Hopf Algebras and Markov Chains - Diaconis, Pang, Ram. (updated version from Amy Pang's homepage)

Here, in reverse chronological order, is a brief description of all mathematical changes since the second version posted on arXiv, which is the same as the version in Journal of Algebraic Combinatoris modulo punctuation. Non-mathematical typographical corrections are not listed here.

30 December 2014
Proof of Proposition 3.25: $\alpha = (\alpha_1, \ldots, \alpha_a)$ (a parts, not n).
First remark after Corollary 4.10: The calculation is of $P_{(n)}\{X_k = 1^n\}$.

20 July 2013 (mostly thanks to Darij Grinberg):
Example 1.4: first paragraph clarified.
Section 2.3: in second paragraph, added an alternative description of Lyndon factorization.
Section 2.4: now using infinitely many variables.
Last paragraph of Example 3.3: some ns are now ls.
Proposition 3.7: positivity assumption added.
Theorem 3.19: last sentence improved.
Proof of Theorem 3.19: last displayed equation, a b becomes a b'.
Proof of Proposition 3.25: just before third displayed equation, "unless $b' = \bullet^{\alpha_a - 1}$", not b; penultimate displayed equation, the sum is over $\bar{\alpha}$ not α.
Paragraph before proof of Proposition 4.3: improved explanation; in the first displayed equation, some λs are now μs; in the second displayed equation, added a sum in μ.
Paragraph before Example 6.3: improved reference.

Hopf algebras and Markov chains: Two examples and a theory

Persi Diaconis[*]
Departments of
Mathematics and Statistics
Stanford University

C. Y. Amy Pang[†‡]
Department of
Mathematics
Stanford University

Arun Ram[§]
Department of
Mathematics and Statistics
University of Melbourne

Abstract

The operation of *squaring* (coproduct followed by product) in a combinatorial Hopf algebra is shown to induce a Markov chain in natural bases. Chains constructed in this way include widely studied methods of card shuffling, a natural "rock-breaking" process, and Markov chains on simplicial complexes. Many of these chains can be explictly diagonalized using the primitive elements of the algebra and the combinatorics of the free Lie algebra. For card shuffling, this gives an explicit description of the eigenvectors. For rock-breaking, an explicit description of the quasi-stationary distribution and sharp rates to absorption follow.

1   Introduction

A Hopf algebra is an algebra $\mathcal{H}$ with a coproduct $\Delta : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$ which fits together with the product $m : \mathcal{H} \otimes \mathcal{H} \to \mathcal{H}$. Background on Hopf algebras is in Section 2.2. The map $m\Delta : \mathcal{H} \to \mathcal{H}$ is called the *Hopf-square* (often denoted Ψ^2 or $x^{[2]}$). Our first discovery is that the coefficients of $x^{[2]}$ in natural bases can often be interpreted as a Markov chain. Specializing to familiar Hopf algebras can give interesting Markov chains: the free associative algebra gives the Gilbert–Shannon–Reeds model of riffle shuffling. Symmetric functions give a rock-breaking model of Kolmogoroff [Kol41]. These two examples are developed first for motivation.

Example 1.1 (Free associative algebra and riffle shuffling)**.** Let $x_1, x_2, \ldots, x_n$ be noncommuting variables and $\mathcal{H} = k\langle x_1, \ldots, x_n \rangle$ be the free associative algebra. Thus $\mathcal{H}$ consists of finite linear combinations of words $x_{i_1} x_{i_2} \cdots x_{i_k}$ in the generators with the concatenation product. The coproduct Δ is an algebra map defined by $\Delta(x_i) = 1 \otimes x_i + x_i \otimes 1$ and extended linearly. Consider

$$\Delta(x_{i_1} \cdots x_{i_k}) = (1 \otimes x_{i_1} + x_{i_1} \otimes 1)(1 \otimes x_{i_2} + x_{i_2} \otimes 1) \cdots (1 \otimes x_{i_k} + x_{i_k} \otimes 1).$$

A term in this product results from a choice of left or right from each factor. Equivalently, for each subset $S \subseteq \{1, 2, \ldots, k\}$, there corresponds the term

$$\prod_{j \in S} x_{i_j} \otimes \prod_{j \in S^C} x_{i_j}.$$

Thus $m\Delta$ is a sum of 2^k terms resulting from removing $\{x_{i_j}\}_{j \in S}$ and moving them to the front. For example,

$$m\Delta(x_1 x_2 x_3) = x_1 x_2 x_3 + x_1 x_2 x_3 + x_2 x_1 x_3 + x_3 x_1 x_2 + x_1 x_2 x_3 + x_1 x_3 x_2 + x_2 x_3 x_1 + x_1 x_2 x_3.$$

Dividing $m\Delta$ by 2^k, the coefficient of a word on the right is exactly the chance that this word appears in a Gilbert–Shannon–Reeds inverse shuffle of a deck of cards labeled by x_i in initial order $x_{i_1} x_{i_2} \ldots x_{i_k}$. Applying $\frac{1}{2^k} m\Delta$ in the dual algebra gives the usual model for riffle shuffling. Background on these models is in Section 5. As shown there, this connection between Hopf algebras and shuffling gives interesting new theorems about shuffling.

Example 1.2 (Symmetric functions and rock-breaking). Let us begin with the rock-breaking description. Consider a rock of total mass n. Break it into two pieces according to the symmetric binomial distribution:

$$P\{\text{left piece has mass } j\} = \binom{n}{j} \Big/ 2^n, \qquad 0 \le j \le n.$$

Continue, at the next stage breaking each piece into $\{j_1, j - j_1\}, \{j_2, n - j - j_2\}$ by independent binomial splits. The process continues until all pieces are of mass one when it stops. This description gives a Markov chain on partitions of n, absorbing at 1^n.

This process arises from the Hopf-square map applied to the algebra $\Lambda = \Lambda(x_1, x_2, \ldots, x_n)$ of symmetric functions, in the basis of elementary symmetric functions $\{e_\lambda\}$. This is an algebra under the usual product. The coproduct, following [Gei77] is defined by

$$\Delta(e_i) = e_0 \otimes e_i + e_1 \otimes e_{i-1} + \cdots + e_i \otimes e_0,$$

extended multiplicatively and linearly. This gives a Hopf algebra structure on Λ which is a central object of study in algebraic combinatorics. It is discussed in Section 2.4. Rescaling the basis elements to $\{\hat{e}_i := i! e_i\}$, a direct computation shows that $m\Delta$ in the $\{e_\lambda\}$ basis gives the rock-breaking process; see Section 4.1.

A similar development works for any Hopf algebra which is either a polynomial algebra as an algebra (for instance, the algebra of symmetric functions, with generators e_n), or is cocommutative and a free associative algebra as an algebra (e.g., the free associative algebra), provided each object of degree greater than one can be broken non-trivially. These results are described in Theorem 3.4.

Our second main discovery is that this class of Markov chains can be explicitly diagonalized using the Eulerian idempotent and some combinatorics of the free associative algebra. This combinatorics is reviewed in Section 2.3. It leads to a description of the left eigenvectors (Theorems 3.15 and 3.16) which is often interpretable and allows exact and asymptotic answers to natural probability questions. For a polynomial algebra, we are also able to describe the right eigenvectors completely (Theorem 3.19).

Example 1.3 (Shuffling). For a deck of n distinct cards, the eigenvalues of the Markov chain induced by repeated riffle shuffling are $1, 1/2, \ldots, 1/2^{n-1}$ [Han90]. The multiplicity of the eigenvalue $1/2^{n-i}$ equals the number of permutations in S_n with i cycles. For example, the second eigenvalue, $1/2$, has multiplicity $\binom{n}{2}$. For $1 \le i < j \le n$, results from Section 5 show that a right eigenvector f_{ij} is given by

$$f_{ij}(w) = \begin{cases} 1, & \text{if } i \text{ and } j \text{ are adjacent in } w \text{ in order } ij, \\ -1, & \text{if } i \text{ and } j \text{ are adjacent in } w \text{ in order } ji, \\ 0, & \text{otherwise.} \end{cases}$$

Summing in $i < j$ shows that $d(w) - \frac{n-1}{2}$ is an eigenvector with eigenvalue $1/2$ (where $d(w)$ is the number of descents in w). Similarly $p(w) - \frac{n-2}{3}$ is an eigenvector with eigenvalue $1/4$ (where $p(w)$ is the number of peaks in w). These eigenvectors are used to determine the mean and variance of the number of carries when large integers are added.

Our results work for decks with repeated values allowing us to treat cases when e.g., the suits do not matter and all picture cards are equivalent to tens. Here, fewer shuffles are required to achieve stationarity. For decks of essentially any composition we show that all eigenvalues $1/2^i$, $0 \leq i \leq n-1$, occur and determine multiplicities and eigenvectors.

Example 1.4 (Rock-breaking). Consider the rock-breaking process of Example 1.2 started at (n), the partition with a single part of size n. This is absorbing at the partition 1^n. In Section 4, this process is shown to have eigenvalues $1, 1/2, \ldots, 1/2^{n-1}$ with the multiplicity of $1/2^{n-l}$ the number of partitions of n into l parts. Thus, the second eigenvalue of $1/2$ has multiplicity 1, as $1^{n-2}2$ is the only partition with $n-1$ parts. The corresponding eigenfunction is

$$f_{1^{n-2}2}(\lambda) = \sum_i \binom{\lambda_i}{2}.$$

This is a monotone function under the refinement partial order on partitions and equals zero if and only if $\lambda = 1^n$. If $X_0 = (n), X_1, X_2, \ldots$ are the successive partitions generated by the Markov chain then

$$E_{(n)}\{f_{1^{n-2}2}(X_k)\} = \frac{1}{2^k} f_{1^{n-2}2}(X_0) = \binom{n}{2}\Big/2^k.$$

Using Markov's inequality,

$$P_{(n)}\{X_k \text{ is not absorbed}\} \leq \binom{n}{2}\Big/2^k.$$

This shows that for $k = 2\log_2 n + c$, the chance of absorption is asymptotic to $1 - 1/2^{c+1}$ when n is large. Section 4 derives all of the eigenvectors and gives further applications.

Section 2 reviews Markov chains (including uses for eigenvectors), Hopf algebras, and some combinatorics of the free associative algebra. Section 3 gives our basic theorems, generalizing the two examples to polynomial Hopf algebras and cocommutative, free associative Hopf algebras. Section 4 treats rock-breaking; Section 5 treats shuffling. Section 6 briefly describes other examples (e.g., graphs and simplicial complexes), counter-examples (e.g., the Steenrod algebra), and questions (e.g., quantum groups).

Two historical notes: The material in the present paper has roots in work of Patras [Pat91, Pat93, Pat94], whose notation we are following, and Drinfeld [Dri89]. Patras studied shuffling in a purely geometric fashion, making a ring out of polytopes in $\mathbb{R}^n$. This study led to natural Hopf structures, Eulerian idempotents, and generalization of Solomon's descent algebra in a Hopf context. His Eulerian idempotent maps decompose a graded commutative or cocommutative Hopf algebra into eigenspaces of the ath Hopf-powers; we improve upon this result, in the case of polynomial algebras or cocommutative, free associative algebras, by algorithmically producing a full eigenbasis. While there is no hint of probability in the work of Patras, it deserves to be much better known. More detailed references are given elsewhere in this paper.

We first became aware of Drinfeld's ideas through their mention in Shnider–Sternberg [SS93]. Consider the Hopf-square, acting on a Hopf algebra $\mathcal{H}$. Suppose that $x \in \mathcal{H}$ is primitive, $\Delta(x) = 1 \otimes x + x \otimes 1$. Then $m\Delta(x) = 2x$ so x is an eigenvector of $m\Delta$ with eigenvalue 2. If x and y are

primitive then $m\Delta(xy + yx) = 4(xy + yx)$ and, similarly, if $x_1, \ldots, x_k$ are primitive then the sum of symmetrized products is an eigenvector of $m\Delta$ with eigenvector 2^k. Drinfeld [Dri89, Prop 3.7] used these facts without comment in his proof that any formal deformation of the cocommutative universal enveloping algebra $\mathcal{U}(\mathfrak{g})$ results already from deformation of the underlying Lie algebra $\mathfrak{g}$. See [SS93, Sect. 3.8] and Section 3.4 below for an expanded argument and discussion. For us, a description of the primitive elements and their products gives the eigenvectors of our various Markov chains. This is developed in Section 3.

Acknowledgements: We thank Marcelo Aguiar, Federico Ardila, Nantel Bergeron, Megan Bernstein, Dan Bump, Gunnar Carlsson, Ralph Cohen, David Hill, Peter J. McNamara, Susan Montgomery and Servando Pineda for their help and suggestions, and the anonymous reviewer for a wonderfully detailed review.

2 Background

This section gives notation and background for Markov chains (including uses for eigenvectors), Hopf algebras, the combinatorics of the free associative algebra and symmetric functions. All of these are large subjects and pointers to accessible literature are provided.

2.1 Markov chains

Let $\mathcal{X}$ be a finite set. A Markov chain on $\mathcal{X}$ may be specified by a transition matrix $K(x, y)$ $(x, y \in \mathcal{X})$ with $K(x, y) \geq 0$, $\sum_y K(x, y) = 1$. This is interpreted as the chance that the chain moves from x to y in one step. If the chain is denoted $X_0, X_1, X_2, \ldots$ and $X_0 = x_0$ is a fixed starting state then

$$P\{X_1 = x_1, \cdots, X_n = x_n\} = \prod_{i=0}^{n-1} K(x_i, x_{i+1}).$$

Background and basic theory can be found in [KT75] or [Bré99]. The readable introduction [LPW09] is recommended as close in spirit to the present paper. The analytic theory is developed in [SC97].

Let $K^2(x, y) = \sum_z K(x, z)K(z, y)$ denote the probability of moving from x to y in two steps. Similarly, K^l is defined. Under mild conditions [LPW09, Sec 1.5] Markov chains have unique stationary distributions $\pi(x)$: thus $\pi(x) \geq 0$, $\sum_x \pi(x) = 1$, $\sum_x \pi(x)K(x, y) = \pi(y)$, so π is a left eigenvector of K with eigenvalue 1. Set

$$L^2(\pi) = \{f : \mathcal{X} \to \mathbb{R}\} \quad \text{with} \quad \langle f_1 | f_2 \rangle = \sum f_1(x) f_2(x) \pi(x).$$

Then K operates as a contraction on L^2 with $Kf(x) = \sum_y K(x, y)f(y)$. The Markov chains considered in this paper are usually not self-adjoint (equivalently reversible), nonetheless, they are diagonalizable over the rationals with eigenvalues $1 = \beta_0 \geq \beta_1 \geq \cdots \geq \beta_{|\mathcal{X}|-1} > 0$. We have a basis of left eigenfunctions $\{g_i\}_{i=0}^{|\mathcal{X}|-1}$ with $g_0(x) = \pi(x)$ and $\sum_x g_i(x)K(x, y) = \beta_i g_i(y)$, and, in some cases, a dual basis of right eigenfunctions $\{f_i\}_{i=0}^{|\mathcal{X}|-1}$ with $f_0(x) \equiv 1$, $Kf_i(x) = \beta_i f_i(x)$, and $\sum_x f_i(x)g_j(x) = \delta_{ij}$. As is customary in discussions of random walks on algebraic structures, we will abuse notation and think of the eigenfunctions f_i both as functions on the state space and as linear combinations of the states - in other words, $\sum_x f_i(x)x$ will also be denoted f_i.

Throughout, we are in the unusual position of knowing β_i, g_i and possibly f_i explicitly. This is rare enough that some indication of the use of eigenfunctions is indicated.

4

Use A For any function $f : \mathcal{X} \to \mathbb{R}$, expressed in the basis of right eigenfunctions $\{f_i\}$ as

$$
(2.1) \qquad f = \sum_{i=0}^{|\mathcal{X}|-1} a_i f_i,
$$

the expectation of f after k steps, having started at x_0, is given by

$$
(2.2) \qquad E_{x_0} \{f(X_k)\} = \sum_{i=0}^{|\mathcal{X}|-1} a_i \beta_i^k f_i(x_0).
$$

For example, for shuffling, the normalized number of descents $d(\pi) - (n-1)/2$ is the sum of the $1/2$-eigenfunctions for riffle shuffling; see Example 5.8. Thus, with $x_0 = id$ and all k, $0 \leq k < \infty$,

$$
E_{id} \{d(X_k)\} = \frac{n-1}{2} \left(1 - \frac{1}{2^k}\right).
$$

In [DF09b, DF09a] it is shown that the number of descents in repeated riffle shuffles has the same distribution as the number of carries when n integers are added. Further, the square of this eigenfunction has a simple eigenfunction expansion leading to simple formulae for the variance and covariance of the number of carries.

Use B If f is a right eigenfunction with eigenvalue β, then the self-correlation after k steps (starting in stationarity) is

$$
E_\pi \{f(X_0)f(X_k)\} = E_\pi \{E \{f(X_0)f(X_k)|X_0 = x_0\}\} = \beta^k E_\pi \{f(X_0)^2\}.
$$

This indicates how certain correlations fall off and gives an interpretation of the eigenvalues.

Use C For f a right eigenfunction with eigenvalue β, let $Y_i = f(X_i)/\beta^i$, $0 \leq i < \infty$. Then Y_i is an $\mathcal{F}_i$ martingale with $\mathcal{F}_i = \sigma(X_0, X_1, \ldots, X_i)$. One may try to use optional stopping, maximal and concentration inequalities and the martingale central limit theorem to study the behavior of the original X_i chain.

Use D One standard use of right eigenfunctions is to prove lower bounds for mixing times for Markov chains. The earliest use of this is the second moment method [Dia03]. Here, one uses the second eigenfunction as a test function and expands its square in the eigenbasis to get concentration bounds. An important variation is Wilson's method [Wil04] which only uses the first eigenfunction but needs a careful understanding of the variation of this eigenfunction. A readable overview of both methods and many examples is in [SC04].

Use E The left eigenfunctions come into computations since $\sum_x g_i(x) f_j(x) = \delta_{ij}$. Thus in (2.1), $a_i = \langle g_i | f/\pi \rangle$. (Here f/π is just the density of f with respect to π.)

Use F A second prevalent use of left eigenfunctions throughout this paper: the dual of a Hopf algebra is a Hopf algebra and left eigenfunctions of the dual chain correspond to right eigenfunctions of the original chain. This is similar to the situation for time reversal. If $K^*(x,y) = \frac{\pi(y)}{\pi(x)} K(y,x)$ is the time-reversed chain (note $K^*(x,y)$ is a Markov chain with stationary distribution π), then g_i/π is a right eigenfunction of K^*.

Use G The left eigenfunctions also come into determining the quasi-stationary distribution of absorbing chains such as the rock-breaking chain. A useful, brief introduction to quasi-stationarity is in [KT75]. The comprehensive survey [vD91] and annotated bibliography [Pol11] are also useful. Consider the case where there is a unique absorbing state $x_\bullet$ and the second eigenvalue β_1 of the chain satisfies $1 = \beta_0 > \beta_1 > \beta_2 \geq \ldots$. This holds for rock-breaking. There are two standard notions of "the limiting distribution of the chain given that it has not been absorbed":

$$(2.3a) \qquad \pi^1(x) = \lim_{k \to \infty} P\{X_k = x | X_k \neq x_\bullet\};$$

$$(2.3b) \qquad \pi^2(x) = \lim_{k \to \infty} \lim_{l \to \infty} P\{X_k = x | X_l \neq x_\bullet\}.$$

In words, $\pi^1(x)$ is the limiting distribution of the chain given that it has not been absorbed up to time k and $\pi^2(x)$ is the limiting distribution of the chain given that it is never absorbed. These quasi-stationary distributions can be expressed in terms of the eigenfunctions:

$$(2.4) \qquad \pi^1(x) = \frac{g_1(x)}{\sum_{y \neq \bullet} g_1(y)}, \qquad \pi^2(x) = \frac{g_1(x) f_1(x)}{\sum_{y \neq \bullet} g_1(y) f_1(y)}.$$

These results follow from simple linear algebra and are proved in the references above. For rock-breaking, results in Section 4 show that $\pi^1 = \pi^2$ is point mass at the partition $1^{n-2}2$.

Use H Both sets of eigenfunctions appear in the formula

$$(2.5) \qquad K^l(x, y) = \sum_{i=0}^{|\mathcal{X}|-1} \beta_i^l f_i(x) g_i(y).$$

This permits the possibility of determining convergence rates. It can be difficult to do for chains with large state spaces. See the examples and discussion in [DF12].

To conclude this discussion of Markov chains we mention that convergence is customarily measured by a few standard distances:

$$(2.6) \qquad \text{Total variation} \qquad \|K_{x_0}^l - \pi\|_{\mathrm{TV}} = \max_{A \subseteq \mathcal{X}} |K_{x_0}^l(A) - \pi(A)| = \frac{1}{2} \sum_y |K^l(x_0, y) - \pi(y)|$$

$$(2.7) \qquad \text{Separation} \qquad \mathrm{sep}_{x_0}(l) = \max_y 1 - \frac{K^l(x_0, y)}{\pi(y)}$$

$$(2.8) \qquad \text{Sup} \qquad l_\infty(l) = \max_y \left| \frac{K^l(x_0, y) - \pi(y)}{\pi(y)} \right|$$

Here $\|K_{x_0}^l - \pi\|_{\mathrm{TV}} \leq \mathrm{sep}_{x_0}(l) \leq l_\infty(l)$ and all distances are computable by determining the maximizing or minimizing values of A or y and using (2.5)-(2.8). See [LPW09, Lemma 6.13] for further discussion of these distances.

2.2 Hopf algebras

A Hopf algebra is an algebra $\mathcal{H}$ over a field k (usually the real numbers in the present paper). It is associative with unit 1, but not necessarily commutative. Let us write m for the multiplication in $\mathcal{H}$, so $m(x \otimes y) = xy$. Then $m^{[a]} : \mathcal{H}^{\otimes a} \to \mathcal{H}$ will denote a-fold products (so $m = m^{[2]}$), formally $m^{[a]} = m(\iota \otimes m^{[a-1]})$ where ι denotes the identity map.

$\mathcal{H}$ comes equipped with a *coproduct* $\Delta : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$, written $\Delta(x) = \sum_{(x)} x_{(1)} \otimes x_{(2)}$ in Sweedler notation [Swe69]. The coproduct is *coassociative* in that

$$(\Delta \otimes \iota)\Delta(x) = \sum_{(x),(x_{(1)})} x_{(1)_{(1)}} \otimes x_{(1)_{(2)}} \otimes x_{(2)} = \sum_{(x),(x_{(2)})} x_{(1)} \otimes x_{(2)_{(1)}} \otimes x_{(2)_{(2)}} = (\iota \otimes \Delta)\Delta(x)$$

so there is no ambiguity in writing $\Delta^{[3]}(x) = \sum_{(x)} x_{(1)} \otimes x_{(2)} \otimes x_{(3)}$. Similarly, $\Delta^{[a]} : \mathcal{H} \to \mathcal{H}^{\otimes a}$ denotes the a-fold coproduct, where Δ is applied $a-1$ times, to any one tensor-factor at each stage; formally $\Delta^{[a]} = (\iota \otimes \cdots \otimes \iota \otimes \Delta)\Delta^{[a-1]}$. The Hopf algebra $\mathcal{H}$ is *cocommutative* if $\sum_{(x)} x_{(1)} \otimes x_{(2)} = \sum_{(x)} x_{(2)} \otimes x_{(1)}$; in other words, an expression in Sweedler notation is unchanged when the indices permute. An element x of $\mathcal{H}$ is *primitive* if $\Delta(x) = 1 \otimes x + x \otimes 1$.

The product and coproduct have to be compatible so Δ is an algebra homomorphism, where multiplication on $\mathcal{H} \otimes \mathcal{H}$ is componentwise; in Sweedler notation this says $\Delta(xy) = \sum_{(x),(y)} x_{(1)} y_{(1)} \otimes x_{(2)} y_{(2)}$. All of the algebras considered here are graded and connected, i.e., $\mathcal{H} = \bigoplus_{i=0}^{\infty} \mathcal{H}_i$ with $\mathcal{H}_0 = k$ and $\mathcal{H}_n$ finite-dimensional. The product and coproduct must respect the grading so $\mathcal{H}_i \mathcal{H}_j \subseteq \mathcal{H}_{i+j}$, and $x \in \mathcal{H}_n$ implies $\Delta(x) \in \bigoplus_{j=0}^{n} \mathcal{H}_j \otimes \mathcal{H}_{n-j}$. There are a few more axioms concerning a counit map and an antipode (automatic in the graded case); for the present paper, the most important is that the counit is zero on elements of positive degree, so, by the coalgebra axioms, $\bar{\Delta}(x) := \Delta(x) - 1 \otimes x - x \otimes 1 \in \bigoplus_{j=1}^{n-1} \mathcal{H}_j \otimes \mathcal{H}_{n-j}$, for $x \in \mathcal{H}_n$. The free associative algebra and the algebra of symmetric functions, discussed in Section 1, are examples of graded Hopf algebras.

The subject begins in topology when H. Hopf realized that the presence of the coproduct leads to nice classification theorems which allowed him to compute the cohomology of the classical groups in a unified manner. Topological aspects are still a basic topic [Hat02] with many examples which may provide grist for the present mill. For example, the cohomology groups of the loops on a topological space form a Hopf algebra, and the homology of the loops on the suspension of a wedge of circles forms a Hopf algebra isomorphic to the free associative algebra of Example 1.1 [BS53].

Joni and Rota [JR79] realized that many combinatorial objects have a natural breaking structure which gives a coalgebra structure to the graded vector space on such objects. Often there is a compatible way of putting pieces together, extending this to a Hopf algebra structure. Often, either the assembling or the breaking process is symmetric, leading to commutative or cocommutative Hopf algebras respectively. For example, the symmetric function algebra is commutative and cocommutative while the free associative algebra is just cocommutative.

