ON THE HOPF ALGEBRA OF GRAPHS

MIODRAG IOVANOV AND JAIUNG JUN

Abstract. The algebra of graphs is defined as the algebra which has a formal basis \( G \) of all isomorphism types of graphs, and multiplication is to take the disjoint union. We explicitly describe here the structure of the Hopf algebra of graphs \( H \). We find an explicit basis \( B \) of the space of primitives, such that each graph is a polynomial with non-negative integer coefficients of the elements of \( B \), and each \( b \in B \) is a polynomial with integer coefficients in \( G \). Using this, we find the cancellation and grouping free formula for the antipode. The coefficients appearing in all these polynomials are, up to signs, numbers counting multiplicities of subgraphs in a graph. We then investigate applications of this to the graph reconstruction conjectures, and rederive some results in the literature on these questions.

1. Introduction

Hopf algebras are now ubiquitous in many fields of mathematics. In combinatorics, Hopf algebras naturally appear in the following way: one has some class of combinatorial objects to study, for instance graphs, matroids, posets, or symmetric functions. One then construct a Hopf algebra (a combinatorial Hopf algebra) whose generators are these objects, and basic operations among combinatorial objects are encoded in the algebraic structure of a Hopf algebra. In this respect, the product of two objects is obtained by a natural combinatorial construction which puts them together (such as taking disjoint union), and coproduct is defined by using all possible ways to “split” the given combinatorial object into two parts in a suitable way. Finally, one hopes to use certain algebraic properties of a Hopf algebra to return to combinatorics and obtain new information, such as new combinatorial identities. For instance, in [CS05], Crapo and Schmitt used this line of idea to give a short proof of Welsh’s conjecture on a lower bound for the number of isomorphism classes of matroids on \( \{1, \ldots, n\} \). We also note that in [EJS18], Eppolito, Jun, and Szczesny provided another interpretation of the combinatorial Hopf algebra arising from matroids as the Hall algebra associated to the category of matroids. For introduction to Hopf algebras in combinatorics, we refer the reader to [GR14].

When one constructs Hopf algebras by the aforementioned recipe, one usually has only explicit description of product and coproduct, which gives one a connected, graded bialgebra. In this case, by classical Hopf algebra theory (see, for example, [Tak71]), one has that such a bialgebra has an antipode and hence is a Hopf algebra.

One of the main questions regarding any such combinatorial algebra, which has received a lot of attention, is to explicitly find the antipode. However, the antipode formulas usually involve massive cancellations, and hence in many cases the antipode formula is not optimal for use in relation to certain combinatorial identities.

Recently, considerable attention has been dedicated to finding cancellation-free formulas for the antipode of various Hopf algebras. In their groundbreaking work [AA17], Aguiar and Ardila provided an elegant unified way to find a cancellation-free and grouping-free antipode formula for various classes of combinatorial Hopf algebras by reducing the question to the case of generalized permutahedra (or polymatroids). In [BEJM18], Bucher, Eppolito, Jun, and Matherne also employed the idea...
of sign-reversing involution, which was introduced in [BS17] by Benedetti and Sagan, and provided a cancellation-free antipode formula for the matroid-minor Hopf algebra. This approach can be also used to provide a cancellation-free antipode formula for Hopf algebras, defined by Eppolito, Jun, and Szczesny in [EJS17], arising from matroids over hyperfields as in Baker-Bowler [BB19].

In contrast to the case of the antipode formulas, primitive elements of combinatorial Hopf algebras seem to have received relatively less attention. In many interesting cases, combinatorial Hopf algebras are connected, commutative, and cocommutative. Hence, by the classical Milnor-Moore theorem, they are isomorphic to a polynomial algebra as Hopf algebras; more precisely, if \( B \) is a linear basis in the space of primitives of such a Hopf algebra \( H \), then \( H \) is the polynomial algebra in the elements of \( B \) (isomorphic to the polynomial algebra in \( |B| \) variables). While the original set of combinatorial objects used to build \( H \) is also often an algebraic monomial basis, those elements are not primitives, and this identification with a polynomial algebra is not a Hopf algebra isomorphism.

In particular, if one knows explicitly a basis of the space of primitive elements, one can translate some combinatorial questions to questions concerning polynomial algebras in an explicit way. For instance, in [Sch95], Schmitt studied invariants of combinatorial objects in this way. To this end, it is beneficial to know an explicit description of a basis of the space of primitive elements of a combinatorial Hopf algebra, which provides a natural isomorphism between combinatorial Hopf algebras and polynomial algebras. In the case of the matroid-minor Hopf algebra, Crapo and Schmitt found explicitly two bases of the space of primitive elements in [CS08] by introducing a new operation (free product) of matroids.

Hence, one can pose the following general problem, which asks to describe the Hopf structure of a combinatorial algebra in a meaningful way which relates to combinatorics:

**Problem.** Given some combinatorial Hopf algebra \( H \), constructed from a certain class of combinatorial objects and their operations, find the precise Hopf algebra structure of \( H \) by giving:

1. a “good” basis of the space of primitives, which relates to the original combinatorial basis in such a way that the coefficients of the algebraic base change have combinatorial significance;
2. a cancellation and grouping-free formula for the antipode in terms of the combinatorial basis.

In this paper, we solve this problem for the Hopf algebra of graphs \( H_G \), and explicitly describe its structure.

