PATTERN HOPF ALGEBRAS

RAUL PENAGUIAO

University of Zurich, Institute of Mathematics

Abstract. This is a study on pattern Hopf algebras in combinatorial structures. We introduce the notion of *combinatorial presheaf*, by adapting the algebraic framework of species to the study of substructures in combinatorics. Afterwards, we consider functions that count the number of patterns of objects and endow the linear span of these functions with a product and a coproduct. In this way, any well behaved family of combinatorial objects that admits a notion of substructure generates a Hopf algebra, and this association is functorial. For example, the Hopf algebra on permutations studied by Vargas in 2014 and the Hopf algebra on symmetric functions are particular cases of this construction.

A specific family of pattern Hopf algebras is of interest, the ones arising from *commutative combinatorial presheaves*. This includes the presheaves on graphs, posets and generalized permutahedra. Here, we show that all the pattern Hopf algebras corresponding to commutative presheaves are free.

We also study a non-commutative presheaf on marked permutations, *i.e.* permutations with a marked element. These objects have an inherent product called inflation, which is an operation motivated by factorization theorems of permutations. In this paper we find new factorization theorems on marked permutations, and use them to show that this is another example of a pattern Hopf algebra that is free.

Contents

1. Introduction 2
1.1. Pattern Hopf algebras from monoids in presheaves 3
1.2. Commutative presheaves 6
1.3. Non-commutative presheaves 7
1.4. Strategy for establishing the freeness of a pattern algebra 9
1.5. Notation and preliminaries 9
2. Substructure algebras 12
2.1. Coinvariants and pattern algebras 12
2.2. Coproducts on pattern algebras 14
2.3. The space of primitive element 15
2.4. The pattern algebra functor 16

E-mail address: raul.penaguiao@math.uzh.ch.
Date: August 11, 2020.
2010 *Mathematics Subject Classification.* 05E05, 16T05, 18D10.
Key words and phrases. marked permutations, presheaves, species, Hopf algebras, free algebras.

2.5. Mangum relations and representability 17
2.6. The Sweedler dual and pattern algebras 18
3. Freeness of commutative presheaves 19
3.1. Relations in general connected associative presheaves 19
3.2. Proof of freeness on commutative presheaves 22
3.3. Marked graphs 23
3.4. Set partitions 24
4. Hopf algebra structure on marked permutations 26
4.1. Unique factorizations 26
4.2. Lyndon factorization on marked permutation 29
4.3. Freeness of the pattern algebra in marked permutations 32
4.4. Proof of unique factorization theorem 33
4.5. Proofs of Lemmas 4.25 to 4.27 38
4.6. Primitive elements, growth rates and asymptotic analysis 41
Aknowledgments 42
References 43

1. Introduction

The notion of substructure is important in mathematics, and particularly in combinatorics. In graph theory, minors and induced subgraphs are the main examples of studied substructures. Substructures of other objects also got some attention: set partitions, trees, paths and, to a bigger extent, permutations, where the study of patterns leads to the concept of permutation class.

A priori unrelated, it has been shown that Hopf algebras are a natural tool in algebraic combinatorics to study discrete objects, like graphs, set compositions and permutations. For instance, the celebrated Hopf algebra on permutations named after Malvenuto and Reutenauer sheds some light on the structure of shuffles in permutations. Other examples of Hopf algebras in combinatorics that are relevant to this work are the Hopf algebra on symmetric functions (described for instance in [Sta86]), and the permutation pattern Hopf algebra introduced by Vargas in [Var14].

With that in mind, we build upon the notion of species, as presented in [AM10] by Aguiar and Mahajan, in order to connect these two areas of algebraic combinatorics. We propose to enrich a species with restriction maps to obtain what we call a *combinatorial presheaf*. With this combinatorial data, we show a functorial construction of a pattern algebra $\mathcal{A}(h)$ from any given combinatorial presheaf h. By further considering an associative product in our objects, we can endow $\mathcal{A}(h)$ with a coproduct that makes it a bialgebra, and under specific circumstances a Hopf algebra. Main examples of combinatorial presheaves are words, graphs and permutations. Examples of associative products on combinatorial objects are the disjoint union on graphs or the direct sum on permutations.

The algebras obtained from a combinatorial presheaf are always commutative. In analyzing Hopf algebras, it is of particular interest to show that such algebras are free commutative (henceforth, we simply say *free*), and to construct free generators of the

algebra structure. The fact that a Hopf algebra is a free algebra has several applications. For instance, in [Foi12], it was shown that any graded free and cofree Hopf algebra is self dual. Moreover, the self dual Hopf algebras are characterized by studying their primitive elements. The freeness of a Hopf algebra also allows us to gain some insight on the character group of the Hopf algebra, see for instance [Sup19]. It can also be used under duality maps to establish cofreeness of Hopf monoids, as described in the methods of Möbius inversion in [San19]. If the Hopf algebra $H = \bigcup_{n\geq 0} H^{(n)}$ is a filtered Hopf algebra with $P(H)$ primitive space, computing the dimension of $H^{(n)} \cap P(H)$ is a classical problem in Hopf algebras and may have applications, for instance in proving that a give Hopf algebra is not a Hopf subalgebra of another one.

In this paper, we show that any commutative combinatorial presheaf gives rise to a pattern algebra that is free commutative. We also study a non-commutative combinatorial presheaf on marked permutations, where we establish the freeness, construct the free elements with help of Lyndon words, and enumerate the primitive elements of the pattern Hopf algebra on marked permutations. In the remaining part of this section, we present these results with more detail, and describe the methods for proving them.

1.1. **Pattern Hopf algebras from monoids in presheaves.** Let $\mathtt{Set}_{\hookrightarrow}$ be the category whose objects are finite sets and morphisms are injective maps between finite sets. Let also $\mathtt{Set}_{\times}$ be the category whose objects are finite sets and morphisms are bijective maps between finite sets. Write $\mathtt{Set}$ for the usual category on finite sets.

A **species** is a functor from $\mathtt{Set}_{\times}$ to itself. Hence, a species h is described by an assignment of each set I to a finite set $h[I]$, together with some relabeling map for each bijection. Species occur very naturally in combinatorics as a way of describing the combinatorial structures on finite sets, for instance graph structures on a vertex set or a poset on a ground set.

Definition 1.1 (Combinatorial presheaves)**.** A *combinatorial presheaf* (or a *presheaf*, for short) is a contravariant functor from $\mathtt{Set}_{\hookrightarrow}$ to $\mathtt{Set}$. A morphism of combinatorial presheaves is simply natural transformation of functors. In this way, we have the category $\mathtt{CPSh}$ of combinatorial presheaves.

In this form, a presheaf is simply a species enriched with restriction maps $\mathbf{res}_J : h[I] \to h[J]$ for each inclusion $J \subseteq I$ is a way that is functorial, that is if $J_1 \subseteq J_2$, then $\mathbf{res}_{J_2} \circ \mathbf{res}_{J_1} = \mathbf{res}_{J_1}$. The notion and the name of presheaves has been around in category theory and geometry for some time, where it generally refers to contravariant functors from the category of open sets of a topology with inclusions as morphisms. Main examples of combinatorial presheaves are graphs, set compositions, and permutations, see Example 1.3. In general, any combinatorial object that admits a notion of restriction admits a presheaf structure.

In presheaves, two objects $a \in h[I], b \in h[J]$ are said to be *isomorphic objects*, or $a \sim b$, if there is a bijection $f : I \to J$ such that $h[f](b) = a$. The equivalence classes are also called *coinvariants*. Let $h[n]$ denote the objects of type h on the set $[n] = \{1, \ldots, n\}$. The collection of coinvariants of a presheaf h is denoted by $\mathcal{G}(h) = \bigcup_{n\geq 0} h[n]/_{\sim}$.

Definition 1.2 (Patterns in presheaves)**.** Let h be a presheaf, and consider two objects $a \in h[I], b \in h[J]$. We say that $J' \subseteq I$ is a *pattern* of a in b if $b|_{J'} \sim a$. We define the

pattern function

$$\mathbf{p}_a(b) := |\{J' \subseteq J \text{ s.t. } \mathbf{res}_{J'}(b) \sim a\}| \ .$$

In Proposition 2.1, we observe that this definition only depends on the isomorphism classes of a and b. Hence, we can consider $\{\mathbf{p}_a\}_{a\in\mathcal{G}(h)}$ as a family of functions from $\mathcal{G}(h)$ to $\mathbb{Q}$, indexed by $\mathcal{G}(h)$.

Example 1.3 (The presheaf on permutations)**.** To fit the framework of presheaves, we use a rather unusual definition of permutations introduced in [ABF18]. There, a permutation on a set I is seen as a pair of orders in I. This relates to the usual notion of a permutation as bijections in the following way: if we order the elements of $I = \{a_1 \leq_P \cdots \leq_P a_k\} = \{b_1 \leq_V \cdots \leq_V b_k\}$, then this defines a bijection via $a_i \mapsto b_i$ in I. Conversely, for any bijection f on I, there are several pairs of orders $(\leq_P, \leq_V)$ that correspond to the bijection f, all of which are isomorphic. By restricting the orders on I to orders on a subset $J \subseteq I$, we obtain a notion of restriction of permutations. The resulting presheaf structure is denoted by `Per`.

It will be useful to represent permutations in I as square diagrams labeled by I. This is done in the following way: we place the elements of I in an $|I| \times |I|$ grid so that the elements are placed horizontally according to the $\leq_P$ order, and vertically according to the $\leq_V$ order. For instance, the permutation $\pi = \{1 <_P 2 <_P 3, 2 <_V 1 <_V 3\}$ in $\{1, 2, 3\}$ can be represented as

(1)
$$\begin{array}{|c|c|c|}\hline & & 3 \\ \hline 1 & & \\ \hline & 2 & \\ \hline \end{array} \ .$$

In this way, there are $(n!)^2$ many elements in $\mathtt{Per}[n]$. Up to relabeling, we can represent a permutation as a diagram with one dot in each column and row. Thus, $\mathcal{G}(\mathtt{Per})$ has $n!$ many isomorphism classes of permutations of size n, as expected.

A crucial observation is that this notion recovers the usual concept of permutation pattern already present in the literature. Specifically, the number of occurrences of a permutation τ in π as described in [Wil02] is precisely $\mathbf{p}_\tau(\pi)$, and a permutation π *avoids* τ in the sense describe in [Knu11] if $\mathbf{p}_\tau(\pi) = 0$.

Denote the family of functions $A \to B$ by $\mathcal{F}(A, B)$. If h is a combinatorial presheaf, then the linear span of the pattern functions is a linear subspace $\mathcal{A}(h) \subseteq \mathcal{F}(\mathcal{G}(h), \mathbb{Q})$ of the space of functions in $\mathcal{G}(h)$ taking rational values.

Theorem 1.4. The vector space $\mathcal{A}(h)$ is closed under pointwise multiplication and has a unit. It forms an algebra, called the *pattern algebra*. More precisely, we have the product rule

(2)
$$\mathbf{p}_a\, \mathbf{p}_b = \sum_c \binom{c}{a, b} \mathbf{p}_c \,,$$

where we define the coefficients $\binom{c}{a,b}$ below in (7) as the number of "quasi-shuffles" of a, b that result in c.

Quasi-shuffles of objects have been studied in several contexts as a vague notion of merging objects together. For details on quasi-shuffles of combinatorial objects, the interested reader can see [Hof00, AM10, FPT16]. We remark that a definition of quasi-shuffles for our purposes is given in Eq. (7).

We introduce now the *Cauchy product* $\odot$ on the category of combinatorial presheaves, that associates to two combinatorial presheaves f, g the combinatorial presheaf

$$f \odot g : I \mapsto \biguplus_{A \uplus B = I} f[A] \times g[B] \,.$$

The unit for this product is the unique presheaf that satisfies $\mathcal{E}[A] = \emptyset$ for $A \neq \emptyset$, and $\mathcal{E}[\emptyset] = \{\diamond\}$. This functor acts as expected on morphisms.

Product structures on categories were debated in [Mac63]. There, it was shown that specifying a *monoidal category* structure, that is, a bifunctor $\otimes : \texttt{CPSh} \times \texttt{CPSh} \to \texttt{CPSh}$ that satisfies some specific associativity properties, is enough to be able to define a *monoidal structure*. In this way, the Cauchy product gives rise to the notion of a monoid structure in the category of presheaves, or simply an *associative presheaf*. This is a triple $(h, *, 1)$, where h is a combinatorial presheaf, $*$ is a natural transformation $h \odot h \Rightarrow h$, and $1 \in h[\emptyset]$ a unit that satisfy classical axioms of associativity and unit. We detail further in Section 1.5 below.

Examples of associative operations on combinatorial presheaves are the disjoint union of graphs and the direct sum of permutations. Another less standard example, which we study in this paper, is the inflation of marked permutations, defined below.

Observe that the associative product $*$ in our combinatorial objects is a natural transformation. This means that the product is stable with respect to relabelings and restrictions, so we can also define the corresponding product on $\mathcal{G}(h)$, which we denote by $\cdot$ for the sake of distinction (see Definition 2.6 for details). With this, we introduce the following coproduct in the pattern algebra $\mathcal{A}(h)$:

$$\Delta\, \mathbf{p}_a = \sum_{\substack{b,c \in \mathcal{G}(h) \\ a = b \cdot c}} \mathbf{p}_b \otimes \mathbf{p}_c \,. \tag{3}$$

where the sum runs over coinvariants b, c such that $a = b \cdot c$. The main property that motivates this operation in $\mathcal{A}(h)$ is that, under the natural identification of the function algebra $\mathcal{F}(\mathcal{G}(h), \mathbb{Q})^{\otimes 2}$ as a subspace of $\mathcal{F}(\mathcal{G}(h) \times \mathcal{G}(h), \mathbb{Q})$, we have

$$\Delta\, \mathbf{p}_a(b, c) = \mathbf{p}_a(b \cdot c) \,. \tag{4}$$

This is shown in Theorem 2.8. The relation (4) is central in establishing that the coproduct Δ is compatible with the product in $\mathcal{A}(h)$.

Theorem 1.5. If $(h, *, 1)$ is an associative presheaf such that $|h[\emptyset]| = 1$, then the pattern algebra of h together with this coproduct, and a natural choice of counit, forms a Hopf algebra.

The reason we need the presheaf to be connected is so we can find an antipote through the so called *Takeuchi formula*, introduced in [Tak71].

Some known Hopf algebras can be constructed as the pattern algebra of a combinatorial presheaf. An example is Sym, the Hopf algebra of *symmetric functions*. This Hopf algebra has a basis indexed by partitions, and corresponds to the pattern Hopf algebra of the presheaf on set partitions (see details in Section 3.4). The pattern Hopf algebra corresponding to the presheaf on permutations described above was introduced by Vargas in [Var14]. Some other Hopf algebras constructed here, like the ones on graphs and on marked permutations below, are new.

We also establish some general properties of the pattern Hopf algebras, like describing its primitive elements, finding the inverse of the so called *pattern action* in $\mathcal{A}(h)$, and relating $\mathcal{A}(h)$ with the Sweedler dual of an algebra generated by $*$.

Remark 1.6. It is common to use category theory tools to construct Hopf algebras in a mechanical way, as it gives us more algebraic tools to understand combinatorial objects. This is the case with the Fock functors (see [AM10, Chapter 15]), where from a Hopf monoid in species we construct four distinct Hopf algebras.

It is then meaningful to compare the construction of a pattern algebra with the Fock functors. In fact, a cocomutative comonoid in set species is precisely a presheaf. Furthermore, an associative presheaf is a cocomutative bimonoid in set species. The coalgebra structure of the pattern Hopf algebras that we construct here is a subcoalgebra of the dual algebra of the so called *bosonic Fock functor* of these comonoids in linearized set species. However, the algebra structure is in general different.

Specifically, on the combinatorial presheaf on graphs introduced below, the corresponding coalgebra structure is the dual of the well known incidence Hopf algebra introduced in [Sch94].

1.2. **Commutative presheaves.** In this paper, we focus on the problem of proving the freeness of some pattern algebras. The first case that we want to explore is the one of commutative presheaves. An associative presheaf $(h, *, 1)$ is called *commutative* if $*$ is commutative, that is for any $a \in h[I], b \in h[J]$ we have that $a * b = b * a$, see Definition 1.16.

As it turns out, this is enough to guarantee the freeness of the pattern Hopf algebra. This is one of the main results of this paper.

Theorem 1.7 (Commutative combinatorial presheaves are free)**.** The pattern Hopf algebra of an associative combinatorial presheaf with a commutative product $*$ is free. The free generators are the pattern functions indexed by the irreducible objects with respect to $*$.

The proof of this result is presented in Section 3. The main ingredient for this result is Corollary 3.4, a surprising structure result on associative presheaves. If one wishes to describe the associative structure of $\mathcal{G}(h)$ under the product $\cdot$, it can be done as it is for groups: by prescribing a set of generators, whose role is played by the *irreducible coinvariants*, and a collection of relations that these satisfy. Corollary 3.4 says that on the case of associative presheaves, the collection of relations is very restricted, allowing only for relations that use the same factors but with different orders.

We remark that the algebraic structure of the pattern Hopf algebra $\mathcal{A}(h)$ is defined independently from its associative product. It follows that the pattern Hopf algebra

of a combinatorial presheaf is free whenever we *can* endow h with a commutative associative product, regardless of whether that is the associative presheaf at hand. In other words, given $(h, *_1, 1)$ and $(h, *_2, 1)$ connected associative presheaf structures on the same presheaf h, such that $*_2$ is commutative, then $\mathcal{A}(h) = \mathcal{A}(h, *_1, 1)$ is a free Hopf algebra. An example of this is presented on the presheaf of marked graphs, in Section 3.3 below. Notwithstanding, all freeness proofs on pattern Hopf algebras uses both the pattern structure and the associative structure. In fact, the role played by the associative structure is streamlined and inflexible, as described in Section 1.4, whereas the role of the pattern structure is somewhat harder to deal with.

Example 1.8 (Presheaf on graphs)**.** There are many notions of patterns on graphs: minors, subgraphs and induced subgraphs are among some of those. The one that forms a presheaf is the one of induced subgraphs.

This forms a combinatorial presheaf $\mathtt{Gr}$ that can be endowed with the associative product of disjoint union of graphs $\uplus$. This is in fact a commutative product, so $\mathcal{A}(\mathtt{Gr})$ is a free algebra, and the pattern functions of connected graphs are the free generators, that is:

$$\mathcal{A}(\mathtt{Gr}) = \mathbb{Q}[\mathbf{p}_G \,|\, G \text{ connected graph}]\,.$$

This was already proved in [Whi32, Theorem 3], where if a function satisfies (2), the connected graphs (there called non-separable) are enough to determine its values, that the remaining values are obtained via polynomial expressions, and that no other polynomial expressions hold for such a generic function.

1.3. **Non-commutative presheaves.** We return to the presheaf on permutations. This is an example of a non-commutative associative presheaf.

Example 1.9 (Permutations and their pattern Hopf algebra)**.** To the presheaf $\mathtt{Per}$ it corresponds a pattern algebra $\mathcal{A}(\mathtt{Per})$ as discussed above. We can further consider $\mathtt{Per}$ with a monoid structure via the direct sum of permutations $\oplus$, defined as follows: Suppose that $\pi \in \mathtt{Per}[I], \tau \in \mathtt{Per}[J]$ are two permutations based on the disjoint sets I, J, respectively. The permutation $\pi \oplus \tau \in \mathtt{Per}[I \sqcup J]$ is the pair of total orders $(\leq_P^{\oplus}, \leq_V^{\oplus})$ extending both of the respective orders from π, τ to $I \sqcup J$ by forcing that $i \leq_P^{\oplus} j$ and $i \leq_V^{\oplus} j$ for any $i \in I, j \in J$. Correspondingly, the diagram of $\pi \oplus \tau$ results from the ones from π, τ as follows"

$$\pi \oplus \tau = \begin{array}{|c|c|} \hline & \tau \\ \hline \pi & \\ \hline \end{array}\ .$$

We note that this is not a commutative presheaf: in general, $\pi \oplus \tau$ is a different permutation than $\tau \oplus \pi$.

