Eigenvectors for a random walk on a left-regular band”. In: *Advances in Applied Mathematics* 48.2 (2012), pp. 306–311. ISSN: 0196-8858. DOI: 10.1016/j.aam.2011.09.002. URL: <http://www.sciencedirect.com/science/article/pii/S0196885811000868>.
- [Sch90] William Schmitt. “Hopf algebras and identities in free partially commutative monoids”. In: *Theoret. Comput. Sci.* 73.3 (1990), pp. 335–340. ISSN: 0304-3975. DOI: 10.1016/0304-3975(90)90184-J. URL: [http://dx.doi.org/10.1016/0304-3975\(90\)90184-J](http://dx.doi.org/10.1016/0304-3975(90)90184-J).
- [Sch93] William R. Schmitt. “Hopf algebras of combinatorial structures”. In: *Canad. J. Math.* 45.2 (1993), pp. 412–428. ISSN: 0008-414X. eprint: 10.4153/CJM-1993-021-5.

- [Sch94] William R. Schmitt. “Incidence Hopf algebras”. In: *J. Pure Appl. Algebra* 96.3 (1994), pp. 299–330. ISSN: 0022-4049. DOI: 10.1016/0022-4049(94)90105-8. URL: [http://dx.doi.org/10.1016/0022-4049\(94\)90105-8](http://dx.doi.org/10.1016/0022-4049(94)90105-8).
- [Sta01] Richard P. Stanley. “Generalized riffle shuffles and quasisymmetric functions”. In: *Ann. Comb.* 5.3-4 (2001). Dedicated to the memory of Gian-Carlo Rota (Tianjin, 1999), pp. 479–491. ISSN: 0218-0006. DOI: 10.1007/s00026-001-8023-7. URL: <http://dx.doi.org/10.1007/s00026-001-8023-7>.
- [Sta95] Richard P. Stanley. “A symmetric function generalization of the chromatic polynomial of a graph”. In: *Adv. Math.* 111.1 (1995), pp. 166–194. ISSN: 0001-8708. DOI: 10.1006/aima.1995.1020. URL: <http://dx.doi.org/10.1006/aima.1995.1020>.
- [Sta97] Richard P. Stanley. *Enumerative combinatorics. Vol. 1.* Vol. 49. Cambridge Studies in Advanced Mathematics. With a foreword by Gian-Carlo Rota, Corrected reprint of the 1986 original. Cambridge: Cambridge University Press, 1997, pp. xii+325. ISBN: 0-521-55309-1; 0-521-66351-2.
- [Sta99] Richard P. Stanley. *Enumerative Combinatorics. Vol. 2.* Vol. 62. Cambridge Studies in Advanced Mathematics. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin. Cambridge: Cambridge University Press, 1999, pp. xii+581. ISBN: 0-521-56069-1; 0-521-78987-7.
- [Str11] Jeffrey Strom. *Modern classical homotopy theory.* Vol. 127. Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2011, pp. xxii+835. ISBN: 978-0-8218-5286-6.
- [Swa12] J.M. Swart. “Advanced Topics in Markov chains”. Lecture notes from a course at Charles University. 2012. URL: <http://staff.utia.cas.cz/swart/chain10.pdf>.
- [Swe69] Moss E. Sweedler. *Hopf algebras.* Mathematics Lecture Note Series. W. A. Benjamin, Inc., New York, 1969, pp. vii+336.

- [Thi12] Jean-Yves Thibon. “Réalizations polynomiales d’algèbres de Hopf combinatoires”. Slides from a talk at Séminaire Philippe Flajolet. 2012. URL: <http://www-igm.univ-mlv.fr/~jyt/TALKS/flaj.pdf>.
- [TO70] John Tate and Frans Oort. “Group schemes of prime order”. In: *Ann. Sci. École Norm. Sup.* 3 (1970), pp. 1–21. ISSN: 0012-9593.
- [Wil94] Herbert S. Wilf. *Generating Functionology*. 2nd. Boston, MA: Academic Press, 1994, pp. ix+228. ISBN: 0127519564.
- [Zab10] Mike Zabrocki. “A zoo of Hopf algebras”. Slides from a talk at Combinatorial Algebra meets Algebraic Combinatorics, Kingston. 2010. URL: <http://garsia.math.yorku.ca/~zabrocki/talks/HopfzooKingston.pdf>.
- [Zel81] Andrey V. Zelevinsky. *Representations of finite classical groups*. Vol. 869. Lecture Notes in Mathematics. A Hopf algebra approach. Berlin: Springer-Verlag, 1981, pp. iv+184. ISBN: 3-540-10824-6.
- [Zho08] Hua Zhou. *Examples of multivariate Markov chains with orthogonal polynomial eigenfunctions*. Thesis (Ph.D.)—Stanford University. ProQuest LLC, Ann Arbor, MI, 2008, p. 168. ISBN: 978-0549-85071-7. URL: http://gateway.proquest.com/openurl?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&res_dat=xri:pqdiss&rft_dat=xri:pqdiss:3332963.