To explain the result, we recall now that this Hopf algebra is constructed as follows: it has as a formal basis (over some field) the set \( \mathcal{G} \) of isomorphism types of finite graphs. The product of \( H_\mathcal{G} \) is the disjoint union of graphs, and coproduct of a graph \( G \) is obtained by splitting \( G \) in all possible ways into two induced subgraphs (see section §2 for the precise definition and formula). For a graph \( G = G(V,E) \) and \( S \subseteq E \), we let \( G - S \) be the subgraph of \( G \) which has the same of vertices \( V = V(G) \) as \( G \), and whose set of edges is \( E - S \). The set of isomorphism types of connected graphs provides a polynomial basis of this algebra.

The key observation is the consideration of the following element in \( H_\mathcal{G} \), which we introduce here:

\[
P_G := \sum_{S \subseteq E} (-1)^{|S|} (G - S)
\]

We first prove the following.

**Theorem A.** (Propositions 3.2 and 3.3) Let \( G \) be a connected graph, then \( P_G \) is a primitive element of \( H_\mathcal{G} \). Moreover, if \( G = G_1 \cdot G_2 \cdots G_r \) (as an element of \( H_\mathcal{G} \)), then \( P_G = P_{G_1} \cdot P_{G_2} \cdots P_{G_r} \) (as an element of \( H_\mathcal{G} \)).

Next, we show that each graph \( G \) is a polynomial with non-negative integer coefficients of the elements \( P_{G - S} \):

\[^2\text{Note that the matroid-minor Hopf algebra is not cocommutative.}\]
**Theorem B.** (Proposition 3.7 and Theorem 3.9) Let $G$ be a graph. Then we have the following:

$$G = \sum_{S \subseteq E(G)} P_{G-S}.$$

In particular, if $\mathcal{G}$ is the set of isomorphism classes of connected graphs in $\mathcal{G}$, then the set $\{P_G\}_{G \in \mathcal{G}}$ forms a linear basis of the space of primitive elements of $H_\mathcal{G}$.

Finally, we utilize the above result to find a cancellation-free and grouping-free antipode formula for the graph Hopf algebra, as well as cancellation-free formulas for the transformations between the bases $\{G\}_{G \in \mathcal{G}}$ and $\{P_G\}_{G \in \mathcal{G}}$. They are given in terms of multiplicity numbers $|G : H|$, defined for every arbitrary pair of graphs $G, H$, to be the number of subgraphs of $G$ isomorphic to $H$. The following statement summarizes below all these results, given in Section §3.

**Theorem C.** For any graph $G$, the following hold.

$$ G = \sum_{H_1 \ldots H_i \text{ connected}} |G : H_1 \ldots H_i| P_{H_1} \ldots P_{H_i}, $$

$$ P_G = (-1)^{|E(G)|} \sum_{H_1 \ldots H_i \text{ connected}} (-1)^{|E(H_1)| + \ldots + |E(H_i)|} |G : H_1 \ldots H_i| H_1 \ldots H_i, $$

$$ S(G) = \sum_{E \subseteq E(G)} (-1)^{c_{G,E}} P_{G-E} = \sum_{H_1 \ldots H_i \text{ connected}} (-1)^i |G : H_1 \ldots H_i| H_1 \ldots H_i, $$

where $c_{G,E}$ is the number of connected components of the graph $G - E$.

We note that the above basis $\{P_G\}_{G \in \mathcal{G}}$ satisfies the requirements of the general problem posed above; in fact, we prove that $\{P_G\}_{G \in \mathcal{G}}$ satisfies certain minimality property and universal property among the bases satisfying two natural conditions.

We also recover some known facts for on the graph reconstruction conjectures using our framework. Finally, we note that our main ideas for finding such a bases, and consequently, cancellation-free formulas for the antipode, might very well be suitable to application in other combinatorial contexts.

This paper is organized as follows. In the interest of a wider audience, in §2, we review basic definitions that we will use throughout the paper; familiarity with Hopf algebra theory is not required. In §3, we introduce a new basis for the space of primitive elements of the graph Hopf algebra, and study basic properties, and give our main results. In §4, we explore some applications following from our explicit description of a primitive basis of the graph Hopf algebra.

**Acknowledgements:** J.J. received support from AMS-Simons travel grant. M.C.I was supported by the Simons Collaboration Grant 637866 for several research visits to J.J.

## 2. Preliminaries

In what follows, all graphs are assumed to be finite unless otherwise stated. For a graph $G$, we let $E(G)$ be the set of edges of $G$ and $V(G)$ the set of vertices of $G$. For each subset $E$ of $E(G)$, we let $G - E$ be the graph which we obtain from $G$ by deleting edges in $E$ while we keep the same vertex set. Finally, for each finite set $A$, we let $|A|$ be the cardinality of $A$.

Let $\mathcal{G}$ be a set of isomorphism classes of graphs, which is closed under taking induced subgraphs and disjoint union. One can naturally impose a (commutative) monoid structure on $\mathcal{G}$ as follows: for $[G_1], [G_2] \in \mathcal{G}$,

$$[G_1] \cdot [G_2] := [G_1 \sqcup G_2],$$

where $\sqcup$ is the disjoint union. Finally, for each finite set $A$, we let $A$ be the cardinality of $A$.
where \([G_i]\) is the isomorphism class of \(G_i\) and \(G_1 \sqcup G_2\) is the disjoint union of \(G_1\) and \(G_2\). In particular, the empty graph \([\emptyset]\) becomes the identity. In what follows, we will interchangeably use \([G]\) and \(G\) to denote the isomorphism class of \(G\) when there is no possible confusion.