As mentioned above, the pattern Hopf algebra on permutations is the one discussed by Vargas in [Var14], where it is shown that it is free. There, free generators were constructed. These generators correspond to Lyndon words of $\oplus$-indecomposable permutations, see [CFL58] for an introduction to combinatorics of Lyndon words.

In this paper we explore other associative presheaves that are non-commutative. Taking the presheaf on permutations as our starting point, we wish to study monoidal

structures that are more complex than the $\oplus$ product, but still allow to establish the freeness property. We suggest the presheaf on marked permutations, which is equipped with the inflation product. This product is motivated by the inflation procedure on permutations described in [AAK03].

In this way, the presheaf on marked permutations introduced below will be one main focus of this paper: Section 4 will be dedicated to the freeness problem on this presheaf, as well as enumerating the dimension of the primitive space of the pattern Hopf algebra.

Example 1.10 (Marked permutations and their pattern Hopf algebra)**.** A marked permutation π^* on I is a pair of orders $(\leq_P, \leq_V)$ on the set $I \sqcup \{*\}$. Intuitively, this gives us a rearrangement of the elements of $I \sqcup \{*\}$, where one element is special and marked. The relabelings and restriction maps are the natural ones borrowed from orders, giving us a combinatorial presheaf, that we call $\texttt{MPer}$. We can represent a marked permutation in a diagram, as we do for permutations. Note that in this case the marked element $*$ never changes position after relabelings. Take for instance the marked permutations $\pi^* = (1 <_P 2 <_P *, 2 <_V 1 <_V *)$, $\tau^* = (* <_P 1 <_P 2, 1 <_V * <_V 2)$, and $\sigma^* = (* <_P 2 <_P 1, 2 <_V * <_V 1)$. Observe that there is no isomorphism between π^* and τ^*, whereas there is one between τ^* and σ^*, via the relabeling $1 \mapsto 2, 2 \mapsto 1$.

$$\pi^* = \begin{array}{|c|c|c|}\hline & & * \\ \hline 1 & & \\ \hline & 2 & \\ \hline\end{array} \qquad \tau^* = \begin{array}{|c|c|c|}\hline & & 2 \\ \hline * & & \\ \hline & 1 & \\ \hline\end{array} \qquad \sigma^* = \begin{array}{|c|c|c|}\hline & & 1 \\ \hline * & & \\ \hline & 2 & \\ \hline\end{array}\ .$$

In this way, there are $((n+1)!)^2$ many elements in $\texttt{MPer}[n]$. Up to relabeling, we can represent a marked permutation as a diagram with one dot in each column and row, where a particular dot is the distinguished element $*$. Therefore, $\mathcal{G}(\texttt{MPer})$ has $(n+1) \times (n+1)!$ many isomorphism classes of marked permutations of size n.

Definition 1.11 (Inflation product)**.** The inflation product $\star$ in marked permutations is defined as follows: Given two marked permutations $\tau^* \in \texttt{MPer}[I]$ and $\pi^* \in \texttt{MPer}[J]$ with I, J disjoint sets, the inflation product $\tau^* \star \pi^* \in \texttt{MPer}[I \sqcup J]$ is a marked permutation resulting from replacing in the diagram of τ^* the marked element with the diagram of π^*. Here is an example:

$$\tau^* = \begin{array}{|c|c|c|}\hline & \odot & \\ \hline & & \bullet \\ \hline \bullet & & \\ \hline\end{array} \qquad \pi^* = \begin{array}{|c|c|}\hline \bullet & \\ \hline & \odot \\ \hline\end{array} \qquad \tau^* \star \pi^* = \begin{array}{|c|c|c|c|}\hline & \bullet & & \\ \hline & & \odot & \\ \hline & & & \bullet \\ \hline \bullet & & & \\ \hline\end{array}\ .$$

Remark 1.12. Note that if $\pi^* \star \tau^* = \pi^* \star \sigma^*$, then $\tau^* = \sigma^*$. Similarly, if $\tau^* \star \pi = \sigma^* \star \pi^*$, then $\tau^* = \sigma^*$.

It is straightforward to observe that this is an associative presheaf with $\texttt{MPer}[\emptyset]$, where the unit is the unique marked permutation on $\emptyset$, denoted by $\bar{1}$. We call it the *presheaf of marked permutations*. Hence, from Theorem 1.13, the algebra $\mathcal{A}(\texttt{MPat})$ is indeed a Hopf algebra. This is not a commutative presheaf, so Theorem 1.7 does not apply. However, we still have the following:

Theorem 1.13 (Freeness of $\mathcal{A}(\texttt{MPat})$)**.** The pattern algebra $\mathcal{A}(\texttt{MPat})$ is free.

To establish the freeness of $\mathcal{A}(\texttt{MPat})$ we present a unique factorization theorem on marked permutations with the inflation product in Corollary 4.12. This is an analogue of [AAK03, Theorem 1] for the inflation product on marked permutations. With it, we find generators of the algebra on marked permutations, and use tools from word combinatorics, specifically the Lyndon factorization of words in [CFL58], to show that these generators are free generators.

In fact, Lyndon words are commonly used to establish the freeness of algebras. Examples are the shuffle algebra in [Rad79] (see also [GR14, Chapter 6]), the algebra of quasisymmetric functions in [Haz01, Theorem 8.1], and the algebra on word quasisymmetric functions in non-commutative variables, in [BZ09].

In Section 4 we also enumerate the dimension of the primitive space of the pattern Hopf algebra $\mathcal{A}(\texttt{MPer})$, which corresponds to enumerating the marked permutations that are irreducible with respect to the inflation product.

1.4. **Strategy for establishing the freeness of a pattern algebra.** We now discuss the general strategy that we employ when establishing the freeness of a pattern Hopf algebra. In particular, we clarify what is the relation between unique factorization theorems and freeness of the algebra of interest. Let $\mathcal{S} \subseteq \mathcal{G}(h)$ be a collection of objects in a presheaf h. Then the set $\{\mathbf{p}_s \,|s \in \mathcal{S}\}$ is a set of free generators of $\mathcal{A}(h)$ if the set

$$\left\{ \prod_{s\in S} \mathbf{p}_s \,\middle|\, S \text{ multiset of elements of } \mathcal{S} \right\},$$

is a basis of $\mathcal{A}(h)$. This is usually established by connecting this set with the set $\{p_a | a \in \mathcal{G}(h)\}$, which is known to be a basis by Remark 2.2. This connection is done with the following ingredients:

- An order $\preceq$ in $\mathcal{G}(h)$.
- A bijection f between $\{\prod_{s\in S} \mathbf{p}_s \,|S$ multiset of elements of $\mathcal{S}\}$ and $\mathcal{G}(h)$, which is usually phrased in terms of a *unique factorization theorem*. See for instance Theorem 4.21.
- The property that, for any S multiset of $\mathcal{S}$,

$$\prod_{s\in S} \mathbf{p}_s = \sum_{t \preceq f(S)} c_{t,S}\, \mathbf{p}_t\,, \tag{5}$$

  with non-negative coefficients $c_{t,s}$ such that $c_{f(S),S} \neq 0$.

These are enough to establish the desired freeness. In the commutative presheaf case, the unique factorization theorem is the one that we naively would expect, see Theorem 3.3. For the case of presheaf on permutations and the presheaf on marked permutations, the construction of the set $\mathcal{S}$ is more technical.

Conjecture 1.14. The pattern Hopf algebra of any associative presheaf is free.

1.5. **Notation and preliminaries.**

1.5.1. *Species and monoidal functors.* If h is a combinatorial presheaf, and if $a \in h[I]$, we define the *size* of a as $|a| := |I|$ and the *indexing set* of a as $\mathbb{X}(a) = I$.

Example 1.15. The species $\mathcal{E}$ with only one object $\diamond$, which has size zero, has a unique presheaf structure. The presheaf of graphs

$$\mathtt{Gr}[I] = \{\text{graphs with vertex set } I\},$$

results from the usual species structure by adding the natural graph restrictions.

In this way, $\mathbb{X}(\diamond) = \emptyset$, and if G is a graph in the vertex set V, we have that $\mathbb{X}(G) = V$ and $|G| = |V|$.

Recall that we are given a bifunctor $\odot$ that endows the category of combinatorial presheaves with a monoidal category structure, as introduced in [AM10].

Definition 1.16 (Associative presheaf)**.** An *associative presheaf* is a monoid in $\mathtt{CPSh}$, that is, is a combinatorial presheaf h together with natural transformations $\eta : h \odot h \Rightarrow h$ and $\iota : \mathcal{E} \Rightarrow h$ that satisfy associativity and unit conditions. We use, for $a \in h[I]$ and $b \in h[J]$, the notation $\eta_{I,J}(a, b) = a * b$. We also denote the unit by $1 := \iota[\emptyset](\diamond) \in h[\emptyset]$.

This is said to be *commutative* if, for any $a \in h[I], b \in h[J]$ with $I \cap J = \emptyset$, we have $a * b = b * a$.

Thus, a product on a presheaf way simply describes how to *merge* objects of a certain type h that are based in disjoint sets.

Observation 1.17 (Naturality axioms in associative presheaves)**.** The naturality of η, the associativity and unit conditions correspond to, respectively,

- For all I, J disjoint sets, all $a \in h[I], b \in h[J]$ and all $A \subseteq I, B \subseteq J$, we have $(a * b)|_{A \sqcup B} = a|_A * b|_B$.
- For all I, J, K disjoint sets and all $a \in h[I], b \in h[J], c \in h[K]$, we have $(a*b)*c = a * (b * c)$.
- For any set I, and $a \in h[I]$, we have $a * 1 = 1 * a = a$.

Remark 1.18 (Monoidal product and quasi-shuffle)**.** In an associative presheaf $(h, *, 1)$, let $a \in h[I], b \in h[J]$ with I, J disjoint sets. Then it may not be the case that $a|_\emptyset = 1$, that $(a * b)|_I = a$ or that $(a * b)|_J = b$.

This is the case, however, when h is a connected presheaf. It follows that in an associative connected presheaf h, we have that $\binom{a*b}{a,b} \geq 1$.

Definition 1.19. Let $(h, *_h, 1_h)$ and $(j, *_j, 1_j)$ be associative presheaves, and let f be a *presheaf morphism* between $(h, *_h, 1_h)$ and $(j, *_j, 1_j)$. This is an *associative presheaf morphism* if it preserves the unit and the associative product of the associative presheaves.

That is, $f : h \Rightarrow j$ is an associative presheaf morphism if it is a presheaf morphism that satisfies $f(1_h) = 1_j$ and $f(b' *_h c') = f(b') *_j f(c')$ for any $b' \in h[I], c' \in h[J]$.

1.5.2. *Preliminaries on permutations and marked permutations.*

Definition 1.20 (Presheaf on permutations - morphisms and notation)**.** If we consider a permutation π on a set I, that is, a pair $(\leq_P, \leq_V)$ of total orders in I, we write $\mathbb{X}(\pi) = I$. If $f : J \to I$ is an injective map, the preimage of each order $\leq_P, \leq_V$ is well defined and is also a total order in J. This defines the permutation $\mathtt{Per}[f](\pi)$.

We can write a permutation in its two-line notation, as $\begin{smallmatrix} a_1,\dots,a_k \\ b_1,\dots,b_k \end{smallmatrix}$ where $a_1 \leq_V a_2 \leq_V \dots$ and $b_1 \leq_P b_2 \leq_P \dots \leq_P b_k$. If we identify $b_1, \dots, b_k$ with $1, \dots, k$, respectively, we can disregard the bottom line. This also disregards the indexing set I, and in fact any two isomorphic permutations have the same representation with the one line notation.

To the unique permutation in the empty set we call the *trivial permutation* and denote it $\emptyset$.

Definition 1.21 (The $\ominus$ operation)**.** Given two permutations, π, σ, we have already introduced the product $\pi \oplus \sigma$. We now define the permutation $\pi \ominus \tau \in \mathtt{Per}[I \sqcup J]$ as the pair of total orders $(\leq_P^{\ominus}, \leq_V^{\ominus})$ extending the respective ones from π, τ to $I \sqcup J$ by forcing that $i \leq_P^{\ominus} j$ and $i \geq_V^{\ominus} j$ for any $i \in I, j \in J$. Correspondingly, the diagram of $\pi \ominus \tau$ results from the ones from π, τ as

$$\pi \ominus \tau = \begin{array}{|c|c|} \hline \pi & \\ \hline & \tau \\ \hline \end{array} \ .$$

It is a routine observation to check that both $\oplus, \ominus$ are associative products on $\mathtt{Per}$, and that $\emptyset$ is the unit of both operations, by simply checking that all properties in Observation 1.17 are fulfilled.

Definition 1.22 (Marked permutations)**.** If we consider a marked permutation π^* on a set I, that is, a pair $(\leq_P, \leq_V)$ of total orders in I^*, we write $\mathbb{X}(\pi^*) = I$. If $f : J \to I$ is an injective map, this can be extended canonically to an injective map $f^* : J^* \to I^*$. Thus, the preimage of each order $\leq_P, \leq_V$ under f^* is well defined and is also a total order in J^*. This defines the marked permutation $\mathtt{MPer}[f](\pi^*)$.

Note that a relabeling of the permutation π^* in I^* is a relabeling of the corresponding marked permutation in I if the relabeling preserves the marked elements.

We can also write marked permutations in a one line notation, where we add a marker over the position of $*$. The resulting notation only disregards the indexing set I, and so any two isomorphic marked permutations have the same one line notation. Note that for each permutation of size n it corresponds n different non-isomorphic marked permutations of size $n-1$, one for each possible marked position.

Example 1.23. If we consider $(1 <_P 2 <_P * <_P 4,\ 1 <_V * <_V 4 <_V 2)$, a marked permutation on $\{1, 2, 4\}$, its representation with the one line notation is $14\bar{2}3$.

The marked permutation $\tau^* = (73 <_P x <_P * <_P 47,\ 73 <_V * <_V x <_V 47)$ is based on the set $I = \{x, 47, 73\}$ and has a one line representation $13\bar{2}4$. Consider now $\pi^* = (1 <_P * <_P 2, 1 <_V * <_V 2)$ and $\sigma^* = (1 <_P * <_P 2, * <_V 1 <_V 2)$ marked permutations in $\{1, 2\}$ So the marked permutations τ^*, π^*, σ^* correspond to the one line notations below

$$\tau^* = 13\bar{2}4 = \begin{array}{|c|c|c|c|} \hline & & & \cdot \\ \hline & \cdot & & \\ \hline & & \odot & \\ \hline \cdot & & & \\ \hline \end{array}, \ \pi^* = 1\bar{2}3 = \begin{array}{|c|c|c|} \hline & & \cdot \\ \hline & \odot & \\ \hline \cdot & & \\ \hline \end{array}, \ \sigma^* = 2\bar{1}3 = \begin{array}{|c|c|c|} \hline & & \cdot \\ \hline \cdot & & \\ \hline & \odot & \\ \hline \end{array}. \tag{6}$$

Then, we have that π^*, σ^* are patterns of τ^*, because $J = \{73, 47\}$ is an occurrence of π^* in τ^*, and $J' = \{x, 47\}$ is an occurrence of σ^* in τ^*.

2. Substructure algebras

In this section, we present the properties of the framework of combinatorial presheaves and associative presheaves, introduced above, and describe the construction of the pattern Hopf algebra mentioned in Theorem 1.5. We establish that for a combinatorial presheaf h, $\mathcal{A}(h)$ is an algebra (see Theorem 2.3). Moreover, if h is connected and endowed with an associative structure, then $\mathcal{A}(h)$ is a Hopf algebra (see Theorem 2.8). We also describe the space of primitive elements in $\mathcal{A}(h)$, clarify that $\mathcal{A}$ is in fact functorial (see Theorem 2.11), and find some identities and properties of the pattern funtions in Propositions 2.14 and 2.15.

2.1. **Coinvariants and pattern algebras.** Given a combinatorial presheaf h, let $[n] = \{1, \dots, n\}$ for n non-negative integer. Then, recall that we define the *coinvariants* of h as the family

$$\mathcal{G}(h) = \biguplus_{n \geq 0} h[n]_{\sim} \,.$$

For an object $a \in \biguplus_J h[J]$, we define the *pattern function* $\mathbf{p}_a : \biguplus_J h[J] \to \mathbb{Q}$ as follows: if $b \in h[I]$, then

$$\mathbf{p}_a(b) := |\{J' \subseteq I \text{ s.t. } b|_{J'} \sim a\}| \,.$$

Proposition 2.1. This definition only depends on the isomorphism type of a, b. That is, if $a_1 \in h[I_1], a_2 \in h[I_2], b_1 \in h[J_1], b_2 \in h[J_2]$ are so that $a_1 \sim a_2, b_1 \sim b_2$ then $\mathbf{p}_{a_1}(b_1) = \mathbf{p}_{a_2}(b_2)$.

This allows us to define the pattern functions $\mathbf{p}_a$ for $a \in \mathcal{G}(h)$ as function in $\mathcal{F}(\mathcal{G}(h), \mathbb{Q})$.

Proof. Because $a_1 \sim a_2$, we have that

$$\{J' \subseteq J_1 \text{ s.t. } b_1|_{J'} \sim a_1\} = \{J' \subseteq J_1 \text{ s.t. } b_1|_{J'} \sim a_2\} \,,$$

so $\mathbf{p}_{a_1}(b_1) = \mathbf{p}_{a_2}(b_1)$. It remains to prove that $\mathbf{p}_{a_2}(b_1) = \mathbf{p}_{a_2}(b_2)$.

Because $b_1 \sim b_2$, there exists some bijective map $f : J_1 \to J_2$ such that $h[f](b_2) = b_1$. This bijection lifts to a bijection between $\{\tilde{J} \subseteq J_2\}$ and $\{J' \subseteq J_1\}$, via $\tilde{J} \mapsto f^{-1}(\tilde{J})$.

We claim that this in fact restricts to a bijection between $\{\tilde{J} \subseteq J_2 \text{ s.t. } b_2|_{\tilde{J}} \sim a_2\}$ and $\{J' \subseteq J_1 \text{ s.t. } b_1|_{J'} \sim a_2\}$. Indeed, let $\tilde{J} \subseteq J_1$, let $\mathrm{inc}_{J_2,\tilde{J}} : \tilde{J} \to J_2$, $\mathrm{inc}_{J_1,f^{-1}(\tilde{J})}$ denote the inclusion maps of $\tilde{J}$ in J_2 and $f^{-1}(\tilde{J})$ in J_1, respectively. Then

$$b_1|_{f^{-1}(\tilde{J})} = h[f^{-1} \circ \mathrm{inc}_{J_2,\tilde{J}}](b_2) = h[\mathrm{inc}_{J_1,f^{-1}(\tilde{J})} \circ f^{-1}](b_2) = h[f^{-1}](b_2|_{\tilde{J}}) \,,$$

so $b_2|_{\tilde{J}} \sim b_1|_{f^{-1}(\tilde{J})}$. This proves that $\mathbf{p}_{a_2}(b_1) = \mathbf{p}_{a_2}(b_2)$. □

Recall that we write

$$\mathcal{A}(h) := \mathrm{span}\{\mathbf{p}_a \,|\, a \in \mathcal{G}(h)\} \subseteq \mathcal{F}(\mathcal{G}(h), \mathbb{Q}) \,,$$

for the linear space spanned by all pattern functions.

Remark 2.2. If two coinvariants a, b are such that $|a| \geq |b|$ and $a \neq b$, then $\mathbf{p}_a(b) = 0$. We also have $\mathbf{p}_b(b) = 1$. Hence, the set $\{\mathbf{p}_a \,|a \in \mathcal{G}(h)\}$ is a basis of $\mathcal{A}(h)$.