The theory developed here is for graded commutative or cocommutative Hopf algebras with one extra condition, that there is a unique way to assemble any given collection of objects. This amounts to the requirement that the Hopf algebra is either a polynomial algebra as an algebra (and therefore commutative) or a free associative algebra as an algebra and cocommutative (and therefore noncommutative). (We write *a* free associate algebra to refer to the algebra structure only, as opposed to *the* free associative algebra which has a specified coalgebra structure - namely, the generating elements are primitive.)

Increasingly sophisticated developments of combinatorial Hopf algebras are described by [ABS06, Sch93, Sch94, Sch87, Sch95] and [AM10]. This last is an expansive extension which unifies many common examples. Below are two examples that are prototypes for their Bosonic Fock functor and Full Fock functor constructions respectively [AM10, Ch. 15]; they are also typical of constructions detailed in other sources.

Example 2.1 (The Hopf algebra of unlabeled graphs)**.** [Sch94, Sec. 12] [Fis10, Sec. 3.2] Let $\bar{\mathcal{G}}$ be the vector space spanned by unlabeled simple graphs (no loops or multiple edges). This becomes a

Hopf algebra with product disjoint union and coproduct

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

where the sum is over subsets of vertices S with G_S, G_{S^c} the induced subgraphs. Graded by number of vertices, $\bar{\mathcal{G}}$ is both commutative and cocommutative, and is a polynomial algebra as an algebra. The associated random walk is described in Example 3.1 below.

Example 2.2 (The noncommutative Hopf algebra of labeled graphs). [Sch94, Sec. 13] [Fis10, Sec. 3.3] Let $\mathcal{G}$ be the vector space spanned by the set of simple graphs where vertices are labeled $\{1, 2, \ldots, n\}$, for some n. The product of two graphs $G_1 G_2$ is their disjoint union, where the vertices of G_1 keep their labels, and the labels in G_2 are increased by the number of vertices in G_1. The coproduct is

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

where we again sum over all subsets S of vertices of G, and G_S, G_{S^c} are relabeled so the vertices in each keep the same relative order. For example,

where 1 denotes the empty graph. $\mathcal{G}$ is noncommutative and cocommutative and a free associative algebra as an algebra; the associated random walk is detailed in Example 3.2. As the notation suggests, $\bar{\mathcal{G}}$ is a quotient of $\mathcal{G}$, obtained by forgetting the labels on the vertices.

Aguiar–Bergeron–Sottile [ABS06] define a combinatorial Hopf algebra as a Hopf algebra $\mathcal{H}$ with a character $\zeta : \mathcal{H} \to k$ which is both additive and multiplicative. They prove a universality theorem: any combinatorial Hopf algebra has a unique character-preserving Hopf morphism into the algebra of quasisymmetric functions. They show that this unifies many ways of building generating functions. When applied to the Hopf algebra of graphs, their map gives the chromatic polynomial. In Section 3.7 we find that their map gives the probability of absorption for several of our Markov chains. See also the examples in Section 6.

A good introduction to Hopf algebras is in [SS93]. A useful standard reference is in [Mon93] and our development does not use much outside of her Chapter 1. The broad-ranging text [Maj95] is aimed towards quantum groups but contains many examples useful here. Quantum groups are neither commutative nor cocommutative and need special treatment; see Example 6.3.

A key ingredient in our work is the Hopf-square map $\Psi^2 = m\Delta$; $\Psi^2(x)$ is also written $x^{[2]}$. In Sweedler notation, $\Psi^2(x) = \sum_{(x)} x_{(1)} x_{(2)}$; in our combinatorial setting, it is useful to think of "pulling apart" x according to Δ, then using the product to put the pieces together. On graded Hopf algebras, Ψ^2 preserves the grading and, appropriately normalized, gives a Markov chain on appropriate bases. See Section 3.2 for assumptions and details. The higher power maps $\Psi^a = m^{[a]} \Delta^{[a]}$ will also be studied, since under our hypothesis, they present no extra difficulty. For example, $\Psi^3(x) = \sum_{(x)} x_{(1)} x_{(2)} x_{(3)}$. In the shuffling example, Ψ^a corresponds to the "a-shuffles" of

[BD92]. A theorem of [TO70] shows that, for commutative or cocommutative Hopf algebras, the power rule holds: $(x^{[a]})^{[b]} = x^{[ab]}$, or $\Psi^a\Psi^b = \Psi^{ab}$. See also the discussion in [LMS06]. In shuffling language this becomes "an a-shuffle followed by a b-shuffle is an ab-shuffle" [BD92]. In general Hopf algebras this power law often fails [Kas00]. Power maps are actively studied as part of a program to carry over to Hopf algebras some of the rich theory of groups. See [GM09, LM00] and their references.

2.3   Structure theory of a free associative algebra

The eigenvectors of our Markov chains are described using combinatorics related to the free associative algebra, as described in the self-contained [Lot97, Chap. 5].

A word in an ordered alphabet is *Lyndon* if it is strictly smaller (in lexicographic order) than its cyclic rearrangements. So 1122 is Lyndon but 21 or 1212 are not. A basic fact [Lot97, Th. 5.1.5] is that any word w has a unique *Lyndon factorization*, that is, $w = l_1 l_2 \cdots l_k$ with each l_i a Lyndon word and $l_1 \geq l_2 \geq \cdots \geq l_k$; an easy way to compute this is to successively remove the longest Lyndon tail. Further, each Lyndon word l has a *standard factorization*: if l is not a single letter, then $l = l_1 l_2$ where l_i is non-trivial Lyndon and l_2 is the longest right Lyndon factor of l. (The standard factorization of a letter is just that letter by definition.) Thus $13245 = 13 \cdot 245$. Using this, define, for Lyndon l, its *standard bracketing* $\lambda(l)$ recursively by $\lambda(a) = a$ for a letter and $\lambda(l) = [\lambda(l_1), \lambda(l_2)]$ for $l = l_1 l_2$ in standard factorization. As usual, $[x, y] = xy - yx$ for words x, y. Thus

$$\lambda(13245) = [\lambda(13), \lambda(245)] = [[1, 3], [2, [4, 5]]]$$
$$= 13245 - 13254 - 13452 + 13542 - 31245 + 31254 + 31452 - 31542$$
$$- 24513 + 25413 + 45213 - 54213 + 24531 - 25431 - 45231 + 54231$$

and

$$\lambda(1122) = [1, \lambda(122)] = [1, [\lambda(12), 2]]$$
$$= 1122 - 2(1212) + 2(2121) - 2211.$$

[GR89, Sec. 2] describes how to visualize the standard bracketing of a Lyndon word as a rooted binary tree: given a Lyndon word l with standard factorization $l = l_1 l_2$, inductively set T_l to be the tree with T_{l_1} as its left branch and T_{l_2} as its right branch. T_{13245} and T_{1122} are shown below.

T$_{13245}$ T$_{1122}$
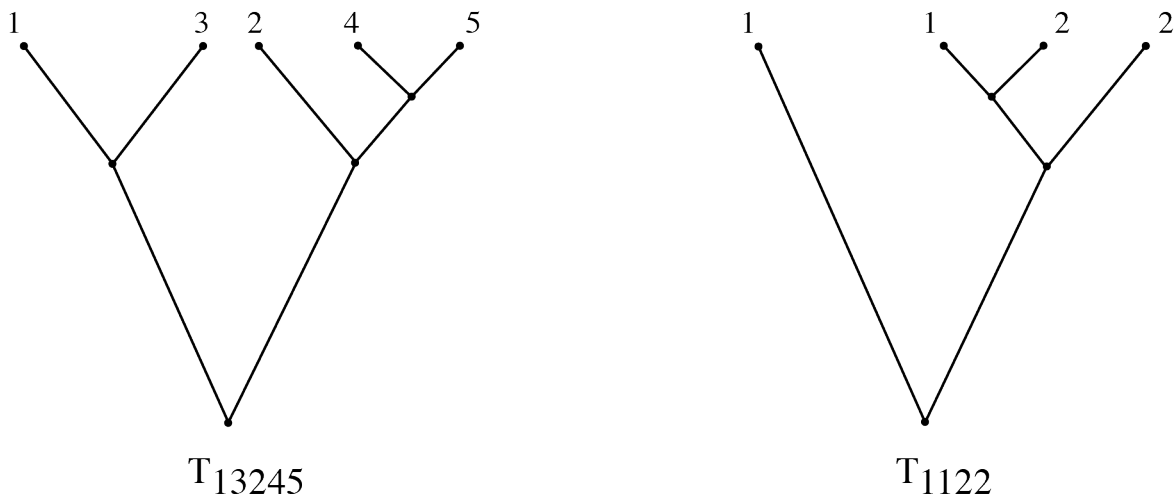
Observe that a word w appears in the expansion of $\lambda(l)$ only if, after exchanging the left and right branches at some vertices of T_l, the leaves of T_l, when read from left to right, spell out w. The coefficient of w in $\lambda(l)$ is then the signed number of ways to do this (the sign is the parity of the number of exchanges required). For example,

- 25413 has coefficient 1 in $\lambda(13245)$ since the unique way to rearrange T_{13245} so the leaves spell 25413 is to exchange the branches at the root and the highest interior vertex;

- 21345 does not appear in $\lambda(13245)$ since whenever the branches of T_{13245} switch, 2 must appear adjacent to either 4 or 5, which does not hold for 21345;

- 1221 has coefficient 0 in $\lambda(1122)$ as, to make the leaves of T_{1122} spell 1221, we can 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.

A final piece of notation is the following symmetrized product: let $w = l_1 l_2 \cdots l_k$ in Lyndon factorization. Then set

$$\text{sym}(w) = \sum_{\sigma \in S_k} \lambda(l_{\sigma(1)})\lambda(l_{\sigma(2)}) \cdots \lambda(l_{\sigma(k)}).$$

Viewing $\text{sym}(w)$ as a polynomial in the letters $w_1, w_2, \ldots, w_l$ will be useful for Theorem 3.16.

Garsia and Reutenauer's tree construction can be extended to visualize $\text{sym}(w)$, using what Barcelo and Bergeron [BB90] call *decreasing Lyndon hedgerows*, which simply consist of $T_{l_1}, T_{l_2}, \ldots, T_{l_k}$ placed in a row. Denote this as T_w also. The example T_{35142} is shown below.

T_{35142}
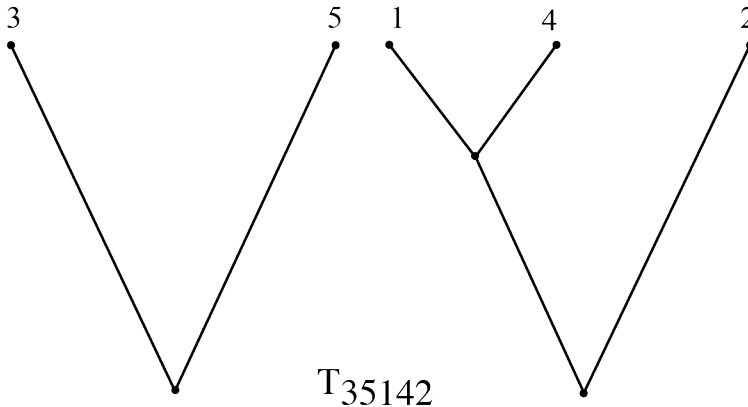
We can again express the coefficient of w' in $\text{sym}(w)$ as the signed number of ways to rearrange T_w so the leaves spell w'. Now there are two types of allowed moves: exchanging the left and right branches at a vertex, and permuting the trees of the hedgerow. The latter move does not come with a sign. Thus 14253 has coefficient -1 in $\text{sym}(35142)$, as the unique rearrangement of T_{35142} which spells 14253 requires transposing the trees and permuting the branches labeled 3 and 5.

It is clear from this pictorial description that every term appearing in $\text{sym}(w)$ is a permutation of the letters in w. [GR89, Th. 5.2] shows that $\{\text{sym}(w)\}$ form a basis for a free associative algebra (over a field of characteristic zero). This will turn out to be a left eigenbasis for inverse riffle shuffling, and similar theorems hold for other Hopf algebras.

2.4 Symmetric functions and beyond

A basic object of study is the (real) vector space Λ_n of homogeneous symmetric polynomials (in infinitely many variables) of degree n. The direct sum $\Lambda = \bigoplus_{n=0}^{\infty} \Lambda_n$ forms a graded algebra

with familiar bases: the monomial (m_λ), elementary (e_λ), homogeneous (h_λ), and power sums (p_λ). For example, $e_2 = \sum_{i<j} x_i x_j$ and for a partition $\lambda = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_l > 0$ with $\lambda_1 + \cdots + \lambda_l = n$, $e_\lambda = e_{\lambda_1} e_{\lambda_2} \cdots e_{\lambda_l}$. As λ ranges over partitions of n, $\{e_\lambda\}$ form a basis for Λ_n, from which we construct the rock-breaking chain of Example 1.2. Splendid accounts of symmetric function theory appear in [Mac95] and [Sta99]. A variety of Hopf algebra techniques are woven into these topics, as emphasized by [Gei77] and [Zel81]. The comprehensive account of noncommutative symmetric functions [GKL+95] and its follow-ups further the deep connection between combinatorics and Hopf algebras. However, this paper will only involve its dual, the algebra of quasisymmetric functions, as they encode informations about absorption rates of our chains, see Section 3.7. A basis of this algebra is given by the monomial quasisymmetric functions: for a composition $\alpha = (\alpha_1, \ldots, \alpha_k)$, define $M_\alpha = \sum_{i_1 < i_2 < \cdots < i_k} x_{i_1}^{\alpha_1} \cdots x_{i_k}^{\alpha_k}$. Further details are in [Sta99, Sec. 7.19].

3 Theory

3.1 Introduction

This section states and proves our main theorems. This introduction sets out definitions. Section 3.2 develops the reweighting schemes needed to have the Hopf-square maps give rise to Markov chains. Section 3.3 explains that these chains are often acyclic. Section 3.4 addresses a symmetrization lemma that we will use in Section 3.5 and Section 3.6 to find descriptions of some left and right eigenvectors respectively for such chains. Section 3.7 determines the stationary distributions and gives expressions for the chance of absorption in terms of generalized chromatic polynomials. Applications of these theorems are in the last three sections of this paper.

As mentioned at the end of Section 2.2, we will be concerned with connected, graded (by non-negative integers) Hopf algebras $\mathcal{H}$ over $\mathbb{R}$ with a distinguished basis $\mathcal{B}$ satisfying one of two "freeness" conditions (in both cases, the number of generators may be finite or infinite):

1. $\mathcal{H} = \mathbb{R}[c_1, c_2, \ldots]$ as an algebra (i.e., $\mathcal{H}$ is a polynomial algebra) and $\mathcal{B} = \{c_1^{n_1} c_2^{n_2} \ldots | n_i \in \mathbb{N}\}$, the basis of monomials. The c_i may have any degree, and there is no constraint on the coalgebra structure. This will give rise to a Markov chain on combinatorial objects where assembling is symmetric and deterministic.

2. $\mathcal{H}$ is cocommutative, $\mathcal{H} = \mathbb{R}\langle c_1, c_2, \ldots \rangle$ as an algebra, (i.e., $\mathcal{H}$ is a free associative algebra) and $\mathcal{B} = \{c_{i_1} c_{i_2} \ldots | i_j \in \mathbb{N}\}$, the basis of words. The c_i may have any degree, and do not need to be primitive. This will give rise to a Markov chain on combinatorial objects where pulling apart is symmetric, assembling is non-symmetric and deterministic.

By the Cartier-Milnor-Moore theorem [MM65, Car07], any (real) graded connected commutative Hopf algebra has a basis which satisfies the first condition. However, we will not make use of this, since the two conditions above are reasonable properties for many combinatorial Hopf algebras and their canonical bases. For example, the Hopf algebra of symmetric functions, with the basis of elementary symmetric functions e_λ, satisfies the first condition.

Write $\mathcal{H}_n$ for the subspace of degree n in $\mathcal{H}$, and $\mathcal{B}_n$ for the degree n basis elements. The *generators* c_i can be identified as those basis elements which are not the non-trivial product of basis elements; in other words, generators cannot be obtained by assembling objects of lower degree. Thus, all basis elements of degree one are generators, but there are usually generators of higher degree; see Examples 3.1 and 3.2 below. One can view the conditions 1 and 2 above as requiring the basis elements to have unique factorization into generators, allowing the convenient view of $b \in \mathcal{B}$ as

a word $b = c_1 c_2 \cdots c_l$. Its *length* $l(b)$ is then well-defined - it is the number of generators one needs to assemble together to produce b. Some properties of the length are developed in Section 3.3. For a noncommutative Hopf algebra, it is useful to choose a linear order on the set of generators refining the ordering by degree: i.e. if $\deg(c) < \deg(c')$, then $c < c'$. This allows the construction of the Lyndon factorization and standard bracketing of a basis element, as in Section 2.3. Example 3.17 demonstrates such calculations.

The ath Hopf-power map is $\Psi^a := m^{[a]}\Delta^{[a]}$, the a-fold coproduct followed by the a-fold product. These power maps are the central object of study of [Pat91, Pat93, Pat94]. Intuitively, Ψ^a corresponds to breaking an object into a pieces (some possibly empty) in all possible ways and then reassembling them. The Ψ^a preserve degree, thus mapping $\mathcal{H}_n$ to $\mathcal{H}_n$.

As noted in [Pat93], the power-map Ψ^a is an algebra homomorphism if $\mathcal{H}$ is commutative:

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

$$= \sum_{(x),(y)} x_{(1)}y_{(1)} \cdots x_{(a)}y_{(a)} = \sum_{(x),(y)} x_{(1)} \cdots x_{(a)} y_{(1)} \cdots y_{(a)} = \Psi^a(x)\Psi^a(y);$$

and a coalgebra homomorphism if $\mathcal{H}$ is cocommutative:

$$(\Psi^a \otimes \Psi^a)(\Delta x) = \sum_{(x)} \Psi^a\left(x_{(1)}\right) \otimes \Psi^a\left(x_{(2)}\right)$$

$$= \sum_{(x)} x_{(1)} \cdots x_{(a)} \otimes x_{(a+1)} \cdots x_{(2a)} = \sum_{(x)} x_{(1)}x_{(3)} \cdots x_{(2a-1)} \otimes x_{(2)}x_{(4)} \cdots x_{(2a)} = \Delta\Psi^a(x).$$

Only the former will be necessary for the rest of this section.

3.2 The Markov chain connection

The power maps can sometimes be interpreted as a natural Markov chain on the basis elements $\mathcal{B}_n$ of $\mathcal{H}_n$.

Example 3.1 (The Hopf algebra of unlabeled graphs, continuing from Example 2.1). The set of all unlabeled simple graphs gives rise to a Hopf algebra $\bar{\mathcal{G}}$ with disjoint union as product and

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

where the sum is over subsets of vertices S with G_S, G_{S^c} the induced subgraphs. Graded by the size of the vertex set, $\bar{\mathcal{G}}$ is a commutative and cocommutative polynomial Hopf algebra with basis $\mathcal{B}$ consisting of all graphs. The generators are precisely the connected graphs, and the length of a graph is its number of connected components.
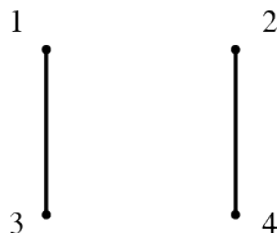
The resulting Markov chain on graphs with n vertices evolves as follows: from G, color the vertices of G red or blue, independently with probability $1/2$. Erase any edge with opposite colored vertices. This gives one step of the chain; the process terminates when there are no edges. Observe that each connected component breaks independently; that Δ is an algebra homomorphism ensures that, for any Hopf algebra, the generators break independently. The analogous Hopf algebra of simplicial complexes is discussed in Section 6.

Example 3.2 (The noncommutative Hopf algebra of labeled graphs, continuing from Example 2.2). Let $\mathcal{G}$ be the linear span of the simple graphs whose vertices are labeled $\{1, 2, \ldots, n\}$ for some n.

The product of two graphs $G_1 G_2$ is their disjoint union, where the vertices of G_1 keep their labels, and the labels in G_2 are increased by the number of vertices in G_1. The coproduct is

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

where the sum again runs over all subsets S of vertices of G, and G_S, G_{S^c} are relabeled so the vertices in each keep the same relative order. An example of a coproduct calculation is in Example 2.2. $\mathcal{G}$ is cocommutative and a free associative algebra; its distinguished basis $\mathcal{B}$ is the set of all graphs. A graph in $\mathcal{G}$ is a product if and only if there is an i such that no edge connects a vertex with label $\leq i$ to a vertex with label $> i$. Thus, all connected graphs are generators, but there are non-connected generators such as

Each step of the associated random walk on $\mathcal{B}_n$, the graphs with n vertices, has this description: from G, color the vertices of G red or blue, independently with probability $1/2$. Suppose r vertices received the color red; now erase any edge with opposite colored vertices, and relabel so the red vertices are $1, 2, \ldots, r$ and the blue vertices are $r+1, r+2, \ldots, n$, keeping their relative orders. For example, starting at the complete graph on three vertices, the chain reaches each of the graphs shown below with probability $1/8$:
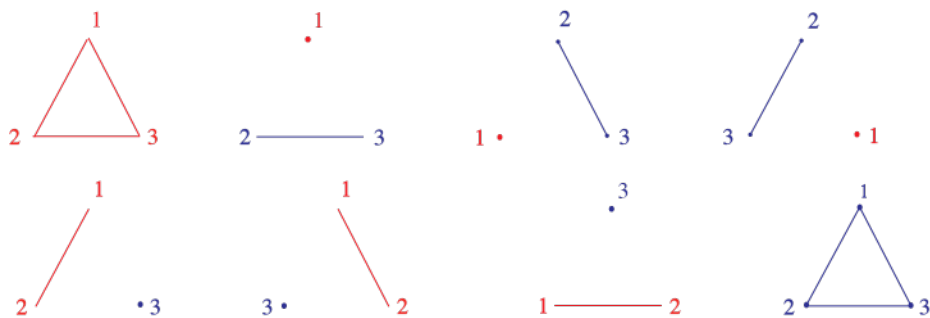
So, forgetting the colors of the vertices,

with probability 1/4

with probability 3/8

with probability 3/8
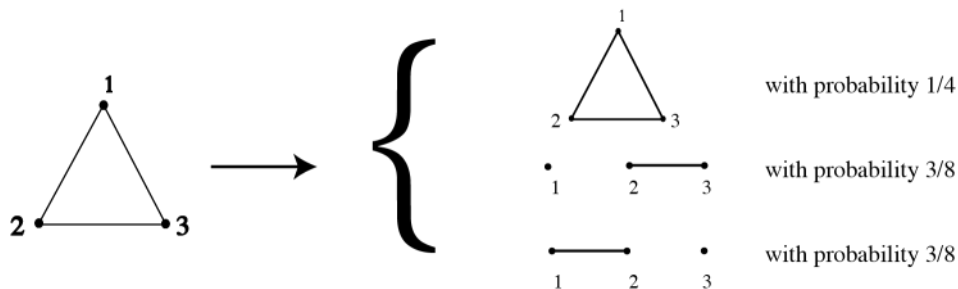
As with $\bar{\mathcal{G}}$, the chain on $\mathcal{G}_n$ stops when all edges have been removed.

When is such a probabilistic interpretation possible? To begin, the coefficients of $m\Delta(b)$ must be non-negative real numbers for $b \in \mathcal{B}$. This usually holds for combinatorial Hopf algebras, but the free associative algebra and the above algebras of graphs have an additional desirable property: for any $b \in \mathcal{B}$, the coefficients of $\Psi^2(b)$ sum to $2^{\deg(b)}$, regardless of b. Thus the operator $\frac{1}{2^n}\Psi^2(b) = \sum_{b'} K(b, b')b'$ forms a Markov transition matrix on basis elements of degree n. Indeed, the coefficients of $\Psi^a(b)$ sum to $a^{\deg(b)}$ for all a, so $\frac{1}{a^n}\Psi^a(b) = \sum_{b'} K_a(b, b')b'$ defines a transition matrix K_a. For other Hopf algebras, the sum of the coefficients in $\Psi^2(b)$ may depend on b, so simply scaling Ψ^2 does not always yield a transition matrix.

Zhou's rephrasing [Zho08, Lemma 4.4.1.1] of the Doob transform [LPW09, Sec.17.6.1] provides a solution: if K is a matrix with non-negative entries and ϕ is a strictly positive right eigenfunction of K with eigenvalue 1, then $\hat{K}(b, b') := \phi(b)^{-1}K(b, b')\phi(b')$ is a transition matrix. Here $\hat{K}$ is the conjugate of K by the diagonal matrix whose entries are $\phi(b)$. Theorem 3.4 below gives conditions for such ϕ to exist, and explicitly constructs ϕ recursively; Corollary 3.5 then specifies a non-recursive definition of ϕ when there is a sole basis element of degree 1. The following example explains why this construction is natural.

Example 3.3 (Symmetric functions and rock-breaking). Consider the algebra of symmetric functions with basis $\{e_\lambda\}$, the elementary symmetric functions. The length $l(e_\lambda)$ is the number of parts in the partition λ, and the generators are the partitions with a single part. The coproduct is defined by

$$\Delta(e_i) = \sum_{j=0}^{i} e_j \otimes e_{i-j} \qquad \text{so}$$

$$\Delta(e_\lambda) = \Delta(e_{\lambda_1})\cdots\Delta(e_{\lambda_l}) = \left(\sum_{i_1=0}^{\lambda_1} e_{i_1} \otimes e_{\lambda_1-i_1}\right)\cdots\left(\sum_{i_l=0}^{\lambda_l} e_{i_l} \otimes e_{\lambda_l-i_l}\right)$$

$$= \sum_{\lambda' \leq \lambda} e_{\lambda'} \otimes e_{\lambda-\lambda'}$$

with the sum over all compositions $\lambda' = \lambda'_1, \lambda'_2, \ldots, \lambda'_l$ with $0 \leq \lambda'_i \leq \lambda_i$, and $\lambda - \lambda'$ is the composition $\lambda_1 - \lambda'_1, \ldots, \lambda_l - \lambda'_l$. When reordered, some parts may be empty and some parts may occur several times. There are $(\lambda_1 + 1)\cdots(\lambda_l + 1)$ possible choices of λ', so the coefficients of $\Psi^2(e_\lambda)$ sum to $(\lambda_1 + 1)\cdots(\lambda_l + 1)$, which depends on λ.