Let \(k\) be a field of characteristic zero and \(k[G]\) be the monoid algebra obtained by considering \(G\) as a monoid as above. Clearly \(k[G]\) is graded by the number of vertices of graphs. The algebra \(k[G]\) has a Hopf algebra structure, which we recall here. Let \(E\) be a finite set and \(A\) be a subset of \(E\). For the notational convenience, we let \(A^c := E - A\). The coproduct \(\Delta : k[G] \to k[G] \otimes_k k[G]\) is defined for each \(G \in G\) by

\[
\Delta(G) := \sum_{T \subseteq V(G)} G_T \otimes G_T,
\]

where \(G_T\) is the induced subgraph (induced by \(T\)) of \(G\). The above formula is linearly extended to \(k[G]\). We further define the counit \(\varepsilon : k[G] \to k\) by letting it be defined on the basis \(G\) as

\[
\varepsilon(G) := \begin{cases} 
1, & \text{if } V(G) = \emptyset, \\
0, & \text{if } V(G) \neq \emptyset.
\end{cases}
\]

With the above coproduct \(\Delta\) and the counit \(\varepsilon\), \(k[G]\) becomes a connected, graded bialgebra and hence a Hopf algebra, which we will denote by \(H_G\). This Hopf algebra is cocommutative and commutative, and by classical Hopf algebra (for example, the more general Cartier-Kostant-Milnor-Moore theorem), it is isomorphic, as a Hopf algebra, to a polynomial algebra in (necessarily) countably many variables (see, for example also [GR14] or [Sch94]). In what follows, we aim to find special explicit bases for the primitives of \(k[G]\).

**Remark 2.1.** We note that formulas for the antipode of \(H_G\) have been obtained by many authors; see for instance Humpert and Martin [HM12], Benedetti and Sagan [BS17], and Aguiar and Ardila [AA17].

### 3. Primitive elements of the Hopf algebra of graphs

In this section, to each graph \(G\), we associate an element \(P_G\) of \(H_G\), and prove that \(\{P_G | G = \text{connected}\}\) is a basis of the space of primitive elements of \(H_G\).

We first rewrite the coproduct (1) of \(H_G\) in terms of partitions of vertices as follows:

\[
\Delta(G) = \sum_{V(G) = P_1 \sqcup P_2} G|_{P_1} \otimes G|_{P_2},
\]

where the sum runs over all partitions of \(V(G)\) and \(G|_{P_1}\) is the isomorphism class of the restriction of \(G\) to \(P_1\).

The following is the key definition of the paper.

**Definition 3.1.** Let \(G\) be a graph. Then we define the following element in \(H_G\):

\[
P(G) = P_G := \sum_{E \subseteq E(G)} (-1)^{|E|} (G - E)
\]

where \(G - E\) is the isomorphism type of the graph obtained from \(G\) by removing the edges in \(E\).

**Proposition 3.2.** Let \(G\) be the set of isomorphism classes of finite graphs, considered as a monoid (with the product given by the disjoint union). Then the function

\[
P : G \to H_G, \quad G \mapsto P_G
\]

is multiplicative. That is, \(P(G_1 \cdot G_2) = P(G_1) \cdot P(G_2)\).
Proof. Let $G = G_1 \sqcup G_2$ (the disjoint union of $G_1$ and $G_2$). For each subset $E$ of $E(G)$, we let $E_i = E \cap E(G_i)$ for $i = 1, 2$. Then one has

$$P(G) = P_G = \sum_{E \subseteq E(G)} (-1)^{|E|} (G - E) = \sum_{E_1 \sqcup E_2 \subseteq E(G)} (-1)^{|E_1| + |E_2|} (G_1 - E_1)(G_2 - E_2)$$

$$= \sum_{E_1 \subseteq E(G_1)} \sum_{E_2 \subseteq E(G_2)} (-1)^{|E_1|} (G_1 - E_1)(-1)^{|E_2|} (G_2 - E_2) = P_{G_1}P_{G_2} = P(G_1)P(G_2).$$

\[\square\]

It follows from Proposition 3.2 that if $G = G_1 \cdot G_2 \cdots G_k$ (as an element of $H_{\delta}$) for some connected graphs $G_i$, then

$$P_G = \prod_{i=1}^{k} P_{G_i}. \quad (4)$$

In fact, the following proposition shows that $P_G$ is a product of primitive elements.

**Proposition 3.3.** Let $G$ be a connected graph. Then $P_G$ (as an element of $H_{\delta}$) is a primitive element.

**Proof.** Since the coproduct is a linear map, with the notation as in (3), we have

$$\Delta(P_G) = \sum_{E \subseteq E(G)} (-1)^{|E|} \sum_{P_1 \sqcup P_2 = V(G - E)} (G - E)|_{P_1} \otimes (G - E)|_{P_2}. \quad (5)$$

For each $E \subseteq E(G)$ and a partition $V(G - E) = V(G) = P_1 \sqcup P_2$, we can write $E$ as a disjoint union

$$E = E_1 \sqcup E_2 \sqcup E',$$

where $E_1$ (resp. $E_2$) is the set of edges of $E$ appearing in $G|_{P_1}$ (resp. $G|_{P_2}$), and $E' := E - (E_1 \sqcup E_2)$.