For a, b objects and $c \in h[C]$, we defined the *quasi-shuffle* number as follows:

$$\binom{c}{a,b} = |\{(I,J)\, s.t.\ I \cup J = C\,,\ c|_I \sim a,\ c|_J \sim b\}|\ . \tag{7}$$

And a *quasi-shuffle* as a pair that contributes to the coefficient above. This is invariant under the equivalence classes of $\sim$, as it can be show in a similar way to Proposition 2.1.

In the following proposition we observe that $\mathcal{A}(h)$ is a subalgebra of $\mathcal{F}(\mathcal{G}(h), \mathbb{Q})$ with the pointwise multiplication structure.

Theorem 2.3. Let h be a presheaf. Then the pattern functions satisfy the following identity:

$$\mathbf{p}_a\, \mathbf{p}_b = \sum_{c \in \mathcal{G}(h)} \binom{c}{a,b} \mathbf{p}_c\ . \tag{8}$$

In particular, the pattern functions of h span a subalgebra of the function algebra $\mathcal{F}(\mathcal{G}(h), \mathbb{Q})$, where the unit is $\sum_{c \in h[\emptyset]} \mathbf{p}_c$. We say that $\mathcal{A}(h)$ is the *pattern algebra* of h.

Proof. Fix $x \in h[I]$, and note that $\mathbf{p}_a(x)\, \mathbf{p}_b(x)$ counts the following

$$\begin{aligned}
\mathbf{p}_a(x)\, \mathbf{p}_b(x) &= |\{A \subseteq I \text{ s.t. } x|_A \sim a\} \times \{B \subseteq I \text{ s.t. } x|_B \sim b\}| \\
&= |\{(A,B) \text{ s.t. } A, B \subseteq I,\ x|_A \sim a,\ x|_B \sim b\}| \\
&= \sum_{C \subseteq I} |\{(A,B) \text{ s.t. } A \cup B = C,\ x|_A \sim a,\ x|_B \sim b, \}| \\
&= \sum_{C \subseteq I} \binom{x|_C}{a,b} = \sum_{c \in \mathcal{G}(h)} \binom{c}{a,b} \mathbf{p}_c(x)\,.
\end{aligned} \tag{9}$$

Hence, the space $\mathcal{A}(h)$ is closed for the product of functions. Further, it is easy to observe that $\sum_{a \in h[\emptyset]} \mathbf{p}_a$ is the constant function equal to one, so this is a unit and $\mathcal{A}(h)$ is an algebra, concluding the proof. □

Corollary 2.4 (Products and quasi-shuffles)**.** Let $c \in h[I]$ and $b_i \in h[J_i]$ for $i = 1, \ldots, k$, and let

$$\binom{a}{b_1, \ldots, b_k} := \left| \left\{ (J_1, \ldots, J_k) \text{ s.t. } \bigcup_{i=1}^{k} J_i = I,\ a|_{J_i} \sim b_i\ \forall i = 1, \ldots, k \right\} \right|\ .$$

Then we have that

$$\prod_{i=1}^{k} \mathbf{p}_{b_i} = \sum_c \binom{c}{b_1, \ldots, b_k} \mathbf{p}_c\ .$$

Definition 2.5 (Shuffles and quasi-shuffles)**.** If an object a is such that $\binom{a}{b_1,\ldots,b_k} > 0$ we say that a is a quasi-shuffle of $b_1, \ldots, b_k$. In addition, if $|a| = \sum_{i=1}^{k} |b_i|$, we say that a is a *shuffle* of $b_1, \ldots, b_k$.

2.2. **Coproducts on pattern algebras.** In this section we consider an associative presheaf $(h, *, 1)$. Concretely, our combinatorial presheaf h is endowed with an associative product $*$ and a unit $1 \in h[\emptyset]$.

Definition 2.6 (Product structure in $\mathcal{G}(h)$)**.** If $(h, *, 1)$ is an associative presheaf, then $\mathcal{G}(h)$ inherits an associative product. If a is an object, we denote its equivalence class under $\sim$ by $\bar{a}$ in this remark. The associative product in $\mathcal{G}(h)$ is defined as follows:

Let $a \in h[n_1], b \in h[n_2]$ and denote $[n_1+1, n_1+n_2] = \{n_1+1, \ldots, n_1+n_2\}$. Consider st the order preserving map $st : [n_1+1, n_1+n_2] \to [n_2]$, and let $b' = h[st](b)$. Then we define the product in $\mathcal{G}(h)$ as $\bar{a} \cdot \bar{b} := \overline{a * b'} \in h[n_1+n_2]_\sim$.

It is a direct computation to see that $\overline{a * b}$ does not depend on the representative chosen for $\overline{a}$ and $\overline{b}$. Thus we have a well defined operation in $\mathcal{G}(h)$.

Definition 2.7. Let $a \in \mathcal{G}(h)$. Then, define

$$\Delta\, \mathbf{p}_a := \sum_{\substack{b,c \in \mathcal{G}(h) \\ b \cdot c = a}} \mathbf{p}_b \otimes \mathbf{p}_c \ .$$

Note that the right hand side is a finite sum, because $b \cdot c = a$ implies that $|b|, |c| \leq |a|$, so this is well defined. Define further the map $\epsilon : \mathcal{A}(h) \to \mathbb{Q}$ that sends $\mathbf{p}_a$ to $\mathbb{1}[a = 1]$, where 1 stands for the unit in the associative presheaf h.

We recall and prove Theorem 1.5 here.

Theorem 2.8 (The pattern Hopf algebra)**.** Let $(h, *, 1)$ be an associative presheaf. Then, the maps Δ and ϵ give $\mathcal{A}(h)$ a structure of a coalgreba. Furthermore, together with pointwise multiplication of functions, this defines a bialgebra structure in $\mathcal{A}(h)$.

Further, $\mathcal{A}(h)$ is a Hopf algebra whenever h is a connected presheaf.

Remark that we refer to the pattern Hopf algebra of $(h, *, 1)$ simply by $\mathcal{A}(h)$ instead of $\mathcal{A}(h, *, 1)$, for simplicity of notation, whenever emphasis on the role of $*$ is not needed.

Proof of Theorem 1.5. First, we note that Δ is trivially coassociative, from the associativity axioms of $*$ described in Observation 1.17. That ϵ is a counit follows from the unit axioms on $(h, *, 1)$.

We first claim that, for $a, x, y \in \mathcal{G}(h)$,

$$\Delta\, \mathbf{p}_a(x, y) = \mathbf{p}_a(x \cdot y)\,, \tag{10}$$

using the natural inclusion $\mathcal{F}(\mathcal{G}(h), \mathbb{Q})^{\otimes 2} \subseteq \mathcal{F}(\mathcal{G}(h)^2, \mathbb{Q})$.

Indeed, take representatives $x \in h[n_1]$ and $y \in h[n_2]$ with no loss of generality, and write $B = [n_1]$, $C = \{n_1+1, \ldots, n_1+n_2\}$. Let st be the order preserving map between

C and $[n_2]$. Then

$$\begin{aligned}
\mathbf{p}_a(x*y) &= |\{J \subseteq B \sqcup C \text{ s.t. } (x*y)|_J \sim a\}| \\
&= |\{J \subseteq B \sqcup C \text{ s.t. } x|_{J\cap B} * y|_{J\cap C} \sim a\}| \\
&= \sum_{\substack{b,c\in\mathcal{G}(h)\\ a\sim b\cdot c}} |\{J \subseteq B \sqcup C \text{ s.t. } x|_{J\cap B} \sim b,\ y|_{J\cap C} \sim c\}| \\
&= \sum_{\substack{b,c\in\mathcal{G}(h)\\ a\sim b\cdot c}} |\{J \subseteq B \text{ s.t. } x|_J \sim b\} \times \{J \subseteq C \text{ s.t. } y|_J \sim c\}| \\
&= \sum_{\substack{b,c\in\mathcal{G}(h)\\ a\sim b\cdot c}} \mathbf{p}_b(x)\,\mathbf{p}_c(y) = \Delta\,\mathbf{p}_a(x,y)\,.
\end{aligned} \tag{11}$$

Since both functions take the same values on $\mathcal{G}(h)^2$, we conclude that (10) holds.

The following are the bialgebra axioms that we wish to establish:

$$\begin{aligned}
\Delta(\mathbf{p}_a\,\mathbf{p}_b) &= \Delta(\mathbf{p}_a)\Delta(\mathbf{p}_b)\,, \\
\Delta\left(\sum_{a\in h[\emptyset]} \mathbf{p}_a\right) &= \left(\sum_{a\in h[\emptyset]} \mathbf{p}_a\right) \otimes \left(\sum_{a\in h[\emptyset]} \mathbf{p}_a\right)\,, \\
\epsilon(\mathbf{p}_a\,\mathbf{p}_b) &= \epsilon(\mathbf{p}_a)\epsilon(\mathbf{p}_b)\,, \\
\epsilon\left(\sum_{a\in h[\emptyset]} \mathbf{p}_a\right) &= 1\,.
\end{aligned}$$

The last three equations are direct computations. For the first equation, we use (10) as follows: take $a, b, x, y \in \mathcal{G}(h)$, then

$$\Delta(\mathbf{p}_a\,\mathbf{p}_b)(x,y) = (\mathbf{p}_a\,\mathbf{p}_b)(x\cdot y) = \mathbf{p}_a(x\cdot y)\,\mathbf{p}_b(x\cdot y) = (\Delta\,\mathbf{p}_a\,\Delta\,\mathbf{p}_b)(x,y)\,.$$

This concludes the first part of the proof.

Now suppose that h is connected, so that the zero degree component $\mathcal{A}(h)_0$ is one dimensional. From [Tak71], because $\mathcal{A}(h)$ is commutative, it suffices to establish that the group-like elements of $\mathcal{A}(h)$ are invertible. Now it is a direct observation that any group-like element is in $\mathcal{A}(h)_0$, which is a one dimensional algebra, so all non-zero elements are invertible. This concludes that $\mathcal{A}(h)$ is a Hopf algebra. □

2.3. **The space of primitive element.** In this section we give a description of the primitive elements of any pattern Hopf algebra. Denote by $P_H := \{a \in H |\, \Delta a = a\otimes 1 + 1\otimes a\}$ the space of *primitive elements*, where 1 stands for the unit of the Hopf algebra.

Definition 2.9 (Irreducible objects)**.** Let $(h, *, 1)$ be a connected associative presheaf. An object $t \in h[I]$ with $t \neq 1$ is called *irreducible* if any two objects $a \in h[A], b \in h[B]$ such that $a * b = t$ and $A \sqcup B = I$ have either $a = 1$ or $b = 1$.

The notion of irreducibility lifts to $\mathcal{G}(h)$. That is, a coinvariant $t \in \mathcal{G}(h)$ with $t \neq 1$ is said to be *irreducible* if any $a, b \in \mathcal{G}(h)$ such that $a \cdot b = t$ have either $a = 1$ or $b = 1$.

We have that $a \in h[I]$ is irreducible if and only if the corresponding equivalence class $\bar{a} \in \mathcal{G}(h)$ is irreducible. The family of irreducible equivalent classes in $\mathcal{G}(h)$ is denoted by $\mathcal{I}(h) \subseteq \mathcal{G}(h)$.

Proposition 2.10 (Primitive space of pattern Hopf algebras)**.** Let $(h, *, 1_h)$ be a connected associative presheaf, and $\mathcal{I}(h)$ the set of irreducible elements in $\mathcal{G}(h)$.

Then

$$P_{\mathcal{A}(h)} = \operatorname{span}\{\mathbf{p}_a \,|a \in \mathcal{I}(h)\}\,.$$

Proof. If f is irreducible, it is straightforward to observe that $\mathbf{p}_f$ is primitive. On the other hand, let $a = \sum_{f\in\mathcal{G}(h)} c_f\, \mathbf{p}_f$ be a generic primitive element from $\mathcal{A}(h)$. Then, the equation $\Delta a = a \otimes \mathbf{p}_{1_h} + \mathbf{p}_{1_h} \otimes a$ becomes

$$\sum_{g_1,g_2\in\mathcal{G}(h)} c_{g_1\cdot g_2}\, \mathbf{p}_{g_1} \otimes \mathbf{p}_{g_2} = \sum_f c_f(\mathbf{p}_f \otimes \mathbf{p}_{1_h} + \mathbf{p}_{1_h} \otimes \mathbf{p}_f)\,.$$

From Remark 2.2, $\{\mathbf{p}_{g_1} \otimes \mathbf{p}_{g_2}\}_{g_1,g_2\in\mathcal{G}(h)}$ is a basis of $\mathcal{A}(h)^{\otimes 2}$, so we have that for any $g_1 \neq 1_h, g_2 \neq 1_h$,

$$c_{g_1\cdot g_2} = 0\,.$$

Thus we conclude that a is a linear combination of the set $\left\{\mathbf{p}_f \,|f \in \mathcal{I}(h)\right\}$, as desired. □

2.4. **The pattern algebra functor.** In this section, we see that the mapping $\mathcal{A}$ is in fact functorial, bringing the parallel between the pattern Hopf algebras and the Fock functors even closer.

Theorem 2.11 (Pattern algebra maps)**.** If $f : h \Rightarrow j$ is a morphism of combinatorial presheaves, then the following formula

$$\mathcal{A}[f](\mathbf{p}_a) := \sum_{f(b)=a} \mathbf{p}_b \in \mathcal{A}(h)\,, \tag{12}$$

defines an algebra map $\mathcal{A}[f] : \mathcal{A}(j) \to \mathcal{A}(h)$.

Further, if f is a morphism of associative presheaves, then $\mathcal{A}[f]$ is a bialgebra morphism. Consequently, if h, j are connected, this is a Hopf algebra morphism.

Remark that the sum in (12) is finite, so this functor is well defined.

Proof. The map $\mathcal{A}[f]$ is linear and sends the unit $\sum_{a\in j[\emptyset]} \mathbf{p}_a$ of $\mathcal{A}(j)$ to

$$\sum_{a\in j[\emptyset]} \sum_{\substack{b\in h[\emptyset]\\ f(b)=a}} \mathbf{p}_b = \sum_{b\in j[\emptyset]} \mathbf{p}_b\,.$$

Hence, to establish that $\mathcal{A}[f]$ is an algebra morphism, it suffices to show that it preserves the product on the basis, *i.e.* that $\mathcal{A}[f](\mathbf{p}_a\, \mathbf{p}_b) = \mathcal{A}[f](\mathbf{p}_a)\mathcal{A}[f](\mathbf{p}_b)$.

It is easy to see that this holds if we have that

$$\binom{f(c')}{a, b} = \sum_{\substack{a',b'\in\mathcal{G}(h)\\ f(a')=a, f(b')=b}} \binom{c'}{a', b'}\,,$$

for any $a, b \in \mathcal{G}(j), c' \in \mathcal{G}(h)$.

Indeed, if $a \in j[A], b \in j[B]$ and $c' \in h[C]$, then for any set I, by naturality of f, we have $f(c')|_I = f(c'|_I)$. Then

$$\begin{aligned}\binom{f(c')}{a,b} &= \left|\left\{(I,J) \text{ s.t. } f(c')|_I \sim a,\ f(c')|_J \sim b,\ I \cup J = C\right\}\right| \\ &= \sum_{\substack{a',b' \in \mathcal{G}(h) \\ f(a')=a,\, f(b')=b}} \left|\left\{(I,J) \text{ s.t. } c'|_I \sim a',\ c'|_J \sim b',\ I \cup J = C\right\}\right| \\ &= \sum_{\substack{a',b' \in \mathcal{G}(h) \\ f(a')=a,\, f(b')=b}} \binom{c'}{a',b'}.\end{aligned}$$

Now suppose further that f is an associative presheaf morphism between $(h, *_h, 1_h)$ and $(j, *_j, 1_j)$. That $\mathcal{A}[f]$ preserves counit follows from $f(1_h) = 1_j$. That $\mathcal{A}[f]$ preserves Δ follows because both $\mathcal{A}[f]^{\otimes 2}(\Delta\, \mathbf{p}_a)$ and $\Delta(\mathcal{A}[f]\, \mathbf{p}_a)$ equal

$$\sum_{\substack{b',c' \in \mathcal{G}(h) \\ f(b' \cdot_h c')=a}} \mathbf{p}_{b'} \otimes \mathbf{p}_{c'}\,,$$

since $f(b' *_h c') = f(b') *_j f(c')$, concluding the proof. □

Definition 2.12 (The pattern algebra functor)**.** Because of Theorem 2.11, we can define the *pattern algebra* contravariant functor $\mathcal{A} : \mathtt{CPSh} \to \mathtt{Alg}_{\mathbb{Q}}$.

This functor when restricted to the subcategory of associative presheaves is also a functor to bialgebras as $\mathcal{A} : Mon(\mathtt{CPSh}) \to \mathtt{BiAlg}_{\mathbb{Q}}$.

Example 2.13 (Graph patterns in permutations)**.** Consider again the associative presheaves $(\mathtt{Gr}, \oplus, \emptyset)$ on graphs and $(\mathtt{Per}, \oplus, \emptyset)$ on permutations. The inversion graph of a permutation is a presheaf morphism $\mathtt{Inv} : \mathtt{Per} \Rightarrow \mathtt{Gr}$ defined as follows: given a permutation $\pi = (\leq_P, \leq_V)$ on the set I, we take the graph $\mathtt{Inv}(\pi)$ with vertex set I and an edge between $i, j \in I$, $i \neq j$ if

$$i \leq_P j \Leftrightarrow j \leq_V i\,.$$

It is a direct observation that this map is indeed an associative presheaf morphism. As a consequence, we have a Hopf algebra morphism $\mathcal{A}(\mathtt{Inv}) : \mathcal{A}(\mathtt{Gr}) \to \mathcal{A}(\mathtt{Per})$.

2.5. **Mangum relations and representability.** The goal of the following two sections is to show that arithmetic properties of the permutation pattern algebra that are established in [Var14] are actually general properties of pattern algebras.

Proposition 2.14 (Magnus inversions)**.** Let h be a combinatorial presheaf, and consider the maps $M, N : \operatorname{span} \mathcal{G}(h) \to \operatorname{span} \mathcal{G}(h)$ given in the basis elements $a \in \mathcal{G}(h)$ by

$$M : a \mapsto \sum_{b \in \mathcal{G}(h)} \mathbf{p}_b(a) b\,,$$

$$N : a \mapsto \sum_{b \in \mathcal{G}(h)} (-1)^{|a|+|b|}\, \mathbf{p}_b(a) b\,,$$

and extended linearly to $\operatorname{span}\mathcal{G}(h)$. Then the maps M, N are inverses of each other.

This result was already known in the context of words in [Hof00] and [AM10] and in the context of permutations in [Var14].