Consider degree 2, where the basis elements are e_{1^2} and e_2. For K such that $\frac{1}{2^2}\Psi^2(b) = \sum_{b'} K(b, b')b'$,

$$K = \begin{bmatrix} 1 & 0 \\ \frac{1}{4} & \frac{1}{2} \end{bmatrix},$$

which is not a transition matrix as the second row does not sum to 1. Resolve this by performing a diagonal change of basis: set $\hat{e}_{1^2} = \phi(e_{1^2})^{-1}e_{1^2}$, $\hat{e}_2 = \phi(e_2)^{-1}e_2$ for some non-negative function $\phi : \mathcal{B} \to \mathbb{R}$, and consider $\hat{K}$ with $\frac{1}{2^2}\Psi^2(\hat{b}) = \sum_{\hat{b}'} \hat{K}(\hat{b}, \hat{b}')\hat{b}'$. Since the first row of K, corresponding to e_{1^2}, pose no problems, set $\phi(e_{1^2}) = 1$. In view of the upcoming theorem, it is better to think of this as $\phi(e_{1^2}) = (\phi(e_1))^2$ with $\phi(e_1) = 1$. Equivalently, $\hat{e}_{1^2} = \hat{e}_1^2$ with $\hat{e}_1 = e_1$. Turning attention to the second row, observe that $\Delta(\hat{e}_2) = \phi(e_2)^{-1}(e_2 \otimes 1 + e_1 \otimes e_1 + 1 \otimes e_2)$, so $\Psi^2(\hat{e}_2) = \hat{e}_2 + \phi(e_2)^{-1}\hat{e}_{1^2} + \hat{e}_2$, which means

$$\hat{K} = \begin{bmatrix} 1 & 0 \\ \frac{1}{4}\phi(e_2)^{-1} & \frac{1}{2} \end{bmatrix},$$

so $\hat{K}$ is a transition matrix if $\frac{1}{4}\phi(e_2)^{-1} + \frac{1}{2} = 1$, i.e. if $\phi(e_2) = \frac{1}{2}$.

Continue to degree 3, where the basis elements are e_{1^3}, e_{12} and e_3. Now define K such that $\frac{1}{2^3}\Psi^2(b) = \sum_{b'} K(b, b')b'$;

$$K = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & 0 \\ 0 & \frac{1}{4} & \frac{1}{4} \end{bmatrix}.$$

Again, look for $\phi(e_{1^3}), \phi(e_{12})$ and $\phi(e_3)$ so that $\hat{K}$, defined by $\frac{1}{2^3}\Psi^2(\hat{b}) = \sum_{\hat{b}'} \hat{K}(\hat{b}, \hat{b}')\hat{b}'$, is a transition matrix, where $\hat{e}_{1^3} = \phi(e_{1^3})^{-1}e_{1^3}$, $\hat{e}_{12} = \phi(e_{12})^{-1}e_{12}$, $\hat{e}_3 = \phi(e_3)^{-1}e_3$. Note that, taking $\phi(e_{1^3}) = (\phi(e_1))^3 = 1$ and $\phi(e_{12}) = \phi(e_2)\phi(e_1) = \frac{1}{2}$, the first two rows of $\hat{K}$ sum to 1. View this as $\hat{e}_{1^3} = \hat{e}_1^3$ and $\hat{e}_{12} = \hat{e}_2\hat{e}_1$. Then, as $\Psi^2(\hat{e}_3) = \phi(e_3)^{-1}(e_3 + e_2e_1 + e_1e_2 + e_3) = \hat{e}_3 + \frac{1}{2}\phi(e_3)^{-1}\hat{e}_{2,1} + \frac{1}{2}\phi(e_3)^{-1}\hat{e}_{2,1} + \hat{e}_3$, the transition matrix is given by

$$\hat{K} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{8}\phi(e_3)^{-1} & \frac{1}{4} \end{bmatrix}$$

and choosing $\phi(e_3) = \frac{1}{6}$ makes the third row sum to 1.

Continuing, we find that $\phi(e_i) = \frac{1}{i!}$, so $\hat{e}_i = i!e_i$, more generally, $\hat{e}_\lambda = \prod(i!)^{a_i(\lambda)}e_\lambda$ with i appearing $a_i(\lambda)$ times in λ. Then, for example,

$$m\Delta(\hat{e}_n) = n!m\Delta(e_n) = n!m\sum_{i=0}^{n} e_i \otimes e_{n-i} = \sum_{i=0}^{n}\binom{n}{i}\hat{e}_i\hat{e}_{n-i}.$$

So, for any partition λ of n with l parts,

$$\begin{aligned} m\Delta(\hat{e}_\lambda) &= m\Delta(\hat{e}_{\lambda_1})\cdots m\Delta(\hat{e}_{\lambda_l}) \\ &= \sum_{\lambda' \le \lambda}\binom{\lambda_1}{\lambda'_1}\binom{\lambda_2}{\lambda'_2}\cdots\binom{\lambda_l}{\lambda'_l}\hat{e}_{\lambda'}\hat{e}_{\lambda-\lambda'} \end{aligned}$$

and the coefficients of $m\Delta(\hat{e}_\lambda)$ sum to $\sum_{\lambda' \le \lambda}\binom{\lambda_1}{\lambda'_1}\cdots\binom{\lambda_l}{\lambda'_l} = 2^{\lambda_1}\cdots 2^{\lambda_l} = 2^n$, irrespective of λ. Thus $\frac{1}{2^n}m\Delta$ describes a transition matrix, which has the rock-breaking interpretation of Section 1.

The following theorem shows that this algorithm works in many cases. Observe that, in the above example, it is the non-zero off-diagonal entries that change; the diagonal entries cannot be changed by rescaling the basis. Hence the algorithm would fail if some row had all off-diagonal entries equal to 0, and diagonal entry not equal to 1. This corresponds to the existence of $b \in \mathcal{B}_n$ with $\frac{1}{2^n}\Psi^2(b) = \alpha b$ for some $\alpha \ne 1$; the condition $\bar{\Delta}(c) := \Delta(c) - 1 \otimes c - c \otimes 1 \ne 0$ below precisely prevents this. Intuitively, we are requiring that each generator of degree greater than one can be broken non-trivially. For an example where this condition fails, see Example 6.5.

Theorem 3.4 (Basis rescaling). *Let $\mathcal{H}$ be a graded Hopf algebra over $\mathbb{R}$ which is either a polynomial algebra or a free associative algebra that is cocommutative. Let $\mathcal{B}$ denote the basis of monomials in the generators. Suppose that, for all generators c with $\deg(c) > 1$, all coefficients of $\Delta(c)$ (in the $\mathcal{B} \otimes \mathcal{B}$ basis) are non-negative and $\bar{\Delta}(c) \ne 0$. Let K_a be the transpose of the matrix of $a^{-n}\Psi^a$ with respect to the basis $\mathcal{B}_n$; in other words, $a^{-n}\Psi^a(b) = \sum_{b'} K_a(b, b')b'$ (suppressing the dependence of K_a on n). Define, by induction on degree,*

$$\hat{c} = c \hspace{4cm} \text{if } \deg(c) = 1$$

$$\hat{c} = \frac{1 - 2^{1-\deg(c)}}{\sum_{b \ne c}\phi(b)K_2(c, b)}c \hspace{2cm} \text{for a generator } c \text{ with } \deg(c) > 1$$

$$\hat{b} = \hat{c}_1 \ldots \hat{c}_l \hspace{3cm} \text{for } b \in \mathcal{B} \text{ with factorization into generators } c_1 \ldots c_l$$

15

where $\phi(b)$ satisfies $b = \phi(b)\hat{b}$. Write $\hat{\mathcal{B}} := \{\hat{b}|b \in \mathcal{B}\}$ and $\hat{\mathcal{B}}_n := \{\hat{b}|b \in \mathcal{B}_n\}$. Then the matrix of the ath power map with respect to the $\hat{\mathcal{B}}_n$ basis, when transposed and multiplied by a^{-n}, is a transition matrix. In other words, the operator $\hat{K}_a$ on $\mathcal{H}_n$, defined by $a^{-n}\Psi^a(\hat{b}) = \sum_{b'} \hat{K}_a(\hat{b}, \hat{b}')\hat{b}' = \sum_{b'} \phi(b)^{-1}K_a(b, b')\phi(b')b'$, has $\hat{K}_a(\hat{b}, \hat{b}') \geq 0$ and $\sum_{b'} \hat{K}_a(\hat{b}, \hat{b}') = 1$ for all $b \in \mathcal{B}_n$, and all $a \geq 0$ and $n \geq 0$ (the same scaling works simultaneously for all a).

Remarks. 1. Observe that, if $b = xy$, then the definition of $\hat{b}$ ensures $\hat{b} = \hat{x}\hat{y}$. Equivalently, ϕ is a multiplicative function.

2. The definition of $\hat{c}$ is not circular: since $\mathcal{H}$ is graded with $\mathcal{H}_0 = \mathbb{R}$, the counit is zero on elements of positive degree so that $\bar{\Delta}(c) \in \bigoplus_{j=1}^{\deg(c)-1} \mathcal{H}_j \otimes \mathcal{H}_{\deg(c)-j}$. Hence $K_2(c, b)$ is non-zero only if $b = c$ or $l(b) > 1$, so the denominator in the expression for $\hat{c}$ only involves $\phi(b)$ for b with $l(b) > 1$. Such b can be factorized as $b = xy$ with $\deg(x), \deg(y) < \deg(b)$, whence $\phi(b) = \phi(x)\phi(y)$, so $\hat{c}$ only depends on $\phi(x)$ with $\deg(x) < \deg(c)$.

Proof. First note that $\hat{K}_2(c, c) = \phi(c)^{-1}K_2(c, c)\phi(c) = K_2(c, c) = 2^{1-\deg(c)}$, since $m\Delta(c) = 2c + m\bar{\Delta}(c)$ and $\bar{\Delta}(c) \in \bigoplus_{j=1}^{\deg(c)-1} \mathcal{H}_j \otimes \mathcal{H}_{\deg(c)-j}$ means no c terms can occur in $m\bar{\Delta}(c)$. So

$$
\begin{aligned}
\sum_{b'} \hat{K}_2(\hat{c}, \hat{b}') &= 2^{1-\deg(c)} + \phi(c)^{-1} \sum_{b' \neq c} K_2(c, b')\phi(b') \\
&= 2^{1-\deg(c)} + \frac{1 - 2^{1-\deg(c)}}{\sum_{b' \neq c} K_2(c, b')\phi(b)} \sum_{b' \neq c} K_2(c, b')\phi(b') \\
&= 1,
\end{aligned}
$$

as desired.

Let η_c^{xy} denote the coefficients of $\Delta(c)$ in the $\mathcal{B} \otimes \mathcal{B}$ basis, so $\Delta(c) = \sum_{x,y \in \mathcal{B}} \eta_c^{xy} x \otimes y$. Then $K_2(c, b) = 2^{-\deg(c)} \sum_{xy=b} \eta_c^{xy}$, and

$$
\begin{aligned}
\hat{K}_2(\hat{c}, \hat{b}) &= 2^{-\deg(c)} \sum_{xy=b} \phi(c)^{-1}\eta_c^{xy}\phi(b) \\
&= 2^{-\deg(c)} \sum_{xy=b} \phi(c)^{-1}\eta_c^{xy}\phi(x)\phi(y).
\end{aligned}
$$

So, if b has factorization into generators $b = c_1 \ldots c_l$, then

$$
\begin{aligned}
\Delta(b) &= \Delta(c_1) \ldots \Delta(c_l) \\
&= \sum_{\substack{x_1,\ldots,x_l \\ y_1,\ldots,y_l}} \eta_{c_1}^{x_1 y_1} \ldots \eta_{c_l}^{x_l y_l} x_1 \ldots x_l \otimes y_1 \ldots y_l,
\end{aligned}
$$

so

$$
K_2(b, b') = 2^{-\deg(b)} \sum_{x_1 \ldots x_l y_1 \ldots y_l = b'} \eta_{c_1}^{x_1 y_1} \ldots \eta_{c_l}^{x_l y_l}.
$$

Thus

$$
\begin{aligned}
\sum_{b'} \hat{K}_2(\hat{b}, \hat{b}') &= 2^{-\deg(b)} \sum_{b'} \phi(b)^{-1} K_2(b, b') \phi(b') \\
&= 2^{-\deg(b)} \sum_{\substack{x_1,\ldots,x_l \\ y_1,\ldots,y_l}} \phi(b)^{-1} \eta_{c_1}^{x_1 y_1} \ldots \eta_{c_l}^{x_l y_l} \phi(x_1 \ldots x_l y_1 \ldots y_l) \\
&= \prod_{i=1}^{l} 2^{-\deg(c_i)} \sum_{x_i, y_i} \phi(c_i)^{-1} \eta_{c_i}^{x_i y_i} \phi(x_i) \phi(y_i) \\
&= \prod_{i=1}^{l} \sum_{b_i} \hat{K}_2(\hat{c}_i, \hat{b}_i) \\
&= 1
\end{aligned}
$$

as desired, where the third equality is due to multiplicativity of ϕ.

The above showed each row of $\hat{K}_2$ sums to 1, which means $(1, 1, \ldots, 1)$ is a right eigenvector of $\hat{K}_2$ of eigenvalue 1. The matrix $\hat{K}_a$ describes Ψ^a in the $\hat{\mathcal{B}}$ basis, which is also a basis of monomials/words, in a rescaled set of generators $\hat{c}$, so, by Theorems 3.19 and 3.20, the eigenspaces of $\hat{K}_a$ do not depend on a. Hence $(1, 1, \ldots, 1)$ is a right eigenvector of $\hat{K}_a$ of eigenvalue 1 for all a, thus each row of $\hat{K}_a$ sums to 1 also.

Finally, to see that the entries of $\hat{K}_a$ are non-negative, first extend the notation η_c^{xy} so $\Delta^{[a]}(c) = \sum_{b_1,\ldots,b_a} \eta_c^{b_1,\ldots,b_a} b_1 \otimes \cdots \otimes b_a$. As $\Delta^{[a]} = (\iota \otimes \cdots \otimes \iota \otimes \Delta) \Delta^{[a-1]}$, it follows that $\eta_c^{b_1,\ldots,b_a} = \sum_x \eta_c^{b_1,\ldots,b_{a-2},x} \eta_x^{b_{a-1},b_a}$, which inductively shows that $\eta_c^{b_1,\ldots,b_a} \geq 0$ for all generators c and all $b_i \in \mathcal{B}$. So, if b has factorization into generators $b = c_1 \ldots c_l$, then

$$
K_a(b, b') = \sum \eta_{c_1}^{b_{1,1},\ldots,b_{1,a}} \ldots \eta_{c_l}^{b_{l,1},\ldots,b_{l,a}} \geq 0,
$$

where the sum is over all sets $\{b_{i,j}\}_{i=1,j=1}^{i=l,j=a}$ such that the product $b_{1,1} b_{2,1} \ldots b_{l,1} b_{1,2} \ldots b_{l,2} \ldots b_{1,a} \ldots b_{l,a} = b'$. Finally, $\hat{K}_a(\hat{b}, \hat{b}') = \phi(b)^{-1} K_a(b, b') \phi(b') \geq 0$. $\qquad \square$

Combinatorial Hopf algebras often have a single basis element of degree 1 - for the algebra of symmetric functions, this is the unique partition of 1; for the Hopf algebra $\mathcal{G}$ of graphs, this is the discrete graph with one vertex. After the latter example, denote this basis element by $\bullet$. Then there is a simpler definition of the eigenfunction ϕ, and hence $\hat{b}$ and $\hat{K}$, in terms of $\eta_b^{b_1,\ldots,b_r}$, the coefficient of $b_1 \otimes \cdots \otimes b_r$ in $\Delta^{[r]}(b)$:

Corollary 3.5. *Suppose that, in addition to the hypotheses of Theorem 3.4, $\mathcal{B}_1 = \{\bullet\}$. Then $\hat{b} = \frac{(\deg b)!}{\eta_b^{\bullet,\ldots,\bullet}} b$, so $\hat{K}_a$ is defined by*

$$
\hat{K}_a(\hat{b}, \hat{b}') = \frac{\eta_{b'}^{\bullet,\ldots,\bullet}}{\eta_b^{\bullet,\ldots,\bullet}} K_a(b, b')
$$

Proof. Work on $\mathcal{H}_n$ for a fixed degree n. Recall that ϕ is a right eigenvector of $\hat{K}_a$ of eigenvalue 1, and hence, by the notation of Section 3.6, an eigenvector of Ψ^{*a} of eigenvalue a^n. By Theorems 3.19 and 3.20, this eigenspace is spanned by f_b for b with length n. Then $\mathcal{B}_1 = \{\bullet\}$ forces $b = \bullet^n$, so $f_{\bullet^n}(b') = \frac{1}{n!} \eta_{b'}^{\bullet,\ldots,\bullet}$ spans the a^n-eigenspace of Ψ^{*a}. Consequently, ϕ is a multiple of $f_{\bullet^n}$. To determine this multiplicative factor, observe that Theorem 3.4 defines $\phi(\bullet)$ to be 1, so $\phi(\bullet^n) = 1$, and $f_{\bullet^n}(\bullet^n) = 1$ also, so $\phi = f_{\bullet^n}$. $\qquad \square$

3.3 Acyclicity

Observe that the rock-breaking chain (Examples 1.2 and 3.3) is *acyclic* - it can never return to a state it has left, because the only way to leave a state is to break the rocks into more pieces. More specifically, at each step the chain either stays at the same partition or moves to a partition which refines the current state; as refinement of partitions is a partial order, the chain cannot return to a state it has left. The same is true for the chain on unlabeled graphs (Example 3.1) - the number of connected components increases over time, and the chain never returns to a previous state. Such behavior can be explained by the way the length changes under the product and coproduct. (Recall that the length $l(b)$ is the number of factors in the unique factorization of b into generators.) Define a relation on $\mathcal{B}$ by $b \to b'$ if b' appears in $\Psi^a(b)$ for some a. If Ψ^a induces a Markov chain on $\mathcal{B}_n$, then this precisely says that b' is accessible from b.

Lemma 3.6. *Let $b, b_i, b_{(i)}$ be monomials/words in a Hopf algebra which is either a polynomial algebra or a free associative algebra that is cocommutative. Then*

(i) $l(b_1 \ldots b_a) = l(b_1) + \cdots + l(b_a)$;

(ii) For any summand $b_{(1)} \otimes \cdots \otimes b_{(a)}$ in $\Delta^{[a]}(b)$, $l(b_{(1)}) + \cdots + l(b_{(a)}) \geq l(b)$;

(iii) if $b \to b'$, then $l(b') \geq l(b)$.

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

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

(iii) follows trivially from (i) and (ii): if $b \to b'$, then $b' = b_{(1)} \ldots b_{(a)}$ for a term $b_{(1)} \otimes \cdots \otimes b_{(a)}$ in $\Delta^{[a]}(b)$. So $l(b') = l(b_{(1)}) + \cdots + l(b_{(a)}) \geq l(b)$. $\square$

If $\mathcal{H}$ is a polynomial algebra, more is true. The following proposition explains why chains built from polynomial algebras (i.e., with deterministic and symmetric assembling) are always acyclic; in probability language, it says that, if the current state is built from l generators, then, with probability a^{l-n}, the chain stays at this state, otherwise, it moves to a state built from more generators. Hence, 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.

Proposition 3.7. *(Acyclicity) Let $\mathcal{H}$ be a Hopf algebra which is a polynomial algebra as an algebra, and $\mathcal{B}$ its monomial basis. Assume also that all coefficients of the coproduct of the generators (in the $\mathcal{B} \otimes \mathcal{B}$ basis) are non-negative. Then the relation $\to$ defines a partial order on $\mathcal{B}$, and the ordering by length refines this order: if $b \to b'$ and $b \neq b'$, then $l(b) < l(b')$. Furthermore, for any integer a and any $b \in \mathcal{B}$ with length $l(b)$,*

$$\Psi^a(b) = a^{l(b)}b + \sum_{l(b') > l(b)} \alpha_{bb'} b'$$

for some $\alpha_{bb'} \geq 0$.

Proof. It is easier to first prove the expression for $\Psi^a(b)$. Suppose b has factorization into generators $b = c_1 c_2 \ldots c_{l(b)}$. As $\mathcal{H}$ is commutative, Ψ^a is an algebra homomorphism, so $\Psi^a(b) = \Psi^a(c_1) \ldots \Psi^a(c_{l(b)})$. 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 $1 \otimes \cdots \otimes 1 \otimes c, 1 \otimes \cdots \otimes 1 \otimes c \otimes 1, \ldots, c \otimes 1 \otimes \cdots \otimes 1$. So $\Psi^a(c) = ac + \sum_{l(b')>1} \alpha_{cb'} b'$ for generators c, and $\alpha_{cb'} \geq 0$ as we assumed that $\Delta(c)$ has non-negative coefficients. As $\Psi^a(b) = \Psi^a(c_1)...\Psi^a(c_l)$, and length is multiplicative (Lemma 3.6 (i)), the expression for $\Psi^a(b)$ follows.

It is then clear that $\to$ is reflexive and antisymmetric. Transitivity follows from the power rule: if $b \to b'$ and $b' \to b''$, then b' appears in $\Psi^a(b)$ for some a and b'' appears in $\Psi^{a'}(b')$ for some a'. So b'' appears in $\Psi^{a'}\Psi^a(b) = \Psi^{a'a}(b)$. (The non-negative coefficient assumption ensures that the b'' term in $\Psi^{a'a}(b)$ cannot cancel out due to contributions from an intermediary different from b'.) $\square$

The same argument applied to a cocommutative free associative algebra shows that all terms in $\Psi^a(b)$ are either a permutation of the factors of b, or have length greater than that of b. The relation $\to$ is only a preorder; the associated chains are not acyclic, as they may oscillate between such permutations of factors. For example, in the noncommutative Hopf algebra of labeled graphs, the following transition probabilities can occur:
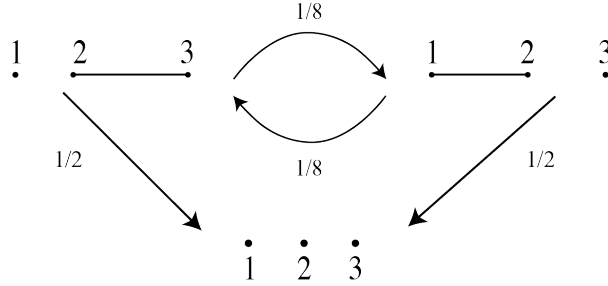
(the bottom state is absorbing). The probability of going from b to some permutation of its factors (as opposed to a state of greater length, from which there is no return to b) is $a^{l(b)-n}$.

Here is one more result in this spirit, necessary in Section 3.5 to show that the eigenvectors constructed there have good triangularity properties and hence form an eigenbasis:

Lemma 3.8. *Let b, b_i, b_i' be monomials/words in a Hopf algebra which is either a polynomial algebra or a free associative algebra that is cocommutative. If $b = b_1 \ldots b_k$ and $b_i \to b_i'$ for each i, then $b \to b_{\sigma(1)}' \ldots b_{\sigma(k)}'$ for any $\sigma \in S_k$.*

Proof. For readability, take $k = 2$ and write $b = xy$, $x \to x'$, $y \to y'$. By definition of the relation $\to$, it must be that $x' = x_{(1)} \ldots x_{(a)}$ for some summand $x_{(1)} \otimes \cdots \otimes x_{(a)}$ of $\bar{\Delta}^{[a]}(x)$. Likewise $y' = y_{(1)} \ldots y_{(a')}$ for some a'. Suppose $a > a'$. Coassociativity implies that $\Delta^{[a]}(y) = (\iota \otimes \cdots \otimes \iota \otimes \Delta^{[a-a']})\Delta^{[a']}(y)$, and $y_{(a')} \otimes 1 \otimes \cdots \otimes 1$ is certainly a summand of $\Delta^{[a-a']}(y_{(a')})$, so $y_{(1)} \otimes \cdots \otimes y_{(a')} \otimes 1 \otimes \cdots \otimes 1$ occurs in $\Delta^{[a]}(y)$. So, taking $y_{(a'+1)} = \cdots = y_{(a)} = 1$, we can assume $a = a'$. Then $\Delta^{[a]}(b) = \Delta^{[a]}(x)\Delta^{[a]}(y)$ contains the term $x_{(1)}y_{(1)} \otimes \cdots \otimes x_{(a)}y_{(a)}$. Hence $\Psi^a(b)$ contains the term $x_{(1)}y_{(1)} \ldots x_{(a)}y_{(a)}$, and this product is $x'y'$ if $\mathcal{H}$ is a polynomial algebra.

If $\mathcal{H}$ is a cocommutative, free associative algebra, the factors in $x_{(1)}y_{(1)} \otimes \cdots \otimes x_{(a)}y_{(a)}$ must be rearranged to conclude that $b \to x'y'$ and $b \to y'x'$. Coassociativity implies $\Delta^{[2a]} = (\Delta \otimes \cdots \otimes \Delta)\Delta^{[a]}$,

19

and $\Delta \left(x_{(i)} y_{(i)}\right) = \Delta \left(x_{(i)}\right) \Delta \left(y_{(i)}\right)$ contains $\left(x_{(i)} \otimes 1\right) \left(1 \otimes y_{(i)}\right) = x_{(i)} \otimes y_{(i)}$, so $\Delta^{[2a]}(b)$ contains the term $x_{(1)} \otimes y_{(1)} \otimes x_{(2)} \otimes y_{(2)} \otimes \cdots \otimes x_{(a)} \otimes y_{(a)}$. As $\mathcal{H}$ is cocommutative, any permutation of the tensor-factors, in particular, $x_{(1)} \otimes x_{(2)} \otimes \cdots \otimes x_{(a)} \otimes y_{(1)} \otimes \cdots \otimes y_{(a)}$ and $y_{(1)} \otimes y_{(2)} \otimes \cdots \otimes y_{(a)} \otimes x_{(1)} \otimes \cdots \otimes x_{(a)}$, must also be summands of $\Delta^{[2a]}(b)$, and multiplying these tensor-factors together shows that both $x'y'$ and $y'x'$ appear in $\Psi^{[2a]}(b)$. $\qquad\square$

Example 3.9 (Symmetric functions and rock-breaking). Recall from Example 3.3 the algebra of symmetric functions with basis $\{e_\lambda\}$, which induces the rock-breaking process. Here, $e_\lambda \to e_{\lambda'}$ if and only if λ' refines λ. Lemma 3.8 for the case $k = 2$ is the statement that, if λ is the union of two partitions μ and ν, and μ' refines μ, ν' refines ν, then $\mu' \amalg \nu'$ refines $\mu \amalg \nu = \lambda$.