Then, we can rewrite (5) as follows:

$$\Delta(P_G) = \sum_{E = E_1 \sqcup E_2 \sqcup E'} (-1)^{|E_1| + |E_2| + |E'|} \sum_{P_1 \sqcup P_2 = V(G)} (G - E)|_{P_1} \otimes (G - E)|_{P_2}$$

Since the set $E$ and partitions $P_1 \sqcup P_2 = V(G)$ range independently, this last sum can be in turn written as follows:

$$= \sum_{P_1 \sqcup P_2 = V(G)} \left( \sum_{E_1 \subseteq E(G|_{P_1})} \sum_{E_2 \subseteq E(G|_{P_2})} (-1)^{|E_1| + |E_2|} \sum_{E' \subseteq E(G - E(G|_{P_1}) - E(G|_{P_2})} (-1)^{|E'|} (G - E)|_{P_1} \otimes (G - E)|_{P_2} \right) \quad (6)$$

However, one can easily see that the sum $\sum_{E'}$ in (6) is independent of the terms $(G - E)|_{P_1} \otimes (G - E)|_{P_2}$. Moreover, since $G$ is connected, whenever both $P_1$, $P_2$ are non-empty, the set $E(G) - E(G|_{P_1}) - E(G|_{P_2})$ is non-empty, and then by the inclusion-exclusion we have that

$$\sum_{E' \subseteq E(G - E(G|_{P_1}) - E(G|_{P_2})} (-1)^{|E'|} = 0.$$

Therefore, the only terms remaining in the formula (6) are the ones where either $P_1$ or $P_2$ is empty, and we obtain

$$\Delta(P_G) = \sum_{E = E_1} (-1)^{|E|} (G - E)|_{P_1} \otimes \emptyset + \sum_{E = E_2} (-1)^{|E|} \emptyset \otimes (G - E)|_{P_2} = P_G \otimes \emptyset + \emptyset \otimes P_G.$$

This proves that $P_G$ is a primitive element. \[\square\]

Applying Möbius inversion, we can immediately prove that any $G$ (as an element of $H_{\delta}$) can be written as a linear combination of $\{P_{G - E} \}_{E \subseteq E(G)}$. 

5
Lemma 3.4. Let $G$ be a connected graph. Then, we have the following:

$$G = \sum_{S \subseteq E(G)} P_{G-S}.$$ 

Proof. We fix the graph $G$, and for notational convenience, we let $E = E(G)$. Let $Q = 2^E$. We consider $Q$ as a partially ordered set with the partial order is given by set-inclusion. Now, consider the incidence algebra structure associated to $Q$ with values in $H_\varphi$. In this case, the Möbius function is given, for each pair of subsets $S \subseteq T$ of $E$, as follows:

$$\mu(S,T) = (-1)^{|T-S|}.$$ 

For each subset $X \subseteq E$, we denote by $G(X)$ the subgraph of $G$ which has the same vertex set as $G$ and edge set $X$. Consider the following functions:

$$g : Q \to H_\varphi, \quad S \mapsto G(S),$$

$$f : Q \to H_\varphi, \quad S \mapsto P_{G(S)},$$

By the definition of $P$, we clearly have

$$f(S) = P_{G(S)} = \sum_{X \subseteq S} (-1)^{|X|} (G(S) - X) = \sum_{Y \subseteq S} (-1)^{|S-Y|} G(Y) = \sum_{Y \subseteq S} \mu(Y,S)G(Y)$$

for all $S \in Q$. It now follows from the Möbius inversion formula that

$$G(S) = \sum_{Y \subseteq S} \zeta(Y,S)f(Y) = \sum_{Y \subseteq S} P_{G(Y)} = \sum_{Y \subseteq S} P_{G(S-Y)}$$

where $\zeta$ is the zeta function of the incidence algebra ($\zeta(X,Y) = 1$ for all $X \subseteq Y$). In particular, we have that

$$G = \sum_{S \subseteq E(G)} P_{G-S}.$$ 

Example 3.5. Let $G = K_3$. Then we have,

$$P_G = G - 3 (\bullet \circ, \circ \bullet) + 3 (\bullet \bullet) - (\bigcirc ^3).$$

Example 3.6. Let $G = \bullet \circ \bullet$.

Then we have the following.

$$\Delta(\bullet \circ \bullet) = (\emptyset \otimes \bullet \circ \bullet) + 2(\bullet \otimes \bullet \circ \bullet) + (\bullet \circ \bullet \otimes \bullet) + 2(\bullet \bullet \otimes \bullet) + (\bullet \bullet \otimes \bullet) + (\bullet \circ \bullet \otimes \emptyset)$$

Let $G_1 = \bullet$ and $G_2 = \circ$. Then we have that

$$P_G = G - 2G_1G_2 + G_1^3.$$ 

Furthermore, we have that

$$\Delta(G) = \emptyset \otimes G + 2G_2 \otimes G_1 + G_2 \otimes G_2 + 2G_1 \otimes G_2 + G_2^2 \otimes G_2 + G \otimes \emptyset.$$ 

$$\Delta(G_1G_2) = \emptyset \otimes G_1G_2 + 2G_2 \otimes G_2^2 + G_2 \otimes G_1 + 2G_2^2 \otimes G_2 + G_1 \otimes G_2 + G_1G_2 \otimes \emptyset.$$ 

$$\Delta(G_2^3) = \emptyset \otimes G_2^3 + 3G_2 \otimes G_2^2 + 3G_2^2 \otimes G_2 + G_2^3 \otimes \emptyset.$$ 

Hence, we have

$$\Delta(P_G) = \Delta(G) - 2\Delta(G_1G_2) + \Delta(G_2^3)$$

$$= \emptyset \otimes P_G + P_G \otimes \emptyset.$$
showing that $P_G$ is a primitive element. Furthermore, in this case, we have

$$P_{G_1} = G_1 - G_2^2, \quad P_{G_2} = G_2.$$ 

It follows from Lemma 3.4 that

$$G = P_G + 2P_{G_1}P_{G_2} + P_{G_2}^3.$$ 

In fact, we have that

$$P_{G_1} + 2P_{G_1}P_{G_2} + P_{G_2}^3 = (G - 2G_1G_2 + G_2^2) + 2(G_1 - G_2^2)G_2 + G_2^3 = G - 2G_1G_2 + G_2^3 + 2G_1G_2 - 2G_2^3 + G_2^3 = G.$$

The next proposition shows that in fact Lemma 3.4 holds for any graph $G$.