Proof. We start by proving a relation on pattern functions. Let $a \in h[I], c \in \mathcal{G}(h)$. Then we claim that

$$\sum_{b\in\mathcal{G}(h)} (-1)^{|b|}\,\mathbf{p}_b(a)\,\mathbf{p}_c(b) = (-1)^{|a|}\mathbb{1}[a \in c]\,.$$

Indeed, for each pair of sets (J, B) such that $J \subseteq B \subseteq I$ and $a|_J \in c$ it corresponds an object $b = a|_B$, and the patterns B of b in a, and J of c in b. Observe that this is a bijective correspondence, so we have

$$\sum_{b\in\mathcal{G}(h)} (-1)^{|b|}\,\mathbf{p}_b(a)\,\mathbf{p}_c(b) = \sum_{\substack{J\subseteq I\\ a|_J\in c}} \sum_{J\subseteq B\subseteq I} (-1)^{|B|} = (-1)^{|I|}\mathbb{1}[a|_I \in c] = (-1)^{|a|}\mathbb{1}[a \in c]\,.$$

It follows that

$$\begin{aligned} N(M(a)) =& N\left(\sum_{b\in\mathcal{G}(h)} \mathbf{p}_b(a)b\right) = \sum_{b\in\mathcal{G}(h)} \mathbf{p}_b(a) \sum_{c\in\mathcal{G}(h)} (-1)^{|b|+|c|}\,\mathbf{p}_c(b) \\ =& \sum_{c\in\mathcal{G}(h)} c(-1)^{|c|} \sum_{b\in\mathcal{G}(h)} (-1)^{|b|}\,\mathbf{p}_b(a)\,\mathbf{p}_c(b) \\ =& \sum_{c\in\mathcal{G}(h)} c(-1)^{|c|}(-1)^{|a|}\mathbb{1}[a = c] = a\,, \end{aligned}$$

and that

$$\begin{aligned} M(N(a)) =& M\left(\sum_{b\in\mathcal{G}(h)} (-1)^{|a|+|b|}\,\mathbf{p}_b(a)b\right) = \sum_{b\in\mathcal{G}(h)} \mathbf{p}_b(a) \sum_{c\in\mathcal{G}(h)} (-1)^{|a|+|b|}\,\mathbf{p}_c(b) \\ =& \sum_{c\in\mathcal{G}(h)} c(-1)^{|a|} \sum_{b\in\mathcal{G}(h)} (-1)^{|b|}\,\mathbf{p}_b(a)\,\mathbf{p}_c(b) \\ =& \sum_{c\in\mathcal{G}(h)} c(-1)^{|a|}(-1)^{|a|}\mathbb{1}[a = c] = a\,, \end{aligned}$$

as desired. □

2.6. **The Sweedler dual and pattern algebras.** Let A be an algebra over a field $\mathbb{K}$. Then the Sweedler dual A° is defined in [Swe69] as

$$A^\circ := \{g \in A^* |\, \ker g \text{ contains a cofinite ideal}\}\,,$$

where a *cofinite ideal* J of A is an ideal such that A/J is a finite dimensional vector space over $\mathbb{K}$. There, it is established that A° is a coalgebra, where the coproduct map is the transpose of the product map.

Consider the following right action of A on A^*: for $f \in A^*$ and $a, b \in A$,

$$(f \cdot b)(a) := f(ab)\,. \tag{13}$$

A description of A° is given in [Swe69, Proposition 6.0.3], as all *representable* elements $f \in A^*$, that is all f such that the vector space $\{f \cdot b\}_{b\in A}$ is finite dimensional.

Let $(h, *, 1)$ now be an associative presheaf, and consider the algebra generated by $\cdot$ in $\mathcal{G}(h)$:

$$\mathrm{Alg}(h) \coloneqq \mathrm{span}(\mathcal{G}(h), \cdot)\,.$$

Then $\mathcal{A}(h) \subseteq \mathrm{Alg}(h)^* = \mathcal{F}(\mathcal{G}(h), \mathbb{Q})$. In fact, the following proposition guarantees that we have $\mathcal{A}(h) \subseteq \mathrm{Alg}(h)^\circ$. Remark that the coproduct in $\mathcal{A}(h)$ is precisely the transpose of the multiplication in $\mathrm{span}(\mathcal{G}(h), \cdot)$, therefore $\mathcal{A}(h) \subseteq \mathrm{Alg}(h)^\circ$ is an inclusion of coalgebras.

Proposition 2.15 (Pattern algebra and the sweedler dual)**.** Let h be an associative presheaf. Then, its pattern algebra satisfies $\mathcal{A}(h) \subseteq \mathrm{Alg}(h)^\circ$.

Proof. We claim that each pattern function is representable, which concludes the proof according to [Swe69, Proposition 6.0.3]. In fact, for $a, b, c \in \mathcal{G}(h)$:

$$\begin{aligned}\mathbf{p}_a \cdot b\,(c) = \mathbf{p}_a(b \cdot c) &= \Delta\,\mathbf{p}_a(b, c)\\ &= \sum_{\substack{a_1,a_2\in\mathcal{G}(h)\\ a=a_1\cdot a_2}} \mathbf{p}_{a_1}(b)\,\mathbf{p}_{a_2}(c)\end{aligned}$$

That is, $\mathbf{p}_a \cdot b = \sum_{a=a_1\cdot a_2} \mathbf{p}_{a_1}(b)\,\mathbf{p}_{a_2}$. It follows that

$$\mathrm{span}\{\mathbf{p}_a \cdot b\}_{b\in\mathcal{G}(h)} \subseteq \mathrm{span}\{\mathbf{p}_{a_2}\}_{\substack{a_2\in\mathcal{G}(h)\\ |a_2|\leq|a|}}\,,$$

which is finite dimensional. □

3. Freeness of commutative presheaves

We start this section with a discussion on factorization theorems on combinatorial presheaves. We will observe that the factorizations of objects in connected associative presheaves into irreducibles is unique up to some possible commutativity. this is a general fact on associative presheaves, and is a central point in establishing freeness of any pattern Hopf algebra so far in the literature.

We also dedicate some attention to commutative presheaves. An almost immediate consequence of the general fact discussed above is that the pattern algebra of a commutative presheaf is free.

We also explore specific combinatorial presheaves that can be endowed with a commutative structure. The main examples are graphs, already studied in [Whi32], marked graphs (see Section 3.3), set partitions (see Section 3.4), simplicial complexes and posets.

3.1. Relations in general connected associative presheaves.

Consider a connected associative presheaf $(h, *, 1)$. In this section we will not assume that h is commutative. Recall that for objects $a \in h[I], b \in h[J]$, if $A \subseteq I \sqcup J$ then $(a*b)|_A = a|_{A\cap I} * a|_{A\cap J}$, as described in Observation 1.17. We recall as well that an object $t \in h[I]$ is called *irreducible* if $t \neq 1$ and if $t = a * b$ only has trivial solutions. We define as well an irreducible coinvariant in $\mathcal{G}(h)$.

Definition 3.1 (Set composition and set partition)**.** Let I be a finite set. A *set composition* of I is a list $(B_1, \dots B_k)$, that can also be written as $B_1|\dots|B_k$, of pairwise disjoint nonempty subsets of I, such that $I = \bigcup_i B_i$. We denote by Π_I the family of set compositions of I. A *set partition* of I is a family $\boldsymbol{\pi} = \{I_1, \dots, I_k\}$ of pairwise disjoint nonempty sets, such that $\bigcup_i I_i = I$. We write Σ_I for the family of set partitions of I.

If $\vec{\boldsymbol{\pi}} \in \Pi_I$ is a set composition, we can define its underlying set partition of I by disregarding the order of the list. We denote it by $\boldsymbol{\lambda}(\vec{\boldsymbol{\pi}})$.

Definition 3.2 (Factorization of objects and coinvariants)**.** Consider an associative presheaf h that is connected, and an object $o \in h[I]$. A factorization of o is a word $(x_1, \dots, x_k)$ of objects such that $x_1 * \cdots * x_k = o$.

A factorization $(x_1, \dots, x_k)$ of o is said to be *into irreducibles* when each x_i is an irreducible object for $i = 1, \dots, l$.

A factorization of a coinvariant $a \in \mathcal{G}(h)$ is a decomposition of the form $a = s_1 \cdots s_k$. This factorization is said to be into *irreducibles* if each s_i is *irreducible*.

It is clear to see that an object is irreducible if and only if its coinvariant is irreducible.

To a factorization $(x_1, \dots, x_k)$ of $o \in h[I]$, it corresponds a set composition $\vec{\boldsymbol{\pi}} = (I_1, \dots, I_k) \models I$, where $x_i \in h[I_i]$. This is indeed a set composition of I by definition of $*$. This correspondence is injective, that is to any two distinct factorizations of o it corresponds distinct underlying set compositions. Indeed, assume otherwise, that to the factorizations $(x_1, \dots, x_k)$ and $(y_1, \dots, y_k)$ it corresponds the same set composition $(I_1, \dots, I_k)$, Then, for any $i = 1, \dots, k$ we have

$$\begin{aligned} o|_{A_i} &= x_1|_{A_1 \cap A_i} * \dots * x_l|_{A_l \cap A_i} \\ &= x_1|_{\emptyset} * \dots * x_i|_{A_i} * \dots * x_l|_{\emptyset} \\ &= 1 * \dots * x_i * \dots * 1 = x_i \,, \end{aligned}$$

and similarly we have that $o|_{A_i} = y_i$, so that $x_i = y_i$.

Conversely, not all set compositions yield a factorization, and the irreducible elements are precisely the ones where only the trivial set composition with one block yields a factorization.

Theorem 3.3. Let h be an associative presheaf and $o \in h[I]$ an object. If $\vec{\boldsymbol{\pi}}_1, \vec{\boldsymbol{\pi}}_2$ are factorizations into irreducibles of o, then their underlying set partitions $\boldsymbol{\lambda}(\vec{\boldsymbol{\pi}}_1)$ and $\boldsymbol{\lambda}(\vec{\boldsymbol{\pi}}_2)$ are the same.

In particular, the number of irreducible factors $j(o)$ and the multiset of irreducible factors $\mathfrak{fac}(o)$ of an object are well defined and do not depend on the factorization into irreducibles at hand.

Proof. Suppose that o has two distinct factorizations

$$o = l_1 * \dots * l_k = r_1 * \dots * r_s \,, \tag{14}$$

where $\vec{\boldsymbol{\pi}}_1 = A_1|\dots|A_k$ and $\vec{\boldsymbol{\pi}}_2 = B_1|\dots|B_s$ are set compositions of X such that $l_i = o|_{A_i}$ and $r_i = o|_{B_i}$. Note that

$$l_j = o|_{A_j} = r_1|_{B_1 \cap A_j} * \dots * r_s|_{B_s \cap A_j} \,.$$

Because l_j is irreducible, for each j there is exactly one i such that $B_i \cap A_j \neq \emptyset$, so $A_j \subseteq B_i$, and $\boldsymbol{\tau}(\vec{\boldsymbol{\pi}}_1)$ is coarser than $\boldsymbol{\tau}(\vec{\boldsymbol{\pi}}_2)$. By a symmetrical argument, we obtain that $\boldsymbol{\tau}(\vec{\boldsymbol{\pi}}_1)$ is finer than $\boldsymbol{\tau}(\vec{\boldsymbol{\pi}}_2)$, so we conclude that $\boldsymbol{\tau}(\vec{\boldsymbol{\pi}}_1) = \boldsymbol{\tau}(\vec{\boldsymbol{\pi}}_2)$. It follows that the number of factors and the multiset of factors are well defined. □

This implies the following for factorizations on $\mathcal{G}(h)$:

Corollary 3.4. Consider an associative presheaf $(h, *, 1)$, together with the usual product in $\mathcal{G}(h)$, denoted by $\cdot$. Consider also $a \in \mathcal{G}(h)$. If $a = x_1 \cdots x_l = s_1 \cdots s_s$ are two factorizations into irreducibles, then the multisets $\{x_1, \ldots, x_l\}$ and $\{s_1, \ldots, s_t\}$ coincide.

In particular, the number of irreducible factors $j(a)$ and the multiset of irreducible factors $\mathfrak{fac}(a)$ of a coinvariant are well defined and do not depend on the factorization into irreducibles at hand.

The algebraic structure of $\mathrm{span}(\mathcal{G}(h), \cdot)$ is determined by the set $\mathcal{I}(h)$ together with the relations between irreducible elements. Corollary 3.4 imposes restrictions on the possible relations. This is different from the case with groups: these can be described by generators and relations, but these relations have more flexibility.

Given an alphabet Ω, denote the set of words on Ω by $\mathcal{W}(\Omega)$.

Problem 3.5 (Factorization theorems in associative presheaves)**.** Given an associative presheaf $(h, *, 1)$, describe $\mathcal{E}(h)$, the collection of fibers of the map

$$\Pi : \mathcal{W}(\mathcal{I}(h)) \to \mathcal{G}(h)\,,$$

defined by taking the product of the letters of a word in $\mathcal{G}(h)$.

The fibers of this map, that is the sets of words $\Pi^{-1}(a)$, correspond to the different factorizations of an object $a \in \mathcal{G}(h)$. In general, according to Corollary 3.4, a fiber consists of a set of words of irreducible elements that result from one another by permuting its letters. Whenever h is a commutative presheaf, Corollary 3.7 tells us that the fibers are as big as possible, restricted to Corollary 3.4. This means that, in this case, each fiber is a set of words resulting from a permutation of a word in $\mathcal{W}(\mathcal{I}(h))$.

Take the example of the combinatorial presheaf `Per`, where no non-trivial rearrangement of the irreducible factors of a factorization of an object yields a distinct factorization of the same object. In this example, the fibers $\Pi^{-1}(a)$ are singletons. In the case of marked permutations, in Theorem 4.8 we show that only transpositions of specific irreducible marked permutations remain factorizations of the same coinvariant a.

Remark 3.6. It can be seen that the freeness proof in Theorem 3.8 depends solely on the corresponding unique factorization theorem, that is on $\mathcal{E}(h)$. That is also the case on the proof given in [Var14] for the presheaf on permutations, and the proof below for marked permutations.

This motivates Conjecture 1.14, as it seems that the freeness of the pattern Hopf algebra only depends on the description of the fibers $\mathcal{E}(h)$. In this way, for instance, we can immediately see that the pattern Hopf algebra $\mathcal{A}(\mathtt{SComp})$, defined below, is free. This follows because it has a unique factorization theorem of the type of the one in the associative presheaf on permutations.

3.2. **Proof of freeness on commutative presheaves.** Recall that a commutative presheaf is an associative presheaf $(h, *, 1)$ such that $*_{A,B} = *_{B,A} \circ \text{ twist}_{A,B}$, see Definition 1.16. In the case of commutative presheaves, we have that any rearrangement of a factorization of an object o yields another factorization of o. For this reason, in the context of commutative presheaves, a set partition is also referred to as a factorization, and we have the following.

Corollary 3.7. Let $(h, *, 1)$ be a connected commutative presheaf, and $a \in \mathcal{G}(h)$ an object. Then, a has a unique factorization into irreducibles $l_1, \dots, l_{j(a)} \in \mathcal{I}(h)$ up to commutativity of factors. Equivalently, if $a \in h[X]$, there is a unique set partition that corresponds to a factorization of o into irreducibles.

Theorem 3.8 (Freeness of pattern algebras with commutative products)**.** Let $(h, *, 1)$ be a connected commutative presheaf. Consider $\mathcal{I}(h) \subseteq \mathcal{G}(h)$ the family of irreducible elements of h.

Then $\mathcal{A}(h)$ is free commutative, and $\{\mathbf{p}_\iota \mid \iota \in \mathcal{I}(h)\}$ is a set of free generators of $\mathcal{A}(h)$.

Proof. We will show that the family

$$\left\{ \prod_{\iota \in \mathcal{L}} \mathbf{p}_\iota \,\middle|\, \mathcal{L} \text{ multiset of elements in } \mathcal{I}(h) \right\}, \tag{15}$$

is a basis for $\mathcal{A}(h)$. The proof follows the strategy described in Section 1.4, by building an order $\leq$ in $\mathcal{G}(h)$ that is motivated in the unique factorization theorem in Theorem 3.3.

Define the following partial strict order $<_p$ in $\mathcal{G}(h)$: we say that $\alpha <_p \beta$ if:

- $|\alpha| < |\beta|$, or;
- $|\alpha| = |\beta|$ and $j(\alpha) < j(\beta)$.

In this way, $\leq_p$ is the order that we use to establish freeness.

Consider $\alpha \in \mathcal{G}(h)$ with $\alpha = \iota_1 \cdots \iota_{j(\alpha)}$ its unique factorization into irreducibles. Then we claim

$$\prod_{i=1}^{k} \mathbf{p}_{\iota_i} = \sum_{\beta \in \mathcal{G}(h)} \binom{\beta}{\iota_1, \dots, \iota_{j(\alpha)}} \mathbf{p}_\beta = \sum_{\beta \leq_p \alpha} c_\beta \, \mathbf{p}_\beta \,, \tag{16}$$

where $c_a \geq 1$. This concludes that (15) is a basis of $\mathcal{A}(h)$, and gives us the result.

Let us prove (16). Pick representatives l_i for ι_i such that $l_i \in h[A_i]$ and let also $a = l_1 * \cdots * l_{j(\alpha)}$ be an object. Observe that the coinvariant of a is α.

First, observe that a is a quasi-shuffle of $l_1, \dots, l_{j(\alpha)}$ by considering the patterns $A_1, \dots, A_{j(\alpha)}$ (see Remark 1.18). Thus we have $c_\alpha \geq 1$.

To show that any term β in (16) with $c_\beta \neq 0$ has $\beta \leq_p \alpha$, consider the maximal $\beta \in \mathcal{G}(h)$ that is a quasi-shuffle of $l_1, \dots, l_{j(\alpha)}$, and let b be a representative in β, such that $b \in h[Y]$. By maximality, we have that $\beta \geq \alpha$. Our goal is to show that $b \sim a$. Let $b = s_1 * \cdots * s_{j(b)}$ be the unique factorization of b into irreducibles, corresponding to the set partition $\{C_1, \dots, C_{j(b)}\}$. Because b is a quasi-shuffle of $l_1, \dots, l_{j(\alpha)}$, we can consider $B_1, \dots, B_{j(\alpha)} \subseteq Y$ sets such that $\bigcup_i B_i = Y$ and

$$b|_{B_i} \sim l_i \quad \text{for } i = 1, \dots, j(\alpha)\,.$$

In particular, we have that

$$l_i \sim b|_{B_i} = s_1|_{C_1 \cap B_i} * \cdots * s_{j(b)}|_{C_{j(b)} \cap B_i} \,,$$

for $i = 1, \ldots, j(\alpha)$.

By indecomposability of l_i, we have that, for each i, there is exactly one j such that $C_j \cap B_i \neq \emptyset$. From $\sqcup_j C_j = \bigcup_i B_i = Y$, we get that each B_i is contained in some C_j. So we can define a map $f : [j(\alpha)] \to [j(b)]$ such that $B_i \subseteq C_{f(i)}$. Thus, we have that $C_i = \bigcup_{k \in f^{-1}(i)} B_k$.

First observe that $|a| = \sum_i |A_i| = \sum_i |l_i| = \sum_i |B_i|$ and $|b| = \sum_j |C_j|$ but $|b| = |Y| \leq \sum_j \sum_{i \in f^{-1}(j)} |B_i| = |a|$. However, from $b \geq_p a$ we have that $|b| \geq |a|$ and thus we have $|b| = |a|$ and that the family $\{B_1, \ldots, B_{j(\alpha)}\}$ is disjoint. Further, f is a surjection, so we immediately have that $j(\alpha) \geq j(b)$ and because $b \geq_p a$, we must have an equality. Thus, f is a bijective map and we conclude that $C_{f(i)} = B_i$ for each $i = 1, \ldots, j(\alpha) = j(b)$.

We then conclude that

$$s_i = b|_{B_i} = b|_{C_{f(i)}} \sim l_{f(i)} \,,$$

and so, by commutativity,

$$b = s_1 * \cdots * s_{j(b)} = l_1 * \cdots * l_{j(\alpha)} = a \,,$$

as desired. □

3.3. **Marked graphs.**

Definition 3.9 (Marked graphs and two products)**.** For a finite set I, a marked graph G^* on I is a graph on the vertex set $I \sqcup \{*\}$. This defines a combinatorial presheaf $\mathtt{MGr}$ via the usual notion of relabeling and induced subgraphs.

We can further endow the combinatorial presheaf $\mathtt{MGr}$ with two different associative presheaf structures.

First, the *joint union*, $\vee$, which is defined as follows: If $G_1^* \in \mathtt{MGr}[I], G_2^* \in \mathtt{MGr}[J]$ with $I \cap J = \emptyset$, then $G_1^* \vee_{I,J} G_2^*$ has no edges between I and J, and the marked vertices are merged.