3.4 The symmetrization lemma

The algorithmic construction of left and right eigenbases for the chains created in Section 3.2 will go as follows:

(i) Make an eigenvector of smallest eigenvalue for each generator c;

(ii) For each basis element b with factorization $c_1 c_2 ... c_l$, build an eigenvector of larger eigenvalue out of the eigenvectors corresponding to the factors c_i, produced in the previous step.

Concentrate on the left eigenvectors for the moment. Recall that the transition matrix K_a is defined by $a^{-n} \Psi^a(b) = \sum_{b'} K_a(b, b') b'$, so the left eigenvectors for our Markov chain are the usual eigenvectors of Ψ^a on $\mathcal{H}$. Step (ii) is simple if $\mathcal{H}$ is a polynomial algebra, because then $\mathcal{H}$ is commutative so Ψ^a is an algebra homomorphism. Consequently, the product of two eigenvectors is an eigenvector with the product eigenvalue. This fails for cocommutative, free associative algebras $\mathcal{H}$, but can be fixed by taking symmetrized products:

Theorem 3.10 (Symmetrization lemma). *Let $x_1, x_2, \ldots, x_k$ be primitive elements of any Hopf algebra $\mathcal{H}$, then $\sum_{\sigma \in S_k} x_{\sigma(1)} x_{\sigma(2)} \ldots x_{\sigma(k)}$ is an eigenvector of Ψ^a with eigenvalue a^k.*

Proof. For concreteness, take $a = 2$. Then

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

$\qquad\square$

In Section 3.5 and Section 3.6, the fact that the eigenvectors constructed give a basis will follow from triangularity arguments based on Section 3.3. These rely heavily on the explicit structure of

a polynomial algebra or a free associative algebra. Hence it is natural to look for alternatives that will generalize this eigenbasis construction plan to Hopf algebras with more complicated structures. For example, one may ask whether some good choice of x_i exists with which the symmetrization lemma will automatically generate a full eigenbasis. When $\mathcal{H}$ is cocommutative, an elegant answer stems from the following two well-known structure theorems:

Theorem 3.11 (Cartier-Milnor-Moore). *[MM65, Car07] If $\mathcal{H}$ is graded, cocommutative and connected, then $\mathcal{H}$ is Hopf 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 3.12 (Poincaré–Birkhoff–Witt). *[Hum72, Lot97]. If $\{x_1, x_2, ...\}$ is a basis for a Lie algebra $\mathfrak{g}$, then the symmetrized products $\sum_{\sigma \in S_k} x_{i_{\sigma(1)}} x_{i_{\sigma(2)}}...x_{i_{\sigma(k)}}$, for $1 \leq i_1 \leq i_2 \leq \cdots \leq i_k$, form a basis for $\mathcal{U}(\mathfrak{g})$.*

Putting these together reduces the diagonalization of Ψ^a on a cocommutative Hopf algebra to determining a basis of primitive elements:

Theorem 3.13 (Strong symmetrization lemma). *Let $\mathcal{H}$ be a graded, cocommutative, connected Hopf algebra, and let $\{x_1, x_2, ...\}$ be a basis for the subspace of primitive elements in $\mathcal{H}$. Then, for each $k \in \mathbb{N}$,*

$$\left\{ \sum_{\sigma \in S_k} x_{i_{\sigma(1)}} x_{i_{\sigma(2)}}...x_{i_{\sigma(k)}} \middle| 1 \leq i_1 \leq i_2 \leq \cdots \leq i_k \right\}$$

is a basis of the a^k-eigenspace of Ψ^a.

Much work [Fis10, AS05, AS06] has been done on computing a basis for the subspace of the primitives of particular Hopf algebras, their formulas are in general more efficient than our universal method here, and using these will be the subject of future work. Alternatively, the theory of good Lyndon words [LR95] gives a Grobner basis argument to further reduce the problem to finding elements which generate the Lie algebra of primitives, and understanding the relations between them. This is the motivation behind our construction of the eigenvectors in Theorem 3.16, although the proof is independent of this theorem, more analogous to that of Theorem 3.15, the case of a polynomial algebra.

3.5 Left eigenfunctions

This section gives an algorithmic construction of an eigenbasis for the Hopf power maps Ψ^a on the Hopf algebras of interest. If K_a as defined by $a^{-n}\Psi^a(b) = \sum_{b'} K_a(b, b')b'$ is a transition matrix, then this eigenbasis is precisely a left eigenbasis of the associated chain, though the results below stand whether or not such a chain may be defined (e.g., the construction works when some coefficients of $\Delta(c)$ are negative, and when there are primitive generators of degree > 1). The first step is to associate each generator to an eigenvector of smallest eigenvalue, this is achieved using the *(first) Eulerian idempotent* map

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

Here $\bar{\Delta}(x) = \Delta(x) - 1 \otimes x - x \otimes 1 \in \bigoplus_{j=1}^{\deg x - 1} \mathcal{H}_j \otimes \mathcal{H}_{\deg x - j}$, as explained in Section 2.2. Then inductively define $\bar{\Delta}^{[a]} = (\iota \otimes \cdots \otimes \iota \otimes \bar{\Delta})\bar{\Delta}^{[a-1]}$, which picks out the terms in $\Delta^{[a]}(x)$ where each tensor-factor has strictly positive degree. This captures the notion of breaking into a non-trivial

pieces. Observe that, if $x \in \mathcal{H}_n$, then $\bar{\Delta}^{[a]}(x) = 0$ whenever $a > n$, so $e(x)$ is a finite sum for all x. (By convention, $e \equiv 0$ on $\mathcal{H}_0$.)

This map e is the first of a series of Eulerian idempotents e_i defined by Patras [Pat93]; he proves that, in a commutative or cocommutative Hopf algebra of characteristic zero where $\bar{\Delta}$ is *locally nilpotent* (i.e. for each x, there is some a with $\bar{\Delta}^{[a]}x = 0$), the Hopf-powers are diagonalizable, and these e_i are orthogonal projections onto the eigenspaces. In particular, this *weight decomposition* holds for graded commutative or cocommutative Hopf algebras. We will not need the full series of Eulerian idempotents, although Example 3.18 makes the connection between them and our eigenbasis.

To deduce that the eigenvectors we construct are triangular with respect to $\mathcal{B}$, one needs the following crucial observation (recall from Section 3.3 that $b \to b'$ if b' occurs in $\Psi^a(b)$ for some a):

Proposition 3.14. *For any generator c,*

$$e(c) = c + \sum_{\substack{c \to b' \\ b' \neq c}} \alpha_{cb'} b' = c + \sum_{l(b') > 1} \alpha_{cb'} b',$$

for some real $\alpha_{cb'}$.

Proof. The summand $\frac{(-1)^{a-1}}{a} m^{[a]} \bar{\Delta}^{[a]}(c)$ involves terms of length at least a, from which the second expression of $e(c)$ is immediate. Each term b' of $e(c)$ appears in $\Psi^a(c)$ for some a, hence $c \to b'$. Combine this with the knowledge from the second expression that c occurs with coefficient 1 to deduce the first expression. $\square$

The two theorems below detail the construction of an eigenbasis for Ψ^a in a polynomial algebra and in a cocommutative free associative algebra respectively. These are left eigenvectors for the corresponding transition matrices. A worked example will follow immediately; it may help to read these together.

Theorem 3.15. *Let $\mathcal{H}$ be a Hopf algebra (over a field of characteristic zero) that is a polynomial algebra as an algebra, with monomial basis $\mathcal{B}$. For $b \in \mathcal{B}$ with factorization into generators $b = c_1 c_2 ... c_l$, set*

$$g_b := e(c_1) e(c_2) ... e(c_l).$$

Then g_b is an eigenvector of Ψ^a of eigenvalue a^l satisfying the triangularity condition

$$g_b = b + \sum_{\substack{b \to b' \\ b' \neq b}} g_b(b') b' = b + \sum_{l(b') > l(b)} g_b(b') b'.$$

Hence $\{g_b | b \in \mathcal{B}_n\}$ is an eigenbasis for the action of Ψ^a on $\mathcal{H}_n$, and the multiplicity of the eigenvalue a^l in $\mathcal{H}_n$ is the coefficient of $x^n y^l$ in $\prod_i (1 - yx^i)^{-d_i}$, where d_i is the number of generators of degree i.

Theorem 3.16. *Let $\mathcal{H}$ be a cocommutative Hopf algebra (over a field of characteristic zero) that is a free associative algebra with word basis $\mathcal{B}$. For $b \in \mathcal{B}$ with factorization into generators $b = c_1 c_2 ... c_l$, set g_b to be the polynomial $\mathrm{sym}(b)$ evaluated at $(e(c_1), e(c_2), \ldots, e(c_l))$. In other words, in the terminology of Section 2.3,*

- *for c a generator, set $g_c := e(c)$;*

22

- *for b a Lyndon word, inductively define $g_b := [g_{b_1}, g_{b_2}]$ where $b = b_1 b_2$ is the standard factorization of b;*

- *for b with Lyndon factorization $b = b_1 \ldots b_k$, set $g_b := \sum_{\sigma \in S_k} g_{b_{\sigma(1)}} g_{b_{\sigma(2)}} \cdots g_{b_{\sigma(k)}}$.*

Then g_b is an eigenvector of Ψ^a of eigenvalue a^k (k the number of Lyndon factors in b) satisfying the triangularity condition

$$g_b = \sum_{b \to b'} g_b(b') b' = \operatorname{sym}(b) + \sum_{l(b') > l(b)} g_b(b') b'.$$

Hence $\{g_b | b \in \mathcal{B}_n\}$ is an eigenbasis for the action of Ψ^a on $\mathcal{H}_n$, and the multiplicity of the eigenvalue a^k in $\mathcal{H}_n$ is the coefficient of $x^n y^k$ in $\prod_i \left(1 - yx^i\right)^{-d_i}$, where d_i is the number of Lyndon words of degree i in the alphabet of generators.

Remarks. 1. If Ψ^a defines a Markov chain, then the triangularity of g_b (in both theorems) has the following interpretation: the left eigenfunction g_b takes non-zero values only on states that are reachable from b.

2. The expression of the multiplicity of the eigenvalues (in both theorems) holds for Hopf algebras that are multigraded, if we replace all xs, ns and is by tuples, and read the formula as multi-index notation. For example, for a bigraded polynomial algebra $\mathcal{H}$, the multiplicity of the a^l-eigenspace in $\mathcal{H}_{m,n}$ is the coefficient of $x_1^m x_2^n y^l$ in $\prod_{i,j} \left(1 - yx_1^i x_2^j\right)^{-d_{i,j}}$, where $d_{i,j}$ is the number of generators of bidegree (i, j). This idea will be useful in Section 5.

3. Theorem 3.16 essentially states that any cocommutative free associative algebra is in fact isomorphic to *the* free associative algebra, generated by $e(c)$. But there is no analogous interpretation for Theorem 3.15; being a polynomial algebra is not a strong enough condition to force all Hopf algebras with this condition to be isomorphic. A polynomial algebra $\mathcal{H}$ is isomorphic to the usual polynomial Hopf algebra (i.e. with primitive generators) only if $\mathcal{H}$ is cocommutative; then $e(c)$ gives a set of primitive generators.

Example 3.17. As promised, here is a worked example of this calculation, in the noncommutative Hopf algebra of labeled graphs, as defined in Example 3.2. Let b be the graph

$$\overset{1}{\cdot} \quad \overset{2}{\underline{\qquad}}\overset{3}{} \quad \overset{4}{\cdot} \;=\; \left(\overset{1}{\cdot} \right)\left(\overset{1}{\underline{\qquad}}\overset{2}{} \right)\left(\overset{1}{\cdot} \right)$$

which is the product of three generators as shown. (Its factors happen to be its connected components, but that's not always true). Since the ordering of generators refines the ordering by degree, a vertex (degree 1) comes before an edge (degree 2), so the Lyndon factorization of b is

$$\left(\overset{1}{\cdot} \quad \overset{2}{\underline{\qquad}}\overset{3}{} \right) \overset{1}{\cdot}$$

So g_b is defined to be

$$\mathbf{g}_b = \mathbf{g}_{\overset{1}{\cdot}\;\overset{2}{\underline{\quad}}\overset{3}{}} \; \mathbf{g}_{\overset{1}{\cdot}} + \mathbf{g}_{\overset{1}{\cdot}} \, \mathbf{g}_{\overset{1}{\cdot}\;\overset{2}{\underline{\quad}}\overset{3}{}}$$

The first Lyndon factor of b has standard factorization

$$\underset{1}{\bullet}\ \underset{2}{\rule{1.2cm}{0.4pt}}\underset{3}{} \ = \ \left(\ \underset{1}{\bullet}\ \right)\left(\ \underset{1}{\rule{1.2cm}{0.4pt}}\underset{2}{}\ \right)$$

so

$$\mathbf{g}\ \underset{1}{\bullet}\ \underset{2}{\rule{1cm}{0.4pt}}\underset{3}{}\ =\ \left[\mathbf{e}\left(\underset{1}{\bullet}\right),\mathbf{e}\left(\underset{1}{\rule{1cm}{0.4pt}}\underset{2}{}\right)\right]$$

The Eulerian idempotent map fixes the single vertex, and

$$\mathbf{e}\left(\underset{1}{\rule{1cm}{0.4pt}}\underset{2}{}\right)\ =\ \underset{1}{\rule{1cm}{0.4pt}}\underset{2}{}\ -\ \frac{1}{2}\ \mathbf{m}\,\overline{\triangle}\ \underset{1}{\rule{1cm}{0.4pt}}\underset{2}{}\ =\ \underset{1}{\rule{1cm}{0.4pt}}\underset{2}{}\ -\ \underset{1}{\bullet}\ \underset{2}{\bullet}$$

thus substituting into the previous equation gives

$$\mathbf{g}\ \underset{1}{\bullet}\ \underset{2}{\rule{1cm}{0.4pt}}\underset{3}{}\ =\ \underset{1}{\bullet}\left(\underset{1}{\rule{1cm}{0.4pt}}\underset{2}{}\ -\ \underset{1}{\bullet}\ \underset{2}{\bullet}\right)\ -\ \left(\underset{1}{\rule{1cm}{0.4pt}}\underset{2}{}\ -\ \underset{1}{\bullet}\ \underset{2}{\bullet}\right)\underset{1}{\bullet}$$

$$=\ \underset{1}{\bullet}\ \underset{2}{}\ \underset{3}{\rule{1cm}{0.4pt}}\ -\ \underset{1}{\rule{1cm}{0.4pt}}\ \underset{2}{}\ \underset{3}{\bullet}$$

Since

$$\mathbf{g}\,\underset{1}{\bullet}\ =\ \mathbf{e}\,(\underset{1}{\bullet})\ =\ \underset{1}{\bullet}$$

returning to the first expression for g_b gives the following eigenvector of eigenvalue a^2:

$$\mathbf{g_b}\ =\ \left(\underset{1}{\bullet}\ \underset{2}{}\ \underset{3}{\rule{1cm}{0.4pt}}\ -\ \underset{1}{\rule{1cm}{0.4pt}}\ \underset{2}{}\ \underset{3}{\bullet}\right)\underset{1}{\bullet}\ +\ \underset{1}{\bullet}\left(\underset{1}{\bullet}\ \underset{2}{}\ \underset{3}{\rule{1cm}{0.4pt}}\ -\ \underset{1}{\rule{1cm}{0.4pt}}\ \underset{2}{}\ \underset{3}{\bullet}\right)$$

$$=\ \underset{1}{\bullet}\ \underset{2}{\bullet}\ \underset{3}{}\ \underset{4}{\rule{1cm}{0.4pt}}\ -\ \underset{1}{\rule{1cm}{0.4pt}}\ \underset{2}{}\ \underset{3}{\bullet}\ \underset{4}{\bullet}$$

Proof of Theorem 3.15 (polynomial algebra). By Patras [Pat93], the Eulerian idempotent map is a projection onto the a-eigenspace of Ψ^a, so, for each generator c, $e(c)$ is an eigenvector of eigenvalue a. As $\mathcal{H}$ is commutative, Ψ^a is an algebra homomorphism, so the product of two eigenvectors is another eigenvector with the product eigenvalue. Hence $g_b := e(c_1)e(c_2)...e(c_l)$ is an eigenvector of eigenvalue a^l.

To see triangularity, note that, by Proposition 3.14,

$$\begin{aligned}
g_b\ &=\ \left(c_1 + \sum_{\substack{c_1 \to c_1' \\ c_1' \neq c_1}} g_{c_1}(c_1')c_1'\right) \cdots \left(c_l + \sum_{\substack{c_l \to c_l' \\ c_l' \neq c_l}} g_{c_l}(c_l')c_l'\right)\\
&=\ b\ +\ \sum_{\substack{c_i \to c_i' \\ c_i' \neq c_i \text{ for some } i}} g_{c_1}(c_1')\ldots g_{c_l}(c_l')c_1'\ldots c_l'.
\end{aligned}$$

24

Lemma 3.8 shows that $b \to c'_1 \ldots c'_l$ in each summand, and the condition $c'_i \neq c_i$ for some i means precisely that $c'_1 \ldots c'_l \neq b$. Also, by Proposition 3.14,

$$
\begin{aligned}
g_b &= \left(c_1 + \sum_{l(c'_1)>1} g_{c_1}(c'_1)c'_1 \right) \ldots \left(c_l + \sum_{l(c'_l)>1} g_{c_l}(c'_l)c'_l \right) \\
&= b + \sum_{l(c'_i)>1 \text{ for some } i} g_{c_1}(c'_1) \ldots g_{c_l}(c'_l) c'_1 \ldots c'_l
\end{aligned}
$$

and thus $l(c'_1 \ldots c'_l) > l$ as length is multiplicative.

The multiplicity of the eigenvalue a^l is the number of basis elements b with length l. The last assertion of the theorem is then immediate from [Wil90, Th. 3.14.1]. $\qquad\square$

Example 3.18. We show that $g_b = e_{l(b)}(b)$, where the higher Eulerian idempotents are defined by

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

By Patras [Pat93], e_i is a projection to the a^i-eigenspace of Ψ^a, so, given the triangularity condition of the eigenbasis $\{g_b\}$, it suffices to show that b is the only term of length $l(b)$ in $e_{l(b)}(b)$. Note that $e_{l(b)}(b)$ is a sum of terms of the form $e(b_{(1)})e(b_{(2)})...e(b_{(l)})$ for some $b_{(i)}$ with $b_{(1)} \otimes \cdots \otimes b_{(l)}$ a summand of $\Delta^{[l]}(b)$. As $e \equiv 0$ on $\mathcal{H}_0$, the b_is must be non-trivial. Hence each term b' of $e_{l(b)}(b)$ has the form $b' = b'_{(1)} \ldots b'_{(l)}$, with $b_{(i)} \to b'_{(i)}$ and $b \to b_{(1)} \ldots b_{(l)}$. It follows from Lemma 3.8 that $b_{(1)} \ldots b_{(l)} \to b'_{(1)} \ldots b'_{(l)}$, so $b \to b'$ by transitivity, which, by Lemma 3.7 means $l(b') > l(b)$ unless $b' = b$.

It remains to show that the coefficient of b in $e_{l(b)}(b)$ is 1. Let $b = c_1 \ldots c_l$ be the factorization of b into generators. With notation from the previous paragraph, taking $b' = b$ results in $b \to b_{(1)} \ldots b_{(l)} \to b'_{(1)} \ldots b'_{(l)} = b$, so $b = b_{(1)} \ldots b_{(l)}$. This forces the $b_{(i)} = c_{\sigma(i)}$ for some $\sigma \in S_l$. As $b_{(i)}$ occurs with coefficient 1 in $e(b_{(i)})$, the coefficient of $b_{(1)} \otimes \cdots \otimes b_{(l)}$ in $(e \otimes \cdots \otimes e)\Delta^{[l]}(b)$ is the coefficient of $c_{\sigma(1)} \otimes \cdots \otimes c_{\sigma(l)}$ in $\Delta^{[l]}(b) = \Delta^{[l]}(c_1) \ldots \Delta^{[l]}(c_l)$, which is 1 for each $\sigma \in S_l$. Each occurrence of $c_{\sigma(1)} \otimes \cdots \otimes c_{\sigma(l)}$ in $(e \otimes \cdots \otimes e)\Delta^{[l]}(b)$ gives rise to a b term in $m^{[l]}(e \otimes \cdots \otimes e)\Delta^{[l]}(b)$ with the same coefficient, for each $\sigma \in S_l$, hence b has coefficient $l!$ in $m^{[l]}(e \otimes \cdots \otimes e)\Delta^{[l]}(b) = l! e_l(b)$.

The same argument also shows that, if $i < l(b)$, then $e_i(b) = 0$, as there is no term of length i in $e_i(b)$. In particular, $e(b) = 0$ if b is not a generator.

Proof of Theorem 3.16 (cocommutative and free associative algebra). Schmitt [Sch94, Thm 9.4] shows that the Eulerian idempotent map e projects a graded cocommutative algebra onto its subspace of primitive elements, so $g_c := e(c)$ is primitive. A straightforward calculation shows that, if $x, y \in \mathcal{H}$ are primitive, then so is $[x, y]$. Iterating this implies that, if b is a Lyndon word, then g_b (which is the standard bracketing of $e(c)$s) is primitive. Now apply the symmetrization lemma (Lemma 3.10) to deduce that, if $b \in \mathcal{B}$ has k Lyndon factors, g_b is an eigenvector of eigenvalue a^k.

To see triangularity, first recall that sym is a linear combination of the permutations of its arguments, hence g_b is a linear combination of products of the form $e(c_{\sigma(1)}) \ldots e(c_{\sigma(l)})$ for some $\sigma \in S_l$. Hence, by Proposition 3.14, each term in g_b has the form $c'_{\sigma(1)} \ldots c'_{\sigma(l)}$ with $c_i \to c'_i$, and by

Lemma 3.8, we have $b \to c'_{\sigma(1)} \ldots c'_{\sigma(l)}$. Also, by Proposition 3.14,

$$
\begin{aligned}
g_b &= \operatorname{sym}\left(\left(c_1 + \sum_{l(c'_1)>1} g_{c_1}(c'_1)c'_1\right), \ldots, \left(c_l + \sum_{l(c'_l)>1} g_{c_l}(c'_l)c'_l\right)\right) \\
&= \operatorname{sym}(b) + \sum_{l(c'_i)>1 \text{ for some } i} \operatorname{sym}\left(g_{c_1}(c'_1)c'_1, \ldots, g_{c_l}(c'_l)c'_l\right),
\end{aligned}
$$

and all terms of the sum have length greater than l, as length is multiplicative, and sym is a linear combination of the permutations of its arguments.

The multiplicity of the eigenvalue a^k is the number of basis elements with k Lyndon factors. The last assertion of the theorem is then immediate from [Wil90, Th. 3.14.1]. $\qquad\square$

3.6 Right eigenvectors

To obtain the right eigenvectors for our Markov chains, consider the graded dual $\mathcal{H}^*$ of the algebras examined above. The multiplication Δ^* and comultiplication m^* on $\mathcal{H}^*$ are given by:

$$
\begin{aligned}
(x^*y^*)(z) &= (\Delta^*(x^* \otimes y^*))(z) = (x^* \otimes y^*)(\Delta z) \\
(m^*x^*)(z \otimes w) &= \quad x^*m(z \otimes w) = x^*(zw)
\end{aligned}
$$

for any $x^*, y^* \in \mathcal{H}^*$, $z, w \in \mathcal{H}$. Then $\Psi^{*a} := \Delta^{*[a]}m^{*[a]}$ is the dual map to Ψ^a. So, if K_a, defined by $a^{-n}\Psi^a(b) = \sum_{b'} K_a(b, b')b'$, is a transition matrix, then its right eigenvectors are the eigenvectors of Ψ^{*a}. The theorems below express these eigenvectors in terms of $\{b^*\}$, the dual basis to $\mathcal{B}$. Dualizing a commutative Hopf algebra creates a cocommutative Hopf algebra, and vice versa, so Theorem 3.19 below, which diagonalizes Ψ^{*a} on a polynomial algebra, will share features with Theorem 3.16, which diagonalizes Ψ^a on a cocommutative free associative algebra. Similarly, Theorem 3.20 and Theorem 3.15 will involve common ideas. However, Theorems 3.19 and 3.20 are not direct applications of Theorems 3.16 and 3.15 to $\mathcal{H}^*$ as $\mathcal{H}^*$ is not a polynomial or free associative algebra - a breaking and recombining chain with a deterministic recombination does not dualize to one with a deterministic recombination. For example, the recombination step is deterministic for inverse shuffling (place the left pile on top of the right pile), but not for forward riffle shuffling (shuffle the two piles together).

The two theorems below give the eigenvectors of Ψ^{*a}; exemplar computations are in Section 4.2. Theorem 3.19 gives a complete description of these for $\mathcal{H}$ a polynomial algebra, and Theorem 3.20 yields a partial description for $\mathcal{H}$ a cocommutative free associative algebra. Recall that $\eta_b^{b_1,\ldots,b_a}$ is the coefficient of $b_1 \otimes \cdots \otimes b_a$ in $\Delta^{[a]}(b)$.