**Proposition 3.7.** Let $G$ be a graph. Then we have the following:

$$G = \sum_{S \subseteq E(G)} P_{G-S}.$$ 

**Proof.** We can uniquely write $G = G_1 \cdots G_k$ for some connected subgraphs $G_i$. It follows from Lemma 3.4 that for each $G_i$, we have

$$G_i = \sum_{S \subseteq E(G_i)} P_{G_i-S}.$$ 

In particular, we have

$$G = G_1 \cdots G_k = \prod_{i=1}^k \left( \sum_{S \subseteq E(G_i)} P_{G_i-S} \right) = \sum_{S \subseteq E(G)} \left( \prod_{i=1}^n P_{G_i-S_i} \right), \quad (7)$$

where $S_i = S \cap E(G_i)$. Furthermore, from the definition of $P_G$ (for an arbitrary graph $G$), we have

$$\sum_{S \subseteq E(G)} \left( \prod_{i=1}^n P_{G_i-S_i} \right) = \sum_{S \subseteq E(G)} P_{G-S}. \quad (8)$$

**Corollary 3.8.** Let $G$ be a graph and $\mathcal{C}$ be the set of isomorphism classes of connected graphs. Then $G$ can be written (as an element of $H_G$) as a polynomial with coefficients in $\mathbb{N} = \{0, 1, 2, \cdots\}$ in elements of $\mathcal{C}$. To be precise, there exist connected graphs $G_1, \ldots, G_k$, and $n_i, a_i \in \mathbb{N}$ such that

$$G = \sum_{i=0}^k a_i P_{G_i}^{n_i}. \quad (9)$$

**Proof.** From Proposition 3.7 we have

$$G = \sum_{S \subseteq E(G)} P_{G-S}. \quad (9)$$

We label each $G-S$ as $H_j$ and rewrite:

$$G = \sum_{j=0}^r P_{H_j} = \sum_{j=0}^r \left( \prod_{i=0}^w P_{H_i} \right), \quad (10)$$

where $P_{H_i}$ are connected components of $P_{H_j}$. Now, regrouping terms appropriately in (10) gives the desired result. \hfill \Box

In particular, we find a basis of the vector space of primitive elements of $H_G$ in the following theorem.

**Theorem 3.9.** Let $\mathcal{C}$ be the set of isomorphism classes of connected graphs in $\mathcal{G}$. The set $\{P_G\}_{G \in \mathcal{C}}$ forms a basis of the space of primitive elements of $H_G$. 

7
In particular, since each $P_{G_i}$ is linearly independent. Suppose that we have
\[ a_1P_{G_1} + \cdots + a_nP_{G_n} = 0. \] (11)
We may assume that any two graphs $G_i$ and $G_j$ are not isomorphic. Suppose that $a_i \neq 0$ for some $a_i$.

From Definition 3.1, for each $i$, we have that

\[ a_iP_{G_i} = \sum_{E \subseteq \varepsilon(G_i)} (-1)^{|E|} (G_i - E). \]

Consider the grading of $H_\varepsilon$ by the number of edges of a graph. Substituting $P_{G_i}$ above into equation (11), we see that the highest degree of the left hand side of that equality is of the form $\sum_{i \in F} a_iP_{G_i}$ for some non-empty subset $F$ of $\{1, \ldots, n\}$. For that to equal zero, we must have two indices $i$ and $j$ such that $G_i = G_j$. But, from Lemma 3.4, this implies that $P_{G_i} = P_{G_j}$ giving a contradiction. Therefore $\{P_G\}_{G \in \varepsilon}$ is linearly independent.

Finally, we note that $\{P_G\}_{G \in \varepsilon}$ spans $V$. From Lemma 3.4, we have that the elements $P_G$ generate $H_\varepsilon$; but a set of primitives which generate the whole Hopf algebra must span the space of primitives. (if $p \in V$, then write $p = \sum_{i=0}^k a_i P_{G_i}$, for some elements $G_i \in \varepsilon$ and $a_i, n_i \in \mathbb{N}$; but $\sum_{i=0}^k a_i P_{G_i}$ is primitive, only when $n_i = 1$ for all $i$).

4. Applications

In this section, we list three applications of our explicit description of the basis $\{P_G\}_{G \in \varepsilon}$. In the first subsection, we provide the cancellation and grouping-free antipode formula for the graph Hopf algebra. We also provide cancellation-free formulas for the linear transformations between the basis $\{G\}_{G \in \varepsilon}$ of $H_\varepsilon$ (as an algebra) and $\{P_G\}_{G \in \varepsilon}$. In the second subsection, we prove that our basis $\{P_G\}_{G \in \varepsilon}$ is the minimal basis in a suitable way. We further prove that $\{P_G\}_{G \in \varepsilon}$ satisfies a certain universal property. Finally, in the third subsection, we list some applications to the graph reconstruction conjectures.

Cancellation and grouping-free formulas. We first use our previous results to derive the antipode formula for $H_\varepsilon$. Let $G$ be a graph and $E$ be a subset of $E(G)$. We define $c_{G,E}$ to be the number of connected components of the graph $G - E$.