The second product, the *inflation product* $\star$ is defined as follows: If $G_1^* \in \mathtt{MGr}[I], G_2^* \in \mathtt{MGr}[J]$ with $I \cap J = \emptyset$, then two vertices $i \in I, j \in J$ are connected in $G_1^* \star_{I,J} G_2^*$ if j and $*$ are connected in G_1^*. The unit of both products is the marked graph 1 with a unique vertex and no edges.

Remark 3.10. The graphs that are irreducible with respect to the $\vee$ product are the graphs G^* such that the graph resulting from removing the marked vertex and its incident edges is a connected graph. In this case, we say that G^* is $\vee$-connected.

In Fig. 1 we have an example of a $\vee$ -connected marked graph and a $\vee$-disconnected marked graph.

Observe that $G_1^* \vee_{I,J} G_2^* = G_2^* \vee_{J,I} G_1^*$, so $\vee$ is a commutative operation, whereas $\star$ is not. It follows from Theorem 3.8 that $\mathcal{A}(\mathtt{MGr})$ is a free algebra. This is something that cannot be done directly with the inflation product of marked graphs. Indeed, a unique factorization theorem on this associative presheaf has not yet been found, and only small irreducible marked graphs can be constructed (see Fig. 2). It is worthwhile

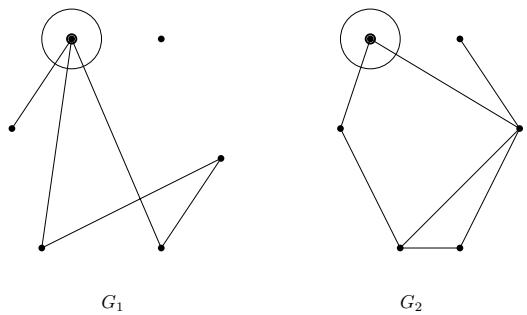

FIGURE 1. A marked graph G_1 with three $\vee$-connected components and a $\vee$-connected graph G_2.

to observe that the fibers under the map Π described in Problem 3.5 are non-trivial, as an example of a non-trivial relation can be seen in Fig. 3

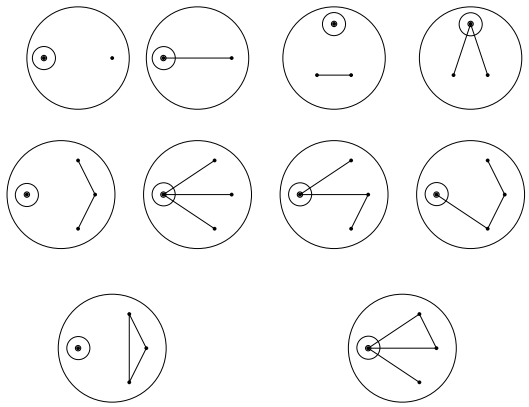

FIGURE 2. The $\star$-irreducible marked graphs of size up to three.

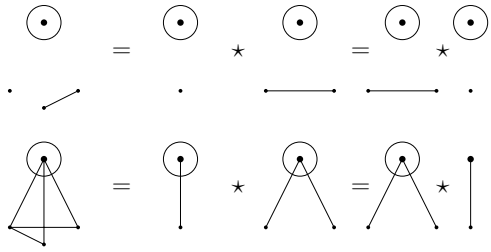

FIGURE 3. The first relations between irreducible marked graphs over $\star$.

3.4. **Set partitions.** We define here associative presheaves on set partitions `SPart` and set compositions `SComp`, and show that Sym, the Hopf algebra of symmetric functions, is the pattern Hopf algebra on set partitions. The functoriality of $\mathcal{A}$ also gives us a Hopf algebra morphism $Sym \to \mathcal{A}(\texttt{SComp})$.

Definition 3.11 (The presheaf on set partitions)**.** If $\boldsymbol{\pi}$ is a set partition of I and $J \subseteq I$, then $\boldsymbol{\pi}|_J = \{I_1 \cap J, \ldots, I_k \cap J\}$ is a set partition of J, after disregarding the empty sets. This defines a presheaf structure $\mathtt{SPart}$ with $\mathtt{SPart}[I] = \Sigma_I$, the family of set partitions of I.

We further endow $\mathtt{SPart}$ with an associative structure $\uplus$ as follows: if $\boldsymbol{\pi} = \{I_1, \ldots, I_q\}$, $\boldsymbol{\tau} = \{J_1, \ldots, J_p\}$ are set partitions of the disjoint sets I, J, respectively, let $\boldsymbol{\pi} \uplus \boldsymbol{\tau} = \{I_1, \ldots, I_q, J_1, \ldots, J_p\}$ be a set partition of $I \sqcup J$. It is straightforward to observe that $(\mathtt{SPart}, \uplus, \emptyset)$ is a commutative connected presheaf.

Note that by Theorem 3.8, the pattern Hopf algebra $\mathcal{A}(\mathtt{SPart})$ is free and the generators correspond to the irreducible elements of $\mathcal{G}(\mathtt{SPart})$. These correspond to the set partitions with only one block, and up to relabeling there is a unique such set partition of each size. We write $\mathcal{I}(\mathtt{SPart}) = \{\{[n]\} | n \geq 1\}$.

Proposition 3.12. Let $\zeta : \mathcal{A}(\mathtt{SPart}) \to Sym$ be the unique algebra morphism mapping $\zeta : \mathbf{p}_{\{[n]\}} \mapsto p_n$, where p_n is the power sum symmetric function. This defines a Hopf algebra isomorphism.

Proof. As ζ sends a free basis to a free basis, it is an isomorphism of algebras. Furthermore, we observe that both $\mathbf{p}_{\{[n]\}}$ and p_n are a primitive element in their respective Hopf algebras, so the described map is a bialgebra morphism. Because the antipode is unique, it must send the antipode of $\mathcal{A}(\mathtt{SPart})$ to the one of Sym. Thus, this is a Hopf algebra isomorphism. □

Definition 3.13 (The presheaf on set compositions)**.** Let I be a finite set, and recall the definition of a set composition in Definition 3.1. If $J \subseteq I$ and $\vec{\boldsymbol{\pi}} = (I_1, \ldots, I_k)$ is a set composition of I, then $\vec{\boldsymbol{\pi}}|_J = (I_1 \cap J, \ldots, I_k \cap J)$ is a set composition of J, after disregarding the empty sets. This defines a presheaf structure $\mathtt{SComp}$ with $\mathtt{SComp}[I] = \Pi_I$, the family of set compositions of I.

We further endow $\mathtt{SComp}$ with an associative structure $\uplus$ as follows: if $\vec{\boldsymbol{\pi}} = (I_1, \ldots, I_q)$, $\vec{\boldsymbol{\tau}} = (J_1, \ldots, J_p)$ are set partitions of the disjoint sets I, J, respectively, let $\vec{\boldsymbol{\pi}} \uplus \vec{\boldsymbol{\tau}} = (I_1, \ldots, I_q, J_1, \ldots, J_p)$ be a set composition of $I \sqcup J$.

It is straightforward to observe that $(\mathtt{SComp}, \uplus, \emptyset)$ is an associative connected presheaf. Further, we can also observe that the map $\boldsymbol{\lambda} : \mathtt{SComp} \Rightarrow \mathtt{SPart}$ is an associative presheaf morphism.

From the map $\boldsymbol{\lambda} : \mathtt{SComp} \Rightarrow \mathtt{SPart}$ we get a Hopf algebra morphism

$$\mathcal{A}(\boldsymbol{\lambda}) : Sym \to \mathcal{A}(\mathtt{SComp}) .$$

Observe that $\mathcal{A}(\mathtt{SComp})$ is a free algebra, because it has a unique factorization theorem of the same type of permutations under the $\oplus$ product, so according to Remark 3.6 the proof in [Var14] holds also in this associative presheaf. This is also the case for the well known Hopf algebra $QSym$, where it was established that it is free in [Haz01], and when we regard both Hopf algebras as filtered Hopf algebras, the enumeration of generators for each degree coincide with the number of Lyndon words of a given size.

Problem 3.14 (The algebraic structure of $\mathcal{A}(\mathtt{SComp})$)**.** Is the Hopf algebra $\mathcal{A}(\mathtt{SComp})$ isomorphic to $QSym$?

4. Hopf algebra structure on marked permutations

In this section we consider the algebra structure of $\mathcal{A}(\texttt{MPer})$, and show that this pattern algebra on marked permutations is freely generated. This will be done using a factorization theorem on marked permutations on the inflation product.

Our strategy is as follows: we describe a unique factorization of marked permutations with the inflation product, in Corollary 4.12. This unique factorization theorem describes all possible factorizations of a marked permutation into irreducibles.

We further consider the Lyndon words on the alphabet of irreducible marked permutations, as introduced in [CFL58]. This leads us to a notion of stable Lyndon marked permutations $\mathcal{L}_{SL}$, in Definition 4.19. Finally, we prove the following result, which is the main theorem of this section, as a corollary of Theorem 4.24:

Theorem 4.1. The algebra $\mathcal{A}(\texttt{MPerm})$ is freely generated by $\{\mathbf{p}_{\iota^*} \,|\iota^* \in \mathcal{L}_{SL}\}$.

In the end of this section we compute the dimension of the space of primitive elements of the pattern Hopf algebra on marked permutations. This, according to Proposition 2.10, can be done by enumerating the irreducible elements in $\mathcal{G}(\texttt{MPer})$.

This section is organized as follows: we start in Section 4.1 and in Section 4.2 by establishing a unique factorization theorem in marked permutations. In Section 4.3 we state and prove the main theorem of this section. The proofs of technical lemmas used in these sections are left to Sections 4.4 and 4.5. Finally, in Section 4.6, we enumerate the irreducible marked permutations.

4.1. Unique factorizations.

We work on the combinatorial presheaf of permutations $(\texttt{Per}, \oplus, \emptyset)$ and on the combinatorial presheaf of marked permutations $(\texttt{MPer}, \star, \bar{1})$, introduced in Definitions 1.20 and 1.22.

Definition 4.2 (Decomposability on the operations $\oplus$ and $\ominus$)**.** We say that a permutation is $\oplus$-indecomposable if it has no non-trivial decomposition of the form $\tau_1 \oplus \tau_2$, and $\oplus$-decomposable otherwise. We say that a marked permutation is $\oplus$-indecomposable if it has no decomposition of the form $\tau \oplus \pi^*$ or $\pi^* \oplus \tau$, where π^* is a marked permutation and τ is a non-trivial permutation, and $\oplus$-decomposable otherwise. Similar definitions hold for $\ominus$. A permutation (resp. a marked permutation) is said to be *indecomposable* if it is both $\oplus$ and $\ominus$-indecomposable, and is said to be *decomposable* otherwise.

We remark that a permutation is $\oplus$-indecomposable whenever it is irreducible on the associative presheaf $(\texttt{Per}, \oplus, \emptyset)$ according to Definition 2.9. For marked permutations, Definition 2.9 specializes to the definition of irreducible marked permutations as follows:

Definition 4.3 (Irreducible marked permutations)**.** A marked permutation π^* is called *irreducible* if any factorization $\pi^* = \tau_1^* \star \tau_2^*$ has either $\tau_1^* = \bar{1}$ or $\tau_2^* = \bar{1}$.

Example 4.4. Examples of irreducible marked permutations include $\bar{1}423, 23\bar{1}$ and $31\bar{4}2$, see Fig. 4. These marked permutations are respectively an $\oplus$-decomposable, $\ominus$-decomposable and an indecomposable irreducible marked permutation.

Remark 4.5 (decomposable irreducible marked permutations)**.** If π is an $\oplus$-indecomposable permutation, then $\bar{1} \oplus \pi$ and $\pi \oplus \bar{1}$ are irreducible marked permutations. Similarly, if

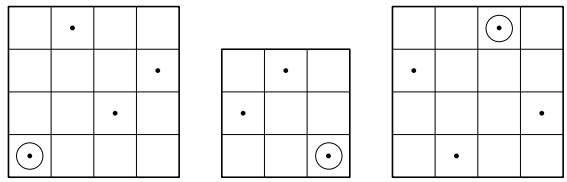

FIGURE 4. Irreducible marked permutations

τ is an $\ominus$-indecomposable permutation, then $\bar{1} \ominus \tau$ and $\tau \ominus \bar{1}$ are irreducible marked permutations.

These are precisely the *decomposable* (resp. $\oplus$-*decomposable*, $\ominus$-*decomposable*) *irreducible marked permutations*.

These decomposable irreducible marked permutations play an important role in the description of a free basis of $\mathcal{A}(\texttt{MPer})$, because they are the only ones that get in the way of a unique factorization theorem for the inflation product. In the following we carefully unravel all these issues.

Remark 4.6 ($\oplus$-relations and $\ominus$-relations)**.** Consider τ_1, τ_2 $\oplus$-indecomposable permutations. Then we have the following relations, called $\oplus$-*relations*

$$(\bar{1} \oplus \tau_1) \star (\tau_2 \oplus \bar{1}) = (\tau_2 \oplus \bar{1}) \star (\bar{1} \oplus \tau_1) = \tau_2 \oplus \bar{1} \oplus \tau_1 \,.$$

Consider now π_1, π_2 permutations that are $\ominus$-indecomposable. Then we have the following relation, called $\ominus$-*relations*

$$(\bar{1} \ominus \pi_1) \star (\pi_2 \ominus \bar{1}) = (\pi_2 \ominus \bar{1}) \star (\bar{1} \ominus \pi_1) = \pi_2 \ominus \bar{1} \ominus \pi_1 \,.$$

We wish to establish in Theorem 4.8 that these generate all the inflation relations between irreducible marked permutations.

We define the alphabet $\Omega := \{\text{irreducible marked permutations}\}$, and consider the set $\mathcal{W}(\Omega)$ of words on Ω. This set forms a monoid under the usual concatenation of words (we denote the concatenation of two words w_1, w_2 as $w_1 \cdot w_2$). When $w \in \mathcal{W}(\Omega)$, we write w^* for the consecutive inflation of its letters. So for instance $(\bar{1}2, \bar{2}1)^* = \bar{1}2 \star \bar{2}1 = \bar{2}13$. We use the convention that the inflation of the empty word on Ω is $\bar{1}$. This defines the *star map*, a morphism of monoids $\star : \mathcal{W}(\Omega) \to \mathcal{G}(\texttt{MPer})$.

For the sake of clarity, we avoid any ambiguity on the notation a^* by using Greek letters for marked permutations, lowercase Latin letters to represent words on any alphabet, and upper case Latin letters to represent sets with an added marked element (see Definition 1.22).

Definition 4.7 (Monoidal equivalence relation on $\mathcal{W}(\Omega)$)**.** We now define an equivalence relation on $\mathcal{W}(\Omega)$. For a word $w \in \mathcal{W}(\Omega)$, if $w = (\xi_1^*, \ldots, \xi_k^*)$ is such that $\xi_i^* \star \xi_{i+1}^* = \xi_{i+1}^* \star \xi_i^*$ is an $\oplus$-relation or an $\ominus$-relation, we say that $w \sim (\xi_1^*, \ldots \xi_{i-1}^*, \xi_{i+1}^*, \xi_i^*, \ldots, \xi_k^*)$.

We further take the transitive and reflexive closure to obtain an equivalence relation on $\mathcal{W}(\Omega)$.

We trivially have that if $w_1 \sim w_2$ and $z_1 \sim z_2$, then $w_1 \cdot z_1 \sim w_2 \cdot z_2$. This means that the quotient $\mathcal{W}(\Omega)/_\sim$ is a monoid. Further, because of Remark 4.6, the star map factors to a monoid morphism $\mathcal{W}(\Omega)/_\sim \to \mathcal{G}(\texttt{MPer})$.

Theorem 4.8. The star map $\star : w \mapsto w^*$ defines an isomorphism from $\mathcal{W}(\Omega)/_\sim$ to the monoid of marked permutations with the inflation product.

We postpone the proof of Theorem 4.8 to Section 4.4, and explore its consequences here.

Informally, this theorem states that any two factorizations of a marked permutation π^* into irreducible marked permutations are related by $\sim$. As a consequence, we recover the following corollary, which was already obtained in Corollary 3.4 in a more general context:

Corollary 4.9. Consider α^* a marked permutation, together with factorizations $\alpha^* = \xi_1^* \star \cdots \star \xi_k^* = \rho_1^* \star \cdots \star \rho_j^*$ into irreducible marked permutations. Then $k = j$ and $\{\xi_1^*, \ldots, \xi_k^*\} = \{\rho_1^*, \ldots, \rho_j^*\}$ as multisets.

We was done in Section 3, we define $j(\alpha^*)$ to be the number of irreducible factors in any factorization of α^* into irreducible factors, and define $\mathfrak{fac}(\alpha^*)$ as the multiset of irreducible factors of α^* in $\mathcal{G}(\texttt{MPer})$. These are well defined as a consequence of Corollary 4.9.

Definition 4.10 (Stability conditions)**.** A factorization of a marked permutation α^* into irreducible marked permutations,

$$\alpha^* = \xi_1^* \star \cdots \star \xi_j^* \,,$$

or the corresponding word $(\xi_1^*, \ldots, \xi_j^*)$ in $\mathcal{W}(\Omega)$, is said to be

- i-$\oplus$**-stable** if there is no π, τ $\oplus$-indecomposable permutations such that
$$\xi_i^* = \bar{1} \oplus \pi \text{ and } \xi_{i+1}^* = \tau \oplus \bar{1}\,;$$
- i-$\ominus$**-stable** if there is no π, τ $\ominus$-indecomposable permutations such that
$$\xi_i^* = \tau \ominus \bar{1} \text{ and } \xi_{i+1}^* = \bar{1} \ominus \pi\,.$$

Such factorization or word is said to be $\oplus$**-stable** (resp. $\ominus$**-stable**) if it is i-$\oplus$-stable (resp. i-$\ominus$-stable) for any $i = 1, \ldots, j-1$. Finally, such a factorization or word is said to be **stable** if it is both $\oplus$-stable and $\ominus$-stable.

Remark 4.11 (Stability reduction)**.** If we are given a factorization of $\alpha^* = \xi_1^* \star \cdots \star \xi_j^*$ that is not i-$\oplus$-stable or i-$\ominus$-stable for some $i = 1, \ldots, j-1$, we can perform an i-*reduction*, that maps $(\xi_1^*, \ldots, \xi_i^*, \xi_{i+1}^*, \cdots \star \xi_j^*)$ to $(\xi_1^*, \ldots, \xi_{i+1}^*, \xi_i^*, \cdots \star \xi_j^*)$.

It is immediate to see that this procedure of finding some i and applying an i-reduction always terminates. The final word is stable and is independent of the order in which we apply the reductions.

Because of the above, any equivalence class in $\mathcal{W}(\Omega)/_\sim$ admits a unique stable word, and a consequence of Theorem 4.8 is the following.

Corollary 4.12 (Unique stable factorization)**.** Let α^* be a marked permutation. Then, α^* has a unique stable factorization into irreducible marked permutations. We refer to it as *the stable factorization of* α^*.

4.2. **Lyndon factorization on marked permutation.** We introduce an order on permutations, two orders on marked permutations and an order in $\mathcal{W}(\Omega)$.

Definition 4.13 (Orders on marked permutations)**.** The *lexicographic order on permutations* is the lexicographic order when reading the one-line notation of permutations, and is written $\pi \leq_{per} \tau$.

Recall that, for a marked permutation $\alpha^* = (\leq_P, \leq_V)$ in I, we define its rank $\mathfrak{rk}(\alpha^*)$ as the rank of $*$ in $I \sqcup \{*\}$ with respect to the order $\leq_P$. We also write α for referring to the corresponding permutation in the set $I \sqcup \{*\}$. We define the *lexicographic order on marked permutations*, also denoted $\leq_{per}$, as follows: we say that $\pi^* \leq_{per} \tau^*$ if $\pi <_{per} \tau$ or if $\pi = \tau$ and $\mathfrak{rk}(\pi^*) \leq \mathfrak{rk}(\tau^*)$.