Theorem 3.19. *Let $\mathcal{H}$ be a Hopf algebra (over a field of characteristic zero) that is a polynomial algebra as an algebra, with monomial basis $\mathcal{B}$. For $b \in \mathcal{B}$ with factorization into generators $b = c_1c_2...c_l$, set*

$$
f_b := \frac{1}{A(b)l!} \sum_{\sigma \in S_l} c^*_{\sigma(1)}c^*_{\sigma(2)} \ldots c^*_{\sigma(l)},
$$

where the normalizing constant $A(b)$ is calculated as follows: for each generator c, let $a_c(b)$ be the power of c in the factorization of b, and set $A(b) = \prod_c a_c(b)!$. Then f_b is an eigenvector of Ψ^{*a} of

eigenvalue a^l, and

$$
\begin{aligned}
f_b(b') &= \frac{1}{A(b)l!} \sum_{\sigma \in S_l} \eta_{b'}^{c_{\sigma(1)}, c_{\sigma(2)}, \ldots, c_{\sigma(l)}} \\
&= \frac{1}{l!} \sum_{\sigma} \eta_{b'}^{c_{\sigma(1)}, c_{\sigma(2)}, \ldots, c_{\sigma(l)}},
\end{aligned}
$$

where the sum on the second line runs over all σ with $\big(c_{\sigma(1)}, \ldots, c_{\sigma(l)}\big)$ distinct (i.e., sum over all coset representatives of the stabilizer of $(c_1, \ldots, c_l)$). The eigenvector f_b satisfies the triangularity condition

$$
f_b = b^* + \sum_{\substack{b' \to b \\ b' \neq b}} f_b(b') b'^* = b^* + \sum_{l(b') < l(b)} f_b(b') b'^*.
$$

Furthermore, $\{f_b\}$ is the dual basis to $\{g_b\}$. In other words, $f_b(g_{b'}) = 0$ if $b \neq b'$, and $f_b(g_b) = 1$. Yet another formulation: the change of basis matrix from $\{f_b\}$ to the basis dual to $\mathcal{B}$, which has f_b as its columns, is the inverse of the matrix with g_b as its rows.

Remarks.   1. *If $\mathcal{H}$ is also cocommutative, then it is unnecessary to symmetrize - just define $f_b = \frac{1}{A(b)} c_1^* c_2^* \ldots c_l^*$.*

    2. *If Ψ^a defines a Markov chain on $\mathcal{B}_n$, then the theorem says $f_b(b')$ may be interpreted as the number of ways to break b' into l pieces so that the result is some permutation of the l generators that are factors of b. In particular, f_b takes only non-negative values, and f_b is non-zero only on states which can reach b. Thus f_b may be used to estimate the probability of being in states that can reach b, see Corollary 4.10 for an example.*

Theorem 3.20. *Let $\mathcal{H}$ be a cocommutative Hopf algebra (over a field of characteristic zero) which is a free associative algebra as an algebra, with word basis $\mathcal{B}$. For each Lyndon word b, let f_b be the eigenvector of Ψ^{*a} of eigenvalue a such that $f_b(g_b) = 1$ and $f_b(g_{b'}) = 0$ for all other Lyndon b'. In particular, $f_c = c^*$ and is primitive. For each basis element b with Lyndon factorization $b = b_1 \ldots b_k$, let*

$$
f_b := \frac{1}{A'(b)} f_{b_1} \ldots f_{b_k},
$$

where the normalizing constant $A'(b)$ is calculated as follows: for each Lyndon basis element b', let $a'_{b'}(b)$ be the number of times b' occurs in the Lyndon factorization of b, and set $A'(b) = \prod_{b'} a'_{b'}(b)!$. Then f_b is an eigenvector of Ψ^{*a} of eigenvalue a^k, and $\{f_b\}$ is the dual basis to $\{g_b\}$. If $b = c_1 c_2 \ldots c_l$ with $c_1 \geq c_2 \geq \cdots \geq c_l$ in the ordering of generators, then

$$
f_b(b') = \frac{1}{A'(b)} \eta_{b'}^{c_1, \ldots, c_l}.
$$

Proof of Theorem 3.19 (polynomial algebra). Suppose $b^* \otimes b'^*$ is a term in $m^*(c^*)$, where c is a generator. This means $m^*(c^*)(b \otimes b')$ is non-zero. Since comultiplication in $\mathcal{H}^*$ is dual to multiplication in $\mathcal{H}$, $m^*(c^*)(b \otimes b') = c^*(bb')$, which is only nonzero if bb' is a (real number) multiple of c. Since c is a generator, this can only happen if one of b, b' is c. Hence c^* is primitive. Apply the symmetrization lemma (Lemma 3.10) to the primitives $c_1^*, \ldots, c_l^*$ to deduce that f_b as defined above is an eigenvector of eigenvalue a^l.

Since multiplication in $\mathcal{H}^*$ is dual to the coproduct in $\mathcal{H}$, $c^*_{\sigma(1)}c^*_{\sigma(2)}\cdots c^*_{\sigma(l)}(b') = c^*_{\sigma(1)} \otimes c^*_{\sigma(2)} \otimes \cdots \otimes c^*_{\sigma(l)}\left(\Delta^{[l]}(b')\right)$, from which the first expression for $f_b(b')$ is immediate. To deduce the second expression, note that the size of the stabilizer of $\left(c_{\sigma(1)}, c_{\sigma(2)}, \ldots, c_{\sigma(l)}\right)$ under S_l is precisely $A(b)$.

It is apparent from the formula that b'^* appears in f_b only if $c_{\sigma(1)}\ldots c_{\sigma(l)} = b$ appears in $\Psi^l(b')$, hence $b' \to b$ is necessary. To calculate the leading coefficient $f_b(b)$, note that this is the sum over S_l of the coefficients of $c_{\sigma(1)} \otimes \cdots \otimes c_{\sigma(l)}$ in $\Delta^{[l]}(b) = \Delta^{[l]}(c_1)\ldots\Delta^{[l]}(c_l)$. Each term in $\Delta^{[l]}(c_i)$ contributes at least one generator to at least one tensor-factor, and each tensor-factor of $c_{\sigma(1)} \otimes \cdots \otimes c_{\sigma(l)}$ is a single generator, so each occurrence of $c_{\sigma(1)} \otimes \cdots \otimes c_{\sigma(l)}$ is a product of terms from $\Delta^{[l]}(c_i)$ where one tensor-factor is c_i and all other tensor-factors are 1. Such products are all $l!$ permutations of the c_i in the tensor-factors, so, for each fixed σ, the coefficient of $c_{\sigma(1)} \otimes \cdots \otimes c_{\sigma(l)}$ in $\Delta^{[l]}(b)$ is $A(b)$. This proves the first equality in the triangularity statement. Triangularity of f_b with respect to length follows, as ordering by length refines the relation $\to$ (Proposition 3.7).

To see duality, first note that, since Ψ^{*a} is the linear algebra map dual to Ψ^a, $f_b(\Psi^a g_{b'}) = (\Psi^{*a} f_b)(g_{b'})$. Now, using that f_b and g_b are eigenvectors, it follows that

$$a^{l(b')}f_b(g_{b'}) = f_b(\Psi^a g_{b'}) = \Psi^{*a} f_b(g_{b'}) = a^{l(b)}f_b(g_{b'}),$$

so $f_b(g_{b'}) = 0$ if $l(b') \neq l(b)$.

Now suppose $l(b') = l(b) = l$. Then

$$f_b(g_{b'}) = \left(b^* + \sum_{l(b_1)<l} f_b(b_1)b_1^*\right)\left(b' + \sum_{l(b_2)>l} g_{b'}(b_2)b_2\right) = b^*(b'),$$

which is 0 when $b \neq b'$, and 1 when $b = b'$. $\qquad\square$

Proof of Theorem 3.20 (cocommutative and free associative algebra). $\mathcal{H}^*$ is commutative, so the power map is an algebra homomorphism. Then, since f_b is defined as the product of k eigenvectors each of eigenvalue a, f_b is an eigenvector of eigenvalue a^k.

For any generator c, c^* is primitive by the same reasoning as in Theorem 3.19, the case of a polynomial algebra. To check that $c^*(g_c) = 1$ and $c^*(g_b) = 0$ for all other Lyndon b, use the triangularity of g_b:

$$c^*(g_b) = c^*(\text{sym}(b)) + \sum_{l(b')>l(b)} g_b(b')c^*(b').$$

Each summand $c^*(b')$ in the second term is 0 as $l(c) = 1 \leq l(b) < l(b')$. As $\text{sym}(b)$ consists of terms of length $l(b)$, $c^*(\text{sym}(b))$ is 0 unless $l(b) = 1$, in which case $\text{sym}(b) = b$. Hence $c^*(g_b) = c^*(\text{sym}(b))$ is non-zero only if $c = b$, and $c^*(g_c) = c^*(\text{sym}(c)) = c^*(c) = 1$.

Turn now to duality. An analogous argument to the polynomial algebra case shows that $f_b(g_{b'}) \neq 0$ only when they have the same eigenvalue, which happens precisely when b and b' have the same number of Lyndon factors. So let $b_1\ldots b_k$ be the decreasing Lyndon factorization of b, and let $b'_1\ldots b'_k$ be the decreasing Lyndon factorization of b'. To evaluate

$$f_b(g_{b'}) = \frac{1}{A'(b)}f_{b_1}\cdots f_{b_k}\left(\sum_{\sigma\in S_k} g_{b'_{\sigma(1)}}\cdots g_{b'_{\sigma(k)}}\right),$$

observe that

$$
\begin{aligned}
f_{b_1}\cdots f_{b_k}\left(g_{b'_{\sigma(1)}}\cdots g_{b'_{\sigma(k)}}\right) &= (f_{b_1}\otimes\cdots\otimes f_{b_k})\,\Delta^{[k]}\left(g_{b'_{\sigma(1)}}\cdots g_{b'_{\sigma(k)}}\right)\\
&= (f_{b_1}\otimes\ldots\otimes f_{b_k})\left(\Delta^{[k]}\left(g_{b'_{\sigma(1)}}\right)\ldots\Delta^{[k]}\left(g_{b'_{\sigma(k)}}\right)\right).
\end{aligned}
$$

28

As $g_{b'_{\sigma(i)}}$ is primitive, each term in $\Delta^{[k]}\left(g_{b'_{\sigma(i)}}\right)$ has $g_{b'_{\sigma(i)}}$ in one tensor-factor and 1 in the $k-1$ others. Hence the only terms of $\Delta^{[k]}\left(g_{b'_{\sigma(1)}}\right)...\Delta^{[k]}\left(g_{b'_{\sigma(k)}}\right)$ without 1s in any tensor factors are those of the form $g_{b'_{\tau\sigma(1)}}\otimes\cdots\otimes g_{b'_{\tau\sigma(k)}}$ for some $\tau\in S_k$. Now $f_{b_i}(1)=0$ for all i, so $f_{b_1}\otimes\cdots\otimes f_{b_k}$ annihilates any term with 1 in some tensor-factor. Hence

$$
\begin{aligned}
f_{b_1}\cdots f_{b_k}\left(\sum_{\sigma\in S_k} g_{b'_{\sigma(1)}}\cdots g_{b'_{\sigma(k)}}\right) &= f_{b_1}\otimes\cdots\otimes f_{b_k}\left(\sum_{\sigma,\tau\in S_k} g_{b'_{\tau\sigma(1)}}\otimes\cdots\otimes g_{b'_{\tau\sigma(k)}}\right) \\
&= \sum_{\sigma,\tau\in S_k} f_{b_1}\left(g_{b'_{\tau\sigma(1)}}\right)\cdots f_{b_k}\left(g_{b'_{\tau\sigma(k)}}\right).
\end{aligned}
$$

As f_b is dual to g_b for Lyndon b, the only summands which contribute are when $b_i = b'_{\sigma\tau(i)}$ for all i. In other words, this is zero unless the b_i are some permutation of the b'_i. But both sets are ordered decreasingly, so this can only happen if $b_i = b'_i$ for all i, hence $b = b'$. In that case, for each fixed $\sigma\in S_k$, the number of $\tau\in S_k$ with $b_i = b_{\sigma\tau(i)}$ for all i is precisely $A'(b)$, so $f_b(g_b) = 1$.

The final statement is proved in the same way as in Theorem 3.19, for a polynomial algebra, since, when $b = c_1 c_2...c_l$ with $c_1 \geq c_2 \geq \cdots \geq c_l$ in the ordering of generators, $f_b = \frac{1}{A'(b)}c_1^* c_2^* \ldots c_l^*$. $\square$

3.7 Stationary distributions, generalized chromatic polynomials, and absorption times

This section returns to probabilistic considerations, showing how the left eigenvectors of Section 3.5 determine the stationary distribution of the associated Markov chain. In the absorbing case, "generalized chromatic symmetric functions", based on the universality theorem in [ABS06], determine rates of absorption. Again, these general theorems are illustrated in the three sections that follow.

3.7.1 Stationary distributions

The first proposition identifies all the absorbing states when $\mathcal{H}$ is a polynomial algebra:

Proposition 3.21. *Suppose $\mathcal{H}$ is a polynomial algebra where K_a, defined by $a^{-n}\Psi^a(b) = \sum_{b'} K_a(b, b')b'$, is a transition matrix. Then the absorbing states are the basis elements $b\in\mathcal{B}_n$ which are products of n (possibly repeated) degree one elements, and these give a basis of the 1-eigenspace of K_a.*

Example 3.22. In the commutative Hopf algebra of graphs in Examples 2.1 and 3.1, there is a unique basis element of degree 1 - the graph with a single vertex. Hence the product of n such, which is the empty graph, is the unique absorbing state. Similarly for the rock-breaking example (symmetric functions) on partitions of n, the only basis element of degree 1 is e_1 and the stationary distribution is absorbing at 1^n (or e_{1^n}).

The parallel result for a cocommutative free associative algebra picks out the stationary distributions:

Proposition 3.23. *Suppose $\mathcal{H}$ is a cocommutative and free associative algebra where K_a, defined by $a^{-n}\Psi^a(b) = \sum_{b'} K_a(b, b')b'$, is a transition matrix. Then, for each unordered n-tuple $\{c_1, c_2, \ldots, c_n\}$ of degree 1 elements (some c_is may be identical), the uniform distribution on $\{c_{\sigma(1)}c_{\sigma(2)}\ldots c_{\sigma(n)}|\sigma\in S_n\}$ is a stationary distribution for the associated chain. In particular, all absorbing states have the form $\bullet^n$, where $\bullet\in\mathcal{B}_1$.*

29

Example 3.24. In the free associative algebra $\mathbb{R}\langle x_1, x_2, \ldots, x_n \rangle$, each x_i is a degree 1 element. So the uniform distribution on $x_{\sigma(1)} \ldots x_{\sigma(n)}$ ($\sigma \in S_n$) is a stationary distribution, as evident from considering inverse shuffles.

Proof of Proposition 3.21. From Theorem 3.15, a basis for the 1-eigenspace is $\{g_b | l(b) = n\}$. This forces each factor of b to have degree 1, so $b = c_1 c_2 \ldots c_n$ and $g_b = e(c_1) \ldots e(c_n)$. Now $e(c) = \sum_{a \geq 1} \frac{(-1)^{a-1}}{a} m^{[a]} \bar{\Delta}^{[a]}(c)$, and, when $\deg(c) = 1$, $m^{[a]} \bar{\Delta}^{[a]}(c) = 0$ for all $a \geq 2$. So $e(c) = c$, and hence $g_b = c_1 c_2 \ldots c_n = b$, which is a point mass on b, so $b = c_1 c_2 \ldots c_n$ is an absorbing state. $\quad\square$

Proof of Proposition 3.23. From Theorem 3.16, a basis for the 1-eigenspace is $\{g_b | b \in \mathcal{B}_n, b \text{ has } n \text{ Lyndon factors}\}$. This forces each Lyndon factor of b to have degree 1, so each of these must in fact be a single letter of degree 1. Thus $b = c_1 c_2 \ldots c_n$ and $g_b = \sum_{\sigma \in S_n} g_{c_{\sigma(1)}} \cdots g_{c_{\sigma(n)}} = \sum_{\sigma \in S_n} c_{\sigma(1)} \cdots c_{\sigma(n)}$, as $g_c = c$ for a generator c. An absorbing state is a stationary distribution which is a point mass. This requires $c_{\sigma(1)} \cdots c_{\sigma(n)}$ to be independent of σ. As $\mathcal{H}$ is a free associative algebra, this only holds when $c_1 = \cdots = c_n =: \bullet$, in which case $g_b = n! \bullet^n$, so $\bullet^n$ is an absorbing state. $\quad\square$

3.7.2 Absorption and chromatic polynomials

Consider the case where there is a single basis element of degree 1; call this element $\bullet$ as in Section 3.2. Then, by Proposition 3.21 and Proposition 3.23, the K_a chain has a unique absorbing basis vector $\bullet^n \in \mathcal{H}_n$. The chance of absorption after k steps can be rephrased in terms of an analog of the chromatic polynomial. Note first that the property $K_a * K_{a'} = K_{aa'}$ implies it is enough to calculate $K_a(b, \bullet^n)$ for general a and starting state $b \in \mathcal{H}_n$. To do this, make $\mathcal{H}$ into a combinatorial Hopf algebra in the sense of [ABS06] by defining a character ζ that takes value 1 on $\bullet$ and value 0 on all other generators, and extend multiplicatively and linearly. In other words, ζ is an indicator function of absorption, taking value 1 on all absorbing states and 0 on all other states. By [ABS06, Th. 4.1] there is a unique character-preserving Hopf algebra map from $\mathcal{H}$ to the algebra of quasisymmetric functions. Define χ_b to be the quasisymmetric function that is the image of the basis element b under this map. (If $\mathcal{H}$ is cocommutative, χ_b will be a symmetric function.) Call this the *generalized chromatic quasisymmetric function* of b since it is the Stanley chromatic symmetric function for the Hopf algebra of graphs [Sta95]. We do not know how difficult it is to determine or evaluate χ_b.

Proposition 3.25. *With notation as above, the probability of being absorbed in one step of K_a starting from b (that is, $K_a(b, \bullet^n)$) equals*

$$\chi_b \left(\tfrac{1}{a}, \tfrac{1}{a}, \ldots, \tfrac{1}{a}, 0, 0, \ldots \right) \qquad \text{(first a arguments are non-zero).}$$

Proof. By definition of K_a, the desired probability $K_a(b, \bullet^n)$ is a^{-n} times the coefficient of $\bullet^n$ in $\Psi^a(b)$. Every occurrence of $\bullet^n$ in $\Psi^a(b) = m^{[a]} \Delta^{[a]}(b)$ must be due to a term of the form $\bullet^{\alpha_1} \otimes \bullet^{\alpha_2} \otimes \cdots \otimes \bullet^{\alpha_a}$ in $\Delta^{[a]}(b)$, for some composition $\alpha = (\alpha_1, \ldots, \alpha_a)$ of n (some α_i may be 0). So, letting $\eta_b^{b_1, \ldots, b_a}$ denote the coefficient of $b_1 \otimes \cdots \otimes b_a$ in $\Delta^{[a]}(b)$,

$$K_a(b, \bullet^n) = a^{-n} \sum_\alpha \eta_b^{\bullet^{\alpha_1}, \bullet^{\alpha_2}, \ldots, \bullet^{\alpha_a}},$$

where the sum runs over all α with a parts. To re-express this in terms of compositions with no parts of size zero, observe that

$$\eta_b^{\bullet^{\alpha_1}, \ldots, \bullet^{\alpha_{a-1}}, 1} = \eta_b^{\bullet^{\alpha_1}, \ldots, \bullet^{\alpha_{a-1}}},$$

because $\Delta^{[a]} = (\iota \otimes \cdots \otimes \iota \otimes \Delta)\Delta^{[a-1]}$ implies $\eta_b^{\bullet^{\alpha_1},\ldots,\bullet^{\alpha_{a-1}},1} = \sum_{b'} \eta_b^{\bullet^{\alpha_1},\ldots,\bullet^{\alpha_{a-2}},b',\bullet^{\alpha_{a-1}},1}$, but $\eta_{b'}^{\bullet^{\alpha_{a-1}},1}$ is zero unless $b' = \bullet^{\alpha_{a-1}}$. Similar arguments show that $\eta_b^{\bullet^{\alpha_1},\ldots,\bullet^{\alpha_a}} = \eta_b^{\bullet^{\bar{\alpha}_1},\ldots,\bullet^{\bar{\alpha}_{l(\bar{\alpha})}}}$, where $\bar{\alpha}$ is α with all zero parts removed. So

$$K_a(b, \bullet^n) = a^{-n} \sum_{\bar{\alpha}} \binom{a}{l(\bar{\alpha})} \eta_b^{\bullet^{\bar{\alpha}_1},\ldots,\bullet^{\bar{\alpha}_{l(\bar{\alpha})}}},$$

summing over all $\bar{\alpha}$ with at most a parts, and no parts of size zero.

Now, for all compositions $\bar{\alpha}$ of n with no zero parts, the coefficient of the monomial quasisymmetric function $M_{\bar{\alpha}}$ in χ_b is defined to be the image of b under the composite

$$\mathcal{H} \xrightarrow{\Delta^{[l(\bar{\alpha})]}} \mathcal{H}^{\otimes l(\bar{\alpha})} \xrightarrow{\pi_{\bar{\alpha}_1} \otimes \cdots \otimes \pi_{\bar{\alpha}_{l(\bar{\alpha})}}} \mathcal{H}_{\bar{\alpha}_1} \otimes \cdots \otimes \mathcal{H}_{\bar{\alpha}_{l(\bar{\alpha})}} \xrightarrow{\zeta^{\otimes l(\bar{\alpha})}} \mathbb{R}$$

where, in the middle map, $\pi_{\bar{\alpha}_i}$ denotes the projection to the subspace of degree $\bar{\alpha}_i$. As ζ takes value 1 on powers of $\bullet$ and 0 on other basis elements, it transpires that

$$\chi_b = \sum_{\bar{\alpha}} \eta_b^{\bullet^{\bar{\alpha}_1},\bullet^{\bar{\alpha}_2},\ldots,\bullet^{\bar{\alpha}_{l(\bar{\alpha})}}} M_{\bar{\alpha}},$$

summing over all compositions of n regardless of their number of parts. Since $M_{\bar{\alpha}}(1, 1, \ldots 1, 0, \ldots) = \binom{a}{l(\bar{\alpha})}$, where a is the number of non-zero arguments, it follows that

$$K_a(b, c_\bullet^n) = a^{-n}\chi_b(1, 1, \ldots, 1, 0, \ldots) = \chi_b\left(\tfrac{1}{a}, \tfrac{1}{a}, \ldots, \tfrac{1}{a}, 0, 0, \ldots\right),$$

with the first a arguments non-zero. $\qquad\square$

Using a different character ζ, this same argument gives the probability of reaching certain sets of states in one step of the $K_a(b, -)$ chain. This does not require $\mathcal{H}$ to have a single basis element of degree 1.

Proposition 3.26. *Let $\mathcal{C}$ be a subset of generators, and $\zeta^{\mathcal{C}}$ be the character taking value 1 on $\mathcal{C}$ and value 0 on all other generators (extended linearly and multiplicatively). Let $\chi_b^{\mathcal{C}}$ be the image of b under the unique character-preserving Hopf map from $\mathcal{H}$ to the algebra of quasisymmetric functions. Then the probability of being at a state which is a product of elements in $\mathcal{C}$, after one step of the K_a chain, starting from b, is*

$$\chi_b^{\mathcal{C}}\left(\tfrac{1}{a}, \tfrac{1}{a}, \ldots, \tfrac{1}{a}, 0, 0, \ldots\right) \qquad \text{(first a arguments are non-zero)}.$$

Example 3.27 (Rock-breaking)**.** Recall the rock-breaking chain of Example 3.3. Let $\mathcal{C} = \{\hat{e}_1, \hat{e}_2\}$. Then $\chi_{\hat{e}_n}^{\mathcal{C}}\left(2^{-k}, 2^{-k}, \ldots, 2^{-k}, 0, 0, \ldots\right)$ measures the probability that a rock of size n becomes rocks of size 1 or 2, after k binomial breaks.

4 Symmetric functions and breaking rocks

This section studies the Markov chain induced by the Hopf algebra of symmetric functions. Section 4.1 presents it as a rock-breaking process with background and references from the applied probability literature. Section 4.2 gives formulae for the right eigenfunctions by specializing from Theorem 3.19 and uses these to bound absorption time and related probabilistic observables. Section 4.3 gives formulae for the left eigenfunctions by specializing from Theorem 3.15 and uses these to derive quasi-stationary distributions.

4.1 Rock-breaking

As in Examples 1.2 and 3.3, the Markov chain corresponding to the Hopf algebra of symmetric functions may be described as a rock-breaking process on partitions of n: at each step, break each part independently with a symmetric binomial distribution. The chain is absorbed when each part is of size one. Let $P_n(\lambda, \mu)$ be the transition matrix or chance of moving from λ to μ in one step for λ, μ partitions of n. For $n = 2, 3, 4$, these matrices are:

$n = 2$

	1^2	2
1^2	1	0
2	$\frac{1}{2}$	$\frac{1}{2}$

$n = 3$

	1^3	12	3
1^3	1	0	0
12	$\frac{1}{2}$	$\frac{1}{2}$	0
3	0	$\frac{3}{4}$	$\frac{1}{4}$

$n = 4$

	1^4	$1^2 2$	2^2	13	4
1^4	1	0	0	0	0
$1^2 2$	$\frac{1}{2}$	$\frac{1}{2}$	0	0	0
2^2	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$	0	0
13	0	$\frac{3}{4}$	0	$\frac{1}{4}$	0
4	0	0	$\frac{3}{8}$	$\frac{1}{2}$	$\frac{1}{8}$

The ath power map Ψ^a yields rock-breaking into a pieces each time, according to a symmetric multinomial distribution. To simplify things, we focus on $a = 2$.