**Proposition 4.1.** The antipode $S(G)$ of $G \in H_\varepsilon$ can be computed as follows:

\[ S(G) = \sum_{E \subseteq E(G)} (-1)^{c_{G,E}} P_{G - E} \] (12)

**Proof.** From Proposition 3.7, we have that $G = \sum_{S \subseteq E(G)} P_{G - S}$, and by taking the antipode map $S$, we have that

\[ S(G) = S(\sum_{S \subseteq E(G)} P_{G - S}) = \sum_{E \subseteq E(G)} S(P_{G - E}). \] (13)

For each graph $G - E$, we can uniquely write $G - E = G_1 \cdots G_n$ for $n = c_{G,E}$. Then we have,

\[ P_{G - E} = P_{G_1} \cdots P_{G_n}. \]

In particular, since each $P_{G_i}$ is primitive by Proposition 3.3,

\[ S(P_{G - E}) = S(P_{G_1} \cdots P_{G_n}) = S(P_{G_1}) \cdots S(P_{G_n}) = (-P_{G_1}) \cdots (-P_{G_n}) = (-1)^{c_{G,E}} P_{G - E}. \] (14)

Therefore, we have that

\[ (13) = \sum_{E \subseteq E(G)} (-1)^{c_{G,E}} P_{G - E}. \]
Example 4.2. Let $G$, $G_1$, and $G_2$ be the graphs as in Example 3.6. Then, following our formula, we have that

$$S(G) = 2P_{G_1 G_2} - P_{G_1 G_2} - P_G = 2(G_1 \cdot G_2 - G_2) - (G_1 - G_1) = 4G_1 \cdot G_2 - G_1^2 - 3G_2.$$

As Example 4.2 shows our antipode formula is not cancellation-free. Similarly, our formulas expressing each graph as a linear combination of primitives are also not cancellation-free. However, we can repackage these to make them cancellation-free. To this end, we need to introduce some notation. Let $G$ and $H$ be graphs. We let $n(G : H)$ be the number of injective graph morphisms from $H$ to $G$, divided by the cardinality of $\text{Aut}(H)$. Equivalently, $n(G : H)$ is the number of non-equivalent embeddings of $H$ into $G$, that is, the number of subgraphs of $G$ which are isomorphic to $H$.

For instance, let $G, G_1, G_2$ be the graphs as in Example 3.6, then we have $n(G : G_2) = 3$. Similarly, $n(G : G_1) = 2$.

Given a graph $G$ and connected graphs $H_1, H_2, \ldots, H_t$, let us denote $[G : H_1 \ldots H_t] = n(G : H_1 \sqcup \cdots \sqcup H_t)$ (the numbers of subgraphs of $G$ isomorphic to $H_1 \ldots H_t = H_1 \sqcup \cdots \sqcup H_t$). We can now re-write the formulas for $P_G$ as well as the formula giving $G$ as a polynomial in the primitives $P_H$ as follows

$$P_G = \sum_{E \subseteq E(G)} (-1)^{|E|} (G - E)$$

(15)

$$= (-1)^{|E(G)|} \sum_{H_1, H_2, \ldots, H_t \text{ connected}} (-1)^{|E(H_1)| + \cdots + |E(H_t)|} [G : H_1 \ldots H_t] H_1 \cdots H_t$$

(16)

where the sum ranges over all $t$-tuples of connected graphs $H_1, \ldots, H_t$ such that the graph $H_1 \sqcup \cdots \sqcup H_t$ has the same number of vertices as $G$. Similarly, we obtain

$$G = \sum_{E \subseteq E(G)} P_{G - E} = \sum_{H_1, H_2, \ldots, H_t \text{ connected}} [G : H_1 \ldots H_t] P_{H_1 \ldots H_t}$$

(17)

$$= \sum_{H_1, H_2, \ldots, H_t \text{ connected}} [G : H_1 \ldots H_t] P_{H_1} \cdots P_{H_t}$$

(18)

Now, we deduce the cancellation-free formula for the antipode; using equations (12) and (18), we have

$$S(G) = \sum_{E \subseteq E(G)} (-1)^{|E|} P_{G - E} = \sum_{H_1, H_2, \ldots, H_t \text{ connected}} (-1)^{|G : H_1 \ldots H_t|} [G : H_1 \ldots H_t] H_1 \cdots H_t$$

We make the remark that this formula expresses the antipode of any graph as a polynomial of the algebraic basis given by graphs, and such a formula is obviously unique (the coefficients are uniquely determined, since this is a polynomial algebra in the $G$’s). Thus, one can regard the multiplicities $[G : H] = [G : H_1 \ldots H_t]$ as having a special meaning in the Hopf algebra of graphs.

**Minimal and Universal Properties.** We now prove that the above defined basis of the space of primitives $\{P_G\}_{G \in \mathcal{G}}$ (where $\mathcal{G}$ is the set of isomorphism types connected graphs) satisfies certain minimality property and uniqueness property.

For two graphs $G, H$ we will write $H \leq G$ if $H$ is a subgraph of $G$, equivalently, $[G : H] > 0$, and $H < G$ if $H \leq G$ but $H \neq G$. Also, for graphs $H \leq G$, we let $G_H$ be the graph with the vertex set $V(G)$ and the edge set $E(H)$. Finally, we note that with respect to the grading on $H_\emptyset$ by the number of vertices, the elements $P_G$ are homogeneous of degree $|V(G)|$.