This in particular endows our alphabet Ω of irreducible marked permutations with an order. When we compare words on $\mathcal{W}(\Omega)$ we order them lexicographically according to $\leq_{per}$, and denote it simply as $\leq$.

We define the *factorization order* $\leq_{fac}$ on marked permutations as follows: Let $\pi^* = \xi_1^* \star \cdots \star \xi_k^*$ and $\tau^* = \tau_1^* \star \cdots \star \tau_j^*$ be the respective unique stable factorizations of π^* and τ^*. Then, we say that $\pi^* \leq_{fac} \tau^*$ if $(\xi_1^*, \ldots, \xi_k^*) \leq (\tau_1^*, \ldots, \tau_j^*)$ in $\mathcal{W}(\Omega)$.

Example 4.14. On permutations we have $12345 \leq_{per} 132 \leq_{per} 231 \leq_{per} 4123$. Observe that the empty permutation is the smallest permutation.

On marked permutations, we have $1\bar{3}2 \leq_{per} 13\bar{2} \leq_{per} \bar{2}31 \leq_{per} 412\bar{3}$. Observe that the trivial marked permutation $\bar{1}$ is the smallest marked permutation.

On words, we have the following examples: $(24\bar{1}3, 31\bar{4}2) \leq (31\bar{4}2, 24\bar{1}3)$, $(\bar{1}32, 21\bar{3}) \leq (\bar{1}432)$ and $(2\bar{4}13, \bar{1}423, 2\bar{4}13) \leq (2\bar{4}13, 2\bar{4}13, \bar{1}423)$.

On marked permutations, because $(24\bar{1}3, 31\bar{4}2)$ and $(31\bar{4}2, 24\bar{1}3)$ are stable, we have that $24\bar{1}3 \star 31\bar{4}2 \leq_{fac} 31\bar{4}2 \star 24\bar{1}3$. Notice however, that $(\bar{1}32, 21\bar{3})$ is not a stable word, as it does not satisfy the 1-$\oplus$-stability condition. Instead, we have $(\bar{1}32, 21\bar{3})^* = 21\bar{3} \star \bar{1}32 \geq_{fac} \bar{1}432$.

Sometimes the orders $\leq_{per}$ and $\leq_{fac}$ on marked permutations do not agree, as exemplified in Fig. 5.

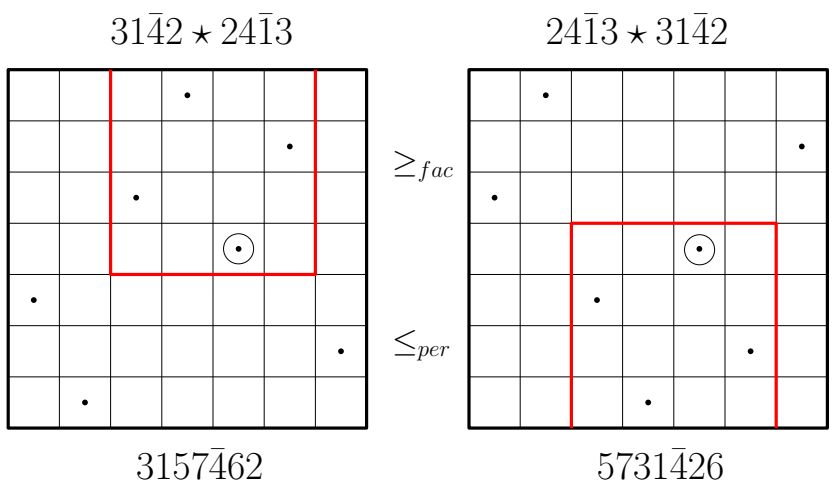

FIGURE 5. Two marked permutations and their order relations.

Remark 4.15. If $w_1 \geq w_2$ are stable words in $\mathcal{W}(\Omega)$, then these correspond precisely to the stable irreducible factorizations of w_1^*, w_2^*. Thus, $w_1^* \geq_{fac} w_2^*$.

Remark 4.16. If a word $w = (\rho_1^*, \dots, \rho_k^*)$ in $\mathcal{W}(\Omega)$ is not i-$\oplus$-stable or i-$\ominus$-stable, then $\rho_i^* <_{per} \rho_{i+1}^*$. Thus, from Theorem 4.8 and Remark 4.11, for a fixed marked permutation α^*, among all words w on irreducible marked permutations such that $w^* = \alpha^*$, the stable factorization is the biggest one in the lexicographical order.

We start the discussion on the topic of Lyndon words. This is useful because the unique factorization theorem obtained in Corollary 4.12 for marked permutations is not enough to establish the freeness of $\mathcal{A}(\texttt{MPer})$ and, as in [Var14], Lyndon words are the tool that allows us to describe an improved unique factorization theorem in Theorem 4.21.

Definition 4.17 (Lyndon words)**.** Given an alphabet A with a total order, a word $l \in \mathcal{W}(A)$ is said to be a *Lyndon word* if, for any concatenation factorization $l = a_1 \cdot a_2$ into non-empty words, we have that $a_2 \geq l$.

Example 4.18 (Examples of Lyndon words)**.** Consider the (usual) alphabet

$$\Omega = \{\bar{1} <_{per} \bar{1}2 <_{per} \bar{1}32 <_{per} \cdots <_{per} 23\bar{1} <_{per} 24\bar{1}3 <_{per} \dots\},$$

then $(\bar{1}2, \bar{1}32, \bar{1}2, 24\bar{1}3)$ is a Lyndon word in this alphabet. Meanwhile, $(\bar{1}, \bar{1})$ is **not** a Lyndon word.

Definition 4.19 (Stable Lyndon marked permutations)**.** A word on irreducible marked permutations $w = (\xi_1^*, \dots, \xi_j^*) \in \mathcal{W}(\Omega)$ is called *stable Lyndon*, or SL for short, if it is a Lyndon word and satisfies the *stability* conditions introduced in Definition 4.10.

A marked permutation π^* is called *stable Lyndon*, or SL for short, if there exists an SL word $l = (\xi_1^*, \dots, \xi_j^*)$ such that $l^* = \pi^*$. We write $\mathcal{L}_{SL}$ for the set of SL marked permutations. Observe that, from Corollary 4.12, if such an SL word exists it is unique.

We see latter in Theorem 4.24 that $\mathcal{L}_{SL}$ is precisely the set that indexes a free basis of $\mathcal{A}(\texttt{MPer})$. To establish the unique factorization theorem in the context of marked permutations, we first recall a classical fact in Lyndon words.

Theorem 4.20 (Unique Lyndon factorization theorem, [CFL58])**.** Consider a finite alphabet A with a total order. Then any word has a unique factorization, in the concatenation product, into Lyndon words $l_1, \dots, l_k$ such that $l_1 \geq \cdots \geq l_k$ for the lexicographical order in $\mathcal{W}(A)$.

This theorem is central in establishing the freeness of the shuffle algebra in [Rad79]. So is the unique factorization into Lyndon marked permutations below, in establishing the freeness of $\mathcal{A}(\texttt{MPer})$.

Theorem 4.21 (Unique stable Lyndon factorization theorem)**.** Let α^* be a marked permutation. Then there is a unique sequence of SL words on Ω, $l_1, \dots, l_k$ such that $l_i \geq l_{i+1}$ and $\alpha^* = l_1^* \star \cdots \star l_k^*$.

To such a sequence of words, we call the *SL factorization* of α^*.

Proof. The existence follows from Corollary 4.12 and Theorem 4.20. Indeed, for any marked permutation α^*, there is a unique stable factorization $\xi_1^*, \dots, \xi_{j(\alpha^*)}^*$, and from the Lyndon factorization theorem, the word $(\xi_1^*, \dots, \xi_{j(\alpha^*)}^*) \in \mathcal{W}(\Omega)$ admits a factorization into Lyndon words $l_1, \dots, l_k$ such that $l_i \geq l_{i+1}$. These words are stable because $(\xi_1^*, \dots, \xi_{j(\alpha^*)}^*)$ is stable, and from $(\xi_1^*, \dots, \xi_{j(\alpha^*)}^*) = l_1 \cdots l_k$ we have that $\alpha^* = l_1^* \star \cdots \star l_k^*$. We then obtain the desired sequence of SL words $l_1, \dots, l_k$.

For the uniqueness of such a factorization, suppose we have SL words $m_1 \geq \cdots \geq m_{k'}$ that form an SL factorization of α^*. We wish to show that this is precisely the sequence $l_1, \dots, l_k$ constructed above. Write $m_k = (\rho_{k,1}^*, \dots, \rho_{k,z_k}^*)$ for $v = 1, \dots, k'-1$, where $z_v = |m_v|$ and, for readability purposes, consider as well the re-indexing $m_1 \cdots m_{k'} = (\rho_1^*, \dots, \rho_z^*)$.

First observe that from $\alpha^* = m_1^* \star \cdots \star m_{k'}^*$ we get that $\rho_1^* \star \cdots \star \rho_z^*$ is a factorization of α^* into irreducibles. Further, because each m_j is stable, the i-$\oplus$-stability and i-$\ominus$-stability of this factorization is given for any i that is not of the form $z_1 + \cdots + z_v$, for some $v = 1, \dots, k'-1$. On the other hand, it follows from the Lyndon property that $\rho_{k,z_k}^* >_{per} \rho_{k,1}^*$; further, because $m_k \geq m_{k+1}$, we have that $\rho_{k,1}^* \geq_{per} \rho_{k+1,1}^*$. We conclude that $\rho_{k,z_k}^* >_{per} \rho_{k+1,1}^*$. Comparing with Remark 4.16, we have the i-$\oplus$-stability and i-$\ominus$-stability condition for any i that is of the form $z_1 + \cdots + z_v$, for some $v = 1, \dots, k'-1$.

Thus, $\rho_1^* \star \cdots \star \rho_z^*$ is the stable factorization of α^*, so that $(\rho_1^*, \dots, \rho_z^*) = (\xi_1^*, \dots, \xi_k^*)$. Further, the sequence $m_1 \geq \cdots \geq m_{k'}$ is the Lyndon factorization of $(\rho_1^*, \dots, \rho_z^*)$, so $(m_1, \dots, m_{k'}) = (l_1, \dots, l_k)$ by the uniqueness in Theorem 4.20. □

We define $k(\alpha^*)$ to be the number of factors in the stable Lyndon factorization of α^*. Note that for any marked permutation α^*, $k(\alpha^*) \leq j(\alpha^*)$ where we recall that $j(\alpha^*)$ is the number of irreducible factors in a factorization of α^* into irreducible marked permutations.

Definition 4.22 (Word shuffle)**.** Consider Ω an alphabet, and let $w, l_1, \dots, l_k \in \mathcal{W}(\Omega)$. We say that $w = (w_1, \dots, w_j)$ is a *word shuffle* of $l_1, \dots, l_k$ if $[j]$ can be partitioned into k many disjoint sets $\{q_1^{(i)} < \cdots < q_{|l_i|}^{(i)}\}$, where i runs over $i = 1, \dots, k$, such that

$$l_i = (w_{q_1^{(i)}}, \dots, w_{q_{|l_i|}^{(i)}}),$$

for all $i = 1, \dots, k$.

The following theorem is proven in [Rad79, Theorem 2.2.2] and is the main property of Lyndon factorizations on words that we wish to use here. This is also the main ingredient in showing that the shuffle algebra is free.

Theorem 4.23 (Lyndon shuffles - [Rad79])**.** Take Ω an ordered alphabet, and $l_1 \geq \cdots \geq l_k$ Lyndon words on $\mathcal{W}(\Omega)$. Consider $w \in \mathcal{W}(\Omega)$, and assume that w is a word shuffle of $l_1, \dots, l_k$. Then $w \leq l_1 \cdots l_k$.

We remark that there are substantial notation differences: particularly, in [Rad79] a *prime factorization* is what we call here the *Lyndon factorization.* Our notation follows [GR14].

4.3. **Freeness of the pattern algebra in marked permutations.** In this section we state the main steps of the proof of Theorem 4.1. The proof of the technical lemmas is postponed to Section 4.5.

We consider the set of SL marked permutations $\mathcal{L}_{SL}$, which play the role of free generators, and consider a multiset of SL marked permutations $\{\iota_1^* \geq_{fac} \cdots \geq_{fac} \iota_k^*\}$ and the marked permutation $\alpha^* = \iota_1^* \star \cdots \star \iota_k^*$.

Then, all the terms that occur in the right hand side of

$$\prod_{i=1}^{k} \mathbf{p}_{\iota_i^*} = \sum_{\beta^*} \binom{\beta^*}{\iota_1^*, \ldots, \iota_k^*} \mathbf{p}_{\beta^*}\,, \tag{17}$$

correspond to *quasi-shuffles* of $\iota_1^*, \ldots, \iota_k^*$. Below we establish that the marked permutation $\alpha^* = \iota_1^* \star \cdots \star \iota_k^*$ is the biggest such marked permutation occurring on the right hand side of (17), with respect to a suitable total order related to $\leq_{fac}$.

Theorem 4.24. Let α^* be a marked permutation, and suppose that $\iota_1^*, \ldots, \iota_k^*$ is its Lyndon factorization. Then there are coefficients $c_{\beta^*} \geq 0$ such that

$$\prod_{i=1}^{k} \mathbf{p}_{\iota_i^*} = \sum_{|\beta^*|<|\alpha^*|} c_{\beta^*}\, \mathbf{p}_{\beta^*} + \sum_{\substack{|\beta^*|=|\alpha^*| \\ j(\beta^*)<j(\alpha^*)}} c_{\beta^*}\, \mathbf{p}_{\beta^*} + \sum_{\substack{|\beta^*|=|\alpha^*| \\ j(\beta^*)=j(\alpha^*) \\ \beta^* \leq_{fac} \alpha^*}} c_{\beta^*}\, \mathbf{p}_{\beta^*}\,. \tag{18}$$

Furthermore, $c_{\alpha^*} \geq 1$.

With this, the linear independence of all products of the form $\prod_{i=1}^{k} \mathbf{p}_{\iota_i^*}$ follows from the linear independence of $\{\mathbf{p}_{\alpha^*} \,|\alpha^* \in \mathcal{G}(\texttt{MPer})\}$, established earlier in Remark 2.2. In other terms, Theorem 4.24 implies Theorem 4.1. The technical lemmas necessary to prove Theorem 4.24 are the following:

Lemma 4.25. Let β^* be a quasi-shuffle of the marked permutations $\iota_1^*, \ldots, \iota_k^*$. Then, $|\beta^*| \leq |\alpha^*|$. Further, if $|\beta^*| = |\alpha^*|$, then $j(\beta^*) \leq j(\alpha^*)$.

Lemma 4.26 (Factor breaking lemma)**.** Let β^* be a quasi-shuffle of the marked permutations $\iota_1^*, \ldots, \iota_k^*$, such that $|\beta^*| = |\alpha^*|$ and $j(\beta^*) = j(\alpha^*)$. Then $\mathfrak{fac}(\beta^*) = \mathfrak{fac}(\alpha^*)$.

Furthermore, if for each $i = 1, \ldots, k$,

$$\iota_i^* = \zeta_{1,i}^* \star \cdots \star \zeta_{j(\iota_i^*),i}^*$$

is the stable factorization of ι_i^*, then there is a marked permutation γ^* with a factorization into irreducibles given by $\gamma^* = \tau_1^* \star \cdots \star \tau_{j(\gamma^*)}^*$ such that

- we have $\mathfrak{fac}(\gamma^*) = \mathfrak{fac}(\beta^*)$. In particular, $|\gamma^*| = |\beta^*| = |\alpha^*|$ and $j(\gamma^*) = j(\beta^*) = j(\alpha^*)$;
- we have $\gamma^* \geq_{fac} \beta^*$;
- the word $(\tau_1^*, \ldots, \tau_{j(\gamma^*)}^*)$ is a word shuffle of the words $\{(\zeta_{1,i}^*, \ldots, \zeta_{j(\iota_i^*),i}^*)\}_{i=1,\ldots,k}$.

These lemmas will be proven below, in Section 4.5. We remark that, here, the chosen ordering $\leq_{per}$ on the irreducible marked permutations plays a role. That is, in

proving that $\gamma^* \geq_{fac} \beta^*$ we use properties of the order $\leq_{per}$ introduced above, like Remark 4.16. This is unlike the work in [Var14], where any order in the $\oplus$-indecomposable permutations gives rise to a set of free generators.

Lemma 4.27 (Factor preserving lemma)**.** Let γ^* be a marked permutation with a factorization into irreducibles given by $\gamma^* = \tau_1^* \star \cdots \star \tau_{j(\gamma^*)}^*$, and let $l_1, \ldots, l_k$ be SL words, such that

- $l_i \geq l_{i+1}$ for each $i = 1, \ldots, k-1$ and;
- The word $(\tau_1^*, \ldots, \tau_{j(\gamma^*)}^*)$ is a word shuffle of the words $\{l_i\}_{i=1,\ldots,k}$.

Then, $\gamma^* \leq_{fac} l_1^* \star \cdots \star l_k^*$.

In the remaining of the section we assume these lemmas and conclude the proof of the freeness of the pattern algebra, by establishing Theorem 4.24.

Proof of Theorem 4.24. Because a product is a quasi-shuffle of its factors, it follows from Remark 1.18 and from $\alpha^* = \iota_1^* \star \cdots \star \iota_k^*$ that $c_{\alpha^*} \geq 1$. We conclude the proof if we show that any β^* that is a *quasi-shuffle* of $\iota_1^* \ldots, \iota_k^*$ satisfies either

- $|\beta^*| < |\alpha^*|$;
- or $|\beta^*| = |\alpha^*|$ and $j(\beta^*) < j(\alpha^*)$;
- or $|\beta^*| = |\alpha^*|$, $j(\beta^*) = j(\alpha^*)$ and $\beta^* \leq_{fac} \alpha^*$.

Suppose that β^* is a quasi-shuffle of $\iota_1^*, \ldots, \iota_k^*$. For each $i = 1, \ldots k$, let l_i be the SL word corresponding to the SL marked permutation ι_i^*. From Lemma 4.25 we only need to consider the case where $|\beta^*| = |\alpha^*|$ and $j(\beta^*) = j(\alpha^*)$.

From Lemma 4.26 we have that in this case, $\mathfrak{fac}(\beta^*) = \mathfrak{fac}(\alpha^*)$, and there is a marked permutation γ^* that satisfies $\mathfrak{fac}(\gamma^*) = \mathfrak{fac}(\beta^*)$, $\gamma^* \geq_{fac} \beta^*$, and also has a factorization into irreducibles that is a word shuffle of $l_1, \ldots, l_k$. Thus, from Lemma 4.27 we have that $\gamma^* \leq_{fac} l_1^* \star \cdots \star l_k^* = \alpha^*$. It follows that $\beta^* \leq_{fac} \gamma^* \leq_{fac} \alpha^*$, as desired. □

4.4. **Proof of unique factorization theorem.** We start this section with the concept of *DC intervals* and *DC chains* of a marked permutation α^*. We will see that these are closely related to factorizations of α^*, and we will exploit this correspondence to prove the unique factorization theorem Theorem 4.8.

Recall that if β^* is a marked permutation in X, and $I \subseteq X$, we denote $I^* := I \sqcup \{*\}$ for simplicity of notation. We also write $\mathbb{X}(\beta^*) = X$.

Definition 4.28. Let β^* be a marked permutation on the set X, *i.e.* $\beta^* = (\leq_P, \leq_V)$ is a pair of orders on the set X^*. A *doubly connected interval* on β^*, or a *DC interval* for short, is a set $I \subseteq X$ such that I^* is an interval on both orders $\leq_P, \leq_V$. A *proper* DC interval is a DC interval I such that $I \neq X$.