The mathematical study of rock-breaking was developed by Kolmogoroff [Kol41] who proved a log normal limit for the distribution of pieces of size at most x. A literature review and classical applied probability treatment in the language of branching processes is in [AN72]. They allow more general distributions for the size of pieces in each break. A modern manifestation with links to many areas of probability is the study of fragmentation processes. Extensive mathematical development and good pointers to a large physics and engineering literature are in [Ber03, Ber06]. Most of the probabilistic development is in continuous time and has pieces breaking one at a time. We have not seen previous study of the natural model of simultaneous breaking developed here.

The rock-breaking Markov chain has the following alternative "balls in boxes" description:

Proposition 4.1. *The distribution on partitions of n induced by taking k steps of the chain P_n starting at the one-part partition (n) is the same as the measure induced by dropping n balls into 2^k boxes (uniform multinomial allocation) and considering the partition induced by the box counts in the non-empty cells. For λ a partition of n, written as $1^{a_1(\lambda)} \cdots n^{a_n(\lambda)}$ with $\sum_{i=1}^n i a_i(\lambda) = n$, set $P_n^k((n), \lambda) = \frac{n! 2^k}{2^{nk} \prod_{i=0}^n (a_i!)^{a_i}} \frac{1}{\prod_{i=1}^n (i!)^{a_i}}$. (Here, $a_0(\lambda)$ is the number of empty cells.) Then, for any partition μ*

$$P_n^k(\mu, \lambda) = P_{\mu_1}^k * P_{\mu_2}^k * \cdots * P_{\mu_l}^k \qquad \text{(usual convolutions of measures on } \mathbb{R}^n\text{).}$$

Remarks. Because of Proposition 4.1, many natural functions of the chain can be understood, as functions of k and n, from known properties of multinomial allocation. This includes the distribution of the size of the largest and smallest piece, the joint distribution of the number of pieces of size i, and total number of pieces. See [KSC78, BHJ92].

Observe from the matrices for $n = 2, 3, 4$ that, if the partitions are written in reverse-lexicographic order, the transition matrices are lower-triangular. This is because lexicographic order refines the ordering of partitions by refinement. Furthermore, the diagonal entries are the eigenvalues $1/2^i$, as predicted in Proposition 3.7. Specializing the counting formula in Theorem 3.15 to this case proves the following proposition.

Proposition 4.2. *The rock-breaking chain $P_n(\lambda, \mu)$ on partitions of size n has eigenvalues $1/2^i$, $0 \le i \le n - 1$, with $1/2^i$ having multiplicity $p(n, i)$, the number of partitions of n into i parts.*

Note. As noted by the reviewer, the same rock-breaking chain can be derived using the basis $\{h_\lambda\}$ of complete homogeneous symmetric functions instead of $\{e_\lambda\}$, since the two bases have the same product and coproduct structures. The fact that the dual basis to $\{h_\lambda\}$ is the more recognizable monomial symmetric functions $\{m_\lambda\}$, compared to $\{e_\lambda^*\}$, is unimportant here as the calculations of eigenfunctions do not explicitly require the dual basis.

4.2 Right eigenfunctions

Because the operator $P_n(\lambda, \mu)$ is not self-adjoint in any reasonable sense, the left and right eigenfunctions must be developed distinctly. The following description, a specialization of Theorem 3.19, may be supplemented by the examples and corollaries that follow. A proof is at the end of this subsection.

Proposition 4.3. *Let* μ, λ *be partitions of* n. *Let* $a_i(\lambda)$ *denote the number of parts of size* i *in* λ *and* $l(\lambda)$ *the total number of parts. Then the* μ*th right eigenfunction for* P_n, *evaluated at* λ, *is*

$$f_\mu(\lambda) = \frac{1}{\prod_i \mu_i!} \sum \prod_j \frac{\lambda_j!}{a_1(\mu^j)! a_2(\mu^j)! \ldots a_{\lambda_j}(\mu^j)!}$$

where the sum is over all sets $\{\mu^j\}$ *such that* μ^j *is a partition of* λ_j *and the disjoint union* $\amalg_j \mu^j = \mu$. *The corresponding eigenvalue is* $2^{l(\mu)-n}$. $f_\mu(\lambda)$ *is always non-negative, and is non-zero if and only if* μ *is a refinement of* λ. *If* $\tilde{\lambda}$ *is any set partition with underlying partition* λ, *then* $f_\mu(\lambda)$ *is the number of refinements of* $\tilde{\lambda}$ *with underlying partition* μ.

Here is an illustration of how to compute with this formula:

Example 4.4. For $n = 5$, $\mu = (2, 1, 1, 1)$, $\lambda = (3, 2)$, the possible $\{\mu^j\}$ are

$$\mu^1 = (2, 1), \qquad \mu^2 = (1, 1)$$
$$\mu^1 = (1, 1, 1), \qquad \mu^2 = (2).$$

Then

$$f_\mu(\lambda) = \frac{1}{2!1!1!1!} \left(\frac{3!}{1!1!} \frac{2!}{2!} + \frac{3!}{3!} \frac{2!}{1!} \right) = 4$$

Example 4.5. For $n = 2, 3, 4$, the right eigenfunctions f_μ are the columns of the matrices:

$n = 2$

1	$\frac{1}{2}$
1	0
1	1

$n = 3$

1	$\frac{1}{2}$	$\frac{1}{4}$
1	0	0
1	1	0
1	3	1

$n = 4$

1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{8}$
1	0	0	0	0
1	1	0	0	0
1	2	1	0	0
1	3	0	1	0
1	6	3	4	1

For some λ, μ pairs, the formula for $f_\mu(\lambda)$ simplifies.

Example 4.6. When $\lambda = (n)$,

$$f_\mu((n)) = \frac{1}{\prod_i \mu_i!} \frac{n!}{a_1(\mu)! a_2(\mu)! \ldots a_n(\mu)!}$$

Example 4.7. $f_{1^n} \equiv 1$ has eigenvalue 1.

Example 4.8. As (n) is not a refinement of any other partition, $f_{(n)}$ is non-zero only at (n). Hence $f_{(n)}(\lambda) = \delta_{(n)}(\lambda)$.

Example 4.9. When $\mu = 1^{n-r} r$ $(r \neq 1)$,

$$f_\mu(\lambda) = \sum_j \binom{\lambda_j}{r}$$

with eigenvalue $1/2^{r-1}$. Thus $f_{1^{n-2}2}(\lambda) = \sum \binom{\lambda_j}{2}$ with eigenvalue $1/2$ is the unique second-largest eigenfunction.

Example 4.9 can be applied to give bounds on the chance of absorption. The following corollary shows that absorption is likely after $k = 2\log_2 n + c$ steps.

Corollary 4.10. *For the rock-breaking chain $X_0 = (n), X_1, X_2, \ldots$,*

$$P_{(n)}\{X_k \neq 1^n\} \leq \frac{\binom{n}{2}}{2^k}.$$

Proof. By Example 4.9, if $\mu = 1^{n-2}2$, then $f_\mu(\lambda) = \sum \binom{\lambda_i}{2}$ is an eigenfunction with eigenvalue $1/2$. Further, $f_\mu(\lambda)$ is zero if and only if $\lambda = 1^n$; otherwise $f_\mu(\lambda) \geq 1$. Hence

$$P_{(n)}\{X_k \neq 1^n\} = P_{(n)}\{f_\mu(X_k) \geq 1\} \leq E_{(n)}\{f_\mu(X_k)\} = \frac{\binom{n}{2}}{2^k},$$

where the last equality is a simple application of Use A in Section 2.1. $\qquad\square$

Remarks.

1. From Proposition 4.1 and the classical birthday problem,

$$P_{(n)}\{X_k = 1^n\} = \prod_{i=1}^{n-1}\left(1 - \frac{i}{2^k}\right) = e^{\sum_{i=1}^{n-1} \log\left(1 - \frac{i}{2^k}\right)} = e^{-\sum_{i=1}^{n-1} \frac{i}{2^k} + O\left(\frac{i^2}{2^{2k}}\right)} = e^{-\frac{\binom{n}{2}}{2^k} + O\left(\frac{n^3}{2^{2k}}\right)}.$$

It follows that, for $k = 2\log_2 n + c$ (or $2^k = 2^c n^2$) the inequality in the corollary is essentially an equality.

2. Essentially the same calculations go through for any starting state λ, since by Use A

$$E_\lambda\{f_\mu(X_k)\} = \frac{f_\mu(\lambda)}{2^k}.$$

3. Other eigenfunctions can be similarly used. For example, when $\mu = 1^{n-r}r$, $f_\mu(\lambda) = \sum_j \binom{\lambda_j}{r} > 0$ if and only if $\max_j \lambda_j \geq r$, and $f_\mu(\lambda) \geq 1$ otherwise. It follows as above that

$$P_{(n)}\{\max_i (X_k)_i \geq r\} = P_{(n)}\{f_\mu(X_k) \geq 1\} \leq E_{(n)}\{f_\mu(X_k)\} = \frac{\binom{n}{r}}{2^{(r-1)k}}.$$

4. The right eigenfunctions with $\mu = 1^{n-r}r$ can be derived by a direct probabilistic argument. Drop n balls into 2^k boxes. Let N_i be the number of balls in box i. Then

$$E_{(n)}\left\{\sum_{i=1}^{2^k} \binom{N_i}{r}\right\} = 2^k E_{(n)}\left\{\binom{N_1}{r}\right\} = \frac{\binom{n}{r}}{2^{(r-1)k}}.$$

The last equality follows because N_1 is binomial $(n, 1/2^k)$ and, if X is binomial (n, p), $E\{X(X-1)\ldots(X-r+1)\} = n(n-1)\ldots(n-r+1)p^r$. The other eigenvectors can be derived using more complicated multinomial moments.

Proof of Proposition 4.3. For concreteness, take $l(\lambda) = 2$ and $l(\mu) = 3$. Then Theorem 3.19 states that $f_\mu(\lambda)$ is the coefficient of $\mu_1 \otimes \mu_2 \otimes \mu_3$ in $\Delta^{[3]}(\lambda)$ (viewing μ_i as a partition of single part), divided by $\prod_i a_i(\mu)!$. Recall that

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

So calculating $f_\mu(\lambda)$ requires summing the coefficients of the terms where $i_1 \amalg i_2 = \mu_1$, $j_1 \amalg j_2 = \mu_2$, $k_1 \amalg k_2 = \mu_3$. As μ_1 only has one part, it must be the case that either $i_1 = \mu_1$ and $i_2 = 0$, or $i_1 = 0$ and $i_2 = \mu_1$, and similarly for μ_2, μ_3. Thus, removing the parts of size 0 from $(i_1, j_1, k_1, i_2, j_2, k_2)$ and reordering gives μ. So $\binom{\lambda_1}{i_1 j_1 k_1}\binom{\lambda_2}{i_2 j_2 k_2} = \frac{\lambda_1!\lambda_2!}{\mu_1!\mu_2!\mu_3!}$. Also, if μ^1 denotes the partition obtained by removing 0s and reordering (i_1, j_1, k_1), and μ^2 from (i_2, j_2, k_2), then the disjoint union $\mu^1 \amalg \mu^2 = \mu$. Given μ^1 and μ^2, the number of different sextuples $(i_1, j_1, k_1, i_2, j_2, k_2)$ it could have come from is

$$\prod_i \binom{a_i(\mu)}{a_i(\mu^1)\,a_i(\mu^2)}.$$

Hence

$$f_\mu(\lambda) = \left(\prod_i a_i(\mu)!\right)^{-1} \frac{\lambda_1!\lambda_2!}{\mu_1!\mu_2!\mu_3!} \prod_i \binom{a_i(\mu)}{a_i(\mu^1)\,a_i(\mu^2)},$$

which simplifies as desired. It is then an easy exercise to check that this is the number of refinements of the set partition $\tilde{\lambda}$ of underlying partition μ. $\qquad\square$

4.3 Left eigenfunctions and quasi-stationary distributions

This subsection gives two descriptions of the left eigenfunctions: one in parallel with Proposition 4.3 and the other using symmetric function theory. Again, examples follow the statement with proofs at the end.

Proposition 4.11. *For the rock-breaking Markov chain P_n on partitions of n, for each partition λ of n, there is a left eigenfunction $g_\lambda(\mu)$ with eigenvalue $1/2^{n-l(\lambda)}$,*

$$g_\lambda(\mu) = \lambda_1!\lambda_2!\ldots\lambda_{l(\lambda)}! \sum \frac{(-1)^{l(\mu)-l(\lambda)}}{\mu_1!\mu_2!\ldots\mu_{l(\mu)}!} \prod_j \frac{(l(\mu^j)-1)!}{a_1(\mu^j)!a_2(\mu^j)!\ldots a_{\lambda_j}(\mu^j)!}$$

where the sum is over sets $\{\mu^j\}$ such that μ^j is a partition of λ_j and $\amalg\mu^j = \mu$. $g_\lambda(\mu)$ is non-zero only if μ is a refinement of λ.

As previously, here is a calculational example.

Example 4.12. When $\mu = (2,1,1,1)$, $\lambda = (3,2)$,

$$g_\lambda(\mu) = 3!2! \frac{(-1)^2}{2!1!1!1!} \left(\frac{1!1!}{2!} + \frac{2!0!}{3!} \right) = 5$$

Example 4.13. For $n = 2,3,4$, the left eigenfunctions g_λ are the rows of the matrices:

$n = 2$				

1	$\frac{1}{2}$
1	0
−1	1

$n = 3$

1	$\frac{1}{2}$	$\frac{1}{4}$
1	0	0
−1	1	0
2	−3	1

$n = 4$

1	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{8}$
1	0	0	0	0
−1	1	0	0	0
1	−2	1	0	0
2	−3	0	1	0
−6	12	−3	−4	1

Observe that these matrices are the inverses of those in Example 4.5, as claimed in Theorem 3.19.

The next three examples give some partitions λ for which the expression for $g_\lambda(\mu)$ condenses greatly:

Example 4.14. If $\lambda = (n)$ then $g_{(n)}$ is primitive and

$$g_{(n)}(\mu) = n! \frac{(-1)^{l(\mu)-1}}{\mu_1!\mu_2!\ldots\mu_{l(\mu)}!} \frac{(l(\mu)-1)!}{a_1(\mu)!a_2(\mu)!\ldots a_{\lambda_j}(\mu)!}$$

with eigenvalue $1/2^{n-1}$.

Example 4.15. If $\lambda = 1^{n-r}r$ $(r \neq 1)$, g_λ has eigenvalue $1/2^{r-1}$ and

$$g_\lambda(\mu) = r! \frac{(-1)^{l(\mu)-n+r-1}}{\mu_1!\mu_2!\ldots\mu_{l(\mu)}!} \frac{(l(\mu)-n+r-1)!}{a_1(\mu-n+r)!a_2(\mu)!\ldots a_{\lambda_j}(\mu)!}$$

if $a_1(\mu) \geq a_1(\lambda)$, and 0 otherwise. In particular, $g_{1^{n-2}2}$ puts equal mass at $\mu = 1^n$ and $\mu = 1^{n-2}2$ with mass 0 for other μ.

Example 4.16. Take $\lambda = 1^n$. As no other partition refines λ, $g_{1^n}(\mu) = \delta_{1^n}(\mu)$, and this is the stationary distribution.

The left eigenfunctions can be used to determine the quasi-stationary distributions π^1, π^2 described in Section 2.1, Use G.

Corollary 4.17. *For the rock-breaking Markov chain P_n on partitions of n,*

$$\pi^1(\mu) = \pi^2(\mu) = \delta_{1^{n-2}2}(\mu) \qquad \text{for } \mu \neq 1^n.$$

Proof. From (2.4), π^1 is proportional to $g_{1^{n-2}2}$ on the non-absorbing states. The Perron–Frobenius theorem ensures that π^1 is non-negative. From Examples 4.15 and 4.16, $\pi^1(\mu) = \delta_{1^{n-2}2}(\mu)$ for $\mu \neq 1^n$. Similarly, π^2 is proportional to the pointwise product $g_{1^{n-2}2}(\mu)f_{1^{n-2}2}(\mu) = \delta_{1^{n-2}2}(\mu)$ for $\mu \neq 1^n$. $\qquad\square$

From [Gei77], the power sum symmetric functions p_n are the primitive elements of the ring of symmetric functions and their products p_λ give the left eigenfunctions of the Hopf-power chains. Up to scaling, p_n is the only primitive element of degree n, so p_n must be a scalar multiple of $g_{(n)}$. By Theorem 3.15, $g_{(n)}$ is normalized so that the coefficient of $\hat{e}_n$ is 1 (recall $\hat{e}_n = n!e_n$), whilst the determinant formula [Mac95, p. 28]:

$$p_n = \det \begin{pmatrix} e_1 & 1 & 0 & \cdots & 0 \\ 2e_2 & e_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ ne_n & e_{n-1} & e_{n-2} & \cdots & e_1 \end{pmatrix}$$

shows that e_n has coefficient $(-1)^{n-1}n$ in p_n. Comparing these shows $g_{(n)} = (-1)^{n-1}(n-1)!p_n$, so the left eigenfunctions g_λ are $(-1)^{n-l(\lambda)}\prod_i(\lambda_i - 1)!p_\lambda$ expressed in the $\{\hat{e}_\lambda\}$ basis. Hence Proposition 4.11 may be rephrased as

$$p_\lambda = (-1)^n \lambda_1 \lambda_2 \dots \lambda_{l(\lambda)} \sum_{\substack{\mu^j \vDash \lambda_j \\ \mu := \amalg \mu^j}} \prod_j \frac{(-1)^{l(\mu)}\left(l\left(\mu^j\right) - 1\right)!}{a_1\left(\mu^j\right)!a_2\left(\mu^j\right)! \dots a_{\lambda_j}\left(\mu^j\right)!}e_\mu$$

For example, for $\lambda = 1^n$, $p_1 = e_1$, and $p_{1^n} = e_{1^n} = \hat{e}_{1^n}$ corresponding to the left eigenvector $(1, 0, \dots, 0)$ with eigenvalue 1. For $\lambda = 21^{n-2}$, $p_2 = e_1^2 - 2e_2 = \hat{e}_1^2 - \hat{e}_2$, so $p_{1^{n-2}2} = \hat{e}_{1^n} - \hat{e}_{1^{n-2}2}$ corresponding to the eigenvector $(1, -1, 0, \dots, 0)$ with eigenvalue $1/2$. For $\lambda = 1^{n-3}3$, $p_3 = \det\begin{pmatrix} e_1 & 1 & 0 \\ 2e_2 & e_1 & 1 \\ 3e_3 & e_2 & e_1 \end{pmatrix} = e_1^3 - 3e_1e_2 + 3e_3 = \hat{e}_{1^3} - \frac{3}{2}\hat{e}_1\hat{e}_2 + \frac{1}{2}\hat{e}_3$. Multiplying by 2 gives the left eigenvector $(2, -3, 1, 0, \dots, 0)$ with eigenvalue $1/4$.

Viewing the duality $\sum_\mu f_\lambda(\mu)g_\nu(\mu) = \delta_{\lambda\nu}$ as a matrix equation implies that $\sum_\mu f_\mu(\lambda)g_\mu(\nu) = \delta_{\lambda\nu}$, so

$$\hat{e}_\lambda = \sum_\mu f_\mu(\lambda)g_\mu = \sum_\mu f_\mu(\lambda)(-1)^{n-l(\mu)}\prod_i(\mu_i - 1)!p_\mu,$$

so Proposition 4.3 gives

$$e_\lambda = \sum_\mu \frac{(-1)^{n+l(\mu)}}{\mu_1 \dots \mu_{l(\mu)}} \sum_{\substack{\mu^j \vDash \lambda_j \\ \amalg\mu^j = \mu}} \prod_j \frac{1}{a_1\left(\mu^j\right)!a_2\left(\mu^j\right)! \dots a_{\lambda_j}\left(\mu^j\right)!}p_\mu,$$

which also follows from [Sta99, Prop. 7.7.1].

Proof of Proposition 4.11. $g_\lambda(\mu)$ is the coefficient of μ in $e\left(\lambda_1\right)e\left(\lambda_2\right)\dots e\left(\lambda_{l(\lambda)}\right)$. Every occurrence of μ in $e\left(\lambda_1\right)\dots e\left(\lambda_{l(\lambda)}\right)$ is a product of a μ^1 term in $e(\lambda_1)$, a μ^2 term in $e(\lambda_2)$, etc., for some choice of partitions μ^j of λ_j with $\amalg_j\mu^j = \mu$. Hence it suffices to show that the coefficient of a fixed μ^j in $e(\lambda_j)$ is

$$\frac{(-1)^{l(\mu^j)-1}\lambda_j!(l(\mu^j) - 1)!}{a_1(\mu^j)!\cdots a_{\lambda_j}(\mu^j)!\mu_1^j!\cdots\mu_{l(\mu)}^j!}.$$

Recall that $e(\lambda_j) = \sum_{a \geq 1} \frac{(-1)^{a-1}}{a} m^{[a]} \bar{\Delta}^{[a]}(\lambda_j)$, and observe that all terms of $m^{[a]} \bar{\Delta}^{[a]}(\lambda_j)$ are partitions with a parts. Hence μ^j only occurs in the summand with $a = l(\mu^j)$. So the number needed is $\frac{(-1)^{l(\mu^j)-1}}{l(\mu^j)}$ multiplied by the coefficient of μ^j in $m^{[a]} \bar{\Delta}^{[a]}(\lambda_j)$. Each occurrence of μ^j in $m^{[a]} \bar{\Delta}^{[a]}(\lambda_j)$ is caused by $\mu^j_{\sigma(1)} \otimes \cdots \otimes \mu^j_{\sigma(a)}$ in $\bar{\Delta}^{[a]}(\lambda_j)$ for some $\sigma \in S_a$. For each fixed σ, $\mu^j_{\sigma(1)} \otimes \cdots \otimes \mu^j_{\sigma(a)}$ has coefficient

$$\binom{\lambda_j}{\mu^j_{\sigma(1)} \cdots \mu^j_{\sigma(a)}} = \binom{\lambda_j}{\mu^j_1 \cdots \mu^j_a}$$

in $\bar{\Delta}^{[a]}(\lambda_j)$, and the number of $\sigma \in S_a$ leading to distinct a-tuples $\mu^j_{\sigma(1)}, \ldots, \mu^j_{\sigma(a)}$ is $\frac{a!}{a_1(\mu^j)! \ldots a_{\lambda_j}(\mu^j)!}$. Hence the coefficient of μ^j in $m^{[a]} \bar{\Delta}^{[a]}(\lambda_j)$ is $\binom{\lambda_j}{\mu^j_1 \cdots \mu^j_{l(\mu^j)}} \frac{l(\mu_j)!}{a_1(\mu^j)! \ldots a_{\lambda_j}(\mu^j)!}$ as desired. □

Note. This calculation is greatly simplified for the algebra of symmetric functions, compared to other polynomial 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 Example 2.2 for an example. 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 cancelations in $e(c)$. Hence much effort [Fis10, AS05, AS06] has gone into developing cancelation-free expressions for primitives, as alternatives to $e(c)$.

5   The free associative algebra and riffle shuffling

This section works through the details for the Hopf algebra $k\langle x_1, x_2, \ldots, x_N \rangle$ and riffle shuffling (Examples 1.1 and 1.3). Section 5.1 gives background on shuffling, Section 5.2 develops the Hopf connection, Section 5.3 gives various descriptions of right eigenfunctions. These are specialized to decks with distinct cards in Section 5.3.1 which shows that the number of descents $- \frac{n-1}{2}$ and the number of peaks $- \frac{n-2}{3}$ are eigenfunctions. The last section treats decks with general composition showing that all eigenvalues $1/2^i$, $0 \leq i \leq n-1$, occur as long as there are at least two types of cards. Inverse riffle shuffling is a special case of walks on the chambers of a hyperplane arrangement and of random walks on a left regular band. [Sal12] and [Den12] give a description of left eigenfunctions (hence right eigenfunctions for forward shuffling) in this generality.

5.1   Riffle shuffles

Gilbert–Shannon–Reeds introduced a realistic model for riffle shuffling a deck of n cards. It may be described in terms of a parameterized family of probability measures $Q_a(\sigma)$ for σ in the symmetric group S_n and $a \in \{1, 2, 3, \ldots\}$ a parameter. A physical description of the a-shuffle begins by cutting a deck of n cards into a piles according to the symmetric multinomial distribution, so the probability of pile i receiving n_i cards is $\binom{n}{n_1 n_2 \cdots n_a}/a^n$. Then, the piles are riffled together by (sequentially) dropping the next card from pile i with probability proportional to pile size; continuing until all cards have been dropped. Usual riffle shuffles are 2-shuffles and [BD92] show that $Q_a * Q_b(\sigma) = \sum_\eta Q_a(\eta) Q_b(\sigma \eta^{-1}) = Q_{ab}(\sigma)$. Thus to study $Q_2^{*k}(\sigma) = Q_{2^k}(\sigma)$ it is enough to understand $Q_a(\sigma)$. They also found the closed formula

$$(5.1) \qquad Q_a(\sigma) = \binom{n + a - (d(\sigma) + 1)}{n} \Big/ a^n, \qquad d(\sigma) = \# \text{ descents in } \sigma.$$

Using this they proved that $\frac{3}{2}\log_2 n + c$ 2-shuffles are necessary and suffice to mix n cards.

The study of $Q_a(\sigma)$ has contacts with other areas of mathematics: to Solomon's descent algebra [Sol68, DFP92], quasisymmetric functions [Sta99, Sta01, Ful01], hyperplane arrangements [BHR99, BD98, AD10], Lie theory [Reu93, Reu03], and, as the present paper shows, Hopf algebras. A survey of this and other connections is in [Dia03] with [ADS11, CH10, DFH11] bringing this up to date. A good elementary textbook treatment is in [GS97].