**Definition 4.3.** Let $q = \{q_G\}_{G \in \mathcal{G}}$ be a basis of the space of primitives of $H_\emptyset$, considered as an algebra over a field of characteristic zero. $q = \{q_G\}_{G \in \mathcal{G}}$ is said to be an integral basis if for each graph
Let \( G = H_1 \cdots H_t \) with connected components \( H_i \), and \( q_G := q_{H_1} \cdots q_{H_t} \), the following two conditions hold:

1. Each \( G \in \mathcal{G} \) can be expressed as \( G = f_G(q_H \mid H \leq G) \), where \( f_G \) is a polynomial with non-negative integers.
2. \( \mathbb{Z}[q_G]_{G \in \mathcal{G}} \) is a subset of \( \mathbb{Z}[G]_{G \in \mathcal{G}} \).

**Proposition 4.4.** Let \( q = \{ q_G \}_{G \in \mathcal{G}} \) be an integral basis. Then, for each graph \( G \), there exist integers \( \alpha_{G,H} \) such that

\[
G = \sum_{H \leq G} \alpha_{G,H} q_H. \tag{19}
\]

**Proof.** Note that the condition (1) of an integral basis for \( G \in \mathcal{G} \) implies the same for all \( G \in \mathcal{G} \). Hence, we may assume that \( G \) is connected. We now prove this by induction on the number of edges \( e = |E(G)| \). For \( e = 0 \) it is obvious. For the induction step, consider the equality

\[
P_G + \sum_{H < G} P_H = G = f_G(q_H \mid H \leq G).
\]

Since for each \( H < G \), we have \( P_H \in \text{Span}_\mathbb{Z}\{q_L \mid L < G\} \subseteq \text{Span}_\mathbb{Z}\{q_L \mid L < G\} \) (the second inclusion follows from the induction hypothesis), we obtain

\[
P_G + \sum_{H < G} \beta_{G,H} q_H = f_G(q_H \mid H \leq G), \quad \text{for some } \beta_{G,H} \in \mathbb{Z}, \tag{20}
\]

and hence \( P_G \) is a polynomial with integer coefficients in the elements \( q_L \) for \( L \leq G \). But since \( P_G \) is primitive, it follows that \( P_G \) is a linear combination of elements \( q_L \) for \( L \) connected, at least one of which must be \( q_G \). Thus, \( P_G \in \gamma q_G + \text{Span}_\mathbb{Z}\{q_H \mid H \leq G\} \) for some integer \( \gamma 
eq 0 \), and now replacing this in (20) gives the desired inductive result. \( \square \)

**Definition 4.5.** An integral basis \( q = \{ q_G \}_{G \in \mathcal{G}} \) is said to be **combinatorial** if for each graph \( G \), there exist positive integers \( \gamma_{G,E} \) for all \( E \subseteq E(G) \) such that

\[
G = \sum_{E \subseteq E(G)} \gamma_{G,E} q_{G - E}. \tag{21}
\]

**Remark 4.6.** Among integral bases, there are ones that are “combinatorial” in the sense that the formula (19) can be written as (21). This is equivalent to saying that all the \( \alpha_{G,H} \) coefficients from \( G = \sum_{H \leq G} \alpha_{G,H} q_H \) satisfy the following inequality:

\[
\alpha_{G,H} \geq [G:H]. \tag{22}
\]

For instance, the basis \( \{ P_G \}_{G \in \mathcal{G}} \) trivially satisfies this condition. In fact, this motivated us to use the term “combinatorial” since in the case of \( \{ P_G \}_{G \in \mathcal{G}} \), the coefficient \( \alpha_{G,H} \) is the number of subgraphs of \( G \) which are isomorphic to \( H \).

Let \( X \) be the set of combinatorial integral bases of the space of primitive elements of \( H_{\mathcal{G}} \). For two elements \( q = \{ q_G \} \) and \( r = \{ r_G \} \) of \( X \), we denote \( q \leq r \) if for any graph \( G \) and two expressions,

\[
G = \sum_{H \leq G} \gamma_{G,H} q_H = \sum_{H \leq G} \beta_{G,H} r_H,
\]

we have that \( \gamma_{G,H} \leq \beta_{G,H} \) for all \( H \leq G \). One can easily see that this relation defines a partial order on the set \( X \); antisymmetry is the only nontrivial part which can be checked inductively. Then, clearly the basis \( \{ P_G \}_{G \in \mathcal{G}} \) is the minimal element in \( X \) because of (22). Hence, we have the following.

**Proposition 4.7.** The basis \( \{ P_G \}_{G \in \mathcal{G}} \) is the minimal element in the set \( X \) of combinatorial integral bases of the space of primitive elements of \( H_{\mathcal{G}} \) with the aforementioned partial order.

Furthermore \( \{ P_G \}_{G \in \mathcal{G}} \) also satisfies a certain universal property as follows.
Proposition 4.8. Let \( \{q_G\} \) be a combinatorial integral basis. Then, \( \{P_G\}_{G \in \mathcal{C}} \) can be written as a linear combination of \( \{q_G\} \) with non-negative coefficients.