Note that X and $\emptyset$ are always DC intervals. Note as well that if I_1, I_2 are DC intervals, then $I_1 \cup I_2$ is also a DC interval.

Remark 4.29. Suppose that $\beta^* = (\leq_P, \leq_V)$ is a marked permutation, and I is a DC interval of β^*. Consider $m_P, M_P \in X^*$ the minimal, respectively maximal, element for the order $\leq_P$. If $m_P, M_P \in I^*$, then $I = X$.

Symmetrically, if $m_V, M_V \in X^*$ are the minimal, respectively maximal, elements for the order $\leq_V$, and $m_V, M_V \in I^*$, then $I = X$.

For a marked permutation β^*, we use the notation m_P, M_P, m_V, M_V in the remaining of the section to refer to its respective extremes, as in Remark 4.29.

Definition 4.30. For a DC interval I of a marked permutation $\alpha^* = (\leq_P, \leq_V)$, we define the permutation $\alpha^* \backslash_{I^*}$ in the set I^c resulting from the restriction of the orders $\leq_P, \leq_V$ to the set I^c. Alternatively, this is the permutation resulting from the removal of the marked element in $\alpha^*|_{I^c}$.

Remark 4.31. If $\alpha^* = a_1^* \star a_2^*$, then $\mathbb{X}(a_2^*)$ is a DC interval in α^*. On the other hand, if I is a DC interval of α^*, then $\alpha^* = \alpha^*|_{I^c} \star \alpha^*|_I$, so right factors of α^* are in bijection with DC intervals of α^*. Furthermore, I is a maximal proper DC interval of α^* if and only if $\alpha^*|_{I^c}$ is irreducible.

To a factorization $\beta^* = \rho_1^* \star \cdots \star \rho_j^*$ of a marked permutation $\beta^* \in \mathtt{MPer}[X]$ it corresponds a chain of DC intervals

$$\emptyset = J_{j+1} \subsetneq J_j \subsetneq \cdots \subsetneq J_1 = X\,.$$

The chain is defined by $J_k = \mathbb{X}(\rho_k^* \star \cdots \star \rho_j^*)$ and satisfies $\beta^*|_{J_k \setminus J_{k+1}} = \rho_k^*$ for any $k = 1, \ldots, j$. This chain of DC intervals is maximal if and only if the original factorization has only irreducible factors.

Remark 4.32. The marked permutation α^* is $\oplus$-decomposable if and only if there is a DC interval I of α^* such that I^* contains both m_P, m_V, or contains both M_P, M_V.

In the first case, α^* factors as $\beta_1^* \oplus \beta_2$, where $\beta_1^* = \alpha^*|_I$ and $\beta_2 = \alpha^* \backslash_{I^*}$. In the second case, α^* factors as $\beta_1 \oplus \beta_2^*$, where $\beta_1 = \alpha^* \backslash_{I^*}$ and $\beta_2^* = \alpha^*|_I$.

Similarly, α^* is $\ominus$-decomposable if and only if there is a DC interval I of α^* such that I^* contains both M_P, m_V, or contains both m_P, M_V.

In the first case, α^* factors as $\beta_1^* \ominus \beta_2$, where $\beta_1^* = \alpha^*|_I$ and $\beta_2 = \alpha^* \backslash_{I^*}$. In the second case, α^* factors as $\beta_1 \ominus \beta_2^*$, where $\beta_1 = \alpha^* \backslash_{I^*}$ and $\beta_2^* = \alpha^*|_I$.

This characterization of $\oplus$-decomposable marked permutations will be useful in the proof of the classification of the factorizations below, specifically in Lemma 4.37. It is also useful to characterize all $\oplus$-decomposable marked permutations that are irreducible in Section 4.6.

Observation 4.33 ($\oplus$-factorization in marked permutations)**.** Let α^* be a marked permutation. Then there are unique $\oplus$-indecomposable permutations $\epsilon_1, \ldots, \epsilon_k$, $\lambda_1, \ldots, \lambda_j$ and β^* an $\oplus$-indecomposable marked permutation such that

$$\alpha^* = \epsilon_1 \oplus \cdots \oplus \epsilon_k \oplus \beta^* \oplus \lambda_j \oplus \cdots \oplus \lambda_1\,.$$

In this case, we say that α^* is q-$\oplus$-decomposable, where $q = k + j$.

Example 4.34 (q-$\oplus$-decomposition)**.** Consider the marked permutation $21\bar{3}54$. This is a 2-decomposable marked permutation, as it admits the $\oplus$-factorization

$$21\bar{3}54 = 21 \oplus \bar{1} \oplus 21\,.$$

A marked permutation is $\oplus$-indecomposable if and only if it is 0-$\oplus$-decomposable.

Lemma 4.35 (Irreducible $\oplus$-factor lemma)**.** Suppose that α^* is a marked permutation that has the following $\oplus$-factorization

$$\alpha^* = \epsilon_1 \oplus \cdots \oplus \epsilon_u \oplus \beta^* \oplus \lambda_v \oplus \cdots \oplus \lambda_1 \,,$$

where $u + v > 0$. Consider any factorization $\alpha^* = \sigma^* \star \gamma^*$, where σ^* is an irreducible marked permutation.

- If both $u > 0$ and $v > 0$, then either $\sigma^* = \bar{1} \oplus \lambda_1$ or $\sigma^* = \epsilon_1 \oplus \bar{1}$.
- If $u = 0$, then $\sigma^* = \bar{1} \oplus \lambda_1$.
- If $v = 0$, then $\sigma^* = \epsilon_1 \oplus \bar{1}$.

Proof. Let us deal with the case where both $u > 0$ and $v > 0$ first. Assume that neither $\sigma^* = \bar{1} \oplus \lambda_1$ nor $\sigma^* = \epsilon_1 \oplus \bar{1}$. Consider the following DC intervals, that are all distinct

$$Y_1 = \mathbb{X}(\epsilon_2 \oplus \cdots \oplus \epsilon_u \oplus \beta^* \oplus \lambda_v \oplus \cdots \oplus \lambda_1) \,,$$
$$Y_2 = \mathbb{X}(\epsilon_1 \oplus \cdots \oplus \epsilon_u \oplus \beta^* \oplus \lambda_v \oplus \cdots \oplus \lambda_2) \,,$$
$$Y = \mathbb{X}(\gamma^*) \,.$$

Note that $m_P \in Y_2^* \setminus Y_1^*$ and $M_P \in Y_1^* \setminus Y_2^*$ by construction. Note also that $\alpha^*|_{Y_1^c} = \epsilon_1 \oplus \bar{1}$ and $\alpha^*|_{Y_2^c} = \bar{1} \oplus \lambda_1$ are irreducible marked permutations, so both Y_1, Y_2 are maximal proper DC intervals. Furthermore, Y is also a maximal proper DC interval. This maximality gives us that $Y \cup Y_1 = X = Y \cup Y_2$.

From $Y \cup Y_1 = X$ and $m_P \notin Y_1^*$, we have that $m_P \in Y^*$, and from $Y \cup Y_2 = X$ and $M_P \notin Y_2^*$ we get that $M_P \in Y^*$. From Remark 4.29, this is a contradiction with the fact that Y is a proper DC interval, thus contradicting the assumption that neither $\sigma^* = \bar{1} \oplus \lambda_1$ nor $\sigma^* = \epsilon_1 \oplus \bar{1}$, as desired.

Now suppose that $u = 0$, and $v > 0$. Assume for the sake of contradiction that $\sigma^* \neq \bar{1} \oplus \lambda_1$, and define the following distinct DC intervals

$$Y_2 = \mathbb{X}(\beta^* \oplus \lambda_v \oplus \cdots \oplus \lambda_2) \text{ and } Y = \mathbb{X}(\gamma^*) \,.$$

Because both σ^* and $\bar{1} \oplus \lambda_1$ are irreducible, the DC intervals Y_2, Y are both maximal proper, so $X = Y \cup Y_2$. Further, $m_P \in Y_2^*$ by construction, so $M_P \in Y^*$ by Remark 4.29. Consider the DC interval $I = \mathbb{X}(\beta^*)$, and notice that $m_P \in I^*$ by construction. Thus we have that $I^* \cup Y^* = X^*$, and $Y^c \subseteq I$.

We have the following decomposition, depicted in Fig. 6:

$$\beta^* = \alpha^* \setminus_{Y^*} \oplus \alpha^*|_{I \cap Y} \,.$$

Notice that $\alpha^* \setminus_{Y^*}$ is not the empty permutation, because Y is a proper DC interval. This is a contradiction with the fact that β^* is $\oplus$-indecomposable, so $\sigma^* = \bar{1} \oplus \lambda_1$.

The case where $u > 0$ and $v = 0$ is done in a similar way, so the result follows. $\square$

Corollary 4.36. Consider a marked permutation α^* that is q-$\oplus$-decomposable, so that $\alpha^* = \epsilon_1 \oplus \cdots \oplus \epsilon_u \oplus \beta^* \oplus \lambda_v \oplus \cdots \oplus \lambda_1$ is the $\oplus$-decomposition of α^*. If $\sigma_1^* \star \cdots \star \sigma_j^*$ is a factorization of α^* into irreducibles, then $j \geq q$ and

(1) $\sigma_{q+1}^* \star \cdots \star \sigma_j^* = \beta^*$.
(2) By applying $\oplus$-relations to $(\sigma_1^*, \ldots, \sigma_q^*)$ we can obtain

$$(\bar{1} \oplus \lambda_1, \ldots, \bar{1} \oplus \lambda_v, \epsilon_1 \oplus \bar{1}, \ldots, \epsilon_u \oplus \bar{1}) \,.$$

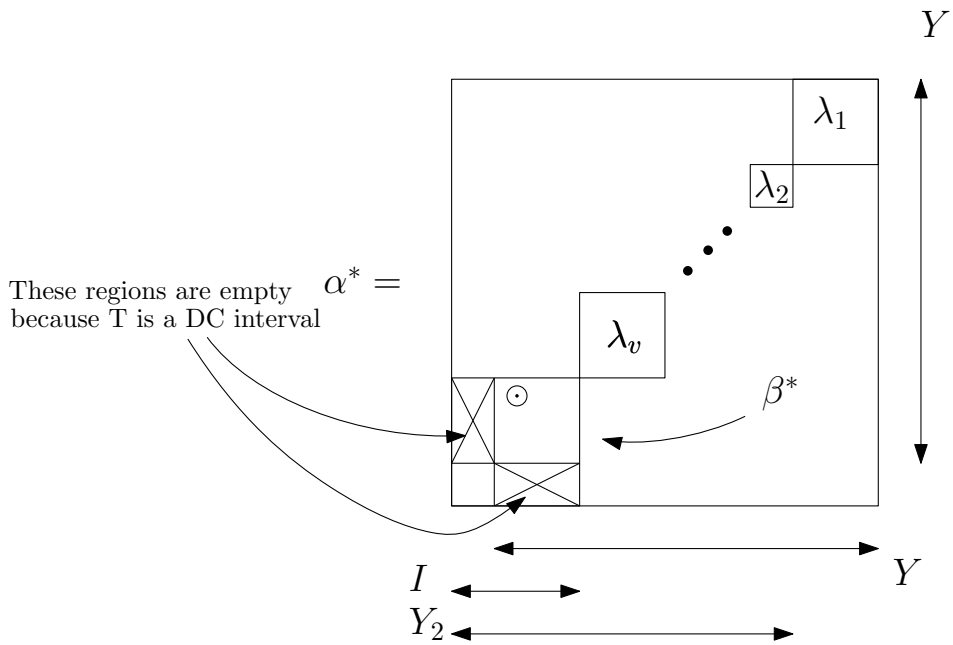

FIGURE 6. A decomposition of β^*, whenever the DC interval Y is proper

Proof. We use induction on q. The base case is $q = 0$, where there is nothing to establish in 2. and we need only to show that

$$\sigma_1^* \star \cdots \star \sigma_j^* = \beta^* ,$$

which follows because $\sigma_1^* \star \cdots \star \sigma_j^* = \alpha^* = \beta^*$.

Now for the induction step, we assume that $q \geq 1$. From Lemma 4.35, σ_1^* is either $\epsilon_1 \oplus \bar{1}$ or $\bar{1} \oplus \lambda_1$, thus according to Remark 1.12 $\zeta^* := \sigma_2^* \star \cdots \star \sigma_k^*$ is either

$$\epsilon_2 \oplus \cdots \oplus \epsilon_u \oplus \beta^* \oplus \lambda_v \oplus \cdots \oplus \lambda_1 ,$$

or

$$\epsilon_1 \oplus \cdots \oplus \epsilon_u \oplus \beta^* \oplus \lambda_v \oplus \cdots \oplus \lambda_2.$$

Without loss of generality, assume the first case. Then ζ^* is $(q-1)$-$\oplus$-decomposable, so the induction hypothesis applies to the factorization of ζ^*. That is, we have that

$$\sigma_{q+1}^* \star \cdots \star \sigma_j^* = \beta^* .$$

Furthermore, the induction case gives us that $(\bar{1} \oplus \lambda_1, \ldots, \bar{1} \oplus \lambda_v, \epsilon_2 \oplus \bar{1}, \ldots, \epsilon_u \oplus \bar{1})$ can be obtained from $(\sigma_2^*, \ldots, \sigma_q^*)$ by a series of $\oplus$-relations.

So, by only using $\oplus$-relations,

$$\begin{aligned} & (\bar{1} \oplus \lambda_1, \ldots, \bar{1} \oplus \lambda_v, \epsilon_1 \oplus \bar{1}, \epsilon_2 \oplus \bar{1}, \ldots, \epsilon_u \oplus \bar{1}) \\ \sim \quad & (\epsilon_1 \oplus \bar{1}, \bar{1} \oplus \lambda_1, \ldots, \bar{1} \oplus \lambda_v, \epsilon_2 \oplus \bar{1}, \ldots, \epsilon_u \oplus \bar{1}) \\ \sim_{ind.hip.} \quad & (\sigma_1^*, \sigma_2^*, \ldots, \sigma_q^*) \end{aligned} \tag{19}$$

concluding the the proof. □

We remark that Observation 4.33, Lemma 4.35, and Corollary 4.36 have counterparts for the $\ominus$ operation.

Lemma 4.37 (Indecomposable irreducible factor lemma)**.** Suppose that α^* is a non-trivial marked permutation that is both $\oplus$-indecomposable and $\ominus$-indecomposable.

Consider two factorizations of α^*, $\sigma_1^* \star \beta_1^*$ and $\sigma_2^* \star \beta_2^*$, where σ_1^*, σ_2^* are irreducible marked permutations. Then $\sigma_1^* = \sigma_2^*$ and $\beta_1^* = \beta_2^*$.

Proof. According to Remark 4.31, it suffices to see that α^* only has one maximal proper DC interval. Suppose otherwise, that there are Y_1, Y_2 distinct proper maximal DC intervals, for the sake of contradiction. Further, consider m_P, M_P, m_V, M_V the usual maximal and minimal elements in $\mathbb{X}(\alpha^*)^*$.

We have that $Y_1 \cup Y_2 = X$, so $m_P, M_P, m_V, M_V \in Y_1^* \cup Y_2^*$. Without loss of generality suppose that $m_P \in Y_1^*$. We know that $\{m_V, M_V\} \not\subseteq Y_2^*$, so there are only two cases to consider:

- We have that $m_P, m_V \in Y_1^*$. Then α^* is $\oplus$-decomposable, according to Remark 4.32.
- We have that $m_P, M_V \in Y_1^*$. Then α^* is $\ominus$-decomposable, according to Remark 4.32.

In either case, we reach a contradiction. It follows that $\beta_1^* = \beta_2^*$ and $\sigma_1^* = \sigma_2^*$. □

Proof of Theorem 4.8. Consider a marked permutation α^*, and take words $w_1 = (\xi_1^*, \dots, \xi_k^*)$ and $w_2 = (\rho_1^*, \dots, \rho_j^*)$ in $\mathcal{W}(\Omega)$ such that

$$\xi_1^* \star \cdots \star \xi_k^* = \rho_1^* \star \cdots \star \rho_j^* =: \alpha^* , \tag{20}$$

and assume for the sake of contradiction that these can be chosen with $w_1 \not\sim w_2$. Further choose such words minimizing $k + j = |w_1| + |w_2|$. We need only to consider four cases:

The marked permutation $\alpha^* = \bar{1}$ is trivial: then by a size argument we have that both words w_1 and w_2 are empty, thus $w_1 = w_2$.

The marked permutation α^* is indecomposable: in this case, from Lemma 4.37 we know that $\xi_1^* = \rho_1^*$ and that $\xi_2^* \star \cdots \star \xi_k^* = \rho_2^* \star \cdots \star \rho_j^*$. Thus, by minimality, $(\xi_2^*, \dots, \xi_k^*) \sim (\rho_2^*, \dots, \rho_j^*)$, which implies $(\xi_1^*, \dots, \xi_k^*) \sim (\rho_1^*, \dots, \rho_j^*)$ and that is a contradiction.

The marked permutation α^* is $\oplus$-decomposable: Assume now that α^* is q-$\oplus$-decomposable, for some $q > 0$. That is, there are $\oplus$-indecomposables such that

$$\alpha^* = \epsilon_1 \oplus \cdots \oplus \epsilon_u \oplus \beta^* \oplus \lambda_v \oplus \cdots \oplus \lambda_1, \quad u + v = q\,.$$

Then Corollary 4.36 tells us that $q \leq k, j$ and that

$$\xi_{q+1}^* \star \cdots \star \xi_k^* = \beta^* = \rho_{q+1}^* \star \cdots \star \rho_j^*\,. \tag{21}$$

Further, we also have that

$$(\xi_1^*, \dots, \xi_q^*) \sim (\bar{1} \oplus \lambda_1, \dots, \bar{1} \oplus \lambda_v, \epsilon_1 \oplus \bar{1}, \dots, \epsilon_u \oplus \bar{1}) \sim (\rho_1^*, \dots, \rho_q^*)\,. \tag{22}$$

Because $q > 0$, from (21) and the minimality of $(\xi_1^*, \dots, \xi_k^*), (\rho_1^*, \dots, \rho_j^*)$ we have that

$$(\xi_{q+1}^*, \dots, \xi_k^*) \sim (\rho_{q+1}^*, \dots, \rho_j^*)\,.$$

This, together with (22) gives us $(\xi_1^*, \dots, \xi_k^*) \sim (\rho_1^*, \dots, \rho_j^*)$ and that is a contradiction.

The marked permutation α^* is $\ominus$-decomposable: This case is similar to the previous one.

We conclude that a factorization of a marked permutation into irreducibles is unique up to the relations in Remark 4.6. □

4.5. **Proofs of Lemmas 4.25 to 4.27.** We start by fixing some notation. Consider a multiset $\{\iota_1^* \geq \cdots \geq \iota_k^*\}$ of SL marked permutations, and let α^* be the marked permutation with Lyndon factorization $(\iota_1^*, \ldots, \iota_k^*)$ which exists and is unique by Theorem 4.21, and has $\alpha^* = \iota_1^* \star \cdots \star \iota_k^*$.

Note that $\mathcal{S} := \biguplus_{i=1}^k \mathfrak{fac}(\iota_i^*) = \mathfrak{fac}(\alpha^*)$ as multisets, from Corollary 4.9.

Proof of Lemma 4.25. Let β^* be a marked permutation in X. By Definition 2.5, there are sets $I_1, \ldots, I_k$ such that:

$$I_1 \cup \cdots \cup I_k = X \text{ and } \beta^*|_{I_i} \sim \iota_i^* \ \ \forall i = 1, \ldots, k\,. \tag{23}$$

Suppose that β^* has stable factorization $\beta^* = \rho_1^* \star \cdots \star \rho_j^*$, where $j = j(\beta^*)$, with a corresponding maximal DC interval chain

$$\emptyset = J_{j+1} \subsetneq J_j \subsetneq \cdots \subsetneq J_1 = X\,,$$

as given in Remark 4.31, so that $\beta^*|_{J_p \setminus J_{p+1}} = \rho_p$ for $p = 1, \ldots, j$.