Of course, shuffling can be treated as a Markov chain on S_n with transition matrix $K_a(\sigma, \pi) = Q_a(\pi\sigma^{-1})$, the chance of moving from σ to π after one a-shuffle. To check later calculations, when $n = 3$, the transition matrix is $1/a^3$ times

	123	132	213	231	312	321
123	$\binom{a+2}{3}$	$\binom{a+1}{3}$	$\binom{a+1}{3}$	$\binom{a+1}{3}$	$\binom{a+1}{3}$	$\binom{a}{3}$
132	$\binom{a+1}{3}$	$\binom{a+2}{3}$	$\binom{a+1}{3}$	$\binom{a}{3}$	$\binom{a+1}{3}$	$\binom{a+1}{3}$
213	$\binom{a+1}{3}$	$\binom{a+1}{3}$	$\binom{a+2}{3}$	$\binom{a+1}{3}$	$\binom{a}{3}$	$\binom{a+1}{3}$
231	$\binom{a+1}{3}$	$\binom{a}{3}$	$\binom{a+1}{3}$	$\binom{a+2}{3}$	$\binom{a+1}{3}$	$\binom{a+1}{3}$
312	$\binom{a+1}{3}$	$\binom{a+1}{3}$	$\binom{a}{3}$	$\binom{a+1}{3}$	$\binom{a+2}{3}$	$\binom{a+1}{3}$
321	$\binom{a}{3}$	$\binom{a+1}{3}$	$\binom{a+1}{3}$	$\binom{a+1}{3}$	$\binom{a+1}{3}$	$\binom{a+2}{3}$

.

It is also of interest to study decks with repeated cards. For example, if suits don't matter, the deck may be regarded as having values $1, 2, \ldots, 13$ with value i repeated four times. Now, mixing requires fewer shuffles; see [CV06, CV07, ADS11] for details. The present Hopf analysis works here, too.

5.2 The Hopf connection

Let $\mathcal{H} = k\langle x_1, x_2, \ldots, x_N \rangle$ be the free associative algebra on N generators, with each x_i primitive. As explained in Examples 1.1 and 1.3, the map $x \to \Psi^a(x)/a^{\deg x}$ is exactly inverse a-shuffling. Observe that the number of cards having each value is unchanged during shuffling. This naturally leads to the following finer grading on the free associative algebra: for each $\nu = (\nu_1, \nu_2, \ldots, \nu_N) \in \mathbb{N}^N$, define $\mathcal{H}_\nu$ to be the subspace spanned by words where x_i appears ν_i times. The ath power map $\Psi^a = m^{[a]}\Delta^{[a]} : \mathcal{H} \to \mathcal{H}$ preserves this finer grading. The subspace $\mathcal{H}_{1^N} \subseteq k\langle x_1, x_2, \ldots, x_N \rangle$ is spanned by words of degree 1 in each variable. A basis is $\{x_\sigma = x_{\sigma^{-1}(1)} \cdots x_{\sigma^{-1}(n)}\}$. The mapping Ψ^a preserves $\mathcal{H}_{1^n}$ and $\frac{1}{a^n}\Psi^a(x_\sigma) = \sum_\pi Q_a(\pi\sigma^{-1})x_\pi$. With obvious modification the same result holds for any subspace $\mathcal{H}_\nu$. Working on the dual space $\mathcal{H}^*$ gives the usual Gilbert–Shannon–Reeds riffle shuffles. Let us record this formally; say that a deck has *composition* ν if there are ν_i cards of value i.

Proposition 5.1. *Let* $\nu = (\nu_1, \nu_2, \ldots, \nu_N)$ *be a composition of n. For any $a \in \{1, 2, \ldots\}$, the mapping $\frac{1}{a^n}\Psi^a$ preserves $\mathcal{H}_\nu$ and the matrix of this map in the monomial basis is the transpose of the transition matrix for the inverse a-shuffling Markov chain for a deck of composition ν. The dual mapping is the Gilbert–Shannon–Reeds measure (5.1) on decks with this composition.*

Note. Since the cards behave equally independent of their labels, any particular deck of interest can be relabeled so that $\nu_1 \geq \nu_2 \geq \cdots \geq \nu_N$. In other words, it suffices to work with $\mathcal{H}_\nu$ for partitions ν.

5.3 Right eigenfunctions

Theorem 3.16 applied to the free associative algebra gives a basis of left eigenfunctions for inverse shuffles, which are right eigenfunctions for the forward GSR riffle shuffles. By Remark 2 after

Theorem 3.16, each word $w \in \mathcal{H}_\nu$ corresponds to a right eigenfunction f_w for the GSR measure on decks with composition ν. As explained in Example 1.3 and Section 2.3, these are formed by factoring w into Lyndon words, standard bracketing each Lyndon factor, then expanding and summing over the symmetrization. The eigenvalue is a^{k-n} with k the number of Lyndon factors of w. The following examples should help understanding.

Example 5.2. For $n = 3$, with $\nu = 1^3$, $\mathcal{H}_\nu$ is 6-dimensional with basis $\{x_\sigma\}_{\sigma \in S_3}$. Consider $w = x_1 x_2 x_3$. This is a Lyndon word so no symmetrization is needed. The standard bracketing $\lambda(x_1 x_2 x_3) = [\lambda(x_1), \lambda(x_2 x_3)] = [x_1, [x_2, x_3]] = x_1 x_2 x_3 - x_1 x_3 x_2 - x_2 x_3 x_1 + x_3 x_2 x_1$. With the labeling of the transition matrix K_a of (5.1), the associated eigenvector is $(1, 1, 0, -1, 0, -1)^T$ with eigenvalue $1/a^2$.

For $n = 3, \nu = (2, 1)$, $\mathcal{H}_\nu$ is three-dimensional with basis $\{x_1^2 x_2, x_1 x_2 x_1, x_2 x_1^2\}$. Consider $w = x_2 x_1^2$. This factors into Lyndon words as $x_2 \cdot x_1 \cdot x_1$; symmetrizing gives the eigenvector $x_1^2 x_2 + x_1 x_2 x_1 + x_2 x_1^2$ or $(1, 1, 1)^T$ with eigenvalue 1.

The description of right eigenvectors can be made more explicit. This is carried foward in the following two sections.

5.3.1   Right eigenfunctions for decks with distinct values

Recall from Section 2.3 that the values of an eigenfunction f_w at w' can be calculated graphically from the decreasing Lyndon hedgerow T_w of w. When w has all letters distinct, this calculation simplifies neatly. To state this, extend the definition of f_l (l a Lyndon word with distinct letters) to words longer than l, also with distinct letters: $f_l(w)$ is f_l evaluated on the subword of w whose letters are those of l, if such a subword exists, and 0 otherwise. (Here, a subword always consist of consecutive letters of the original word.) Because w has distinct letters, there is at most one such subword. For example, $f_{35}(14253) = f_{35}(53) = -1$.

Proposition 5.3. *Let w be a word with distinct letters and Lyndon factorization $l_1 l_2 \ldots l_k$. Then, for all w' with distinct letters and of the same length as w, $f_w(w') = f_{l_1}(w') f_{l_2}(w') \ldots f_{l_k}(w')$ and f_w takes only values 1,-1 and 0.*

Example 5.4. $f_{35142}(14253) = f_{35}(14253) f_{142}(14253) = -1 \cdot 1 = -1$ as calculated in Section 2.3.

Proof. Recall from Section 2.3 that $f_w(w')$ is the signed number of ways to permute the branches and trees of T_w to spell w'. When w and w' each consist of distinct letters, such a permutation, if it exists, is unique. This gives the second assertion of the proposition. This permutation is precisely given by permuting the branches of each T_{l_i} so they spell subwords of w'. The total number of branch permutations is the sum of the number of branch permutations of each T_{l_i}. Taking parity of this statement gives the first assertion of the proposition. $\square$

The example above employed a further shortcut that is worth pointing out: the Lyndon factors of a word with distinct letters start precisely at the *record minima* (working left to right, thus $\underline{3}5\underline{1}42$ has minima $3, 1$ in positions $1, 3$), since a word with distinct letters is Lyndon if and only if its first letter is minimal. This leads to

Proposition 5.5. *The multiplicity of the eigenvalue a^{n-k} on $\mathcal{H}_{1,1,\ldots 1}$ is $c(n, k)$, the signless Stirling number of the first kind.*

Proof. By the above observation, the multiplicity of the eigenvalue a^{n-k} on $\mathcal{H}_{1,1,\ldots,1}$ is the number of permutations with k record minima, which is also the number of permutations with k cycles, by [Sta12, Prop. 1.3.1]. $\square$

Example 5.6 (Invariance under reversal). Let $\bar{w}$ denote the reverse of w, then, for any σ, switching branches at every node shows that $f_\sigma(\bar{w}) = \pm f_\sigma(w)$ where the parity is $n - \#$ Lyndon factors in σ. Thus, each eigenspace of Ψ^a is invariant under the map $w \to \bar{w}$. For example, $f_{35142}(35241) = f_{35}(35241) f_{142}(35241) = 1 \cdot 1 = -f_{35142}(14253)$ when compared with Example 5.4 above. The quantity $(n - \#$ Lyndon factors in 35142) is $5 - 2 = 3$, hence the change in sign.

Example 5.7 (Eigenvalue 1). $\sigma = n, n - 1, \ldots, 1$ is the only word with n Lyndon factors, so $f_\sigma(w) = 1$ spans the 1-eigenspace.

Example 5.8 (Eigenvalue $1/a$ and descents). There are $\binom{n}{2}$ permutations σ which have $n - 1$ Lyndon factors. They may be realized by choosing $i < j$ and taking $\sigma = n, n - 1, \ldots, j + 1, j - 1, \ldots, i + 1, i, j, i - 1, \ldots, 1$. Then, all but j are record minima and the corresponding eigenfunctions are (in the notation at the start of this subsection),

$$f_{ij}(w) = \begin{cases} 1, & \text{if } ij \text{ occurs as a subword of } w, \\ -1, & \text{if } ji \text{ occurs as a subword of } w, \\ 0, & \text{otherwise.} \end{cases}$$

Their sum is $f(w) := \sum_{i<j} f_{ij}(w) = n - 1 - 2d(w)$ with $d(w)$ the number of descents in w. This is thus an eigenfunction with eigenvalue $1/a$, as claimed in Example 1.4. This eigenfunction appears in the Markov chain recording the number of descents in successive a-shuffles, which is the same as the Markov chain of carries when n numbers are added in base a. The transition matrix of this Markov chain is Holte's [Hol97] "amazing matrix." See [DF09b, DF09a, DF12]. The theory developed there shows that, with $s(n, k)$ the Stirling numbers $(x(x - 1) \cdots (x - n + 1) = \sum_{k \geq 0} s(n, k) x^k)$,

$$(5.2) \qquad h_j(w) = n! \sum_{k \geq 0} \frac{s(k, n - j)}{k!} \binom{n - d(w) - 1}{n - k}$$

is a right eigenfunction with eigenvalue $1/a^j$, $0 \leq j \leq n - 1$, and the eigenfunction f above is $\frac{2}{n} h_1$.

Example 5.9 (Eigenvalue $1/a^2$ and peaks). Recall that a permutation w has a *peak* at position i, $1 < i < n$, if $w(i-1) < w(i) > w(i+1)$, and a *trough* at position i, $1 < i < n$, if $w(i-1) > w(i) < w(i+1)$. Call the remaining case a *straight*: $w(i-1) < w(i) < w(i+1)$ or $w(i-1) > w(i) > w(i+1)$. The number of peaks and the peak set have been intensively investigated [Ste97, WS96]. The following development shows that $\#$ peaks$(w) - \frac{n-2}{3}$ is an eigenfunction with eigenvalue $1/a^2$, as is $\#$ troughs$(w) - \frac{n-2}{3}$. Indeed, a basis of this eigenspace is $\{f_\sigma\}$ where σ is obtained from $n, n - 1, \ldots, 1$ by removing j and inserting it after i for $i < j$, and then removing $k > j$ and inserting it further down also. Then, all but j, k are record minima. There are three places to insert k (in the examples, $i = 1, j = 3, k = 5$):

Case 1 after l with $l \neq i$, $l \neq j, l < k$ (e.g., 42513);

Case 2 after j, i.e., $\sigma = n, n - 1, \ldots, k + 1, k - 1, \ldots, j + 1, j - 1, \ldots, i + 1, i, j, k, i - 1, \ldots, 1$ (e.g., 42135);

Case 3 after i, i.e., $\sigma = n, n - 1, \ldots, k + 1, k - 1, \ldots, j + 1, j - 1, \ldots, i + 1, i, k, j, i - 1, \ldots, 1$ (e.g., 42153).

Then $f_\sigma(w)$ is, respectively

41

Case 1 1 if ij and kl both occur as subwords of w, or if ji and lk both occur; -1 if ji and kl both occur, or if ij and lk both occur; 0 if i is not adjacent to j in w or if k is not adjacent to l (this is $f_{ij}f_{lk}$);

Case 2 1 if ijk or kji occur as subwords of w; -1 if ikj or jki occur; 0 otherwise (this is $f_{ij}f_{jk} + f_{ik}f_{jk}$);

Case 3 1 if ikj or jki occur as subwords of w; -1 if kij or jik occur; 0 otherwise (this is $-f_{ik}f_{jk} + f_{ij}f_{ik}$).

Proposition 5.10. $f_\wedge(w) := \#$ *peaks in* $w - \frac{n-2}{3}$, $f_\vee(w) := \#$ *troughs in* $w - \frac{n-2}{3}$ *and* $f_-(w) := \#$ *straights in* $w - \frac{n-2}{3}$ *are right eigenfunctions with eigenvalue* $1/a^2$.

Proof. Let s_r denote the sum of all eigenfunctions arising from case r, as defined in Example 5.9 above. Note that

$$s_2 = \sum_{i<j<k} f_{ij}f_{jk} + f_{ik}f_{jk} = \#\text{ straights} - \#\text{ peaks}$$

$$s_3 = \sum_{i<j<k} -f_{ik}f_{jk} + f_{ik}f_{ij} = \#\text{ peaks} - \#\text{ troughs}.$$

Since each successive triple in w either forms a straight, a peak or a trough,

$$\#\text{ straights} + \#\text{ peaks} + \#\text{ troughs} = n - 2.$$

Hence $f_\wedge = \frac{1}{3}(s_3 - s_2)$, $f_\vee = \frac{-1}{3}(s_2 + 2s_3)$, $f_- = \frac{1}{3}(2s_2 + s_3)$ are all in the $1/a^2$-eigenspace. $\qquad\square$

It may be possible to continue the analysis to patterns of longer length; in particular, one interesting open question is which linear combinations of patterns and constant functions give eigenfunctions.

5.3.2 Right eigenfunctions for decks with general composition

Recall from Proposition 5.1 that, for a composition $\nu = (\nu_1, \nu_2, \cdots, \nu_N)$ of n, the map $\frac{1}{a^n}\Psi^a$ describes inverse a-shuffling for a deck of composition ν, i.e., a deck of n cards where ν_i cards have value i. Theorem 3.16 applies here to determine a full left eigenbasis (i.e., a right eigenbasis for forward shuffles). The special case of $\nu = (n-1, 1)$ (follow one labeled card) is worked out in [Ciu98] and used to bound the expected number of correct guesses in feedback experiments. His work shows that the same set of eigenvalues $\{1, 1/a, 1/a^2, \ldots, 1/a^{n-1}\}$ occur.

This section shows that this is true for all deck compositions (provided $N > 1$). It also determines a basis of eigenfunctions with eigenvalue $1/a$ and constructs an eigenfunction which depends only on an appropriately defined number of descents, akin to Example 5.8.

The following proposition finds one "easy" eigenfunction for each eigenvalue of $1/a^k$. The examples that follow the proof show again that eigenfunctions can correspond to natural observables.

Proposition 5.11. *Fix a composition* ν *of* n. *The dimension of the* $1/a^k$-*eigenspace for the* a-*shuffles of a deck of composition* ν *is bounded below by the number of Lyndon words in the alphabet* $\{1, 2, \ldots, N\}$ *of length* $k+1$ *in which letter* i *occurs at most* ν_i *times. In particular,* $1/a^k$ *does occur as an eigenvalue for each* k, $0 \le k \le n-1$.

Proof. By remark 2 after Theorem 3.16, the dimension of the $1/a^k$-eigenspace is the number of monomials in $\mathcal{H}_\nu$ with $n - k$ Lyndon factors. One way of constructing such monomials is to choose a Lyndon word of length $k + 1$ in which letter i occurs at most ν_i times, and leave the remaining $n - k$ letters of ν as singleton Lyndon factors. The monomial is then obtained by putting these factors in decreasing order. This shows the lower bound.

To see that $1/a^k$ is an eigenvalue for all k, it suffices to construct, for each k, a Lyndon word of length $k + 1$ in which letter i occurs at most ν_i times. For $k > \nu_1$, this may be achieved by placing the smallest $k + 1$ values in increasing order. For $k \leq \nu_1$, take the word with k 1s followed by a 2. $\qquad\square$

Example 5.12. For $\nu = (3, 2, 1, 2)$, the eight eigenfunctions constructed in the last step of the proof correspond to the words shown below in order $1/a^k$, $0 \leq k \leq 7$. The bracketed term is the sole non-singleton Lyndon factor:

44322111, 443(12)11, 4432(112)1, 4432(1112), 443(11122), 44(111223), 4(1112234), (11122344).

Example 5.13. For an n-card deck of composition ν, the second largest eigenvalue is $1/a$. Our choice of eigenvectors correspond to words with $n - 1$ Lyndon factors. Each such word must have $n - 2$ singleton Lyndon factors and a Lyndon factor of length 2. Hence the bound in Proposition 5.11 is attained; furthermore it can be explicitly calculated: the Lyndon words of length 2 are precisely a lower value followed by a higher value, so the multiplicity of eigenvalue $1/a$ is $\binom{N}{2}$. This doesn't depend on ν, only on the number N of distinct values.

Summing these eigenfunctions and arguing as in Example 5.8 gives

Proposition 5.14. *For any n-card deck of composition ν, let $a(w)$, $d(w)$ be the number of strict ascents, descents in w respectively. Then $a(w) - d(w)$ is an eigenfunction of Ψ^a with eigenvalue $1/a$.*

Proof. Fix two values $i < j$. Order the deck in decreasing order, then take a card of value j and put it after the first card of value i. In other words, let ij be the only non-singleton Lyndon factor. By inspection, the corresponding eigenvector is (up to scaling)

$$f_{ij}(w) = \{\# \text{ subwords } ij \text{ in } w\} - \{\# \text{ subwords } ji \text{ in } w\};$$

summing f_{ij} over $1 \leq i < j \leq N$ shows that $\{\# \text{ ascents in } w\} - \{\# \text{ descents in } w\}$ is an eigenfunction with eigenvalue $1/a$. $\qquad\square$

Remarks. Under the uniform distribution, the expectation of $a(w) - d(w)$ is zero. If initially the deck is arranged in increasing order w^0, $a(w^0) - d(w^0) = N - 1$. If w^k is the permutation after k a-shuffles, the proposition gives $E\{a(w^k) - d(w^k)\} = \frac{1}{a^k}(N - 1)$. Thus for $a = 2$, $k = \log_2(N - 1) + \theta$ shuffles suffice to make this expected value $2^{-\theta}$. On the other hand, consider a deck with n cards labeled 1 and n cards labeled 2. If the initial order is $w^0 = 11 \cdots 12 \cdots 21$, $a(w^0) - d(w^0) = 0$ and so $E\{a(w^k) - d(w^k)\} = 0$ for all k.

Central limit theorems for the distribution of descents in permutations of multi-sets are developed in [CV07].

Example 5.15. Specialize Example 5.13 to $\nu = (1, n-1)$, so there is one exceptional card of value 1 in a deck of otherwise identical cards of value 2. Then there is a unique eigenfunction f_{12} of eigenvalue $1/a$:

$$f_{12}(w) = \begin{cases} 1, & \text{if 1 is the top card,} \\ -1, & \text{if 1 is the bottom card,} \\ 0, & \text{otherwise.} \end{cases}$$

6 Examples and counter-examples

This section contains a collection of examples where either the Hopf-square map leads to a Markov chain with a reasonable "real world" interpretation — Markov chains on simplicial complexes and quantum groups — or the constructions do not work out to give Markov chains — a quotient of the symmetric functions algebra, Sweedler's Hopf algebra and the Steenrod algebra. Further examples will be developed in depth in future work.

Example 6.1 (A Markov chain on simplicial complexes). Let $\mathcal{X}$ be a finite set and $\mathcal{C}$ a simplicial complex of subsets of $\mathcal{X}$. Recall that this means that $\mathcal{C}$ is a collection of non-empty subsets of $\mathcal{X}$ such that $c \in \mathcal{C}$ implies that all non-empty subsets of c are in $\mathcal{C}$. As an example, consider the standard triangulation of the torus into 18 triangles:
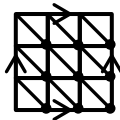
Here the top and bottom edges are identified as are the left and right sides. This identifies several vertices and edges and $\mathcal{X}$ consists of nine distinct vertices. The complex $\mathcal{C}$ contains these nine vertices, the 24 distinct edges, and the 18 triangles.

The set of all simplicial complexes (on all finite sets $\mathcal{X}$) is a basis for a Hopf algebra, under disjoint union as product and coproduct

$$\Delta(\mathcal{C}_{\mathcal{X}}) = \sum_{S \subseteq \mathcal{X}} \mathcal{C}_S \otimes \mathcal{C}_{S^c}$$

with the sum over subsets $S \subseteq \mathcal{X}$, and $\mathcal{C}_S = \{A \subseteq S : A \in \mathcal{C}\}$. By convention $\mathcal{C}_\emptyset = 1$ in this Hopf algebra so $S = \emptyset$ is allowed in the sum. Graded by $|\mathcal{X}|$, this gives a commutative, cocommutative Hopf algebra with basis given by all complexes $\mathcal{C}$; the generators are the connected complexes.

The associated Markov chain, restricted to complexes on n vertices, is simple to describe: from a given complex $\mathcal{C}$, color the vertices red or blue, independently, with probability $1/2$. Take the disjoint union of the complex induced by the red vertices with the complex induced by the blue vertices. As usual, the process terminates at the trivial complex consisting of n isolated vertices.

This Markov chain is of interest in quantifying the results of a "topological statistics" analysis. There, a data set (n points in a metric space) gives rise to a family of complexes $\mathcal{C}_\epsilon$, $0 \le \epsilon < \infty$, where the vertices of each $\mathcal{C}_\epsilon$ are the data points, and k points form a simplex if the intersection of the ϵ balls around each point is non-empty in the ambient space. For ϵ small, the complex is trivial. For ϵ sufficiently large, the complex is the n-simplex. In topological statistics [CCdS06, Car09] one studies things like the Betti numbers of $\mathcal{C}_\epsilon$ as a function of ϵ. If these are stable for a range of ϵ this indicates interpretable structure in the data.

Consider now a data set with large n and ϵ fixed. If a random subset of k points is considered (a frequent computational ploy) the induced sub-complex essentially has the distribution of the "painted red" sub-complex (if the painting is done with probability k/n). Iterating the Markov chain corresponds to taking smaller samples.

If the Markov chain starts out at the n-simplex, every connected subset of the resulting Markov chain is a simplex. Thus, at each stage, all of the higher Betti numbers are zero and β_0 after k steps is $2^k - X_k$, where X_k is distributed as the number of empty cells if n balls are dropped into 2^k boxes. This is a thoroughly studied problem [KSC78]. The distribution of the Betti numbers for more interesting starting complexes is a novel, challenging problem. Indeed, consider the triangulation of the torus with $2n^2$ initial triangles. Coloring the vertices red or blue with probability $1/2$, the edges with red/red vertices are distributed in the same way as the "open sites" in site percolation on a triangular lattice. Computing the Betti number β_0 amounts to computing the number of connected components in site percolation. In the infinite triangular lattice, it is known that $p = 1/2$ is the critical threshold and at criticality, the chance that the component containing the origin has size greater than k falls off as $k^{-5/48}$. These and related facts about site percolation are among the deepest results in modern probability. See [Gri99, Wer09, SSG11] for background and recent results. Iterates of the Markov chain result in site percolation with p below the critical value but estimating β_0 is still challenging.

It is natural to study the absorption of this chain started at the initial complex $\mathcal{C}$. This can be studied using the results of Section 3.7.

Proposition 6.2. *Let the simplicial complex Markov chain start at the complex $\mathcal{C}$. Let G be the graph of the 1-skeleton of $\mathcal{C}$. Suppose that the chromatic polynomial of G is $p(x)$. Then the probability of absorption after k steps is $p(2^k)/2^{nk}$ (with $n = |\mathcal{X}|$).*

For example, if $\mathcal{C}$ is the n-simplex, $p(x) = x(x-1) \cdots (x - n + 1)$ and $P\{$absorption after k steps$\} = \prod_{i=1}^{n-1}(1 - i/2^k) \sim e^{-2^{-(2c-1)}}$ if $k = 2(\log_2 n + c)$ for n large. If $\mathcal{C}$ is a tree, $p(x) = x(x-1)^{n-1}$ and $P\{$absorption after k steps$\} = (1 - 1/2^k)^{n-1} \sim e^{-2^{-c}}$ if $k = \log_2 n + c$ for n large.

Using results on the birthday problem in non-standard situations [BHJ92, CDM05] it is possible to do similar asymptotics for variables such as the number of l-simplices remaining after k steps for more interesting starting $\mathcal{C}$ such as a triangulation of the torus into $2(n-1)^2$ triangles.

As a final remark, note that the simplicial complex Markov chain induces a Markov chain on the successive 1-skeletons. The P basis beautifully developed in [Fis10] are eigenvectors of this Markov chain (though not the eigenbasis produced by the algorithm of our Section 3.5). These all lift to eigenvectors of the complex chain, so much is known.