Proof. We prove the following equation by induction on the number of edges \( e = E(G) \).

\[
P_G = q_G + \sum_{H < G} b_{G,H} q_H, \quad \text{for some} \ b_{G,H} \in \mathbb{N}, \tag{23}
\]

When \( e = 0 \), it is clear, since the one point graph \( G = \cdot \) is primitive and we must have \( q_G = P_G \). Now, for the induction step, consider the following equation:

\[
G = P_G + \sum_{H < G} [G : H] P_H. \tag{24}
\]

We can replace each \( P_H \) with \( \{q_G\} \) from (23). In particular, we obtain

\[
G = q_G + \sum_{H < G} c_{G,H} P_H, \tag{25}
\]

for some integers \( c_{G,H} \). We note that \( c_{G,H} \geq [G : H] \) since the formula (23) has the term \( q_G \). Finally, then we have

\[
q_G + \sum_{H < G} c_{G,H} P_H = G = P_G + \sum_{H < G} [G : H] P_H,
\]

showing that

\[
P_G = q_G + \sum_{H < G} (c_{G,H} - [G : H]) P_H.
\]

Since \( (c_{G,H} - [G : H]) \) are non-negative integers, this proves our proposition. \( \square \)

The reconstruction conjectures. Let \( G \) be a graph. The vertex deck of the \( G \) is defined to be the multiset of (isomorphism type of) graphs obtained by removing some subset \( X \subseteq V(G) \) and taking the induced graph of \( G \) graph \( V(G) - X \). The edge deck of \( G \) is defined to be the multiset of (isomorphism type of) graphs \( (G - E)_{E \subseteq E(G)} \). The following two conjectures are well known in combinatorics:

Conjecture [Vertex Reconstruction Conjecture] If two graphs have the same vertex decks, then they are isomorphic.

Conjecture [Edge Reconstruction Conjecture] If two graphs have the same edge decks, then they are isomorphic.

The vertex reconstruction conjecture is known to hold for some classes of graphs, such as trees. It is also well-known that the vertex reconstruction conjecture implies the edge reconstruction conjecture (see, for instance, \[Hem69\]). We note now that this also follows as a consequence of our algebraic setup:

Corollary 4.9. The vertex reconstruction conjecture implies the edge reconstruction conjecture.

Proof. Let \( G, G' \) be two graphs. If \( G \) and \( G' \) have the same edge deck, then the cancellation-free expression of a graph in (18) shows that \( G - G' = P_G - P_{G'} \). The same formula now implies that \( G \) and \( G' \) have the same vertex deck. Indeed, if not, there exists a graph \( H = G - \{v\} \) for some vertex \( v \) of \( G \) such that \( H \) does not appear in the vertex deck of \( G' \). But then the graph \( L = H \cup \{v\} \) has the property that \( [G : L] > 0 \) and \( [G' : L] = 0 \), which means that in (18) the term \( [G : L]L \) does not cancel, a contradiction. \( \square \)

We can also prove the following case of the vertex reconstruction conjecture by using the Hopf algebras of graphs.

Proposition 4.10. The vertex reconstruction conjecture is true for disconnected graphs.
Proof. Let $G = G_1 \cdots G_k$, where $G_i$ are connected components of $G$. Suppose that $H = H_1 \cdots H_t$ is a graph with the same vertex deck as $G$. Let $G_i = f_i(P)$, a polynomial of the primitives $P = \{P_L\}_L$ as in Corollary 3.8. Similarly, we write $H_i = g_i(P)$. As $G$ and $H$ have the same vertex deck, we know that $G - H$ is a primitive element; but $G - H = f_1(P) \cdots f_k(P) - g_1(P) \cdots g_t(P)$, and as $G$ and $H$ are disconnected, neither of $f_1(P) \cdots f_k(P)$ or $g_1(P) \cdots g_t(P)$ can have a monomial of degree 1 (nor a constant term), showing that $G - H = 0$, and hence $G = H$. \qed

References

[AA17] Marcelo Aguiar and Federico Ardila. Hopf monoids and generalized permutahedra. arXiv preprint arXiv:1709.07504, 2017.

[ABS06] Marcelo Aguiar, Nantel Bergeron, and Frank Sottile. Combinatorial Hopf algebras and generalized Dehn–Sommerville relations. Compositio Mathematica, 142(1):1–30, 2006.

[BB19] Matthew Baker and Nathan Bowler. Matroids over partial hyperstructures. Advances in Mathematics, 343:821–863, 2019.

[BEJM18] Eric Bucher, Chris Eppolito, Jaiung Jun, and Jacob P. Matherne. Matroid-minor hopf algebra: a cancellation-free antipode formula and other applications of sign-reversing involutions. arXiv preprint arXiv:1811.01687, 2018.

[BS17] Carolina Benedetti and Bruce E Sagan. Antipodes and involutions. Journal of Combinatorial Theory, Series A, 148:275–315, 2017.

[CS05] Henry Crapo and William Schmitt. The free product of matroids. European Journal of Combinatorics, 26(7):1060–1065, 2005.

[CS08] Henry Crapo and William Schmitt. Primitive elements in the matroid-minor Hopf algebra. Journal of Algebraic Combinatorics, 28(1):43–64, 2008.

[EJS17] Christopher Eppolito, Jaiung Jun, and Matt Szczesny. Hopf algebras for matroids over hyperfields. arXiv preprint arXiv:1712.08903, 2017.

[EJS18] Christopher Eppolito, Jaiung Jun, and Matt Szczesny. Proto-exact categories of matroids, Hall algebras, and K-theory. arXiv preprint arXiv:1805.02281, to appear in Math. Z., 2018.

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

[Hem69] Robert L. Hemminger. On reconstructing a graph. Proceedings of the American Mathematical Society, 20(1):185–187, 1969.

[HM12] Brandon Humpert and Jeremy L Martin. The incidence Hopf algebra of graphs. SIAM Journal on Discrete Mathematics, 26(2):555–570, 2012.

[LR10] Jean-Louis Loday and Maria Ronco. Combinatorial Hopf algebras. Quantum of Maths, pages 347–383, 2010.

[Sch94] William R Schmitt. Incidence Hopf algebras. Journal of Pure and Applied Algebra, 96(3):299–330, 1994.

[Sch95] William R Schmitt. Hopf algebra methods in graph theory. Journal of Pure and Applied Algebra, 101(1):77–90, 1995.

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