Since β^* is a quasi-shuffle of $\iota_1^*, \ldots, \iota_k^*$, the sets $I_1, \ldots, I_k$ cover X and so

$$|\beta^*| = |X| \leq \sum_{i=1}^k |I_i| = \sum_{i=1}^k |\iota_i^*| = |\alpha^*|\,, \tag{24}$$

so we conclude that $|\beta^*| \leq |\alpha^*|$.

Now, assume that we have $|\beta^*| = |\alpha^*|$, so we have equality in (24), thus the sets I_i are mutually disjoint. Recall that $j(\alpha^*) = \sum_i j(\iota_i^*)$, as shown in Corollary 4.9. Therefore, we only need to establish that $\sum_i j(\iota_i^*) \geq j$. For each $i = 1, \ldots, k$, the following is a DC interval chain of ι_i^*, which is not necessarily a maximal one:

$$\emptyset = J_{j+1} \cap I_i \subseteq J_j \cap I_i \subseteq \cdots \subseteq J_1 \cap I_i = I_i\,, \tag{25}$$

so let us consider the set $U_p := \{i \in [k] | J_{p+1} \cap I_i \neq J_p \cap I_i\}$.

First it is clear that each U_p is non empty, as otherwise we would have

$$J_{p+1} = \bigcup_i J_{p+1} \cap I_i = \bigcup_i J_p \cap I_i = J_p\,.$$

On the other hand, the length of the DC chain is given precisely by the number of strict inequalities in (25), that is by $|\{p \in [j] | i \in U_p\}|$. From Corollary 4.9 and Remark 4.31, this is at most $j(\beta^*|_{I_i}) = j(\iota_i^*)$, so

$$\sum_i j(\iota_i^*) \geq \sum_i |\{p \in [j] | i \in U_p\}| = \sum_{p=1}^j |U_p| \geq j\,. \tag{26}$$

So we conclude that $j(\alpha^*) \geq j(\beta^*)$. □

Proof of Lemma 4.26. Assume now that $|\beta^*| = |\alpha^*|$ and $j(\beta^*) = j(\alpha^*)$. As a corollary of the proof above, and using the same notation, we have that the sets I_i are pairwise disjoint, and we obtain equality all through (26), so that each U_p is in fact a singleton.

We wish to split the indexing set $[j(\gamma^*)]$ into k many disjoint increasing sequences $q_1^{(i)} < \cdots < q_{j(\iota_i^*)}^{(i)}$ for $i = 1, \ldots, k$ such that

$$\iota_i^* = \tau^*_{q_1^{(i)}} \star \cdots \star \tau^*_{q_{j(\iota_i^*)}^{(i)}}\,, \tag{27}$$

is precisely the stable factorization of each ι_i^*.

Define the map $\zeta : [j] \to [k]$ that sends p to the unique element in U_p. This map will give us the desired increasing sequences. It satisfies

$$|\zeta^{-1}(i)| = |\{p \in [j] | i \in U_p\}| = j(\iota_i^*) \text{ for } i = 1, \ldots, k\,.$$

Write $\zeta^{-1}(i) = \{q_1^{(i)} < \cdots < q_{j(\iota_i^*)}^{(i)}\} \subseteq [j]$ and set $q_{j(\iota_i^*)+1}^{(i)} = j+1$. In the following, we identify the marked permutations ι_i^* and $\beta^*|_{I_i}$ in order to find a factorization of ι^* into irreducibles. Specifically, we get that

$$\iota_i^* = \beta^*|_{I_i} = \rho_1^*|_{I_i \cap J_1 \setminus J_2} \star \cdots \star \rho_j^*|_{I_i \cap J_j \setminus J_{j+1}} \tag{28}$$

is a factorization of ι_i^*. Because each U_p is a singleton, $I_i \cap J_p \setminus J_{p+1}$ is either $J_p \setminus J_{p+1}$ or $\emptyset$, thus $\rho_p^*|_{I_i \cap J_p \setminus J_{p+1}}$ is either ρ_p^* of $\bar{1}$, and the factorization in (28), after removing the trivial terms, becomes the following factorization into irreducibles:

$$\iota_i = \rho^*_{q_1^{(i)}} \star \cdots \star \rho^*_{q_{j(\iota_i^*)}^{(i)}}\,. \tag{29}$$

Then, we indeed have that $\mathfrak{fac}(\beta^*) = \biguplus_{i=1}^k \mathfrak{fac}(\iota_i^*)$. We now construct the desired marked permutation γ^*.

According to Corollary 4.9, for each i, the unique stable factorization of ι_i^* results from $\iota_i^* = \rho^*_{q_1^{(i)}} \star \cdots \star \rho^*_{q_{j(\iota_i^*)}^{(i)}}$ by a permutation of the factors. Thus, for each i we can find indices $p_1^{(i)}, \ldots, p_{j(\iota_i^*)}^{(i)}$ such that $\{p_1^{(i)}, \ldots, p_{j(\iota_i^*)}^{(i)}\} = \{q_1^{(i)}, \ldots, q_{j(\iota_i^*)}^{(i)}\}$ and

$$\iota_i^* = \rho^*_{q_1^{(i)}} \star \cdots \star \rho^*_{q_{j(\iota_i^*)}^{(i)}} = \rho^*_{p_1^{(i)}} \star \cdots \star \rho^*_{p_{j(\iota_i^*)}^{(i)}}$$

is the stable factorization of ι_i^*.

For each $s \in [j]$, if $s = q_u^{(i)}$ for some integers i, u, define $\tau_s := \rho^*_{p_u^{(i)}}$ and finally define

$$\gamma^* = \tau_1^* \star \cdots \star \tau_j^*\,.$$

A graphical description of this construction for $j = 8$ is given in Fig. 7. We claim that γ^* satisfies the three conditions described in the lemma. First, it is clear that $\mathfrak{fac}(\beta^*) = \biguplus_{i=1}^k \mathfrak{fac}(\iota_i^*) = \mathfrak{fac}(\gamma)$, and that the disjoint increasing sequences $q_1^{(i)} < \cdots < q_{j(\iota_i^*)}^{(i)}$ are such that

$$\tau^*_{q_1^{(i)}} \star \cdots \star \tau^*_{q_{j(\iota_i^*)}^{(i)}} = \rho^*_{p_1^{(i)}} \star \cdots \star \rho^*_{p_{j(\iota_i^*)}^{(i)}}$$

Stable Factorizations

$$\alpha^* = \iota_1^* \star \iota_2^* \star \iota_3^*$$

$$\beta^* = \rho_1^* \star \rho_2^* \star \rho_3^* \star \rho_4^* \star \rho_5^* \star \rho_6^* \star \rho_7^* \star \rho_8^*$$

$$\iota_1^* = \rho_5^* \star \rho_2^* \star \rho_8^*$$

$$\iota_2^* = \rho_3^* \star \rho_4^* \star \rho_1^*$$

$$\boxed{\iota_3^* = \rho_6^* \star \rho_7^*}$$

$$\gamma^* = \boxed{\rho_3^*} \star \boxed{\rho_5^*} \star \boxed{\rho_4^*} \star \boxed{\rho_1^*} \star \boxed{\rho_2^*} \star \boxed{\rho_6^*} \star \boxed{\rho_7^*} \star \boxed{\rho_8^*}$$

FIGURE 7. Example of construction of γ^* for $j = 8$.

is precisely the stable factorization of ι_i^*. Thus, we need only to establish that $\beta^* \leq_{fac} \gamma^*$.

We claim that $(\tau_1^*, \ldots, \tau_j^*) \geq (\rho_1^*, \ldots, \rho_j^*)$. Indeed, $(\tau_1^*, \ldots, \tau_j^*)$ is obtained from $(\rho_1^*, \ldots, \rho_j^*)$ by the stabilization procedure in its subwords. From the stabilization procedure we obtain bigger words in the lexicographical order, according to Remark 4.16. This still holds true even if only applied to a subword, thus the resulting word $(\tau_1^*, \ldots, \tau_j^*)$ is lexicographically bigger than $(\rho_1^*, \ldots, \rho_j^*)$.

Now, if $\gamma^* = \epsilon_1^* \star \cdots \star \epsilon_j^*$ is the stable factorization of γ^*, because $\gamma^* = \tau_1^* \star \cdots \star \tau_j^*$ we have that from Remark 4.16 that

$$(\tau_1^*, \ldots, \tau_j^*) \leq (\epsilon_1^*, \ldots, \epsilon_j^*)\,,$$

Thus

$$(\rho_1^*, \ldots, \rho_j^*) \leq (\epsilon_1^*, \ldots, \epsilon_j^*)\,,$$

and so $\beta^* \leq_{fac} \gamma^*$, as desired. □

In the following proof, we will start by showing that the given factorization of γ^* can be assumed to be the stable factorization. Then, using Theorem 4.20 and the fact that the factorization of γ^* is a shuffle of Lyndon words to establish the desired inequality.

Proof of Lemma 4.27. Write $j = j(\gamma^*)$. We first see that if γ^* has some factorization that is a word shuffle of stable words $l_1, \ldots, l_k$, then the stable factorization is also a word shuffle of these words.

Indeed, take a factorization $\tau_1^* \star \cdots \star \tau_j^*$ such that $(\tau_1^*, \ldots, \tau_j^*)$ is a word quasi-shuffle of $l_1, \ldots, l_k$, and say that there is some $u \in \{1, \ldots, j-1\}$ such that this factorization is not u-$\oplus$-stable or is not u-$\ominus$-stable. According to Remark 4.11, to show that the stable factorization is also a word shuffle of $l_1, \ldots, l_k$, it suffices to show that the word resulting from swapping τ_u^* and τ_{u+1}^* in $(\tau_1^*, \ldots, \tau_j^*)$ is still a word quasi-shuffle of $l_1, \ldots, l_k$.

When we apply a u-stability reduction (see Remark 4.11), we can still find a suitable partition of $[j]$ into k many disjoint increasing sequences. Say that $[j]$ is partitioned into the blocks $\{q_1^{(i)} < \cdots < q_{j(\iota_i^*)}^{(i)}\}$, then there are integers i_1, i_2, v_1, v_2 such that $u = q_{v_1}^{(i_1)}$

and $u+1 = q_{v_2}^{(i_2)}$. Because $(\tau^*_{q_1^{(i)}}, \dots, \tau^*_{q^{(i)}_{j(i_i^*)}})$ is stable for each i, we cannot have $i_1 = i_2$. Therefore, by swapping the elements $u, u+1$ we obtain a new partition for the new factorization, thus showing that it is a word quasi-shuffle of $l_1, \dots, l_k$.

Now let $\rho_1^* \star \cdots \star \rho_j^*$ be the stable factorization of γ^*. Since $(\rho_1^*, \dots, \rho_j^*)$ is a word shuffle of $l_1, \dots, l_k$, which are Lyndon words in $\mathcal{W}(\Omega)$, we have from Theorem 4.23 that $(\rho_1^*, \dots, \rho_j^*) \leq l_1 \cdots l_k$, and so $\gamma^* \leq_{fac} l_1^* \star \cdots \star l_k^*$. □

4.6. **Primitive elements, growth rates and asymptotic analysis.** Recall that in Section 2.3 we define the space of primitive elements $P(H)$ of a Hopf algebra H is the subspace of H given by $\{a \in H | \Delta a = a \otimes 1 + 1 \otimes a\}$. In the particular case of the pattern Hopf algebra $\mathcal{A}(\texttt{MPer})$, its primitive space is spanned by $\{\mathbf{p}_{\pi^*} | \pi^* \text{ is irreducible marked permutation}\}$, according to Proposition 2.10. So, we are interested in enumerating the irreducible marked permutations. We consider some generating functions:

- The power series $P^*(x) = \sum_{\pi^* \text{ marked permutation}} x^{|\pi^*|} = \sum_{n \geq 1} n \cdot n!\, x^{n-1}$ counts marked permutations.
- The power series $P(x) = \sum_{\pi \text{ permutation}} x^{|\pi|} = \sum_{n \geq 0} n!\, x^n$ counts permutations.
- The power series $S^*(x) = \sum_{k \geq 0} s_k x^k = \sum_{\pi^* \text{ irreducible marked permutation}} x^{|\pi^*|}$ counts irreducible marked permutations. This is the main generating function that we aim to enumerate here.
- The power series $S_o^*(x) = \sum_{k \geq 0} so_k x^k = \sum_{\substack{\pi^* \text{ irreducible and indecomposable} \\ \text{marked permutation}}} x^{|\pi^*|}$ counts irreducible marked permutations that are indecomposable, that is neither $\oplus$-decomposable nor $\ominus$-decomposable.
- The power series $P^{\oplus}(x) = \sum_{\pi \text{ is } \oplus- \text{ decomposable}} x^{|\pi|}$ counts permutations that are $\oplus$-indecomposable. This also counts the permutations that are $\ominus$-indecomposable.

Because we have a unique factorization theorem on permutations under the $\oplus$ product, the coefficients of $P^{\oplus}(x)$ can be easily extracted via the following power series relation

$$\frac{1}{1 - P^{\oplus}(x)} = P(x) = \sum_{n \geq 0} n! x^n \,, \text{ which implies } P^{\oplus}(x) = 1 - P(x)^{-1} \,.$$

Observation 4.38. Any $\oplus$-decomposable irreducible marked permutation is either of the form $\bar{1} \oplus \pi$ or of the form $\pi \oplus \bar{1}$ for π an $\oplus$-indecomposable permutation, and symmetrically for $\ominus$-decomposable irreducible marked permutations.

Thus, we have

$$S^*(x) - S_o^*(x) = 4P^{\oplus}(x) \,. \tag{30}$$

The following proposition allows us to enumerate easily the irreducible marked permutations and the irreducible indecomposable marked permutations, as done in Table 1.

Proposition 4.39 (Power series of irreducible marked permutations)**.**

$$S^*(x) = 3 + \frac{2}{P(x)^2} - \frac{1}{P'(x)} - \frac{4}{P(x)} \,.$$

$$S_o^*(x) = -1 + \frac{2}{P(x)^2} - \frac{1}{P'(x)}\,.$$

We compare this result with the enumeration of simple permutations in [AAK03, Equation 1], where the power series enumerating simple permutations is given as the solution of a functional equation that is not rational. Thus, we expect it to be simpler to compute the coefficients explicitly.

n	0	1	2	3	4	5	6	7	8	9
so_n	0	0	0	8	78	756	7782	85904	1016626	12865852
s_n	0	4	4	20	130	1040	9626	99692	1132998	13959224

TABLE 1. First elements of the sequences so_n and s_n.

Lemma 4.40. Let π^* be a marked permutation. Then, there are four cases

- There are unique marked permutations σ^* and α^* such that σ^* is indecomposable and irreducible, and $\pi^* = \sigma^* * \alpha^*$.
- The marked permutation π^* is $\oplus$-decomposable.
- The marked permutation π^* is $\ominus$-decomposable.
- $\pi^* = \bar{1}$.

In particular, we have the following equation

$$P^*(x) = S_o^*(x)P^*(x) + 2(P(x) - P^{\oplus}(x))' + 1\,. \tag{31}$$

Proof. If π^* is indecomposable, then, either $\pi^* = \bar{1}$, or there are unique marked permutations σ^* and α^* such that σ^* is indecomposable and irreducible, and $\pi^* = \sigma^* * \alpha^*$, according to Lemma 4.37. This concludes the first part of the lemma.

Further, we observe that

$$(P(x) - P^{\oplus}(x))'$$

counts the marked permutations that are $\oplus$-decomposable (and by symmetry, the ones that are $\ominus$-decomposable). □

Proof of Proposition 4.39. From (31) along with the fact that $P^*(x) = P'(x)$, we have that

$$S_o^*(x) = -1 + 2P(x)^{-2} - P'(x)^{-1}\,,$$

$$S^*(x) = 3 + 2P(x)^{-2} - P'(x)^{-1} - 4P(x)^{-1}\,,$$

as desired. □

AKNOWLEDGMENTS

The author would like to thank the support of the SNF grant number 200020-172515. The author is also grateful for the conversations with Yannic Vargas about pattern Hopf algebras and his insightful comments.

References

[AAK03] Michael H Albert, Mike D Atkinson, and Martin Klazar. The enumeration of simple permutations. *J. Integer Seq*, 6(03.4):4, 2003.

[ABF18] Michael Albert, Mathilde Bouvel, and Valentin Féray. Two first-order logics of permutations. *preprint arXiv:1808.05459*, 2018.

[AM10] Marcelo Aguiar and Swapneel Arvind Mahajan. *Monoidal functors, species and Hopf algebras*. American Mathematical Society Providence, RI, 2010.

[BZ09] Nantel Bergeron and Mike Zabrocki. The Hopf algebras of symmetric functions and quasi-symmetric functions in non-commutative variables are free and co-free. *Journal of Algebra and Its Applications*, 8(04):581–600, 2009.

[CFL58] Kuo Tsai Chen, Ralph H Fox, and Roger C Lyndon. Free differential calculus, IV. The quotient groups of the lower central series. *Annals of Mathematics*, 68:81–95, 1958.

[Foi12] Loïc Foissy. Free and cofree Hopf algebras. *Journal of Pure and Applied Algebra*, 216(2):480–494, 2012.

[FPT16] Loïc Foissy, Frédéric Patras, and Jean-Yves Thibon. Deformations of shuffles and quasi-shuffles. *Annales de l'Institut Fourier*, 66(1):209–237, 2016.

[GR14] Darij Grinberg and Victor Reiner. Hopf algebras in combinatorics. *preprint arXiv:1409.8356*, 2014.

[Haz01] Michiel Hazewinkel. The algebra of quasi-symmetric functions is free over the integers. *Advances in Mathematics*, 164(2):283–300, 2001.

[Hof00] Michael E Hoffman. Quasi-shuffle products. *Journal of Algebraic Combinatorics*, 11(1):49–68, 2000.

[Knu11] Donald E Knuth. *Art of Computer Programming, Volumes 1-4A Boxed Set*. Addison-Wesley Professional, 2011.

[Mac63] Saunders MacLane. Natural associativity and commutativity. *Rice Institute Pamphlet-Rice University Studies*, 49(4), 1963.

[Rad79] David E Radford. A natural ring basis for the shuffle algebra and an application to group schemes. *Journal of Algebra*, 58(2):432–454, 1979.

[San19] Mario Sanchez. Möbius inversion as duality for Hopf monoids. volume 80B, 2019.

[Sch94] William R. Schmitt. Incidence Hopf algebras. *J. Pure Appl. Algebra 96*, pages 299–330, 1994.

[Sta86] Richard P Stanley. Enumerative combinatorics. vol. I, The Wadsworth & Brooks/Cole Mathematics Series, Wadsworth & Brooks, 1986.

[Sup19] Mariel Supina. The Hopf monoid of orbit polytopes. *preprint arXiv:1904.08437*, 2019.

[Swe69] M.E. Sweedler. *Hopf algebras*. Mathematics lecture note series. W. A. Benjamin, 1969.

[Tak71] Mitsuhiro Takeuchi. Free Hopf algebras generated by coalgebras. *Journal of the Mathematical Society of Japan*, 23(4):561–582, 1971.

[Var14] Yannic Vargas. Hopf algebra of permutation pattern functions. In *Discrete Mathematics and Theoretical Computer Science*, pages 839–850. Discrete Mathematics and Theoretical Computer Science, 2014.

[Whi32] Hassler Whitney. The coloring of graphs. *Annals of Mathematics*, 33:688–718, 1932.

[Wil02] Herbert S Wilf. The patterns of permutations. *Discrete Mathematics*, 257(2-3):575–583, 2002.