Example 6.3 (Quantized shuffle algebras)**.** It is natural to seek useful deformations of processes like riffle shuffling. One route is via the quantized shuffle algebras of [Gre95, Gre97] and [Ros95, Ros97, Ros98]. These have become a basic object of study [Lec04, KR11]. Consider the vector space $k\langle x_1, \ldots, x_n \rangle$, and equip its degree 1 subspace with a symmetric $\mathbb{Z}$ form $x_i \cdot x_j$. Turn this into an algebra with the product of concatentation. Take as coproduct $\Delta(x_i) = 1 \otimes x_i + x_i \otimes 1$. However, Δ is to be multiplicative with respect to the twisted tensor product $(x_1 \otimes x_2)(y_1 \otimes y_2) = q^{x_2 \cdot y_1}(x_1 x_2 \otimes y_1 y_2)$. Green translates this into shuffling language. The upshot is if $w = x_{i_1} x_{i_2} \cdots x_{i_k}$ is a word in $k\langle x_1, \ldots, x_n \rangle$ then

$$m\Delta(w) = \sum_{S \subseteq \{1,2,\ldots,k\}} q^{\mathrm{wt}(S,w)} w_S w_{S^c}.$$

Here the sum is over all subsets (including the empty set), $w_S w_{S^c}$ is the inverse shuffle moving the letters in the positions marked by S to the front (keeping them in the same relative order). The

weight $\text{wt}(S, w)$ is the sum of $x_{j'} \cdot x_j$ for $j' \in S^{\mathcal{C}}$, $j \in S$, $j' < j$. Thus if $w = ijklm$ and $S = \{2, 4\}$, we have $w_S w_{S^{\mathcal{C}}} = jlikm$, and $\text{wt}(S, w) = i \cdot j + i \cdot l + k \cdot l$.

When $q = 1$ this shuffle product gives Ree's shuffle algebra and general values of q lead to elegant combinatorial formulations of quantum groups. For general $q > 0$, there is also a naturally associated Markov chain. Work on the piece with multi-grading 1^n, so each variable appears once and we may work with permutations in S_n. For a starting permutation π, and $0 \le j \le n$, let $\theta(j) = \sum_{|S|=j} q^{\text{wt}(S, \pi)}$. Set $\theta = \sum_{j=0}^{n} \theta(j)$. Choose j with probability $\theta(j)/\theta$ and then S (with $|S| = j$) with probability $q^{\text{wt}(S, \pi)}/\theta(j)$. Move to $\pi_S \pi_{S^{\mathcal{C}}}$. This defines a Markov transition matrix $K_q(\pi, \pi')$ via $\frac{1}{\theta} m \Delta(\pi)$. Note that the normalization θ depends on π.

We have not seen our way through this to nice mathematics. There is one case where progress can be made: suppose $x_i \cdot x_j \equiv 1$. Then $\text{wt}(S, \pi) = \text{inv}(S)$, the minimum number of pairwise adjacent transpositions needed to move S to the left. (When $n = 5$ and $S = \{2, 4\}$, $\text{inv}(S) = 3$.) Since $\text{wt}(S, \pi)$ doesn't depend on π, neither does θ and the Markov chain becomes a random walk on S_n driven by the measure

$$\mu(\sigma) = \begin{cases} q^{\text{inv}(\sigma)}/z_n, & \text{if } \sigma \text{ has a single descent}, \\ 0, & \text{otherwise}, \end{cases}$$

where z_n is a normalizing constant.

The preceding description gives inverse riffle shuffles. It is straightforward to describe the q-analog of forward riffle shuffles by taking inverses. Let $[j]_q = 1 + q + \cdots + q^{j-1}$, $[j]_q! = [j]_q [j - 1]_q \ldots [1]_q$, and $\begin{bmatrix} n \\ k \end{bmatrix}_q = \frac{[n]_q!}{[k]_q! [n-k]_q!}$, the usual q-binomial coefficient. We write $I(w)$ for the number of inversions of the permutation w and $R(w) = d(w^{-1}) + 1$ for the number of rising sequences.

Proposition 6.4. *For $q > 0$, the q-riffle shuffle measure has the following description on S_n:*

(6.1)
$$Q_q(w) = \begin{cases} q^{I(w)}/z_n, & \text{if } R(w) \le 2, \\ 0, & \text{otherwise}, \end{cases}$$

where z_n is the normalizing constant $\sum_{w:R(w)\le 2} q^{I(w)}$. To generate w from Q_q, cut off j cards with probability $\begin{bmatrix} n \\ j \end{bmatrix}_q / z_n$ and drop cards sequentially according to the following rule: if at some stage there are A cards in the left pile and B cards in the right pile, drop the next card

(6.2) *from left with probability* $\dfrac{q^B [A]_q}{[A + B]_q}$ *and from right with probability* $\dfrac{[B]_q}{[A + B]_q}.$

Continue, until all cards have been dropped.

Proof. Equation (6.1) follows from the inverse description because $I(w) = I(w^{-1})$. For the sequential description, it is classical that $\begin{bmatrix} n \\ j \end{bmatrix}_q$ is the generating function for multi-sets containing j ones and $n - j$ twos by number of inversion [Sta12, Sect. 1.7]. For two piles with say, $1, 2, \ldots, j$ in the left and $j + 1, \ldots, k$ in the right, in order, dropping j induces $k - j$ inversions. Multiplying the factors in (6.2) results in a permutation with $R(w) \le 2$, with probability $q^{I(w)}/\begin{bmatrix} n \\ j \end{bmatrix}_q$. Since the cut is made with probability $\begin{bmatrix} n \\ j \end{bmatrix}_q / z_n$, the two-stage procedure gives (6.1). □

Remarks. When $q = 1$, this becomes the usual Gilbert–Shannon–Reeds measure described in the introduction and in Section 5. In particular, for the sequential version, the cut is j with probability $\binom{n}{j}/2^n$ and with A in the left and B in the right, drop from left or right with probability $\frac{A}{A+B}$, $\frac{B}{A+B}$

respectively. For general q, as far as we know, there is no closed form for z_n. As $q \to \infty$, the cutting distribution is peaked at $n/2$ and most cards are dropped from the left pile. The most likely permutation arises from cutting off $n/2$ cards and placing them at the bottom. As $q \to 0$, the cutting distribution tends to uniform on $\{0, 1, \ldots, n\}$ and most cards are dropped from the right pile. The most likely permutation is the identity. There is a natural extension to a q-a-shuffle with cards cut into a piles according to the q-multinomial distribution and cards dropped sequentially with probability "q-proportional" to packet size.

We hope to analyze this Markov chain in future work. See [DR00] for related q-deformations of a familiar random walk.

Example 6.5 (A quotient of the algebra of symmetric functions)**.** Consider $\bar{\Lambda} = \Lambda/e_1 = k[e_2, e_3, \ldots]$, with

$$\Delta(e_n) = 1 \otimes e_n + \sum_{j=2}^{n-2} e_j \otimes e_{n-j} + e_n \otimes 1.$$

This is the Weyl group invariants of type A. It is a polynomial algebra, so the theory in Section 3 generates an eigenbasis of Ψ^a and Ψ^{*a}. One hopes this will induce a rock-breaking process where pieces of size one are not allowed; however, we cannot rescale the basis via Theorem 3.4 to obtain a transition matrix, as both e_2 and e_3 are primitive basis elements of degree greater than one. Hence e_2^3, e_3^2 have the same degree, but $m\Delta(e_2^3) = 8e_2^3$ and $m\Delta(e_3^2) = 4e_3^2$, so no rescaling can make the sum of the cofficients of $m\Delta(e_2^3)$ equal to that of $m\Delta(e_3^2)$.

Example 6.6 (Sweedler's example)**.** The four-dimensional algebra

$$H_4 = k(1, g, x, gx : g^2 = 1, \ x^2 = 0, \ xg = -gx)$$

becomes a Hopf algebra with $\Delta(g) = g \otimes g$, $\Delta(x) = x \otimes 1 + g \otimes x$, $\epsilon(g) = 1$, $\epsilon(x) = 0$. The antipode is $s(g) = g^{-1}$, $s(x) = -gx$. It is discussed in [Mon93] as an example of a Hopf algebra that is neither commutative nor cocommutative. With the given basis $\{1, g, x, gx\}$, $m\Delta(1) = 1$, $m\Delta(g) = 1$, $m\Delta(x) = x - gx$, $m\Delta(gx) = -x + gx$. The negative coefficients forstall our efforts to find a probabilistic interpretation. The element $v = x - gx$ is an eigenvector for $m\Delta$ with eigenvalue 2 and high powers of $\frac{1}{2}m\Delta$ applied to a general element $a + bg + cx + dgx$ converge to $(c - d)v$. Of course, this example violates many of our underlying assumptions. It is not graded and is neither a polynomial algebra nor a free associative algebra.

Example 6.7 (The Steenrod algebra)**.** Steenrod squares (and higher powers) are a basic tool of algebraic topology [Hat02]. They give rise to a Hopf algebra A_2 over $\mathbb{F}_2$. Its dual A_2^* is a commutative, noncocommutative Hopf algebra over $\mathbb{F}_2$ with a simple description. As an algebra, $A_2^* = \mathbb{F}_2[x_1, x_2, \ldots]$ (polynomial algebra in countably many variables) graded with x_i of degree $2^i - 1$. The coproduct is $\Delta(x_n) = \sum_{i=0}^{n} x_{n-i}^{2^i} \otimes x_i$ ($x_0 = 1$). Alas, because the coefficients are mod 2, we have been unable to find a probabilistic interpretation of $m\Delta$. For example, $(A_2^*)_3$ has basis $\{x_1^3, x_2\}$ and $m\Delta(x_1^3) = 0$, $m\Delta(x_2) = x_1^3$ so $(m\Delta)^2 \equiv 0$. Of course, high powers of operators can be of interest without positivity [DSC12, GM09].

References

[ABS06]   M. Aguiar, N. Bergeron, and F. Sottile. Combinatorial Hopf algebras and generalized Dehn–Sommerville relations. *Compos. Math.*, 142(1):1–30, 2006.

[AD10]   C. A. Athanasiadis and P. Diaconis. Functions of random walks on hyperplane arrangements. *Adv. in Appl. Math.*, 45(3):410–437, 2010.

[ADS11]   S. Assaf, P. Diaconis, and K. Soundararajan. A rule of thumb for riffle shuffling. *Ann. Appl. Probab.*, 21:843–875, 2011.

[AM10]   M. Aguiar and S. Mahajan. *Monoidal functors, species and Hopf algebras*, volume 29 of *CRM Monograph Series*. American Mathematical Society, Providence, RI, 2010. With forewords by Kenneth Brown and Stephen Chase and André Joyal.

[AN72]   K. B. Athreya and P. E. Ney. *Branching Processes*. Springer-Verlag, New York, 1972. Die Grundlehren der mathematischen Wissenschaften, Band 196.

[AS05]   M. Aguiar and F. Sottile. Structure of the Malvenuto-Reutenauer Hopf algebra of permutations. *Adv. Math.*, 191(2):225–275, 2005.

[AS06]   M. Aguiar and F. Sottile. Structure of the Loday-Ronco Hopf algebra of trees. *J. Algebra*, 295(2):473–511, 2006.

[BB90]   H. Barcelo and N. Bergeron. The Orlik-Solomon algebra on the partition lattice and the free Lie algebra. *J. Combin. Theory Ser. A*, 55(1):80–92, 1990.

[BD92]   D. Bayer and P. Diaconis. Trailing the dovetail shuffle to its lair. *Ann. Appl. Probab.*, 2(2):294–313, 1992.

[BD98]   K. S. Brown and P. Diaconis. Random walks and hyperplane arrangements. *Ann. Probab.*, 26(4):1813–1854, 1998.

[Ber03]   J. Bertoin. The asymptotic behavior of fragmentation processes. *J. Eur. Math. Soc.*, 5(4):395–416, NOV 2003.

[Ber06]   J. Bertoin. *Random Fragmentation and Coagulation Processes*, volume 102 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2006.

[BHJ92]   A. D. Barbour, L. Holst, and S. Janson. *Poisson Approximation*, volume 2 of *Oxford Studies in Probability*. The Clarendon Press Oxford University Press, New York, 1992. Oxford Science Publications.

[BHR99]   P. Bidigare, P. Hanlon, and D. Rockmore. A combinatorial description of the spectrum for the Tsetlin library and its generalization to hyperplane arrangements. *Duke Math. J.*, 99(1):135–174, 1999.

[Bré99]   Pierre Brémaud. *Markov chains*, volume 31 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1999. Gibbs fields, Monte Carlo simulation, and queues.

[BS53]   R. Bott and H. Samelson. On the Pontryagin product in spaces of paths. *Comment. Math. Helv.*, 27:320–337 (1954), 1953.

[Car07]   P. Cartier. A primer of Hopf algebras. In *Frontiers in Number Theory, Physics, and Geometry. II*, pages 537–615. Springer, Berlin, 2007.

[Car09]   G. Carlsson. Topology and data. *Bull. Amer. Math. Soc. (N.S.)*, 46(2):255–308, 2009.

[CCdS06] E. Carlsson, G. Carlsson, and V. de Silva. An algebraic topological method for feature identification. *Internat. J. Comput. Geom. Appl.*, 16(4):291–314, 2006.

[CDM05] S. Chatterjee, P. Diaconis, and E. Meckes. Exchangeable pairs and Poisson approximation. *Probab. Surv.*, 2:64–106 (electronic), 2005.

[CH10] M. A. Conger and J. Howald. A better way to deal the cards. *Amer. Math. Monthly*, 117(8):686–700, 2010.

[Ciu98] M. Ciucu. No-feedback card guessing for dovetail shuffles. *Ann. Appl. Probab.*, 8(4):1251–1269, 1998.

[CV06] M. Conger and D. Viswanath. Riffle shuffles of decks with repeated cards. *Ann. Probab.*, 34(2):804–819, 2006.

[CV07] M. Conger and D. Viswanath. Normal approximations for descents and inversions of permutations of multisets. *J. Theoret. Probab.*, 20(2):309–325, 2007.

[Den12] G. Denham. Eigenvectors for a random walk on a hyperplane arrangement. *Advances in Applied Mathematics*, 48(2):312 – 324, 2012.

[DF09a] P. Diaconis and J. Fulman. Carries, shuffling, and an amazing matrix. *Amer. Math. Monthly*, 116(9):788–803, 2009.

[DF09b] P. Diaconis and J. Fulman. Carries, shuffling, and symmetric functions. *Adv. in Appl. Math.*, 43(2):176–196, 2009.

[DF12] P. Diaconis and J. Fulman. Foulkes characters, eulerian idempotents, and an amazing matrix. *Journal of Algebraic Combinatorics*, pages 1–16, 2012. 10.1007/s10801-012-0343-7.

[DFH11] P. Diaconis, J. Fulman, and S. Holmes. Analysis of Casino Shelf Shuffling Machines. *ArXiv e-prints*, July 2011.

[DFP92] P. Diaconis, J. A. Fill, and J. Pitman. Analysis of top to random shuffles. *Combin. Probab. Comput.*, 1(2):135–155, 1992.

[Dia03] P. Diaconis. Mathematical developments from the analysis of riffle shuffling. In *Groups, Combinatorics & Geometry (Durham, 2001)*, pages 73–97. World Sci. Publ., River Edge, NJ, 2003.

[DR00] P. Diaconis and A. Ram. Analysis of systematic scan Metropolis algorithms using Iwahori-Hecke algebra techniques. *Michigan Math. J.*, 48:157–190, 2000. Dedicated to William Fulton on the occasion of his 60th birthday.

[Dri89] V. G. Drinfel′d. Quasi-Hopf algebras. *Algebra i Analiz*, 1(6):114–148, 1989.

[DSC12] P. Diaconis and L. Saloff-Coste. Convolution powers of complex functions on $\mathbb{Z}$. *ArXiv e-prints*, 2012.

[Fis10] F. Fisher. *CoZinbiel Hopf algebras in combinatorics*. PhD thesis, The George Washington University, 2010.

[Ful01]   J. Fulman. Descent algebras, hyperplane arrangements, and shuffling cards. *Proc. Amer. Math. Soc.*, 129(4):965–973, 2001.

[Gei77]   L. 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)*, volume 579 of *Lecture Notes in Math.*, pages 168–181. Springer, Berlin, 1977.

[GKL⁺95] I. M. Gelfand, D. Krob, A. Lascoux, B. Leclerc, V. S. Retakh, and J.-Y. Thibon. Noncommutative symmetric functions. *Adv. Math.*, 112(2):218–348, 1995.

[GM09]   R. Guralnick and S. Montgomery. Frobenius–Schur indicators for subgroups and the Drinfel′d double of Weyl groups. *Trans. Amer. Math. Soc.*, 361(7):3611–3632, 2009.

[GR89]   A. M. Garsia and C. Reutenauer. A decomposition of Solomon's descent algebra. *Adv. Math.*, 77(2):189–262, 1989.

[Gre95]   J. A. Green. Hall algebras, hereditary algebras and quantum groups. *Invent. Math.*, 120(2):361–377, 1995.

[Gre97]   J. A. Green. Quantum groups, Hall algebras and quantized shuffles. In *Finite reductive groups (Luminy, 1994)*, volume 141 of *Progr. Math.*, pages 273–290. Birkhäuser Boston, Boston, MA, 1997.

[Gri99]   G. Grimmett. *Percolation*, volume 321 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 1999.

[GS97]   C. M. Grinstead and J. L. Snell. *Introduction to Probability*. American Mathematical Society, Providence, RI, 2nd revised edition, 1997.

[Han90]   P. Hanlon. The action of S_n on the components of the Hodge decomposition of Hochschild homology. *Michigan Math. J.*, 37(1):105–124, 1990.

[Hat02]   A. Hatcher. *Algebraic Topology*. Cambridge University Press, Cambridge, 2002.

[Hol97]   J. M. Holte. Carries, combinatorics, and an amazing matrix. *Amer. Math. Monthly*, 104(2):138–149, 1997.

[Hum72]   J. E. Humphreys. *Introduction to Lie Algebras and Representation Theory*, volume 9 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1972.

[JR79]   S. A. Joni and G.-C. Rota. Coalgebras and bialgebras in combinatorics. *Stud. Appl. Math.*, 61(2):93–139, 1979.

[Kas00]   Y. Kashina. A generalized power map for Hopf algebras. In *Hopf Algebras and Quantum Groups (Brussels, 1998)*, volume 209 of *Lecture Notes in Pure and Appl. Math.*, pages 159–175. Dekker, New York, 2000.

[Kol41]   A. N. Kolmogoroff. Über das logarithmisch normale Verteilungsgesetz der Dimensionen der Teilchen bei Zerstückelung. *C. R. (Doklady) Acad. Sci. URSS (N. S.)*, 31:99–101, 1941.

[KR11]   A. Kleshchev and A. Ram. Representations of Khovanov-Lauda-Rouquier algebras and combinatorics of Lyndon words. *Math. Ann.*, 349(4):943–975, 2011.

[KSC78]   V. F. Kolchin, B. A. Sevast′yanov, and V. P. Chistyakov. *Random Allocations*. V. H. Winston & Sons, Washington, D.C., 1978. Translated from the Russian, Translation edited by A. V. Balakrishnan, Scripta Series in Mathematics.

[KT75]   S. Karlin and H. M. Taylor. *A First Course in Stochastic Processes*. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, second edition, 1975.

[Lec04]   B. Leclerc. Dual canonical bases, quantum shuffles and q-characters. *Math. Z.*, 246(4):691–732, 2004.

[LM00]   V. Linchenko and S. Montgomery. A Frobenius–Schur theorem for Hopf algebras. *Algebr. Represent. Theory*, 3(4):347–355, 2000. Special issue dedicated to Klaus Roggenkamp on the occasion of his 60th birthday.

[LMS06]   R. Landers, S. Montgomery, and P. Schauenburg. Hopf powers and orders for some bismash products. *J. Pure Appl. Algebra*, 205(1):156–188, 2006.

[Lot97]   M. Lothaire. *Combinatorics on Words*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1997. With a foreword by Roger Lyndon and a preface by Dominique Perrin; corrected reprint of the 1983 original, with a new preface by Perrin.

[LPW09]   D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson.

[LR95]   P. Lalonde and A. Ram. Standard Lyndon bases of Lie algebras and enveloping algebras. *Trans. Amer. Math. Soc.*, 347(5):1821–1830, 1995.

[Mac95]   I. G. Macdonald. *Symmetric Functions and Hall Polynomials*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, second edition, 1995. With contributions by A. Zelevinsky, Oxford Science Publications.

[Maj95]   S. Majid. *Foundations of Quantum Group Theory*. Cambridge University Press, Cambridge, 1995.

[MM65]   J. W. Milnor and J. C. Moore. On the structure of Hopf algebras. *Ann. Math.*, 81:211–264, 1965.

[Mon93]   S. Montgomery. *Hopf algebras and their actions on rings*, volume 82 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1993.

[Pat91]   F. Patras. Construction géométrique des idempotents Eulériens. Filtration des groupes de polytopes et des groupes d'homologie de Hochschild. *Bull. Soc. Math. France*, 119(2):173–198, 1991.

[Pat93]   F. Patras. La décomposition en poids des algèbres de Hopf. *Ann. Inst. Fourier (Grenoble)*, 43(4):1067–1087, 1993.

[Pat94]   F. Patras. L'algèbre des descentes d'une bigèbre graduée. *J. Algebra*, 170(2):547–566, 1994.

[Pol11]   P. K. Pollett. Quasi-stationary distributions: A bibliography. 2011.

[Reu93]   C. Reutenauer. *Free Lie Algebras*, volume 7 of *London Mathematical Society Monographs. New Series*. The Clarendon Press Oxford University Press, New York, 1993. Oxford Science Publications.

[Reu03]   C. Reutenauer. Free lie algebras. In *Handbook of Algebra*, volume 3, pages 887–903. North-Holland, Amsterdam, 2003.

[Ros95]   M. Rosso. Groupes quantiques et algèbres de battage quantiques. *C. R. Acad. Sci. Paris Sér. I Math.*, 320(2):145–148, 1995.

[Ros97]   M. Rosso. Some applications of quantum shuffles. In *Deformation Theory and Symplectic Geometry (Ascona, 1996)*, volume 20 of *Math. Phys. Stud.*, pages 249–258. Kluwer Acad. Publ., Dordrecht, 1997.

[Ros98]   M. Rosso. Quantum groups and quantum shuffles. *Invent. Math.*, 133(2):399–416, 1998.

[Sal12]   F. Saliola. Eigenvectors for a random walk on a left-regular band. *Advances in Applied Mathematics*, 48(2):306 – 311, 2012.

[SC97]   L. Saloff-Coste. Lectures on finite Markov chains. In *Lectures on Probability Theory and Statistics (Saint-Flour, 1996)*, volume 1665 of *Lecture Notes in Math.*, pages 301–413. Springer, Berlin, 1997.

[SC04]   L. Saloff-Coste. Total variation lower bounds for finite Markov chains: Wilson's lemma. In *Random walks and geometry*, pages 515–532. Walter de Gruyter GmbH & Co. KG, Berlin, 2004.

[Sch87]   W. R. Schmitt. Antipodes and incidence coalgebras. *J. Combin. Theory Ser. A*, 46(2):264–290, 1987.

[Sch93]   W. R. Schmitt. Hopf algebras of combinatorial structures. *Canad. J. Math.*, 45(2):412–428, 1993.

[Sch94]   W. R. Schmitt. Incidence Hopf algebras. *J. Pure Appl. Algebra*, 96(3):299–330, 1994.

[Sch95]   W. R. Schmitt. Hopf algebra methods in graph theory. *J. Pure Appl. Algebra*, 101(1):77–90, 1995.

[Sol68]   L. Solomon. A decomposition of the group algebra of a finite Coxeter group. *J. Algebra*, 9:220–239, 1968.

[SS93]   S. Shnider and S. Sternberg. *Quantum Groups: From Coalgebras to Drinfel′d Algebras, A Guided Tour*. Graduate Texts in Mathematical Physics, II. International Press, Cambridge, MA, 1993.

[SSG11]   O. Schramm, S. Smirnov, and C. Garban. On the scaling limits of planar percolation. *Ann. Probab.*, 39(5):1768–1814, 2011.

[Sta95]   R. P. Stanley. A symmetric function generalization of the chromatic polynomial of a graph. *Adv. Math.*, 111(1):166–194, 1995.

[Sta12]   Richard P. Stanley. *Enumerative combinatorics. Volume 1*, volume 49 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2012.

[Sta99]   R. P. Stanley. *Enumerative Combinatorics. Vol. 2*, volume 62 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1999. With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin.

[Sta01]   R. P. Stanley. Generalized riffle shuffles and quasisymmetric functions. *Ann. Comb.*, 5(3-4):479–491, 2001. Dedicated to the memory of Gian-Carlo Rota (Tianjin, 1999).

[Ste97]   J. R. Stembridge. Enriched P-partitions. *Trans. Amer. Math. Soc.*, 349(2):763–788, 1997.

[Swe69]   M. E. Sweedler. *Hopf algebras*. Mathematics Lecture Note Series. W. A. Benjamin, Inc., New York, 1969.

[TO70]   J. Tate and F. Oort. Group schemes of prime order. *Ann. Sci. École Norm. Sup.*, 3:1–21, 1970.

[vD91]   E. A. van Doorn. Quasi-stationary distributions and convergence to quasi-stationarity of birth-death processes. *Adv. in Appl. Probab.*, 23(4):683–700, 1991.

[Wer09]   W. Werner. Lectures on two-dimensional critical percolation. In *Statistical Mechanics*, volume 16 of *IAS/Park City Math. Ser.*, pages 297–360. Amer. Math. Soc., Providence, RI, 2009.

[Wil90]   H. S. Wilf. *Generatingfunctionology*. Academic Press, Boston, MA, 1990.

[Wil04]   D. B. Wilson. Mixing times of Lozenge tiling and card shuffling Markov chains. *Ann. Appl. Probab.*, 14(1):274–325, 2004.

[WS96]   D. Warren and E. Seneta. Peaks and Eulerian numbers in a random sequence. *J. Appl. Probab.*, 33(1):101–114, 1996.

[Zel81]   A. V. Zelevinsky. *Representations of Finite Classical Groups: A Hopf Algebra Approach*, volume 869 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1981.

[Zho08]   H. Zhou. *Examples of multivariate Markov chains with orthogonal polynomial eigenfunctions*. ProQuest LLC, Ann Arbor, MI, 2008. Thesis (Ph.D.)–Stanford University.